

“Seeing” shapes that are almost totally occluded: A new look at Parks’s camel

SHINSUKE SHIMOJO and WHITMAN RICHARDS
Massachusetts Institute of Technology, Cambridge, Massachusetts

“Seeing” shapes that are almost totally occluded: A new look at Parks’s camel

SHINSUKE SHIMOJO and WHITMAN RICHARDS
Massachusetts Institute of Technology, Cambridge, Massachusetts

When a shape moves behind a very narrow window, or “slit,” the entire shape can be recognized, although at any instant only meaningless fragments are visible. In some manner, the human visual system has the ability to piece together these sequential fragments into a coherent shape that is seen as moving across the largely occluded field. As a first step toward understanding how this feat may be accomplished, we show some minimal conditions for recovering the direction and speed of motion of an unknown shape moving behind a slit. Two important conditions are that the object be rigid and that its motion be a simple transition in the image plane.

When animals and objects move about in our visual world, they are often partially occluded by other objects, such as trees, streetlights, buildings, and so forth. Rarely do we see for long periods a shape in its entirety. In the extreme case, when an animal creeps through a forest, it may be almost totally occluded except for fleeting “slit-like” views. In 1965, Parks mimicked this extreme case by passing a shape behind a narrow window, or slit (Figure 1). Surprisingly, the shape could be recognized although at any instant only a sequence of meaningless fragments were visible—an observation first reported by Zöllner in 1862. The strength of this perception can be seen by viewing a person or object that moves behind a door that is slightly ajar. How can a visual system recover a shape that is almost totally occluded?

Several computational problems are raised by Parks’s slit paradigm (Rock, 1981). The most obvious is that the fragmented views must somehow be stored in memory and integrated into a whole. However, such an integration process cannot proceed successfully by using only the available image information. Quite simply, the space-time locus of integration must be specified. In particular, we must know in which direction and at what speed the occluded object is moving in order to piece together correctly the fragments of the outline. If integration is achieved by a smooth pursuit movement of the eye, as proposed by the “retinal painting theory” (Haber, 1968; Haber & Nathanson, 1968), then to recover the correct shape, the eye must move in the same direction as the object, and at the object’s angular speed. An eye movement at any other rate or direction will produce a distortion of the true shape. Similarly, internal scan models will suffer the same fate, unless the object’s velocity is known.

William Gilson provided technical help in the preparation of the manuscript. This research was supported by the NSF-Information Science and Technology Division, Grant IST-8312240.

S. Shimojo’s current address is: Smith-Kettlewell Institute of Visual Science, 2200 Webster Street, San Francisco, CA 94115. W. Richards’s mailing address is: Natural Computation Group, E10-120, MIT, Cambridge, MA 02139.

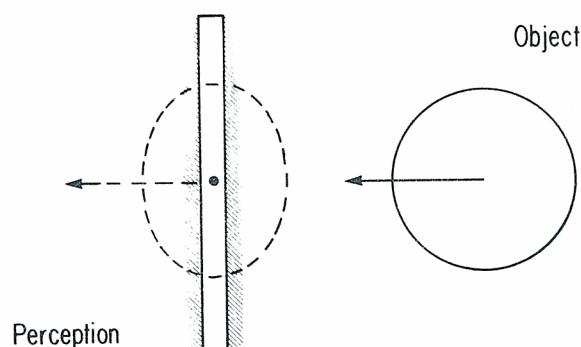


Figure 1. Recovery of a moving shape from slit views.

In general, regardless of one’s model for the integration process, without knowledge of an object’s speed and direction, there is a many-to-one mapping between the observed sequence of fragments and possible shapes. For example, Figure 2 shows three different shapes and motions that will produce identical slit views. [Our perception, however, is as in (A), of a tented shape translating horizontally.] This simple result suggests that two important constraints upon the recovery process are that the shape be rigid and that it move at constant speed and direction. Our first immediate computational problem is thus to explore the conditions under which the direction and speed of motion of a rigid shape can be recovered from slit viewing.

COMPUTATIONAL ANALYSIS

Let a rigid shape move behind a vertical slit with speed V_ϕ and at an angle ϕ , as shown in Figure 3. Depending upon the width of the slit, the observed contour will have an orientation θ_1 and curvature k_1 at time t_1 , or simply appear as a “point” undergoing up-and-down motion if the slit is very narrow. We will show first that this last extreme case, in which only points in motion are seen, will not permit the correct recovery of the shape and its

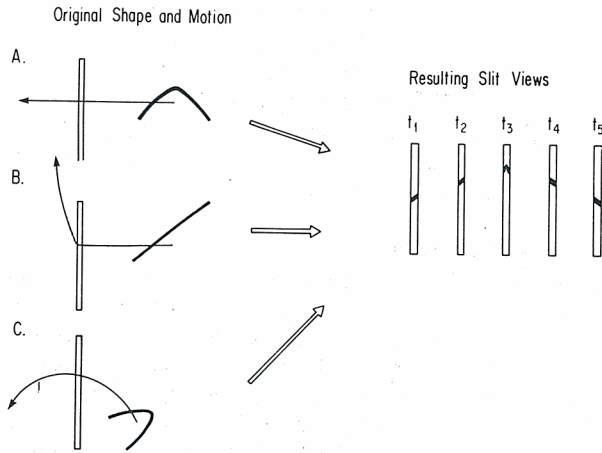


Figure 2. Ambiguity in slit views: original shape and motion.

