

The geometry of the occluding contour and its effect on motion interpretation

Josh McDermott

Department of Brain and Cognitive Science,
MIT, Cambridge, MA, USA



Edward H. Adelson

Department of Brain and Cognitive Science,
MIT, Cambridge, MA, USA



Form information related to occlusion is needed to correctly interpret image motion. This work describes one of a series of investigations into the form constraints on motion perception. In the present study, we focus specifically on the geometry of the occluding contour, and in particular on whether its influence on motion can be accounted for merely by its effect on perceived occlusion. We used an occluded square moving in a circle, holding the T-junctions at points of occlusion constant while manipulating the occluding contour. We found evidence for two main influences of occluding contour geometry on motion interpretation and occlusion: the convexity of the occluding contour and additional static T-junctions that are formed elsewhere on the occluding contour. Our results suggest that convex occluding contours are more occlusive than concave ones, and that T-junctions along the contour increase or decrease the strength of occlusion depending on their orientation. Motion interpretation is influenced by both factors, but their effect on motion appears to be dominated by interactions occurring at an intermediate "semilocal" scale, which is larger than the scale at which junctions are defined, but smaller than the scale of the whole moving figure. We propose that these computations are related to occlusion but are not identical to the computations that mediate static occlusion judgments.

Keywords: motion, aperture problem, occlusion, junction, convexity

Introduction

The aperture problem is the well-known geometric ambiguity that results from sampling a moving edge through a local aperture such as a receptive field. As shown in [Figure 1a](#), the image motion of an edge constrains its motion in the world to a line in velocity space, but does not narrow it down to a single velocity (Wallach, 1935; Adelson & Movshon, 1982). Local motion measurements thus do not fully specify the direction that objects in the world are moving, and it is necessary to combine measurements across space.

One approach involves using the unambiguous motion of two-dimensional (2D) features, such as a corner of one of the diamonds in [Figure 1a](#) (Wallach, 1935; Nakayama & Silverman, 1988). Some 2D features, however, are the spurious products of occlusion (e.g., the T-junctions in [Figure 1a](#)). Such features must be discounted to avoid faulty motion estimates, and in human vision they apparently are, as we rarely if ever mistake the motion at points of occlusion for object motion. Distinguishing spurious features from real ones appears to necessitate the use of form information, because the motion generated by such features does not in itself distinguish them.

A second approach to the aperture problem involves integrating motion information from multiple edges. Although individual moving edges provide ambiguous motion information, they do narrow down the range of possible velocities to a line in velocity space. Multiple edges pro-

duce multiple constraint lines, and their intersection can yield a single unambiguous velocity (Adelson & Movshon, 1982). In scenes with multiple objects, however, such an approach cannot be applied blindly - integration will produce the correct object velocities only if the motions that are integrated arise from the same object. As shown in [Figure 1b](#), if edge motions from two diamonds moving in opposite horizontal directions are combined, the resulting intersection of constraints is in an erroneous vertical direction. Thus prior to integrating motion, the visual system must segregate local motion measurements into groups that are likely to be due to the same object. This seems to necessitate form information as well, because in the motion domain it is not obvious which local motions belong together.

Attempting to solve the aperture problem by integrating motion across space thus results in two further problems, both of which seem to require the use of form information. Numerous motion illusions confirm the importance of form constraints. Consider, for instance, the square stimulus introduced by Lorenceau and Shiffrar (1992), shown in [Figure 2](#). The outline of a square translates in a circle, its corners hidden by occluders. The only moving features are the T-junctions that occur where the occluders overlap the square; these oscillate sinusoidally in the direction normal to the orientation of the bars composing the square. Despite the fact that no local feature is moving in a circle, observers generally report seeing the coherent circular motion of the square rather than the sinusoidal motions of the bar endpoints, indicating that the

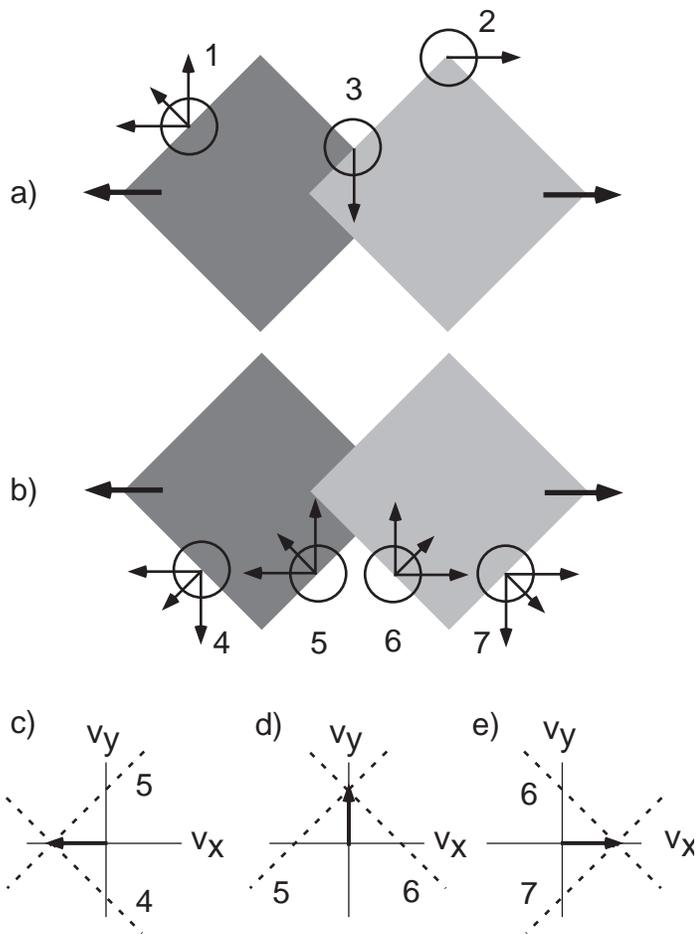


Figure 1. Example illustrating two problems that occur in motion interpretation, both of which implicate the use of form information. In (a) and (b), two squares translate horizontally. The edge motions (e.g., 1) are ambiguous, whereas the corner motions (e.g., 2) are unambiguous. The T-junction motions (e.g., 3) are also unambiguous, but their vertical motion is spurious (no object is moving vertically) and must somehow be discounted. Integration also poses a problem: (c), (d), and (e) show the velocity-space representations of the motion constraints provided by edges 4 and 5, 5 and 6, and 6 and 7, respectively. If the motion constraints from two edges of the same object are combined via intersection of constraints, as in (c) and (e), the correct horizontal motions result. If, however, motion constraints from edges of different objects are combined, as in (d), an erroneous upward motion is obtained. Click link to view [demo](#).

T-junctions are discounted and the edge motions integrated to yield the circular motion. When the occluders are removed, however, as in the stimulus of Figure 2b, the percept is quite different. The stimulus breaks up into separate motions, with each bar appearing to move sinusoidally in the direction of its endpoints. The motion that is perceived seems to depend on the presence of the occluders. Of course, this makes sense; for there to be a square executing a circular trajectory, something must hide the corners, and the presence of visible occluders in the image obviously

makes this scenario in the world more likely. But how do the occluders exert their effects? Somehow, form information is extracted from the occluders and used to interpret the image motion. By manipulating these sorts of stimuli, we can study the form computations that are involved.

