Reciprocal interactions between occlusion and motion computations

Barton L. Anderson* and Pawan Sinha

Massachusetts Institute of Technology, Cambridge, MA 02139

ABSTRACT The "aperture problem" refers to the inherent ambiguity of the motion generated by an untextured contour moving within an aperture. The limited spatial extent of the receptive fields of neurons in cortical areas like V1 and MT render them susceptible to this problem. Most psychophysical experiments have probed how the visual system overcomes the aperture problem by presenting moving contours behind one or more simulated apertures. The assumption has been that the computational ambiguities that arise in resolving these displays are equivalent to the computational problems created by receptive fields that sample a small region of visual space. Evidence is presented here that challenges this view. We demonstrate that a fundamental computational difference in the interpretation of contour terminators arises in these two variants of the aperture problem. When the aperture is a receptive field, and a moving contour extends beyond its boundaries, the contour "terminators" delimit the boundaries of the receptive field, not the ends of the contour. In contrast, when a moving contour is viewed through a simulated aperture, the contour terminators are generated by the occluding edges of the aperture. In a series of experiments, we show that reciprocal interactions arise between computations of occlusion and those of motion direction and integration. Our results demonstrate that the visual system solves the aperture problem by decomposing moving contours into moving segments, and unpaired terminators that arise from the accretion and deletion of contours behind occluding edges, generating both coherent motion and illusory occluding surfaces.

Fig. 1a depicts the "aperture problem," i.e., the fact that the direction of motion of an untextured contour is locally ambiguous; any one of the depicted motions is a possible trajectory of the contour (1–4). This ambiguity exists whether the aperture under consideration is a receptive field (5–7) or an aperture in the environment. Most psychophysical experiments have probed how the visual system overcomes the aperture problem by presenting moving contours behind one or more simulated apertures (8–15). However, if the aperture being considered is a hole in an occluding surface, a second problem also arises: Contour discontinuities are generated, and the visual system is faced with the task of sensing and classifying these discontinuities. Image discontinuities generated by occluding contours present a unique problem for computations involving multiple views such as stereopsis and motion perception. Although both binocular disparity and motion rely on identifying corresponding features in the multiple views, differential occlusion generates features that cannot be matched and therefore do not generate either

disparity or motion signals.† Fig. 1b depicts this problem for a contour moving behind an aperture. Note that every possible motion direction will generate both accreted and deleted contour segments behind the edges of the occluding aperture (dashed and dotted contour segments in Fig. 1, respectively). Thus, when a contour moves behind an aperture, its direction of motion and the segments that are accreted and/or deleted are coupled. This suggests the possibility that the computations of motion direction used by the visual system and the recovery of occlusion geometry also are coupled. We performed a number of experiments to assess this possibility.‡

We began by creating motion analogues of patterns that have been reported to generate illusory contours in binocular viewing (16, 17). These displays are shown in Fig. 2a. Two vertically oriented contours were displaced horizontally underneath an opaque, wedge-shaped surface with invisible boundaries. As the contours translated to the right, they changed length in a manner consistent with an occluding surface translated to the left at an equal speed. Two small dots were attached to the vertical contours to eliminate the percept of the lines looming in depth. To all observers, the lines simply appeared to grow and shrink while translating; no percept of occlusion was observed. But if a few squares were attached to the occluding surface, the perceived organization was transformed dramatically. The motion parallax of the squares generated a strong percept of an occluding surface sliding over the lines, and vivid illusory contours were reported by all observers.

We suggest that the illusory contours arose in this display from the visual system interpreting the contour terminators as

*To whom reprint requests should be addressed. e-mail: bart@ psyche.mit.edu.

†The claim that occlusion generates unmatchable features has been discussed extensively previously for stereopsis (16–19), but comparatively little has been said about the same problem as it applies to motion. The simplest way to see how the same problem arises in motion is to consider discrete time slices of a motion stimulus (cf. Fig. 1), generated either by the parallax of a moving observer or by differential object motion that leads to partial occlusion. In such cases, there will be some features that are present in one temporal interval that are absent in the other. Although it is theoretically possible to compute motion for these features by matching them with other (spurious) image features, these motion signals would only interfere with the recovery of the true object motion.