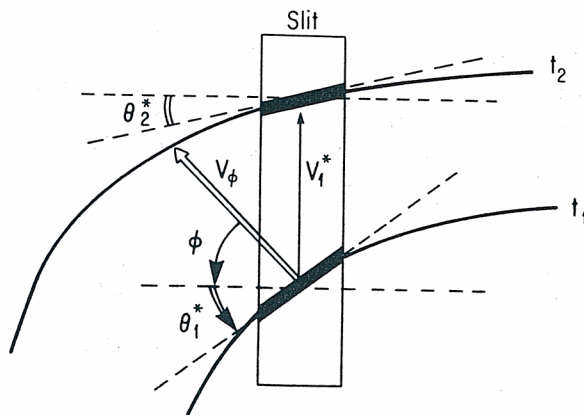


Figure 3. Geometrical relationships and definitions. V_ϕ , Velocity of translational motion in direction ϕ ; V_1^* , perceived velocity; θ_1^* , angle of contour at time t_1 ; θ_2^* , contour angle at t_2 . The asterisks indicate an observable value.

motion. However, if the slit width is increased so the slope of the contour is available, then two slit views of two tangents will be sufficient to recover the direction of translational motion (Fennema & Thompson, 1979). Finally, we show that contour curvature does not provide much additional information of use unless the curvature is very high (i.e., a "cusp").

Points in Motion

We are given as visual input a stream of points and their instantaneous velocities, namely $y(t)$ and $\dot{y}(t)$. We will represent these continuous functions in a discrete form,¹ as if the visual system took a sequence of "snapshots" of the object seen through a narrow slit. Our data are thus a sequence of positions, $y_1^*, y_2^*, y_3^*, \dots$, and their associated velocities, $\dot{y}(t_1) = V_1^*, \dot{y}(t_2) = V_2^*, \dot{y}(t_3) = V_3^*, \dots$ (where the asterisk indicates an observable—see Figure 3). From this input, the task is to recover the shape that generated the sequence, assuming the shape is rigid and translates at constant velocity. A simple counterexample, depicted

in Figure 4A, will show that this cannot be done. Let the shape be a parallelogram with sides at a 60° angle to the vertical slit and moving horizontally at 1 unit/sec. Then the vertical motion seen through the slit will be $\pm 1/\sqrt{3}$ units/sec. This motion can be mimicked, however, by constructing another (two) parallelograms with sides separated by 25° and moving 1.9 units/sec at either $\pm 45^\circ$ to the slit, as illustrated in Figure 4B. Specifically, the observed velocity, V^* , of the contour seen as a point through the slit will be

$$V^* = V_\phi \sin(\phi - \theta) / \cos \theta, \quad (1)$$

where V_ϕ is the object's speed in direction ϕ and θ is the orientation of the contour. Because V_ϕ , ϕ , and θ are all unknowns, we can offset speeds V_ϕ with angles ϕ and θ to keep the image velocity, V^* , the same. Not surprisingly, this result agrees with psychophysical observations that when the slit becomes so narrow that only points in motion are seen, the shape (and direction) of the occluded object cannot be recovered.

Tangents in Motion

As the slit widens, the orientation of the contour, θ , becomes visible. Most psychophysical and physiological evidence provide support for using this feature as an input (Adelson & Movshon, 1983; Andrews, 1967; Frisby, 1980; Hubel & Wiesel, 1962, 1968; Movshon, Adelson, Gizzi, & Newsome, 1985; Schiller, Finley, & Volkman, 1976a, 1976b). Not so clear, however, is whether the instantaneous motion perpendicular to the contour is known. Several investigators (Hildreth, 1984a, 1984b; Marr & Ullman, 1981; Waxman & Wohn, 1985) assume that the magnitude of this perpendicular component is available. However, we will elect the more conservative option that it is not, and that only the speed parallel to the slit is available. (This is as if the receptive fields for speed were

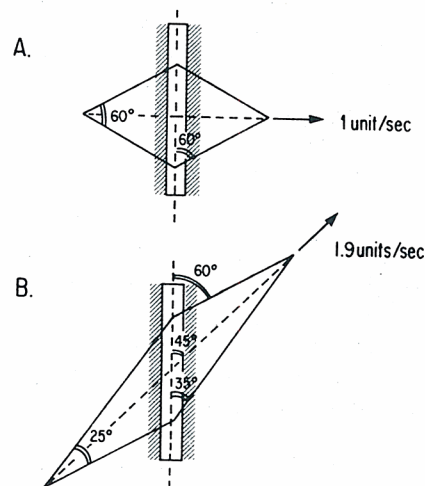


Figure 4. Two different parallelograms whose motion can generate the same sequence of points seen through a narrow slit. A mirror reflection of case B about the horizontal axis will also create identical point motions in the slit.

larger than the receptive fields that detect orientation.) Our givens are therefore the position of the contour at a given instant, $y(t_i) = y_i$, and its orientation, $\theta(t_i) = \theta_i$. To calculate the observed speed V_1^* , we will consider discrete position samples y_1 and y_2 taken at t_1 and t_2 . Associated with each sample will be two independent orientations, θ_1^* and θ_2^* (see Figure 3). Our knowns, or observables, are thus V_1^* , θ_1^* , and θ_2^* . Appendix 1 shows that the observed speed V_1^* will be given by

$$V_1^* = V_\phi \cos \phi \left[-\tan \phi + \tan \left(\frac{\theta_1^* + \theta_2^*}{2} \right) \right], \quad (2)$$

where V_ϕ and ϕ are the unknown speeds and directions of motion of the object. For simple translation, we obtain two such equations from two "views" (or two sets of samples), which can be solved for ϕ and V_ϕ (see Appendix 1):

$$\phi = \arctan \left[\frac{V_1^* \tan \left(\frac{\theta_{21}^* + \theta_{22}^*}{2} \right) - V_2^* \tan \left(\frac{\theta_{11}^* + \theta_{12}^*}{2} \right)}{V_1^* - V_2^*} \right] \quad (3)$$

$$V_\phi = \frac{V_1^*}{\cos \phi^* \left[-\tan \phi^* + \tan \left(\frac{\theta_{11}^* + \theta_{12}^*}{2} \right) \right]}, \quad (4)$$

where θ_{11}^* , θ_{12}^* are the two orientations associated with V_1^* and θ_{21}^* , θ_{22}^* are the two contour orientations associated with V_2^* . Note that this solution for the velocity (speed and direction) of the translating object does not require explicit knowledge of the contour position, but only its observed orientation, θ^* , and speed, V^* , along the slit.²

To obtain two independent equations (2), we can use either the top and bottom portions of the shape as they simultaneously pass through the slit or, alternatively, two "snapshots" of only one contour seen at different times. Because the translating shape of a single contour can be recovered correctly by the human observer, even if the contour is not continuous, the latter scheme is physiologically plausible. Computationally, this might be expected, for several "snapshots" allow the observer to check whether his assumption about a constant direction of motion of the shape is valid.