Given that motion interpretation has often been thought to be mostly independent of form analysis, any form constraints on motion might be assumed to be simple in nature. One simple explanation of many form and motion phenomena is that T-junctions are detected and their motions simply ignored in the process of motion interpretation (e.g., Nowlan & Sejnowski, 1995). Most previous work on these issues is consistent with this sort of a theory (Anstis, 1990; Stoner, Albright, & Ramachandran, 1990; Vallortigara & Bressan, 1991; Lorenceau & Shiffrar, 1992; Bressan, Ganis, & Vallortigara, 1993; Trueswell & Hayhoe, 1993; Shiffrar, Li, & Lorenceau, 1995; Lindsay & Todd, 1996; Shiffrar & Lorenceau, 1996; Castet & Wuerger, 1997; Liden & Mingolla, 1998; Stoner & Albright, 1998; Rubin, 2001; see also Anderson & Sinha, 1997; Lorenceau, 1999). Indeed, a simple junction-based account can readily explain the effects of adding occluders to the square stimulus of Figure 2. Because the motion of the bar endpoints is the only thing inconsistent with a single coherent motion, if the end points are ignored when they form T-junctions with the occluders, coherence could plausibly become the preferred interpretation. Form constraints based on T-junctions can thus account for the basic effect of occluders on the square, but are junctions in fact driving the effect?

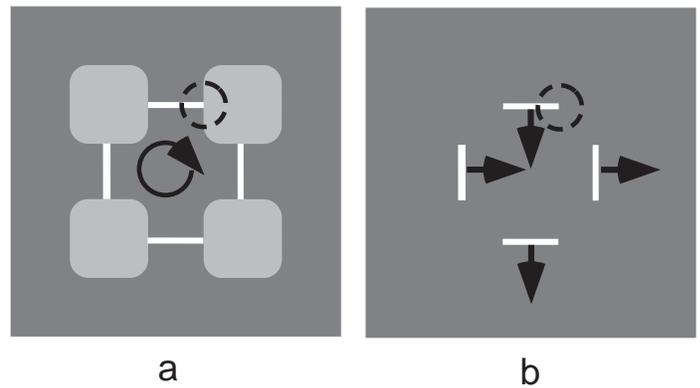


Figure 2. The basic square stimulus, generated by moving a square in a circle behind occluders, which can either be visible (a) or invisible (b). Click link to view [demo](#).

We have tested the importance of junctions with stimuli such as those in Figure 3 (McDermott, Weiss, & Adelson, 2001). The stimuli of Figure 3a and 3b have identical junctions at the bar endpoints, but differ globally in the extent to which the bars appear to be occluded. If T-junctions play a dominant role in the form constraints governing motion interpretation, the two stimuli should cohere to similar extents. As we have reported elsewhere (McDermott et al., 2001), we find that observers report the

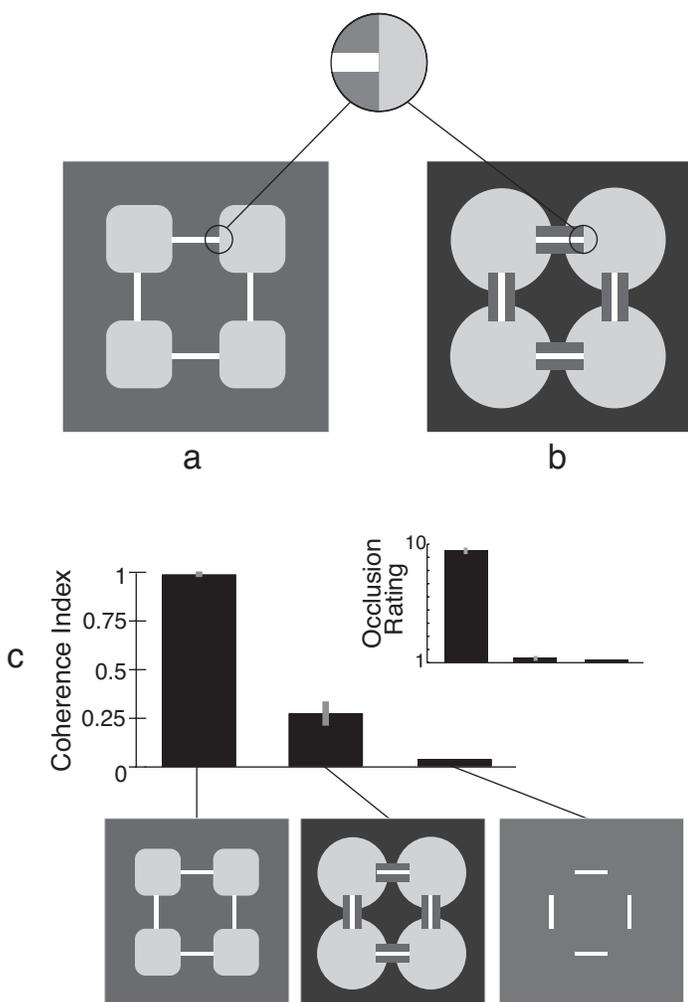


Figure 3. Stimuli and results of Experiment 1. (a) and (b). Experimental stimuli. Stimuli are identical in the local vicinity of the square contours but differ globally in the extent to which the contours look occluded. (c). Observed coherence levels (for six naïve subjects) and occlusion ratings for each stimulus. Error bars in this and all other graphs denote standard errors. Click link to view [demo](#).

second stimulus to be far less coherent than the first, consistent with the weaker impression of occlusion that it conveys (Figure 3c). The second stimulus is still more coherent than the bars alone, indicating that the T-junctions may be doing something. But the T-junctions alone do a poor job of predicting motion interpretation; evidently more complex and nonlocal constraints are at work. The remainder of this work is devoted to exploring the nature of these constraints.

We focused on the occlusion cues provided by the occluding contour, and sought to characterize the effect of various cues on perceived motion and on perceived occlusion. We were particularly interested in whether the effects of occlusion cues on motion could just be due to their effect on perceived occlusion (i.e., if anything that affected perceived occlusion would also affect perceived motion in

the expected manner, and vice versa). Accordingly, we measured perceived motion and perceived occlusion for a variety of displays.

Methods

Naïve subjects participated in all experiments. All had normal or corrected-to-normal vision. Stimuli were presented on a Hitachi monitor controlled by a Silicon Graphics Indy R4400. Viewing distance was approximately 95 cm. Subjects were instructed to freely view the experimental stimuli while confining their gaze to the central region of the display. This policy was adopted because our untrained subjects found it unnatural and difficult to maintain fixation while attending to the moving bars. Informal observation by the authors suggests that maintaining fixation would not have qualitatively changed any of the effects.