‡The notion that the contour segments depicted in Fig. 1 are unmatched is contingent upon the assumption that the contours do not stretch or shrink when translating. In some of the displays described in this paper, such stretching transformations were observed (e.g., experiment 1). However, when such (divergence) transformations were not observed and there was some perceived component of motion along the contour's orientation, vivid illusory contours were perceived. We argue that these illusory contours were generated by the interpretation of the contour terminators as accreted and deleted, i.e., features that do not have correspondences in the multiple views and therefore do not generate motion signals. Recent work has demonstrated that unpaired contour terminators in stereopsis can generate vivid illusory contours (16, 17, 19), which provides conceptual support for our suggestion that the classification of contour terminators as unpaired in a motion sequence could generate similar subjective contours.
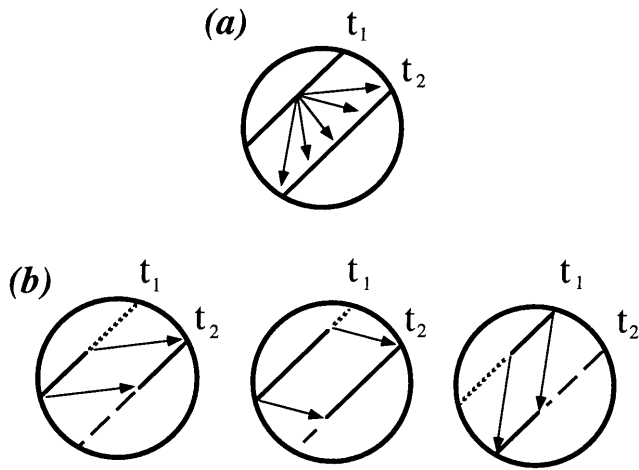
FIG. 1. A schematic of two aperture problems and the computations that they entail. (*a*) The ambiguity in motion speed and direction generated when an untextured contour is either sampled by a receptive field smaller than the contour's length or when a contour moves behind an aperture in the environment. Two discrete time frames are presented (t_1 and t_2), which may be regarded as either the temporal sampling limit of the visual system or a limit on the presentation rate of the display device (e.g., a computer monitor). (*b*) The unique computational problem that arises when a contour moves behind an occluding surface in the environment. Three possible interpretations of motion direction are shown (arrows and solid black contour segments). Note that regions of the contour will either be accreted (dashed segments) or deleted (dotted segments) in all possible directions of motion and therefore will not have corresponding elements in the two views. (Also see † footnote)

unmatchable features, i.e., as accreted and deleted contour segments. No accretion or deletion was involved when the contours were perceived simply as growing and shrinking, and consequently, no illusory contours were perceived in that condition. However, an alternative explanation of this result is that the illusory contours were formed by mechanisms that simply integrated the trajectory of the terminator's motion with the relative depth signal generated by the motion parallax of the small squares. To test this alternative, we created displays similar to those described above, but in which the length of the lines was constant during its horizontal translation (see Fig. 2*c*). In this display, an interpretation of occlusion would not entail either accretion or deletion of the contour terminators. Therefore, if the percept of occlusion was simply the product of integrating the contour terminator's motion with the relative depth signal generated by motion parallax, then the percept of occlusion should be as vivid in this display as it was in the display depicted in Fig. 2*b*. However, we found that no observers reported the presence of subjective contours or occlusion. The display simply appeared as two lines translating in depth behind the translating squares. Hence, the interpretation of the contour terminators as accreted and/or deleted (i.e., spatiotemporally unmatched) seems to be critical for generating the appearance of occlusion and subjective contours.