Curvature

As the slit widens, the curvature of the contour becomes apparent (Watt & Andrews, 1982). Because curvature is the rate of change of the tangent, in principle the velocity of the occluded translating object could be recovered from a single view of a curved segment plus its speed along the two edges of the slit. Rock (1981) reports that under these conditions the correct shape may be inferred. However, both our observations and those of S. Tsuzaki (personal communication, 1985) suggest that the human visual system does not have this capability, probably because the sampling rate and spatial resolution for curvature are too low for the given V_ϕ and slit width (Nakayama & Silverman, 1983; Wilson, 1985). The exception is when

a "corner" passes a wide slit (Andrews, Butcher, & Buckley, 1973). Then two tangents are visible and the solution is the same as in the previous section. We suggest that Rock's observations really relied on large curvature changes that could be viewed as a "two-tangent" solution. Further experimental work is needed here to clarify the role, if any, of curvature.

Rotation Constraint

Translation is simply a degenerate case of rotation, with the center of rotation at infinity. Is it possible, then, to recover correctly a rigid shape when it is rotated behind the slit? Now our unknowns are an angular velocity, ω , and the center of rotation, x_0, y_0 . Appendix 2 shows that given only the orientation of the contour θ_i^* and speeds V_i^* , this recovery is difficult, although possible in some special cases. Adding knowledge about the local shape of the contour will help, but then the slit must be widened so extensively that the local, instantaneous velocity field along the contour can be measured (Hildreth, 1984a, 1984b; Waxman & Wohn, 1985).

Psychophysical observations support our conclusion that a shape undergoing rotation behind a slit will not be perceived correctly (Anstis & Atkinson, 1967; Roget, 1824, cited in Boring, 1942). Figure 5 illustrates one example. When the pattern is under rotation behind the slit, as shown in Figure 5A, a distorted pattern would be predicted if: (1) the translation constraint was used, and (2) the positions of the line segments were simply replotted

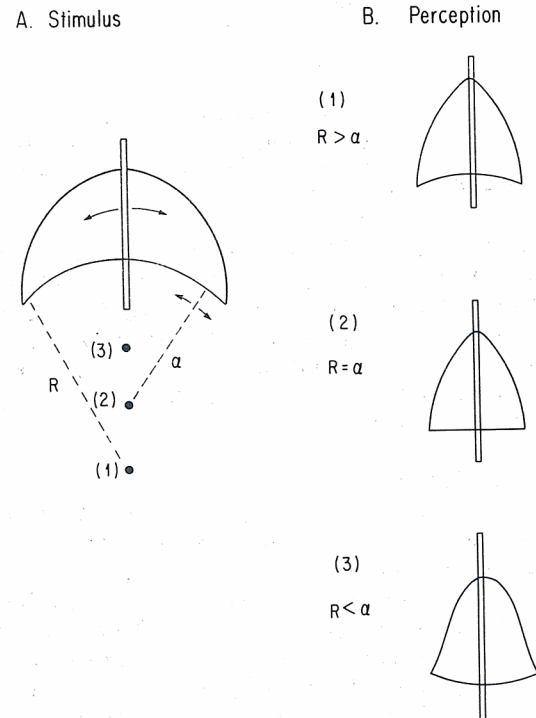


Figure 5. Rotation. The stimulus is shown in A. Three different centers of rotation of this shape give slightly different perceptions, as illustrated in B. The observations are consistent with the assumption that the shape behind the slit is undergoing a simple translation.

ted along the recovered direction of the motion to reconstruct the shape. Figure 5B shows a series of qualitative observations by the authors and four naive subjects. The topmost figure is the perception when the radius of the arc is shorter than the radius of rotation [(1) in Figure 5A]. There is an obvious distortion, but the arc still has the same sign of curvature. The middle figure illustrates the perception when the radius of the arc is exactly the same as that of the rotation [(2) in Figure 5A]. If the local curvature is completely neglected and the shape is recovered only by the stored positional information and the computationally recovered ϕ and V_ϕ , then the arc would be expected to appear as a straight line. Since the perceived line still has a slight curvature, this cannot be the case. The most interesting condition occurs when the radius of the arc is made longer than that of rotation [(3) in Figure 5A]. Now, as observed by all 6 subjects, the sign of the curvature becomes reversed, as shown in the bottommost figure of 5B. Thus, even if it is not totally negligible, the local curvature is not a major cue to the recovery of shape.³ The higher the velocity, the larger the shrinkage effect, just as in the case of translation (Zöllner, 1862). However, the shrinkage does not seem to change the qualitative aspects of the above-mentioned observations. These demonstrations have two implications:

1. The direction, ϕ , and the speed, V_ϕ , of the original motion are recovered computationally first, and then the stored positional information is used to reconstruct the shape. [If the shape is recovered independently of ϕ and V_ϕ , and the constraints ($\phi_1 = \phi_2$, $V_1 = V_2$) are required only for the recovery of motion, then the distortion would not be expected.]