In all our experiments, observers were shown short (3 s) clips of each stimulus, and were asked to judge whether it looked coherent, incoherent, or somewhere in between, which they indicated by pressing 1, 2, or 3, respectively, on the keyboard number pad following each trial. Coherence judgments were used instead of the more objective direction of rotation judgments used in several previous studies (e.g., Lorenceau & Shiffrar, 1992) because pilot experiments revealed that some subjects could learn to perform the rotation judgments even for conditions that appeared entirely incoherent (these subjects were presumably learning to discriminate the phase relationships between the bars rather than integrating the bar motions). For such subjects, judgments of rotation direction are clearly not a suitable measure of motion integration. The use of three choices in describing perceived coherence is generous; most previous studies have used only two (e.g., Adelson & Movshon, 1982). In practice our subjects rarely used the intermediate response choice. Subjects' responses were normalized to yield a coherence index ranging from 0 to 1. A coherence index of 0 corresponds to a percept of completely incoherent motion on every single trial, whereas 1 indicates consistently coherent motion. Subjects completed several practice trials before beginning the experimental trials.

In all experiments we plot data averaged across subjects, for the sake of clarity. Data from individual subjects were qualitatively similar, though, and the qualitative patterns of most of the results that we report have been confirmed in many observers during conference presentations. Many of the effects are large enough that they can be confirmed informally in demos, such as those we have available online: <http://web.mit.edu/persci/demos/Motion&Form/master.html>.

Although we find that the ordinal relationships between coherence levels for different displays are almost always the same across subjects, the overall degree of coherence can vary substantially from subject to subject, which can occasionally result in ceiling and floor effects if the stimuli are not adjusted. In all our experiments, the con-

trast of the moving bars was adjusted for each subject in an effort to avoid ceiling and floor effects. The coherence of this stimulus tends to decrease as the bar contrast is increased (Lorenceau & Shiffrar, 1992), and so by changing the contrast we could partially shift overall coherence levels up or down. In separate experiments (unpublished), we have found that the effect of contrast does not interact with the effect of the different stimulus configurations explored in this work, making it a suitable variable to manipulate for such purposes. However, because the contrast was adjusted separately for each experiment and because the subjects in each experiment were not identical, coherence levels cannot be compared across experiments.

To investigate the qualitative relationship between perceived occlusion and motion interpretation, we conducted a separate set of experiments in which subjects rated perceived occlusion in our stimuli. Subjects were asked to view static versions of the stimuli and judge the extent to which the bars of the square appeared to be occluded, rating each stimulus on a scale of 1-10. We were surprised to find that subjects were quite comfortable making these judgments, and that the ratings were quite consistent from subject to subject. As for the coherence ratings, we plot the results averaged across subjects. Note that all the occlusion ratings were collected in a single experimental session, so the ratings can be compared across figures in the paper. The subjects for the occlusion experiments were distinct from those for the motion experiments.

The experimental parameters for the initial experiment of Figure 3 are as follows (many remain the same in the other experiments). The background luminance of the stimulus of Figure 3a was 9.4 cd/m^2 (this was also the luminance of the rectangles in Figure 3b to keep the junctions identical); the luminance of the occluders of Figure 3a (and of the circles of Figure 3b) was 30.1 cd/m^2 ; the background luminance of Figure 3b was 2.4 cd/m^2 . The Michelson contrast of the bars was set individually for each subject to help avoid ceiling and floor effects, but was always between 0.4 and .75. The speed of the square was 1.67 deg/s , the range of motion was 0.25 deg , and the stimulus was displayed for 2 s on each trial. The length of the moving bars was 38 pixels (0.6 deg). The width of the T-junction that was held constant across stimuli was 25 pixels (0.4 deg). Note that only the bars moved in the stimuli; the rest of each stimulus was static. The bars approached the borders of the rectangles at the extremes of their trajectories, but never touched. All six subjects, who were naive to the purposes of the experiment, completed 15 trials per condition in a single block. This block included 5 additional conditions, some of which are reported in Experiments 4 and 10.

The parameters of the experiments that follow were identical to those of this experiment unless otherwise mentioned.

Results

Convexity

Apertures and occluders

The stimuli of Figure 3 differ in a number of ways, but one obvious difference stems from the geometry of the occluding contour. Note that in the stimulus of Figure 3a, the occluding contour abutting each moving bar is convex, whereas in Figure 3b, it is concave. Contour convexity is a well-known cue to border ownership (Stevens & Brooks, 1988; Pao, Geiger, & Rubin, 1999), so it seemed possible that this might play a role in the form constraints governing motion perception. As a first test of the importance of convexity, we compared the coherence obtained for the occluded square with that for an identical square viewed through apertures with the same occluding contours as the occluders, as shown in Figure 4. Six naïve subjects completed 15 trials in a single block that included the conditions of Experiment 1.

As shown in Figure 4, we found that the apertures produced substantially lower levels of coherence than did the occluders. The occlusion ratings mirror those for coher-

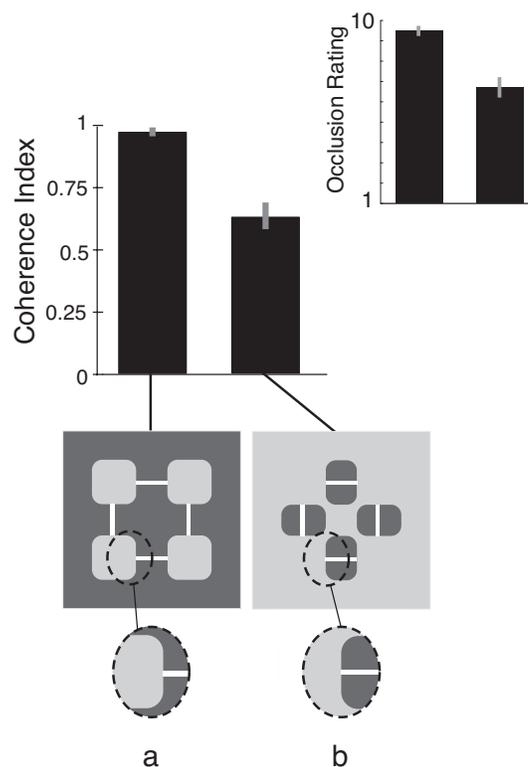


Figure 4. Stimuli and results for Experiment 2. (a). With occluders the square is highly coherent. (b). Apertures with the same occluding contour produce lower coherence and occlusion ratings, perhaps because the occluding contour is concave. Click link to view [demo](#).

ence, consistent with the notion that convexity influences both the strength of border ownership and the strength of motion integration.

Aperture shape

In two other experiments we altered the nature of the contour concavity by parametrically varying the width and curvature of the apertures, as shown in Figure 5. All five naïve subjects completed 15 trials per condition in a single block which included the conditions for both experiments.