Our next set of studies focused on displays that have been extensively used to evaluate the conditions that support the integration of multiple, ambiguous moving contour segments (12–15). An outline figure of a diamond was translated horizontally behind three vertical occluding strips such that only the sides of the diamond's edges were visible (i.e., the vertices of the diamond were occluded; Fig. 3). The vertical strips were the same color as the background, and hence, only the moving sides of the diamond were visible. As reported by others (12–15), no observer reported the four moving sides as a single, partially occluded diamond. The parallel diagonal sides were simply seen to oscillate vertically in counterphase to the other
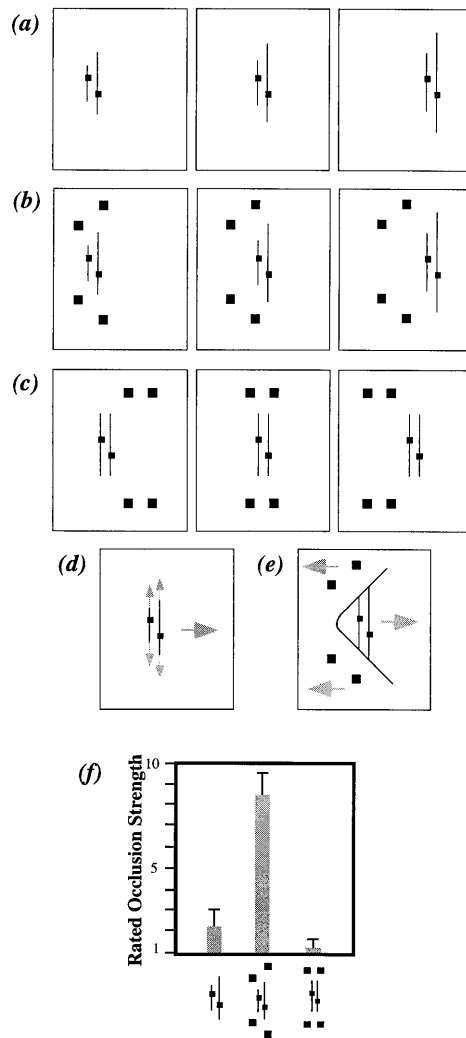

FIG. 2. A schematic of the motion displays used in experiment 1. (*a*) Three frames of the motion sequence used in the experiment. As the lines and dots translated to the right, the length of the lines increased, and the small dots on the lines translated to the right in phase with the lines. This display appeared as two lines that grew in length while they translated (depicted schematically in *d*). (*b*) Same as *a*, except that small squares were placed in the surround and translated in the opposite direction of the lines. This display generated vivid illusory contours, and the lines appeared to be progressively uncovered behind an occluding surface (depicted schematically in *e*). (*c*) If the lines did not change in length during their translation, no illusory contours were observed. (*f*) Mean ratings of 10 observers of the strength of apparent occlusion generated by the displays depicted in *a*–*c* (where more vivid occlusion is represented as larger ratings).

two parallel sides. However, if small dots were placed along the contour and translated horizontally in phase with the diamond's motion, the coherent motion of the diamond was recovered. More significantly, the coherent motion of the diamond also gave rise to vivid illusory contours, and the presence of occluding surfaces was now apparent. Thus, two nearly identical displays did or did not generate illusory contours depending on the perceived motion direction of their constituents.

