Max-Planck-Institut
für biologische Kybernetik
Arbeitsgruppe Bülthoff

Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 27 February 1996

Active Kinetic Depth Effect

Pascal Mamassian and **Heinrich H. Bülthoff**

Abstract

We investigated the motion perception of an occluding contour by an active observer. The task of the observer was to decide whether the occluding contour of an object was a sharp edge in depth, or else belonged to a smooth surface. In a first experiment using real objects, we found a strong bias to label the contour smooth. We then replicated the experiment with virtual objects in order to carefully control for the appearance of the objects, and found an advantage for active over passive viewing. Finally, we did not find any improvement of performance when the scene was observed binocularly.

We would like to thank Andrew Blake and Daniel Kersten for thoughtful discussions. PM was supported by a stipend from the Max-Planck Society. A preliminary report of this study was presented at ARVO Annual Meeting, in May 1994, in Sarasota, Florida. Please direct all correspondence to Pascal Mamassian, 6 Washington Place (8th Floor), New York University, New York NY 10003, USA. Email: `pascal@cns.nyu.edu`, `hhb@mpik-tueb.mpg.de`. Manuscript to be submitted to *Perception & Psychophysics*. Please do not cite without permission.

1 Introduction

Since the original work of Wallach and O'Connell (1953), the research on the perception of three-dimensional structure-from-motion has often relied on two working hypotheses. The first one is that motion is produced by identifiable features in the scene which can be tracked over time (cf. Sperling et al., 1989). These features are taken to be surface markings, such as texture elements (Ullman, 1979) or contour edges (Hildreth, 1983). One can then define an optical flow as the velocity field of these features in the image. The second hypothesis prevailing structure-from-motion research consists in disregarding the distinction between object moving and observer motion. As a consequence, the optical flow confounds two kinds of information: about the structure of the environment and about the trajectory of the observer (e.g. Waxman and Ullman, 1985). In this paper, we are interested in the perception of structure-from-motion in case neither of these two working hypotheses are satisfied.

To start with the first hypothesis, the lack of correspondence from one view to the next presents a real challenge for both computational and biological systems. This problem arises naturally for smooth textureless objects, or for glossy objects whose highlights are not fixed features on the surface (cf. Blake and Bülthoff, 1991). ¿From a computational standpoint, the motion of smooth objects has been seriously considered only recently (Giblin and Weiss, 1987; Cipolla and Blake, 1992; Vaillant and Faugeras, 1992; Kutulakos and Dyer, 1994). This issue has also received a recent interest in psychophysics (Todd, 1985; Cortese and Andersen, 1991; Pollick et al., 1992). Wallach and O'Connell (1953) first noticed that the shadow of a rotating rigid object appeared like a flat deforming figure whenever the object did not provide any trackable features (e.g., an ellipsoid). The common explanation for this percept is illustrated in Figure 1. While a few feature points are sufficient to infer a rigid surface which would support them (Figure 1a), the occluding boundary of a smooth object is unreliable since it depends on the viewpoint (Figure 1b). In general, one point belonging to the occluding boundary for one object's orientation will no longer belong to the occluding boundary for a slightly different object orientation. Therefore, an ellipsoid in rotation will not produce a three-dimensional percept because, it is believed, the occluding boundary is not a fixed curve on the object.

Concerning the second hypothesis, the knowledge of the observer's motion is now believed to present a clear advantage to infer the structure of the environment. In particular, the recovery of the absolute distance of objects is impossible if one is ignorant about his own motion. There is now good evidence that several animal species displace their head up and down (Goodale et al., 1990), sideways (Sobel, 1990), or even their whole body (Lehrer et al., 1988) in order to measure the absolute distance of objects. For human observers, the computation of certain environment attributes from optical flow can be more accurate if the motion of the observer is known. For instance, relative depth from motion parallax appears to be more veridically estimated when the observer is actively moving than when he is passively looking at a similar optical flow (Rogers and Graham, 1979). This facilitation seems to be the consequence of some proprioceptive feedback (e.g. Roll et al., 1991).

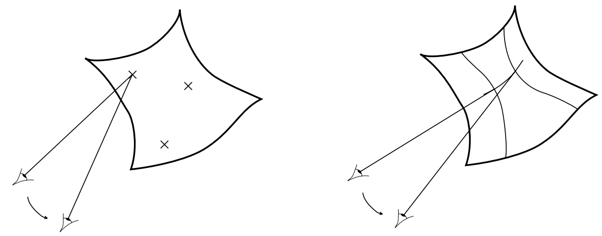

Figure 1: Recovering the structure of an object when it is in motion relative to the observer is a different problem whether or not the object has a texture. For textured objects, one can rely on the fact that the texture elements are fixed on the surface (a). For textureless surfaces however, the only information available is the deformation of the occluding contour (b).