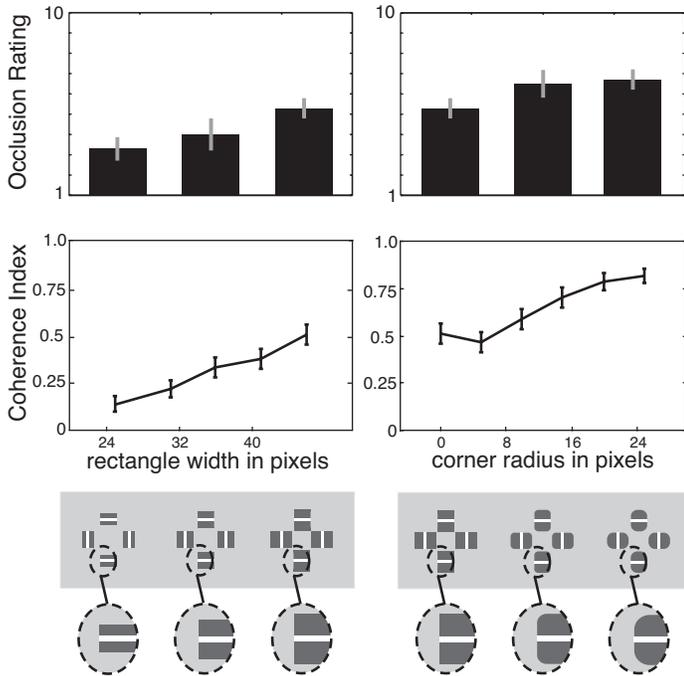


Figure 5. Stimuli and results for Experiment 3, a parametric study of the effects of aperture width and roundedness. Increasing aperture width increases coherence and perceived occlusion, as does increasing the roundedness. Click link to view demo.

Increasing the aperture width increased the degree of coherence, as shown in Figure 5a. This is consistent with the popular idea that figure size serves as a cue to border ownership, with small regions more likely to be seen as figure rather than ground. The static occlusion ratings also support this notion.

There was also an effect of how round the aperture was. Rectangular apertures the same width as the round ones produced lower levels of coherence, and parametrically varying the amount of curvature systematically changed the degree of coherence, as shown in Figure 5b. This effect was again also reflected in the static occlusion ratings – round apertures are seen as more occlusive than rectangular ones. To our knowledge it is the first time curvature sharpness has been documented as an occlusion cue. Both width and roundedness affect perceived motion and occlusion in much the same way.

Lines

To further probe the role of occluding contour shape, we conducted some experiments with outline stimuli, which allow one to isolate the effect of the local contour geometry.

The lines in the stimuli were 2 pixels in width. Their luminance was 30.1 cd/m². The segment composing the basic T-junction was 25 pixels (0.4 deg) in length. The segments added to form the convexity were 5 pixels in length; those added to form the concavities were 10 pixels in length. Six naïve subjects participated in the experiments, completing 15 trials in a single block that included other conditions not reported in this study.

To first make sure that outline stimuli behave in much the same way as stimuli composed of filled regions, we replicated the results of Figures 3-5 with outline versions of the same stimuli, as shown in Figure 6. Both the coherence and occlusion ratings are similar to those of the previous experiments for all the principle stimuli, suggesting that the line stimuli are tapping the same mechanisms.

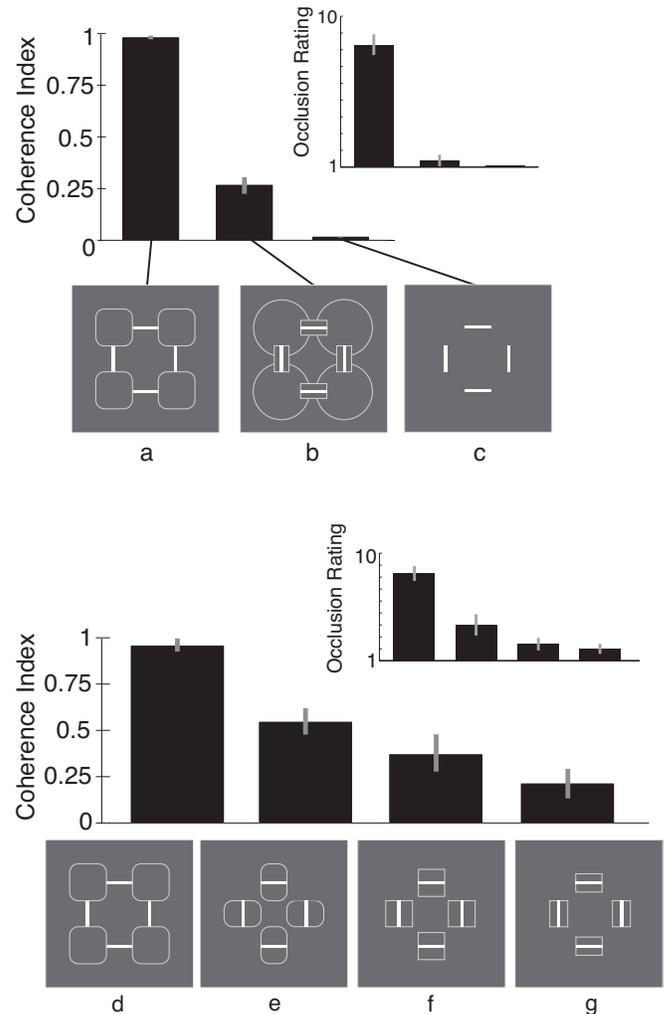


Figure 6. Stimuli and results for Experiment 4, which replicates the effects of Experiments 1, 2, and 3 with stimuli composed of lines. Click link to view demo.

We then took the outline occluders of Figure 6a and removed most of the occluding contour, leaving just the T-junctions at the bar endpoints, shown in Figure 7a. This stimulus generates intermediate levels of coherence. In the stimuli of Figure 7b and 7c, we added short line segments to the T-junctions to produce local convexities and concavities, respectively. We found that the convexities increased the level of coherence relative to the T-junctions alone, whereas the concavities decreased it. Note that no occluders are visible in these stimuli; there are just isolated pieces of contour. Although the mean occlusion rating for the convex condition is somewhat higher than for the T-junction and concave conditions, all three stimuli produced quite low ratings of perceived occlusion in our subjects. Nonetheless, manipulating the local concavity produced a sizeable effect on perceived motion. This is the first substantial lack of correspondence between perceived occlusion and perceived motion that we have documented thus far in this work. It seems that what we will term the semilocal neighborhood around a moving occlusion point - larger than the junctions at the point in question, but smaller than the entire stimulus - is at least somewhat predictive of perceived motion even when perceived occlusion is not much affected. One interpretation is that the global context plays an important role in determining perceived occlusion, but is less important for determining perceived motion.