It is important to note that the coherent motion of the diamond was not the only percept that could be observed in this display. A number of observers also reported that the line segments could appear to slide through the dots with prolonged viewing, i.e., the motion of the dots failed to capture the motion of the line segments. However, when this motion was
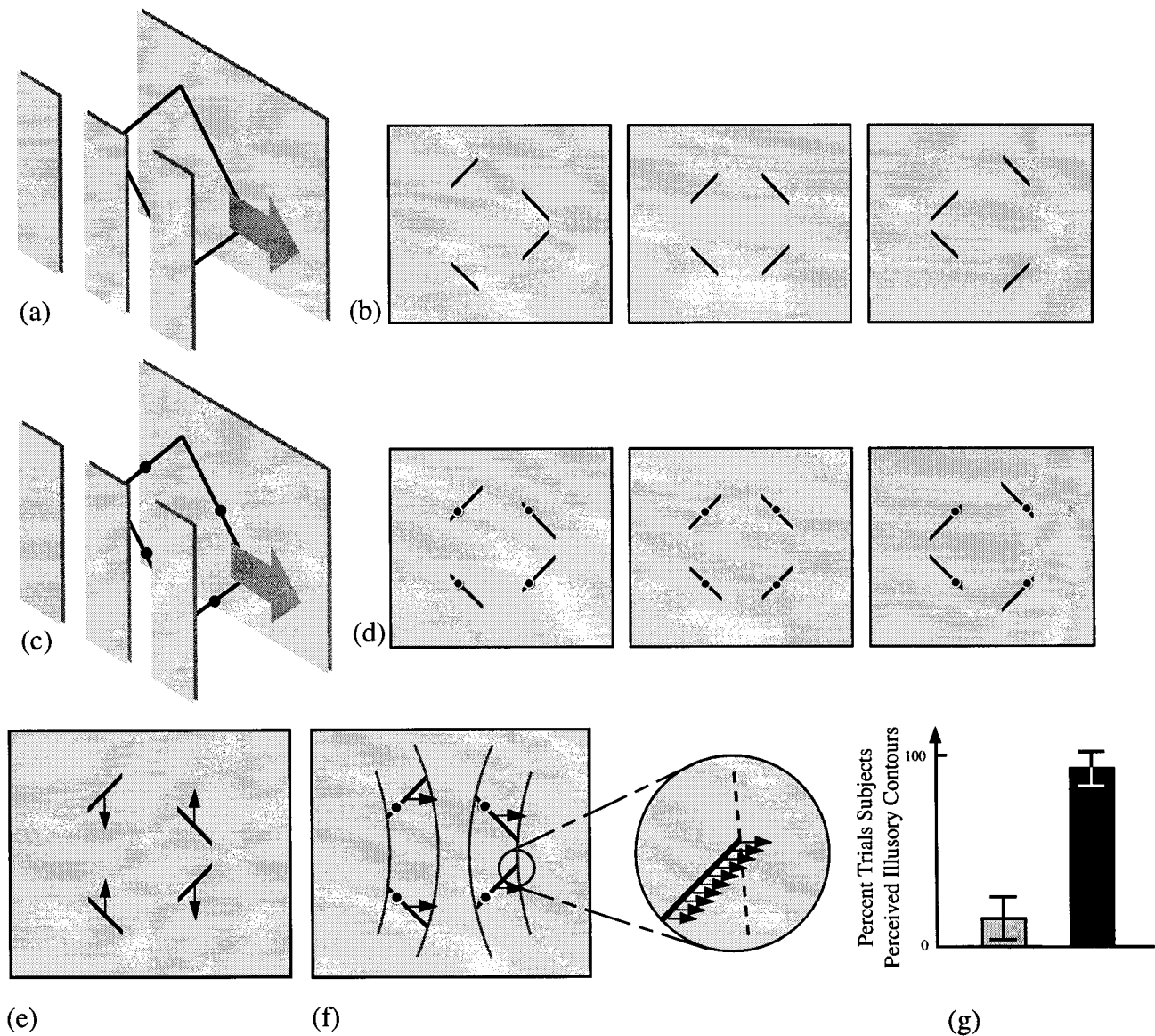
FIG. 3.   (*a*) The setup used to generate the sequence shown in *b*. The diamond oscillated horizontally behind occluders that had the same gray level as the background. There is no physical significance to the depth differences among the occluders, diamond, and background shown in this figure. They are indicative only of the relative depth ordering used to generate the sequences. The displays subtended 7° of visual angle at a viewing distance of 80 cm. The diamond's speed of oscillation was 0.7°/s. (*c*) The setup used to generate the sequence shown in *d*. The four point features attached to the moving diamond serve to disambiguate its motion. (*e*) The percept with sequence *b*—four lines oscillating vertically. No occluding contours are perceived. (*f*) The percept with sequence *d*. The four lines are seen as being part of a partially occluded diamond translating horizontally; strong subjective contours are reported at the indicated locations. As shown in the enlarged *Inset*, the illusory contours coincide with the discontinuities in the horizontal motion field. The reported contours were not precisely vertical but were slightly curved, possibly due to a bias toward orthogonality to the inducing contour. (*g*) A summary of the results obtained with the two sequences.