2. The human system assumes the constraint of "constant direction and velocity" to recover ϕ and V_ϕ .

Thus, our basic idea is that ϕ and V_ϕ are recovered before the reconstruction of the shape, and that the "translation" constraint is imposed in the recovery process, as

suggested by Rock (1981) in his "problem solving" interpretation of the phenomenon. Note that this same constraint was also proposed by Fahle and Poggio (1984) to explain spatiotemporal hyperacuity.

ALGORITHM AND MECHANISM

Linearization

It is quite unlikely that a biological visual system will implement the trigonometric functions necessary to solve Equations 3 and 4 to obtain the translational motion of a rigid shape seen through a slit. Instead, some shortcut must be used to obviate the need for arc tangent, tangent, or cosine functions. (Division and simple arithmetic operations seem plausible, however.) A simple step would be to linearize the trigonometric functions. This linearization can be accomplished in several ways. For example, we could simply linearize the present trigonometric functions as they now appear in the two equations. Alternatively, we might break down the tangent function into a ratio of sines and cosines. Our choice is dictated in part by what seems physiologically plausible and by what will introduce as little error as possible. These issues are algorithmic and are independent of the preceding theoretical analysis showing the minimal conditions for recovering the hidden object's motion (Marr, 1982).

Figure 6 shows three approximations to the sine, cosine, and tangent functions we will use. In Figures 6A and 6B, the sine and cosine are approximated by simple linear functions (dotted curves). In Figure 6C, the tangent and arc-tangent approximations are ratios of the previous linear approximations. Our choice for linearization is to reformulate Equations 3 and 4 entirely in terms of sines and cosines, and then to linearize these functions. The new, linearized equations are now:

$$\phi = 90x/(1+x) \quad (5)$$

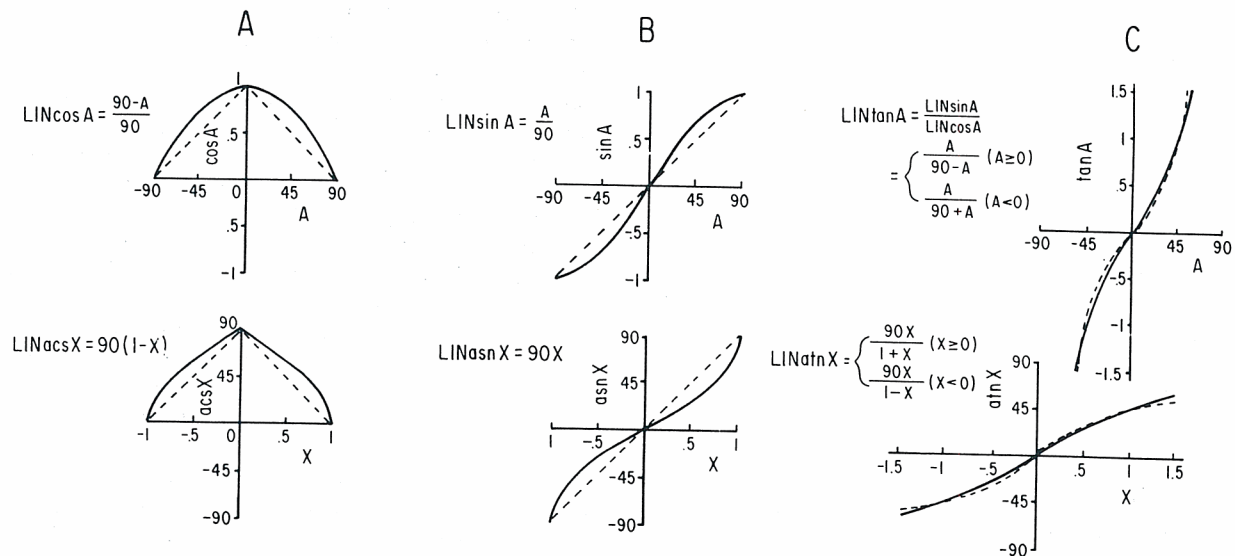


Figure 6. Linearization of sine, cosine, and tangent functions by simple angular approximations (A and B) or their ratios (C).

$$V_{\phi} = V_1^* \left/ \left[\frac{90 - \phi^*}{90} \left(\frac{\phi_*}{90 + |\phi^*|} + \frac{\theta_1^*}{90 - |\theta_1^*|} \right) \right] \right|, \quad (6)$$

where

$$x = \left[\frac{V_2^* \theta_1^*}{90 - \theta_1^*} - \frac{V_1^* \theta_2^*}{90 - \theta_2^*} \right] / (V_2^* - V_1^*). \quad (7)$$

A simulation using the above shows that the approximations do not degrade the calculation of the direction of the object motion by more than 5° (see Figure 7). When these linearized values for ϕ are then used to calculate V_{ϕ} according to Equation 6, the estimated speed ranges from 70% to 160% of the true value in most cases.⁴ On the average, there is about a 20% overestimation of V_{ϕ} , consistent with our observations for foveal viewing. (Note, however, that peripheral viewing can cause V_{ϕ} to be underestimated, suggesting sampling complications in the biological system.)

ERROR ANALYSIS

Sensitivity to Errors in Contour Orientation θ^*

No measuring system can provide perfect data. How robust are the solutions to Equations 3 and 4, or their linearized versions, Equations 5 and 6, in the presence of noise in inputs? What is required by our theory is a high sensitivity to the angle θ^* and to the contour rotation ($\theta_1^* - \theta_2^*$). If we perturb the observed contour angles, θ_1^* , or the observed speeds, V_1^* , how seriously will these errors affect the estimates of ϕ and V_{ϕ} ? The following simulation shows that perturbations of the order of psychophysical resolution do not affect the estimates of the hidden object's velocity much worse than the linearization.

A circle passing behind a slit at constant speed, V_{ϕ} , was used as the test object. The direction, ϕ , was 0°, 30°, or 60°. Because the noise effects are similar at all these

orientations, only the 30° results are shown here. The noise was added by perturbing θ^* by various angular amounts ranging from 0° to 40°.