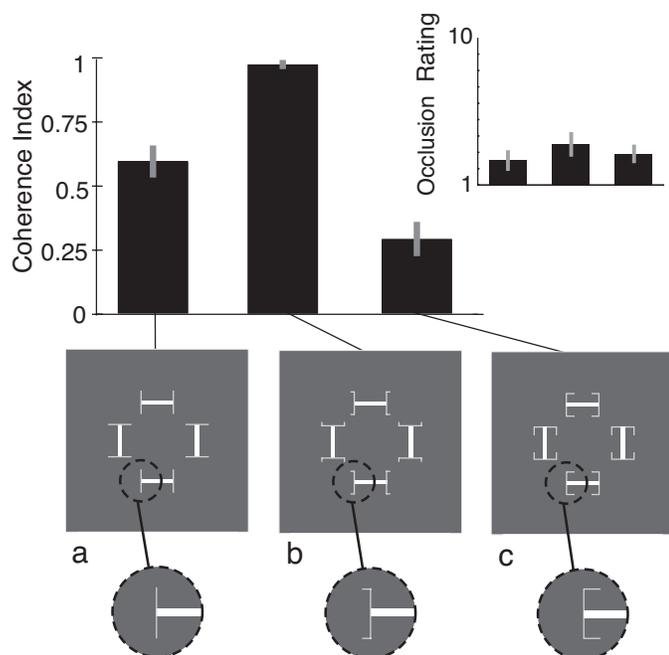


Figure 7. Stimuli and results for Experiment 5. (a). T-junctions alone produce intermediate levels of coherence, which is increased by adding convexities (b) and decreased by adding concavities (c). All three stimuli produce low occlusion ratings. Click link to view [demo](#).

Compound contours

To further test the extent to which the semilocal region surrounding each terminator could predict the effect of occluding contour geometry on perceived motion, we concocted the compound stimuli of Figure 8. As can be seen from the zoom-ins, stimuli (b) and (c) result from taking part of the occluding contour of the round occluders and part of the occluding contour from the rectangular apertures. As shown in the graphs of Figure 8, the levels of coherence obtained for the compound stimuli were intermediate between those of the original stimuli (shown in Figure 8a and 8d), and similar for the two stimuli (results are from six naïve subjects) ($t[200] = 0.53$; $p = .59$). This is consistent with the results of the local convexity experiment of Figure 5; the motion in the stimuli can be predicted from the local convexities and concavities, which are the same in 8b and 8c. Notably, the static occlusion ratings did not follow the same pattern. The bars of Figure 8c were seen as less occluded than those of Figure 8b ($t[18] = 2.31$; $p = .03$), even though the coherence levels were comparable. This is another instance in which the static occlusion ratings and perceived coherence apparently do not display the same dependencies on occlusion cues. There may be some small difference in perceived coherence between the stimuli of Figure 8b and 8c that is hidden by the noisiness of our measurements, but any such difference is small, again suggesting that motion interpretation may be mostly determined by the semilocal neighborhood.

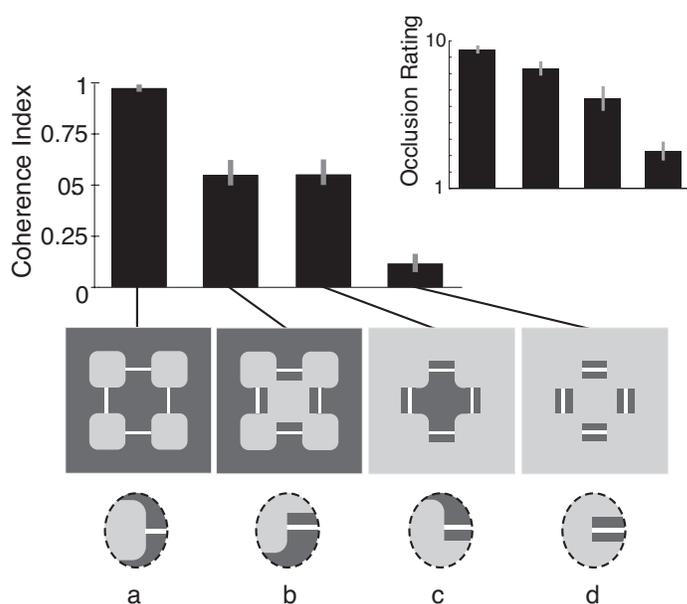


Figure 8. Stimuli and results for Experiment 6. Compound stimuli were generated, the occluding contours of which are the result of taking part of the contour from the occluders and part of the contour from the rectangular apertures. Coherence levels for the compound stimuli are similar, and fall between those produced by the source stimuli, even though the occlusion ratings for (b) and (c) are different. Click link to view [demo](#).

In sum, our experiments manipulating occluding contour shape suggest that contour convexity is an important constraint on both border ownership (i.e., the strength of perceived occlusion) and motion interpretation. However, motion interpretation seems to be strongly influenced by the contour geometry within a semilocal neighborhood around a moving terminator, more so than is perceived occlusion. This suggests that the mechanisms driving coherence are related but not identical to those driving the perception of static occlusion.

T-junctions along the occluding contour

We next wondered whether additional T-junctions along the occluding contour might influence border ownership and perhaps motion interpretation. The stimuli of Figure 9 were designed to address this issue. The round apertures of Figure 9a alone produced fairly high levels of coherence and perceived occlusion in our six naïve subjects, as did the oddly shaped occluders of Figure 9b. But when combined in the stimulus of Figure 9c, coherence was substantially lower than in either stimulus alone, consistent with the weak percept of occlusion that most observers reported. Here, though, the weak coherence cannot be attributed merely to the shape of the occluding contour. Something happens specifically when the two contours are combined. One explanation is that the T-junctions serve to modulate the strength of border ownership, and also influence motion interpretation. The control of Figure 9d is further consistent with this notion, in that the small

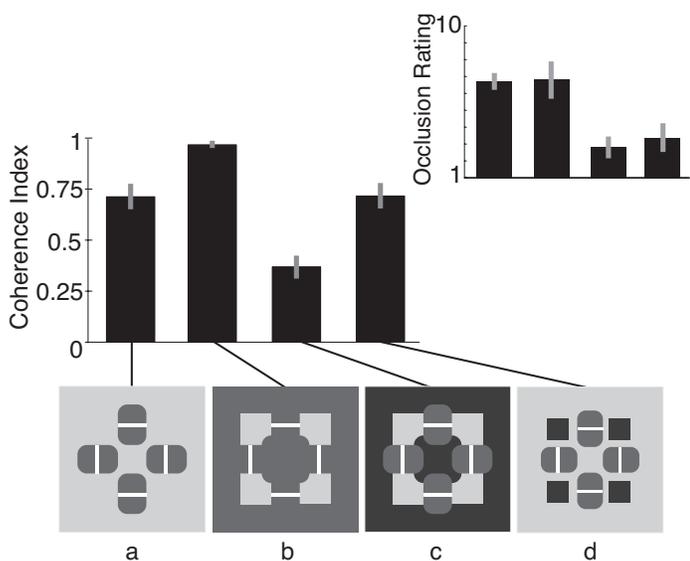


Figure 9. Stimuli and results for Experiment 7, exploring the role of static T-junctions along the occluding contour. When the round apertures of (a) and the occluders of (b) are combined in (c), coherence is lower than it is for either stimulus alone, as is the occlusion rating. The control condition in (d) suggests the T-junctions created in (c) are key. However, the dark squares introduced in (d) seem to interfere with the percept of occlusion. Click link to view [demo](#).

squares that do not generate T-junctions had no effect on perceived coherence. Note, however, that the squares did have an effect on perceived occlusion, which is lower for the stimulus of Figure 9d than for that of 9a. Apparently, the squares interfere with perceived occlusion but have little effect on motion perception, perhaps because they are removed from the semilocal neighborhood.