perceived, the percept of illusory contours was transformed dramatically. The illusory contours previously generated along the terminators of the line segments were greatly diminished or completely absent. Note, however, that, because small gaps were placed between the dots and the lines to form "extra" contour terminators, this percept implies that the line segments were being accreted and deleted behind the dots. Remarkably, a number of observers spontaneously reported the formation of new illusory contours when the percept of the line segments sliding through the dots was dominant. In addition to the vertical motion of the contour segments, a faint, illusory cross appeared to overlay and occlude the moving contour segments. The limbs that formed the cross appeared at ±45°, i.e., orthogonal to the orientation of the vertically moving line segments. Note that this is entirely consistent with the thesis developed herein, namely, that the illusory contours

orthogonal to the orientation of the contour terminators are generated when the terminators are unmatched (accreted and deleted). Thus, our theoretical assertions are not undermined by the existence of alternative percepts in these motion displays. Indeed, the alternative organizations lend further support to our basic thesis.

The preceding experiments reveal that computations of motion direction are intrinsically tied to the recovery of occlusion geometry. To this point, the evidence presented to support this thesis has been the formation of vivid illusory contours by contour terminators that are putatively unpaired. Our last set of experiments sought to evaluate the reciprocity of these interactions. We modified the displays depicted in Fig. 3 by imposing static figural information to suggest the presence of occluding surfaces, either with visible contrast or with monocular illusory contours (see Fig. 4). The presence of these
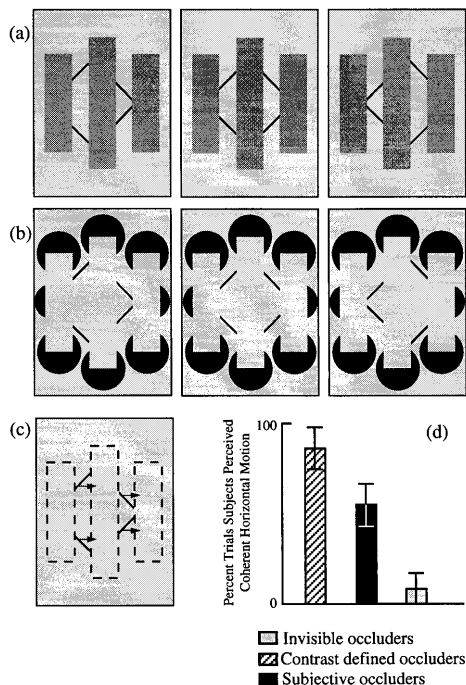
FIG. 4. (*a*) The same sequence as the one shown in Fig. 3*b*, except the occluders are explicitly defined by contrast differences. (*b*) A sequence with subjectively defined occluders. (*c*) The dominant percept with sequences *a* and *b*: a partially obscured diamond translating horizontally behind three opaque strips. (*d*) A comparison of the results obtained with the sequence shown in Fig. 3*b* and sequences *a* and *b* shown here. The explicit indication of occluding surfaces (contrast defined or subjective) strongly biased the percept from being one of motion in the vertical direction to one of coherent horizontal motion. Both the contrast and subjectively defined occluders generated a much stronger bias for perceiving coherent horizontal motion when compared with the display that contained only the moving contour segments (pairwise *t* test, $P < 0.001$).

static form cues greatly enhanced the percept of a single translating diamond, in stark contrast to the case in which the occluders were not visible (see Fig. 4*d*).

In a related paradigm, Shimojo *et al.* (11) demonstrated that the perceived direction and speed of moving contours were influenced by the presence of stereoscopically specified occluding surfaces. They argued that their results showed that extrinsic contour terminators created by an occluding surface were processed differently from intrinsic terminators that were caused by the contours actually ending. The findings described here support and extend this classification scheme in novel ways. Specifically, Shimojo *et al.* (11) demonstrated that imposing stereoscopic information about the presence of an occluding surface can cause a terminator to be interpreted as extrinsic (i.e., occluded), similar to the motion experiments depicted in Fig. 3. However, we have shown that the critical property needed to classify moving contour terminators as extrinsic is the perception of some component of motion along the orientation of the contour accompanied by a change in the contour's length, at least in stimuli for which motion is the only information present. We suggest that this perceived motion direction causes contours to be decomposed into matchable (moving) and unmatchable segments (accreted and deleted terminators), which in turn generates vivid illusory contours. Although this is consistent with the intrinsic/extrinsic classification scheme described by Shimojo *et al.* (11), it extends this dichotomy by underscoring the critical role of unpaired con-