Figure 7A shows the effect of angular errors in θ^* upon the recovered value of ϕ , where errors in ϕ are plotted on the abscissa. Two curves are given, one for the linearized equation (5) (dashed) and the other for the correct equation (3) (solid). For no noise, $\delta\theta = 0$, our discrete sampling technique gave an error of 0° without linearization and a 3° error with linearization, as illustrated by comparing the solid and dashed curves in Figure 7A. Over the entire range explored, the errors in θ^* affect the ϕ calculation with linearization in much the same way as the exact solution. (For $\phi = 0$, the linearized result was about 5° higher over this same range.) Also shown in Figure 7A, by the dotted curve, is the range of the errors in estimating ϕ . (This range results from the different orientations of the circle's contour.) Again, for a reasonable error in θ^* of 20° or less, the error in ϕ is inconsequential.

Figure 7B shows the compound effect of errors in θ^* upon the estimates of V_{ϕ} when the estimated ϕ^* value is used. Again, the solid curve shows the exact solution (Equation 4), with varying degrees of error in θ used as input, to be compared with the dashed curve, which is the linearized solution (Equation 6). Once again, the effects of linearization *on the average* are negligible considering the sensitivity of the human observer. However, the error in V_{ϕ} may become substantial for certain contour orientations, as shown by the dotted curve which gives the range of the V_{ϕ} error. For a 20° error in θ , we can expect certain orientations of the contour to result in as much as a 50% overestimate of V_{ϕ} . This analysis predicts that when the slit is narrower, the error in θ is larger and, therefore, the more V_{ϕ} should be overestimated, independent of whether the observer fixates the slit or is allowed to pursue the moving shape (Fujii, 1982).

Sensitivity to Elongation

As an object becomes increasingly elongated, the sample of contour orientations available for estimating the translation vector V_{ϕ} becomes limited. Consider a simple ellipse. As illustrated in Figure 8, as the ellipse becomes more and more elongated, eventually it degenerates into a line. But a line moving across a slit will appear to move perpendicularly to its orientation (Wallach, 1976). Thus, because the human visual system has limited temporal sampling of the object's contour as it moves behind the slit, we expect that an ellipse becomes more and more eccentric and the observer's estimate of the direction of movement, ϕ , will be increasingly biased to the minor axis of the ellipse. Figure 8 shows the result of running our simulation on ellipses of various eccentricities. Also shown are data obtained from 3 observers. The fact that the human data deteriorate more rapidly than our simulation suggests that the sampling used to generate Figure 7 is a bit too fine. Alternatively, the system may adopt a compromise between the computational output and a

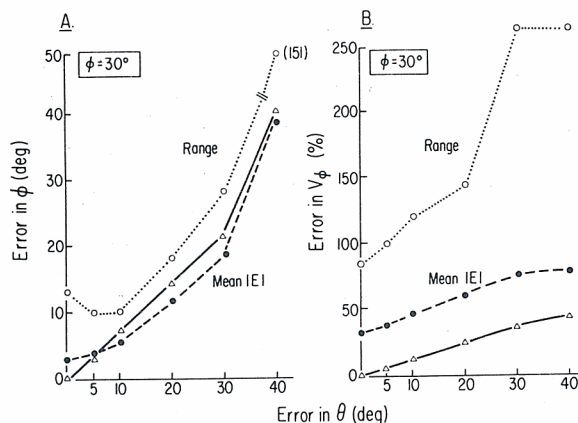


Figure 7. Error analysis. Solid curves are errors in θ introduced into the exact equations 3 and 4. The dashed curve introduces the same errors into the linearized equations 5 and 6. Linearization has little effect on the result. The dotted curve shows the range of errors for linearized results.

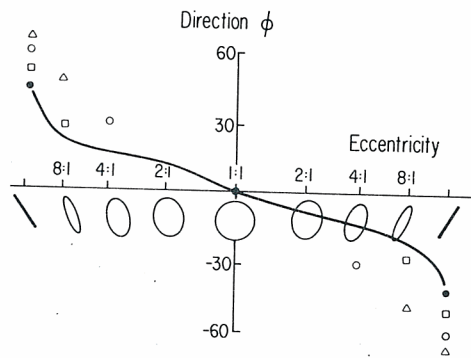


Figure 8. An ellipse with its major axis oriented at 45° to the vertical is moved in the horizontal direction behind a vertical slit. As the outline becomes more and more elongated, the predicted direction of apparent motion moves toward 45° , as shown by the curve. (The sampling distance was 0.5 for a disk of radius 10.) Data obtained from 3 subjects are shown for comparison.

Wallach-like solution along the slit. The predicted range of errors in V_ϕ and ϕ may therefore be underestimated in some cases.

In summary, we conclude that both linearization and errors less than 20° in estimating θ^* will have negligible effects on recovering ϕ and V_ϕ provided that the aspect ratio of elongation of an object is less than 2:1.

DISCUSSION

Our analysis has several implications for the recovery of shape from motion. (1) Strict local correspondence is not required, since the (two) tangents needed for the recovery process appear at different parts of the contour. (2) A velocity field need not be constructed, since only contour speed parallel to the slit is required. (3) The human observer imposes the constraint that unseen object motion is the translation of a rigid body at constant speed (see Fahle & Poggio, 1984). (4) The human mechanism can use a simple linearization of the explicit nonlinear, mathematical solution. (5) Temporal sampling along a single contour may also be used. These points will be elaborated below.

Local Correspondence

A most significant result of our analyses is that local correspondence of identical pieces of the contour is not required to recover the shape's motion. This is counterintuitive and against many traditional theories of the phenomenon (but see comments by Hochberg, 1968). Figure 9 shows three types of sampling. Local feature points (e.g., a "corner") are presented more than once at slightly different positions in the slit in the case illustrated in Figure 9A. Thus, local correspondence cues are given. When the temporal sampling rate is coarser, as shown in Figures 9B and 9C, the feature points are presented only once in the slit, and therefore the same contour is not available for matching. In Figures 9C and 9C', some parts of the contour are never presented in fact.