Can T-junctions along the occluding contour also augment the strength of occlusion, and, perhaps, motion coherence? Comparing the stimuli of Figure 10 provided some insight. If occluders are added to the thin rectangles of 10a to produce the new stimulus of 10c, T-junctions are formed that might be thought to increase the likelihood of occlusion. To assess whether these T-junctions affect border ownership and/or motion interpretation, we compared the coherence and perceived occlusion of this stimulus to that of the combination of the same occluders with the thick rectangles of 10b. We know from the experiment described earlier (Figure 5) that the thick rectangles produce higher occlusion and coherence ratings. However, their combination with occluders, shown in Figure 10d, lacks the T-junctions of its counterpart in 10c, and so it might be predicted to produce lower degrees of coherence and occlusion. The thin and thick rectangles were 26 and 50 pixels in width, respectively, and all six naïve subjects completed 20 trials per condition. Even though the thick rectangles alone produce higher levels of coherence ($t[240] = 3.85$; $p < 10^{-4}$) and perceived occlusion ($t[18] = 2.92$; $p = .0046$) than the thin ones, when occluders are added the effect reverses – the combination with the thick rectangles (Figure

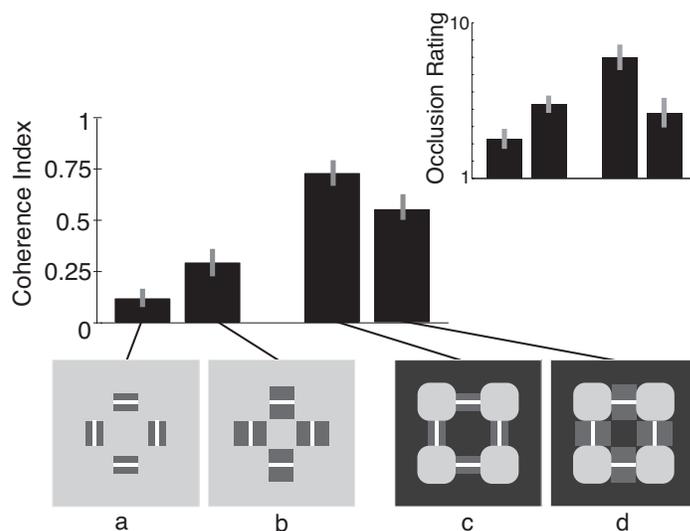


Figure 10. Stimuli and results for Experiment 8, again exploring the role of static T-junctions. The thin rectangles of (a) produce lower levels of coherence and perceived occlusion than the thick rectangles of (b), but when occluders are added in (c) and (d), the effect reverses. One explanation is that the T-junctions created in (c) serve to increase the strength of occlusion, which serves to increase the tendency to cohere. Click link to view [demo](#).

10d) is less coherent than that with the thin (Figure 10c) ($t[240] = 2.66$; $p = .004$), and is perceived to be less occluded ($t[18] = 3.17$; $p = .0026$). This is consistent with the idea that the T-junctions augment the strength of border ownership, and also somehow play a role in determining coherence.¹

Given the importance of the semilocal scale suggested by the occluding contour experiments and by the results of Figure 9, we wondered whether the effect of static T-junctions along the occluding contour would depend on their distance from the moving T-junction. The stimulus of Figure 11b is one attempt to test this notion. A dark gray cross has been added behind the round apertures of Figure 11a, generating T-junctions identical to those of Figure 11c, but situated further along the occluding contour. As shown in Figure 11, the effect of such T-junctions on perceived motion seems to be weaker; the coherence of the stimulus in Figure 11b is only slightly reduced relative to the apertures alone (and is significantly greater than that for the stimulus of Figure 11c; $t[178] = 2.52$; $p = .0125$). In contrast, the occlusion ratings for this stimulus are just as low as those for the stimulus of Figure 11c ($t[18] = .12$; $p = .9$). This is again suggestive of a semilocal region of influence that affects perceived motion more than perceived occlusion.

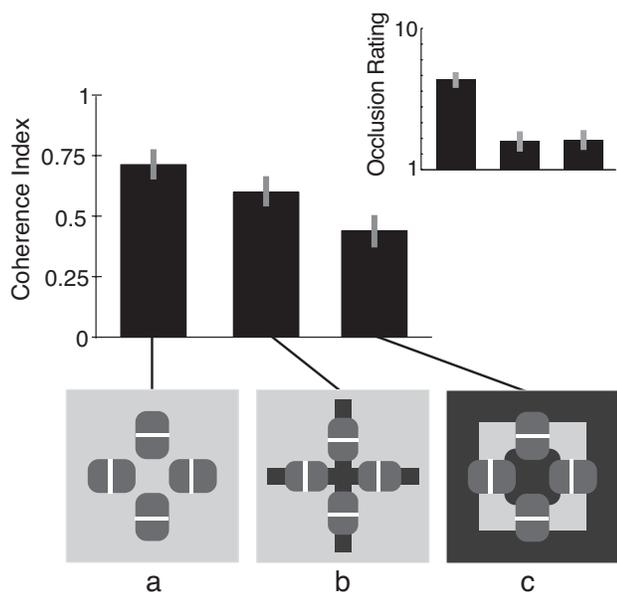


Figure 11. Stimuli and results (for six naïve subjects) for Experiment 9, testing the effect of distance on the T-junctions' influence. The coherence of the round apertures (a) is reduced with T-junctions near to the moving terminators (b), but is virtually unaffected when the junctions are moved further along the occluding contour (c). Occlusion ratings are similar for the two stimuli. Click link to view [demo](#).

In our original stimulus, shown again in Figure 12b, circles were drawn behind the rectangular apertures because they seemed to enhance the sense that the moving

bars are not occluded. The occlusion ratings for this stimulus and for the rectangular apertures alone confirm that this is the case – the apertures look more occlusive alone than with the circles added ($t[18] = 3.78$; $p = .001$). But the results of Figure 11 suggest that the circles and the T-junctions they produce ought to play little role in how the motion of the bars is interpreted, because they are “around the corner” from the closest moving terminator. This in fact seems to be the case. As shown in Figure 12c, if the circles in the original stimulus are removed, eliminating the T-junctions, coherence is no higher than it was before ($t[168] = .73$; $p = .46$). This suggests that the T-junctions in this stimulus, although apparently affecting our percept of occlusion, have little effect on the occlusion computations that influence motion interpretation, perhaps because they are outside the semilocal region of influence.