In this paper, we have investigated how an active observer would perceive the local structure of an object from the motion at one of its edges. By active observer, we refer to the viewing condition where one is actively engaged in the exploration of his environment. On the contrary, passive vision corresponds to a static observer with objects moving around him. We report an experiment which suggests that the moving silhouette of smooth objects can be interpreted as the projection of three-dimensional objects, under the condition that the observer is actively moving. Before detailing the design of the experiment, we outline the theoretical background which motivated this study.

2 Background

In this section, we describe how a surface described only from its occluding contour differ from a textured surface. We briefly determine what the information in the image is, and how this information can be retrieved. We also discuss how motion and binocular cues can be combined for these textureless surfaces.

2.1 Surface Curvature from Motion

We consider here the case of a textureless surface for which only the occluding contour is visible. Surface shape and curvature at the occluding contour can still be computed when the observer is in motion around the object (Giblin and Weiss, 1987; Cipolla and Blake, 1992). To understand why the knowledge of the observer's motion is so critical, we first outline the information available from a static image.

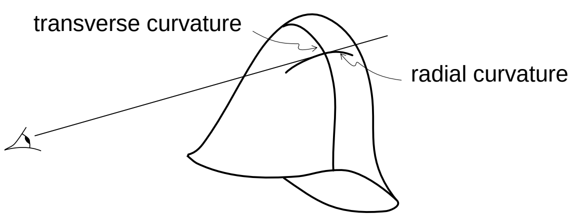

Figure 2: The surface curvature at one point belonging to the occluding contour can be decomposed into radial and transverse curvatures.

The local solid shape at the occluding boundary can be decomposed as illustrated in Figure 2. First, the occluding boundary on the surface projects as the occluding contour in the image, and the curvature of one point on this contour is called *transverse curvature*. Second, the viewing direction and the surface normal at one point of the occluding boundary define a plane and the curvature of the intersecting curve of this plane with the surface is called the *radial curvature*. As shown by Koenderink (1984), the product of the radial curvature by the transverse curvature determines the Gaussian curvature, a qualitative descriptor of local solid shape.

While the transverse curvature is directly available from the image, the radial curvature is completely ambiguous, except for its sign which determines on which side of the occluding contour the object lies. The radial curvature is therefore the critical attribute to estimate in order to recover the local solid shape of the surface. Cipolla and Blake (1992) showed that the radial curvature could be recovered if the observer motion is known. More precisely, the computation of the radial curvature from the occluding contour motion requires the knowledge of the observer's acceleration.

2.2 Motion Parallax of the Occluding Contour

The computation of the radial curvature from the occluding contour motion is very sensitive to small errors in the estimate of the observer's motion. This is due to the fact that the computation relies on the knowledge of the observer's acceleration, and that high-order derivations are ill-conditioned. However, under certain conditions, it is possible to significantly improve the robustness of the computation by relying only on the observer's velocity and no longer on his acceleration (Cipolla and Blake, 1992). The idea underlying such an alternative computation is based on the well-known *motion parallax* principle. In its classical formulation, motion parallax provides a way to compute the relative depth of nearby points by comparing their velocities (Koenderink, 1986). Similarly, the motion parallax of the occluding contour enables one to compute the surface curvature by comparing the velocities of the contour and of a fixed feature on the surface.

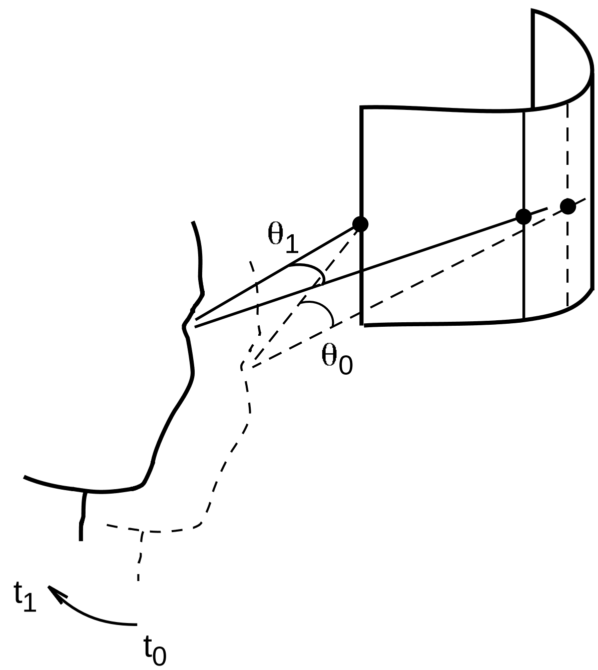

Figure 3: Motion parallax can be used to enhance the robustness of the surface curvature computation. As shown on this figure, the presence of a sharp edge should theoretically improve the estimation of the radial curvature of a nearby occluding contour.