Yet a human observer will perceive the two-dimensional moving object even in such cases as B, C, and C'. (Interestingly enough, the perceived shapes were not only shrunk, but also broken in C and C'.)

The Velocity Field

Slit viewing can be regarded as a problem of recovering a two-dimensional shape and motion from limited, almost one-dimensional shape and motion information. When a two-dimensional moving pattern is presented in a limited visual field, a similar correspondence problem is raised, and the ambiguity of the shape and motion is left unsolved (the "aperture problem"). Hildreth (1984a, 1984b) and Waxman and Wohn (1985) have suggested that the constraint of "smoothness" enables the human system to compute a unique instantaneous two-dimensional velocity field from local velocity measurements. This can be done iteratively by minimizing the velocity variations along the contour, or by examining the deformation of the contour. Our study of slit viewing suggests that the inference process may also be accomplished by a noniterative computation based on speed in a fixed direction (direction of the slit) and orientation changes of the contour.

The Translation Constraint

Eye-movement and "retinal painting" theories have been proposed to explain the slit view phenomenon (Haber, 1968; Haber & Nathanson, 1968; Morgan, Findlay, & Watt, 1982). The "retinal painting" theory basically states that when the eyes move to follow the shape behind the slit, the successive "painting" of the image on the retina is sufficient to explain the phenomenon, and

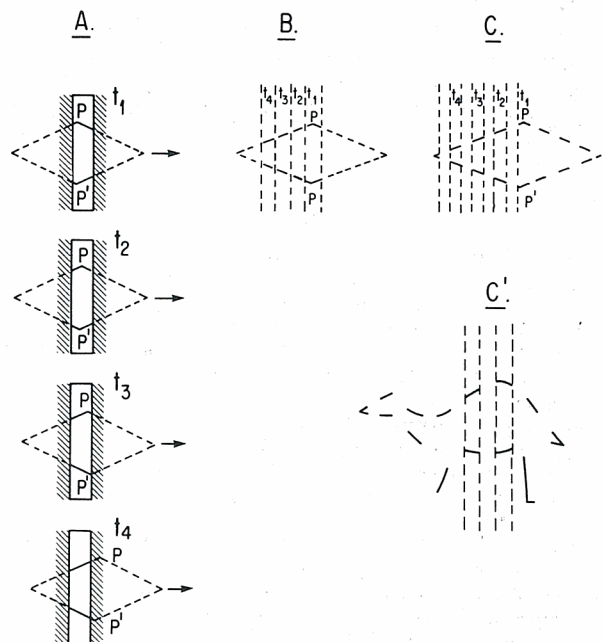


Figure 9. Three types of sampling.

that any further computation or processing is not required. Thus, it assumes implicitly that the direction and speed of motion is somehow known and used to recover the shape of the object. This cannot be true in most cases (Rock, 1981). Even to move the eyes appropriately for "retinal painting," the direction and speed of motion should be recovered beforehand. The primary problem for slit viewing is the recovery of motion, for without this the shape of the object cannot be recovered. We show that a minimal condition for the recovery of the occluded shape is that it undergo a translation, in which case only two "views" of tangents to the rigid contour are required.

Shape Distortion and Integration

One aspect of slit motion not previously discussed is the distortion in the perceived shape (Anstis & Atkinson, 1967; Zöllner, 1862). Usually, this distortion is a shrinkage, such as would result if the speed were underestimated (Rock, 1981; Rock & Sigman, 1973). However, we have noted that peripheral viewing can lead to an elongation—a new observation, as far as we know. At present, we do not have a satisfactory explanation for these effects. Figure 8 suggests that they arise at least in part from sampling artifacts. However, biases in the orientation or motion detection could also contribute.

Finally, nothing has been said about how the contour segments, once assigned their correct directions of motion, are integrated into a coherent whole shape. Instead, we have dealt only with the relatively simple first stage of the recovery process, namely the recovery of the direction and speed of motion of an unknown shape moving behind a slit. However, now that the conditions for recovery of the object's motion have been specified more exactly, perhaps possible schemes for assembling the image segments into a whole shape will become clearer. For example, at least we know now that any theory of the integration process—be it retinal painting, postretinal storage, or whatever—should proceed as if the object was undergoing a simple translation.