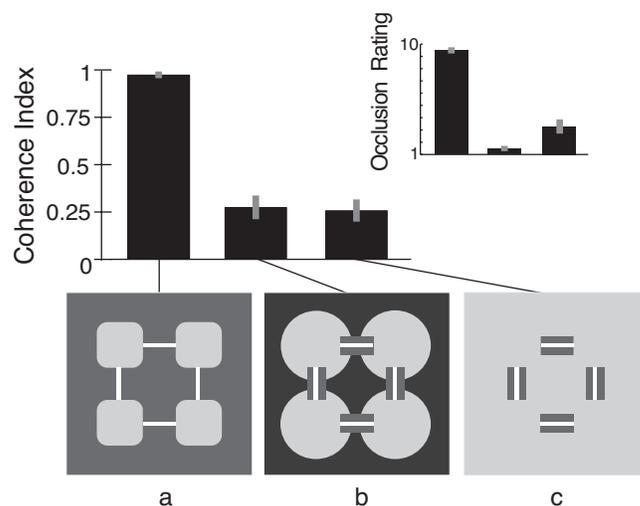


Figure 12. Stimuli and results (from six naïve subjects) for Experiment 10. Removing the circles of (b) has little effect on coherence (c), although it has a significant effect on the occlusion ratings. Click link to view [demo](#).

Discussion

The goal of the experiments described here was to characterize the form constraints that influence motion interpretation. Our strategy, as in previous studies (McDermott et al., 2001), was to hold the terminator T-junctions in our stimuli constant, and manipulate other aspects of the stimulus related to occlusion. In our previous work, we presented several demonstrations that nonlocal form constraints can exert a dominant influence on motion perception. The present work was devoted to exploring the nonlocal constraints that are presumably related to border ownership (i.e., the strength or probability of occlusion). The results suggest that although the local junction effects can be substantially modulated by nonlocal cues, most of the effects arise from nearby information, which we have termed “semi-local.” A relatively small set of factors seems to be most important, including the curvature of the oc-

cluding contour and additional T-junctions that occur along the contour. To a large extent, perceived motion and perceived occlusion judgments exhibit similar dependencies on static occlusion cues, although static occlusion judgments do not seem to be dominated by the semilocal neighborhood, at least not to the same extent. This raises the possibility that the form constraints on motion may derive from a mechanism distinct from static occlusion analysis.

The first few experiments focused on the shape of the occluding contour and its effect on the coherence of our stimuli. We found that whether the occluding contour curves toward or away from a moving edge has a large effect on perceived motion. Coherence was substantially higher for convex occluding contours than concave. This effect works even for the minimal stimuli of [Figure 7](#), which have only small effects on perceived occlusion, suggesting that the effect on motion is driven mostly by the contour geometry in the semilocal neighborhood. The sharpness of the curvature also matters (which to our knowledge has not been noted before in discussions of occlusion), as does the distance of the curvature from the moving junction.

The last few experiments tested for effects of additional static T-junctions along the occluding contour. We find that such T-junctions can either increase or decrease the strength of occlusion, depending on which side of the occluding contour they lie, but their influence on motion seems to fall off rapidly with distance. The results of these experiments are again consistent with the idea that the strength of occlusion at a particular point, as it influences motion perception, is mainly determined from a semilocal image neighborhood surrounding that point, even though the perception of occlusion is itself influenced by more global factors.

Overall, we found there to be a close correspondence between the occlusion ratings and motion coherence, which is strong evidence that our motion percepts are due to occlusion-related computations. However, these may not be exactly the same occlusion computations that subserve occlusion judgments in static images. We observed several qualitative discrepancies between the static occlusion ratings our subjects made and their ratings of perceived coherence - [Figures 7](#) and [8](#) with the effects of local convexities and concavities, [Figure 9d](#) with the control stimulus for the T-junction manipulation, and [Figures 11](#) and [12](#), which document the effects of T-junction distance on their influence. [Figure 7](#) shows that local convexities are sufficient to induce large changes in motion interpretation even when they do not have an equivalent effect on perceived occlusion. [Figure 8](#) shows that the global context can influence perceived occlusion but does not seem to affect motion interpretation to the same extent. [Figure 9](#) shows that the static squares a short distance away from the moving terminator can impair the percept of occlusion, but have no effect on perceived motion, consistent with the idea that perceived motion is not much affected by image events outside the semilocal neighborhood. And [Figures 11](#) and [12](#) show

that moving a T-junction away from a moving terminator decreases its influence on motion interpretation but not on perceived occlusion. All of these discrepancies support the importance of a semilocal region of influence in motion interpretation.

We do not claim that all the form constraints on motion interpretation are spatially limited, at least not at the scale that the effects of this paper seem to be. Our work on amodal completion (McDermott et al., 2001), for instance, demonstrates dependencies on the closure of the occluding contour, and on whether the occluding contour is the border of a solid surface, which implicate processes that analyze a much larger region of the image. Nonetheless, our present results suggest that the effects of occluding contour geometry on motion interpretation depend most strongly on what happens within a semilocal region surrounding each moving terminator.

In the case of convexity, it is worth noting that even the local convexities in the stimuli of [Figure 7](#) might stimulate a process sensitive to the presence of an occluding surface, and their influence need not indicate that the process that acts on them is spatially limited. We also cannot rule out the possibility that some sort of long range grouping process acts to group the pairs of contour pieces together, as is apparently the case in the visual search stimuli of Elder and Zucker (1993), and that these grouped contours are driving the differences in coherence. Nonetheless, the convexity manipulation of [Figure 7](#) is about as minimal as it could be, and produced robust effects on motion without inducing large changes in perceived occlusion. Moreover, the compound contours of [Figure 8](#) produce similar degrees of coherence when they are locally similar, even when the global shapes of the occluding contours look very different. Our results thus suggest that semilocal form analysis plays a large role in motion interpretation but may be less important for determining occlusion.

It is possible that the dependence of occlusion judgments on global stimulus properties may in part be a function of the nature of the task. For instance, the subjects who gave occlusion ratings viewed the stimuli for as long as they wanted before giving their rating. In practice this may not have been much longer than the motion trial duration of 3 s, but longer viewing times could conceivably place an emphasis on more global and sophisticated stimulus properties, and it would be interesting to obtain occlusion ratings for briefly presented stimuli. Brief presentations might tap a precursor to the global occlusion representations that our subjects were apparently basing their reports on, and perhaps it is this precursor that influences motion interpretation.

It is also known that the stimulus motion itself can serve as an occlusion cue, and one might therefore expect that occlusion judgments for static stimuli would display different stimulus dependencies than motion judgments, as the occlusion cues are not the same.² However, the differences we observed between motion and occlusion judgments were systematic: Motion judgments seem to be more

dependent on semilocal cues than do occlusion judgments. It is not obvious how any additional motion-dependent occlusion cues might cause such a pattern of results.

In other work (McDermott & Adelson, 2004), we have looked, unsuccessfully, for the presence of strictly local constraints based on junctions. We have found that in some cases changing the junctions in a stimulus produces large changes in motion, but in others it does not. Whether or not the change to the junction has an effect appears to depend on whether it induces a change in a global cost function governing the motion percept. At present we thus have ample evidence for fairly global form constraints on motion, some evidence for semilocal constraints, and no evidence for strictly local constraints.