tour segments in the genesis of illusory contours in moving displays.§

These results lead us to conclude that the aperture problem—historically treated as a single computational problem—actually entails two distinct but coupled computations: the direction and speed of contour motion; and the recovery of occlusion geometry. We have shown that, when observers perceive a component of motion along the orientation of a contour, the visual system is biased to interpret the contour terminator as accreted or deleted, which gives rise to vivid illusory contours. Reciprocally, the presence of occluding surfaces, contrast defined or subjective, greatly improves the ability of the visual system to interpret contour terminators as accreted and deleted, which leads to a strong enhancement in the ability to integrate contour segments into a single object (cf. 11). Previous work on the aperture problem has focused on explicating the computations that underlie the recovery of contour speed and direction (8–15). We argue that this information is insufficient to understand the nature of the aperture problem or the computations that are needed to solve it. The information contained in classifying a terminator as unmatched (accreted or deleted) is complementary to the information contained in motion signals. Motion signals provide information about surface regions that are visible in the multiple views, allowing for the preservation of object identity over a displacement in space/time. Accreted and deleted surface regions provide information about the presence of an occluding surface, which plays a critical role in image segmentation (20). Both processes are needed to preserve the identity of an object that is partially occluded during its motion.

1. Wohlgemuth, A. (1911) *Br. J. Psychol. Monogr.* **1,** Suppl.
2. Stumpf, P. (1911) *Z. Psychol.* **59,** 321–330.
3. Wallach, H. (1935) *Psychol. Forsch.* **20,** 325–380.
4. Ullman, S. (1979) *The Interpretation of Visual Motion* (MIT Press, Cambridge, MA).
5. Albright, T. D. (1984) *J. Neurophysiol.* **52,** 1106–1130.
6. Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. (1985) in *Pattern Recognition Mechanisms*, eds. Chagas, C., Gattass, R. & Gross, C. (Pontifical Academy of Science, Vatican City, Italy), pp. 117–151.
7. Rodman, H. R. & Albright, T. D. (1989) *Exp. Brain Res.* **75,** 53–64.
8. Adelson, E. H. & Movshon, J. A. (1982) *Nature (London)* **300,** 523–525.
9. Welch, L. (1989) *Nature (London)* **337,** 734–736.
10. Hildreth, E. (1984) *The Measurement of Visual Motion* (MIT Press, Cambridge, MA).
11. Shimojo, S., Silverman, G. H. & Nakayama, K. (1989) *Vision Res.* **29,** 619–626.
12. Shiffrar, M. & Pavel, M. (1991) *J. Exp. Psychol. Hum. Percept. Perform.* **17,** 44–54.
13. Lorenceau, J. & Shiffrar, M. (1992) *Vision Res.* **32,** 263–273.
14. Lorenceau, J., Shiffrar, M., Wells, N. & Castet, E. (1993) *Vision Res.* **33,** 1921–1936.
15. Shiffrar, M., Li, X. & Lorenceau, J. (1995) *Vision Res.* **35,** 2137–2146.
16. Anderson, B. L. (1994) *Nature (London)* **367,** 365–368.
17. Anderson, B. L. & Julesz, B. (1995) *Psychol. Rev.* **102,** 705–743.
18. Nakayama K. & Shimojo, S. (1990) *Vision Res.* **30,** 1811–1825.
19. Malik, J. (1996) in *Proceedings of the European Conference on Computer Vision* (Cambridge, MA), pp. 167–174.
20. Gibson, J. J. (1950) *The Perception of the Visual World* (Houghton–Mifflin, Boston).

§One of us (B.L.A.) has also recently completed experiments that reveal that the original result using stereoscopic "barber pole" stimuli reported by Shimojo *et al.* (11) only occurs for stimuli that contain contour terminators that are unpaired interocularly. In other words, even when depth specifies that a contour is occluded by a near surface, the contour terminators will still capture the ambiguous motion of the contours unless the contour terminators are also interocularly unpaired. This result provides further support for our conjecture that a critical property needed to classify contour terminators as extrinsic is that these features be unmatchable, either in space or in space/time.