REFERENCES

- ADELSON, E. H., & MOVSHON, J. A. (1983). The perception of coherent motion in two-dimensional patterns. In *Proceedings, ACM Workshop on Motion: Representation and Perception* (pp. 11-14). Toronto: Association for Computer Machinery, New York.
- ANDREWS, D. P. (1967). Perception of contour orientation in the central fovea. Part I: Short lines. *Vision Research*, 7, 975-997.
- ANDREWS, D. P., BUTCHER, A. K., & BUCKLEY, B. R. (1973). Acuties for spatial arrangement in line figures: Human and ideal observers compared. *Vision Research*, 13, 599-620.
- ANSTIS, S. M., & ATKINSON, J. (1967). Distortions in moving figures viewed through a stationary slit. *American Journal of Psychology*, 80, 572-585.
- BORING, E. G. (1942). *Sensation and perception in the history of psychology*. New York: Appleton-Century-Crofts.
- FAHLE, M., & POGGIO, T. (1984). Visual hyperacuity: Spatiotemporal interpolation in human vision. In S. Ullman & W. Richards (Eds.), *Image understanding 1984*. Norwood, NJ: Ablex.
- FENNEMA, C. I., & THOMPSON, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics & Image Processing*, 9, 301-315.
- FRISBY, J. P. (1980). *Seeing: Illusion, brain and mind*. Oxford: Oxford University Press.
- FUJII, T. (1982, May). *Phenomenal distortion of moving shapes*. Paper presented at the First Study Meeting of the Japanese Psychonomic Society, at Keio-Gijyuba University, Tokyo.
- HABER, R. N. (1968). Perceptual reports vs. recognition responses: A reply to Parks. *Perception & Psychophysics*, 4, 374.
- HABER, R. N., & NATHANSON, L. S. (1968). Post-retinal storage? Some further observations on Parks' camel as seen through the eye of a needle. *Perception & Psychophysics*, 3, 349-355.
- HILDRETH, E. C. (1984a). Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23, 309-354.
- HILDRETH, E. C. (1984b). The computation of the velocity field. *Proceedings of the Royal Society, London, B*, 221, 189-220.
- HOCHBERG, J. (1968). In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception* (pp. 309-331). New York: Holt, Rinehart and Winston.
- HUBEL, D. H., & WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* (London), 166, 106-154.
- HUBEL, D. H., & WIESEL, T. N. (1968). Receptive fields and functional architecture. *Journal of Physiology* (London), 195, 215-243.
- MARR, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.
- MARR, D., & ULLMAN, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society, London, B*, 211, 151-180.
- MORGAN, M. J., FINDLAY, J. M., & WATT, R. J. (1982). Aperture viewing: A review and a synthesis. *Quarterly Journal of Experimental Psychology*, 34A, 211-233.
- 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 (Eds.), *Pattern recognition mechanisms*. Rome: Vatican Press.
- NAKAYAMA, K., & SILVERMAN, G. H. (1983). Perception of moving sinusoidal lines. *Journal of the Optical Society of America*, 73, 1862A.
- PARKS, T. E. (1965). Post-retinal visual storage. *American Journal of Psychology*, 78, 145-147.
- POGGIO, T., & TORRE, V. (1984). Ill-posed problems and regularization analysis in early vision (*AI Memo 773*). Cambridge, MA: MIT.
- RICHARDS, W. A., RUBIN, J. M., & HOFFMAN, D. D. (1983). Equation counting and the interpretation of sensory data. *Perception*, 11, 557-576.
- ROCK, I. (1981). Anorthoscopic perception. *Scientific American*, 244(3), 145-153.
- ROCK, I., & SIGMAN, E. (1973). Intelligence factors in the perception of form through a moving slit. *Perception*, 2, 357-369.
- SCHILLER, P. H., FINLEY, B. L., & VOLKMAN, S. F. (1976a). Quantitative studies of single cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *Journal of Neurophysiology*, 39, 1288-1319.
- SCHILLER, P. H., FINLEY, B. L., & VOLKMAN, S. F. (1976b). Quantitative studies of single cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance. *Journal of Neurophysiology*, 39, 1320-1333.
- WALLACH, H. (1976). On perceived identity: 1. The direction of motion of straight lines. In H. Wallach (Ed.), *On perception*. New York: Quadrangle.
- WATT, R. J., & ANDREWS, D. P. (1982). Contour curvature analysis: Hyperacuties in the discrimination of shape. *Vision Research*, 22, 449-460.
- WAXMAN, A. M., & WOHN, K. (1985). Contour evolution, neighborhood deformation and image flow: Planar surfaces in motion. *International Journal of Robotics Research*, 4, No. 3, 95-108.

WILSON, H. R. (1985). Discrimination of contour curvature: Data and theory. *Journal of the Optical Society of America A*, 2, 1191-1199.
 ZÖLLNER, F. (1862). Über eine neue Art anorthoskopischer Zerrbilder. *Annalen der Physik*, 117, 477-484.

NOTES

1. There is no loss in generality by discretizing the continuous function as long as the sampling rate is twice that of the Nyquist frequency of the measuring device. In this case, a rate of 60 Hz will suffice for the human observer.

2. For motion parallel to the slit, $\phi = 90^\circ$ and Equation 4 degenerates to $V_\phi = V^*$.

3. This observation is consistent with Waxman's analysis of the recovery of shape from "evolving" contours (Waxman & Wohn, 1985).

4. Instead of solving first for ϕ and then using this value to compute V_ϕ , we could have used a regularization procedure to solve for both simultaneously (Hildreth, 1984; Poggio & Torre, 1984). Although we favor our scheme, it remains an empirical issue to test between these two methods in finding ϕ and V_ϕ .

APPENDIX 1

Derivation of the speed, V_ϕ , and direction, ϕ , of object motion (refer to Figures 3 and 10 for definitions):

Let V_1^* be the observed velocity of a contour having an orientation θ_1^* at t_1 and θ_2^* at t_2 , a short time, $\Delta t = (t_2 - t_1)$, later. Then, if the contour appears to move a vertical distance, ΔD , in the slit, the observed velocity, V_1^* , will be

$$V_1^* = \Delta D / \Delta t. \quad (A1)$$

In this same period, the original piece of contour will move from point O to point A through a horizontal distance ΔS_1 as shown in Figure 10. Its (unknown) horizontal rate of motion will be

$$\frac{\Delta S}{\Delta t} = V_\phi \cos \phi. \quad (A2)$$

Combining Equations A1 and A2,

$$V_1^* = (V_\phi \cos \phi) (\Delta D / \Delta S). \quad (A3)$$

Our strategy for recovering the ratio $\Delta D / \Delta S$ is to solve for the two segments BE and EO which constitute ΔD . To begin, we approximate the piece of contour by the arc of a circle. (This approximation becomes exact as $\Delta t \rightarrow 0$, but will lead to er-