Despite the importance of nonlocal constraints, most of the results described here would appear to lend themselves to implementation with interactions between local cues. Many of the stimulus properties that seem to matter, such as the direction and sharpness of contour curvature in the vicinity of a moving terminator, or the presence of T-junctions along the occluding contour, could plausibly be detected with simple and local operations (although see McDermott [2004], for evidence that even T-junctions may not be so easy to detect given only local information). One can envision signals from these local cues propagating along the occluding contour to determine the probability of occlusion at each terminator. Elsewhere (McDermott & Adelson, 2004) we have argued that it may sometimes be possible to provide a computational description of motion and form interactions without resorting to junction labels and instead describing the computation with a cost function based on layered surface interpretations (Weiss & Adelson, 2000). It may be possible to describe the occlusion computations of this work in similar terms, even though junctions and other local features provide a natural language with which to envision their implementation.

Conclusions

An extensive literature has documented the influence of form on motion, but most studies are consistent with a simple, junction-based account of the form processes that are involved. The findings we report here and elsewhere (McDermott et al., 2001; McDermott & Adelson, 2004) demonstrate that the computations are considerably more complex, going beyond strictly junction-based mechanisms to include a variety of other nonlocal computations. In the present work, we explored the effects of various nonlocal cues to border ownership, notably the convexity of occluding contours and the presence of static T-junctions in the neighborhood of the occluding contour. We find that local junction structure, per se, has relatively little explanatory power. It is necessary to consider the junctions in the context of the intermediate scale neighborhood in which they are embedded. Because these computations are neither local nor global, we refer to them as semilocal. Our results

suggest that these computations are related to, but are not identical to, the computations that mediate static occlusion judgments.

Acknowledgments

This work was funded by National Institutes of Health Grants EY11005-04 and EY12690-02 and ONR/MURI contract N00014-01-0625 to EA. JM was supported by the Gatsby Charitable Foundation and a Marshall Scholarship. The authors would like to thank Janice Chen for programming the flash demos linked to this paper.

Commercial relationships: none.

Corresponding author: Josh McDermott.

Email: jhm@mit.edu.

Address: NE20-444, MIT, 3 Cambridge Center, Cambridge, MA 02139.

Footnotes

¹Note that both perceived occlusion and perceived coherence are lower in the stimulus of Figure 10c than for the basic occluded stimulus (e.g., of Figure 3a, the data for which were recollected in the present experiment to allow comparison, although they are not displayed). One explanation is that the convexity of the rectangular aperture continues to exert an effect despite the presence of the T-junctions.

²Unfortunately we could not ask for occlusion ratings for moving stimuli, as occlusion percepts for the moving stimuli are largely determined by the motion percept – if the stimulus coheres it generally looks occluded. Occlusion judgments for moving stimuli thus do not provide a measure of occlusion representation independent of motion perception.

References

- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, 300, 523-525. [[PubMed](#)]
- Anderson, B. L., & Sinha, P. (1997). Reciprocal interactions between occlusion and motion computations. *Proceedings of the National Academy of Sciences*, 94, 3477-3480. [[PubMed](#)][[Article](#)]
- Anstis, S. (1990). *AI and the eye* (A. Blake & T. Troscianko, Eds.). New York: Wiley.
- Bressan, P., Ganis, G., & Vallortigara, G. (1993). The role of depth stratification in the solution of the aperture problem. *Perception*, 22, 215-228. [[PubMed](#)]
- Castet, E., & Wuerger, S. (1997). Perception of moving lines: Interactions between local perpendicular signals and 2D motion signals. *Vision Research*, 37, 705-720. [[PubMed](#)]

- Elder, J. H., & Zucker, S. W. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, 33, 981-991. [PubMed]
- Liden, L., & Mingolla, E. (1998). Monocular occlusion cues alter the influence of terminator motion in the barber pole phenomenon. *Vision Research*, 38, 383-3898. [PubMed]
- Liden, L., & Pack, C. (1999). The role of terminators and occlusion cues in motion integration and segmentation: A neural network model. *Vision Research*, 39, 3301-3320. [PubMed]
- Lindsey, D. T., & Todd, J. T. (1996). On the relative contributions of motion energy and transparency to the perception of moving plaids. *Vision Research*, 36, 207-222. [PubMed]
- Lorenceanu, J. (1999). Cooperative and competitive spatial interactions in motion integration. *Visual Neuroscience*, 16, 755-770. [PubMed]
- Lorenceanu, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, 32, 263-273. [PubMed]
- McDermott, J. (2004). Psychophysics with junctions in real images. *Perception*, 33, 1101-1127.
- McDermott, J., & Adelson, E. H. (2004). Junctions and cost functions in motion interpretation. *Journal of Vision*, 4(7), 552-563, <http://journalofvision.org/4/7/3/>, doi:10.1167/4.7.3. [PubMed][Article]
- McDermott, J., Weiss, Y., & Adelson, E. H. (2001). Beyond junctions: Nonlocal form constraints on motion interpretation. *Perception*, 30, 905-923. [PubMed]
- Nakayama, K., & Silverman, G. H. (1988). The aperture problem-II: Spatial integration of information along contours. *Vision Research*, 28, 747-753. [PubMed]
- Nowland, S., & Sejnowski, T. (1995). A selection model for motion processing in area MT of primates. *Journal of Neuroscience*, 15, 1195-1214. [PubMed]
- Pao, H., Geiger, D., & Rubin, N. (1999). Measuring convexity for Figure/Ground separation. *Proceedings of the 7th IEEE International Conference on Computer Vision*, II, 948-955.
- Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception*, 30, 339-366. [PubMed]
- Shiffrar, M., Li, X., & Lorenceanu, J. (1995). Motion integration across differing image features. *Vision Research*, 35, 237-2146. [PubMed]
- Shiffrar, M., & Lorenceanu, J. (1996). Increased motion linking across edges with decreased luminance contrast, edge width and duration. *Vision Research*, 36, 2061-2067. [PubMed]
- Stevens, K. A., & Brookes, A. (1988). The concave cusp as a determiner of figure-ground. *Perception*, 17, 35-42. [PubMed]
- Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual motion: Evidence for surface segmentation mechanisms. *Vision Research*, 36, 1291-1310. [PubMed]
- Stoner, G. R., & Albright, T. D. (1998). Luminance contrast affects motion coherency in plaid patterns by acting as a depth-from occlusion cue. *Vision Research*, 38, 387-401. [PubMed]
- Stoner, G. R., Albright, T. D., & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, 344, 153-155. [PubMed]
- Trueswell, J. C., & Hayhoe, M. M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research*, 33, 313-328. [PubMed]
- Vallortigara, G., & Bressan, P. (1991). Occlusion and the perception of coherent motion. *Vision Research*, 31, 1967-1978. [PubMed]
- Wallach, H. (1935). Ueber visuell whargenommene bewegungrichtung. *Psychologische Forschung*, 20, 325-380.
- Weiss, Y., & Adelson, E. H. (2000). Adventures with gelatinous ellipses: Constraints on models of human motion analysis. *Perception*, 29, 543-566. [PubMed]