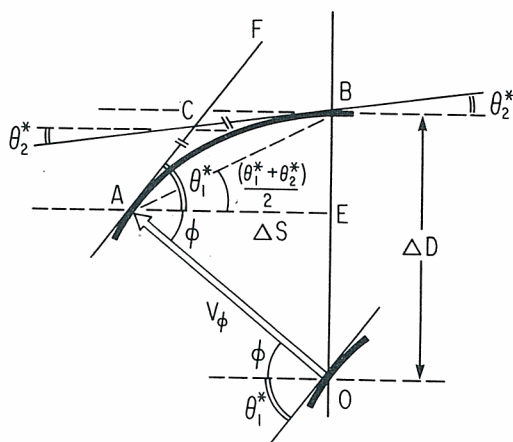


Figure 10. Geometrical relationship and definition of angles and distances used in the derivation of Equation A7 in Appendix 1.

rors if the observer's sampling rates are too low.) Note from Figure 10 that this approximation forces $AC = CB$. Because $\angle FCB = (\theta_1^* - \theta_2^*)$, we see that

$$\angle CAB = \angle CBA = [\pi - (\pi - \theta_1 + \theta_2)] / 2 = \frac{\theta_1 - \theta_2}{2}. \quad (A4)$$

Consequently, we can find $\angle BAE$ because $\angle BAE = \theta_1^* - \angle CBA$:

$$\begin{aligned} \angle BAE &= \theta_1^* - (\theta_1^* - \theta_2^*) / 2 \\ &= (\theta_1^* + \theta_2^*) / 2. \end{aligned} \quad (A5)$$

The segments BE and EO that compose ΔD are now easily calculated:

$$\Delta D = \Delta S \tan \phi + \Delta S \tan (\theta_1^* + \theta_2^*) / 2$$

or

$$\frac{\Delta D}{\Delta S} = -\tan \phi + \tan (\theta_1^* + \theta_2^*) / 2. \quad (A6)$$

Hence,

$$V_1^* = V_\phi \cos \phi [-\tan \phi + \tan (\theta_1^* + \theta_2^*) / 2], \quad (A7)$$

which is the text Equation 2.

We now proceed to solve for V_ϕ and ϕ . Because we have one equation in two unknowns, we require at least two such "views." Let the contour orientations, θ_j^* , be indexed to indicate both the view (j) and the contour position (j). We then have two equations with knowns V_1^* , θ_1^* , θ_2^* , and V_2^* , θ_1^* , θ_2^* , which can be solved for V_ϕ and ϕ to yield text Equations 3 and 4.

We reiterate that these equations are valid only for short Δt 's. Also, these solutions only demonstrate that *in principle*, the angle ϕ and speed V_ϕ are recoverable using the "rigidity" and "constant speed" constraints. It is unlikely that our visual system "solves for" ϕ using the equations as presented. This is an algorithmic issue, addressed by the linearization scheme presented in the Algorithm and Mechanism section.

APPENDIX 2

If an object may be rotated behind the slit, then our unknowns are the angular speed, ω , and the center of rotation, x_0 , y_0 . Clearly, with three unknowns, we will need at least three "views." Consider, however, the unknown speed, V_ϕ , which is now a function of how far the contour is from the center of rotation. From Figure 11, we see that

$$V_{\phi 1} = L_1 \omega$$

or, in general, the rotation constraint requires

$$V_{\phi i} = L_i \omega. \quad (A8)$$

Our text Equation 2 thus becomes

$$V_1^* = L_1 \omega \cos \phi_1 \left[-\tan \phi_1 + \tan \frac{\theta_1^* + \theta_2^*}{2} \right], \quad (A9)$$

where the unknowns are L_1 , ω , and ϕ_1 . Unfortunately, the angle ϕ_i , which the velocity vector $V_{\phi i}$ makes with the slit, is also not constant. A second "view" will thus add unknowns L_2 and ϕ_2 . However, the displacement D_2^* is known and can be related to ϕ_1 , ϕ_2 and L_1 , L_2 .

Note that, because $V_{\phi 2}$ is perpendicular to the radial line $L_2 = OB$, $\angle ABO = \pi/2 - \phi_2$. Hence $\angle AOB = \phi_2$, as shown.

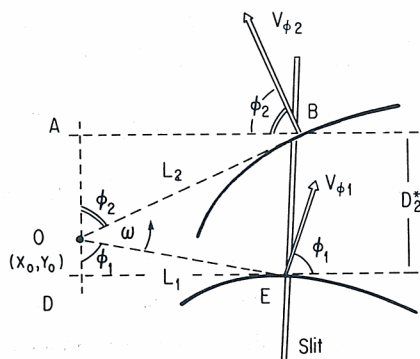


Figure 11. Geometrical relationship and definition of angles and distances used in the derivation of Equations A9 and A10.

Similarly, $\angle DOE = \phi_1$. Hence, we have a new relationship, namely,

$$D_2^* = L_1 \cos \phi_1 + L_2 \cos \phi_2. \quad (\text{A10})$$

But, as the horizontal distance from the center of rotation O to the slit is fixed, we also have

$$L_1 \sin \phi_1 = L_2 \sin \phi_2. \quad (\text{A11})$$

Hence, with two views, we have four equations in five unknowns, L_1 , L_2 , ϕ_1 , ϕ_2 , and ω . Adding one more "view" adds two more unknowns, L_3 and ϕ_3 , and three equations. Hence, three "views" are required to "solve for" the rotation center and angular speed ω .

Alternatively, we might have considered measuring three tangents to the contours θ_1^* , θ_2^* , and θ_3^* , for then two "views" would result in four equations of the form of Equation A9 in unknowns L_1 , L_2 , ϕ_1 , ϕ_2 , and ω . Equations A10 and A11 would then cause the number of equations to exceed the number of unknowns. Such a scheme fails the Jacobian test, however, and will not yield a solution in this case (Richards, Rubin, & Hoffman, 1983). The reason is that Equation A9 assumes an arc of constant curvature, and hence the third tangent provides no new information. To succeed, the arc must be a biquadratic (Waxman & Wohn, 1985), in which case the number of unknowns increases by two for each view.

(Manuscript received December 23, 1985,
revision accepted for publication May 5, 1986.)