The motion parallax of the occluding contour relative to a fixed feature is illustrated in Figure 3. The fixed feature on the surface is here taken to be a sharp edge (a discontinuity of the surface orientation). As opposed to the occluding boundary, the location of a sharp edge is independent of the viewer position, and can therefore be used as a reference landmark. As the observer moves, the rate at which the sharp edge and the occluding boundary approach each other in the image is inversely related to the radial curvature of the surface, and this relation no longer involves the observer's acceleration (Cipolla and Blake, 1992). More intuitively, the faster the occluding contour approaches the sharp edge, the smaller is the surface curvature.

2.3 Best Observer's Motion

When the observer is actively moving around an object, one can wonder whether there is an optimal trajectory to extract as accurately as possible the structure of the object. Consider the simple case of a vertically-oriented cylinder whose cross-section is elliptical. The best observer's trajectory to recover the curvature of the cylinder depends on whether it is textured or not. Droulez and Cornilleau-Pérès (1990) defined the *spin variation* which represents the amount of stretch in the image of a textured surface patch when it undergoes a particular motion. The curvature of our cylinder will be most accurately computed when the spin variation is large, and this was confirmed by experimental evidence (Cornilleau-Pérès and Droulez, 1989). In order to maximize the spin variation, the best strategy of an observer will then be to translate along the axis of the cylinder, or equivalently, to rotate about the horizontal axis (Figure 4a).

The best strategy to recover the structure of a textureless surface is markedly different. Since the information available is now the motion of the occluding contour, the curvature of the surface will then be most accurately computed when the displacement of this contour is large. To take the example of our cylinder, the best trajectory of the observer is a translation perpendicularly to the edge of the cylinder, or equivalently, a rotation about the cylinder axis (Figure 4b). Two complementary strategies are therefore underlying the recovery of the structure of moving objects when these are textured or not.

2.4 Interaction of Motion and Stereo

A final difference between textured and textureless surfaces concerns the way motion and stereo interact. While there is now good evidence that a textured surface is more accurately perceived when both motion and stereo are available to the observer (cf. Cornilleau-Peres and Droulez, 1993; Landy et al., 1995), such an interaction might be absent for textureless surfaces.

One reason why motion and stereo would not cooperate is that these two cues will usually provide inconsistent information about the structure of the surface. It is easy to see that from the two images of binocular viewing, a contour can always be interpreted as originating from a sharp edge (Figure 5a). On the other hand, we have already discussed how the motion of an occluding contour (taking into account three or more views) can be used to infer the correct surface curvature at that location (Figure 5b). Consequently, a smooth occluding boundary is likely to be interpreted differently from stereopsis and motion parallax. This is one of the rare instances of inconsistent information provided by two cues in a noiseless situation.

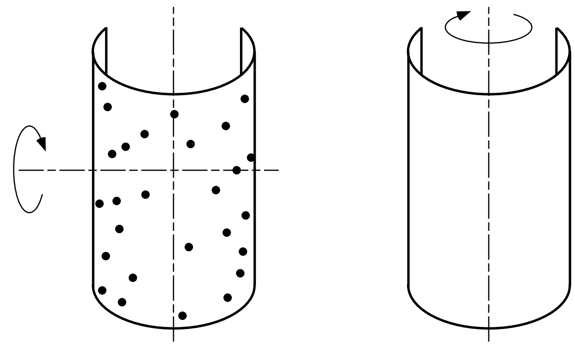

Figure 4: The best observer's motion to determine the curvature of a cylinder (with a non-circular cross-section) depends on whether it is textured. If the cylinder is textured, the proper observer's motion is a rotation perpendicular to the axis of the cylinder (a). If only the silhouette of the cylinder is visible, the proper observer's motion is a rotation about the axis of the cylinder (b). (We assume that neither the top nor the bottom of the cylinder are visible.)

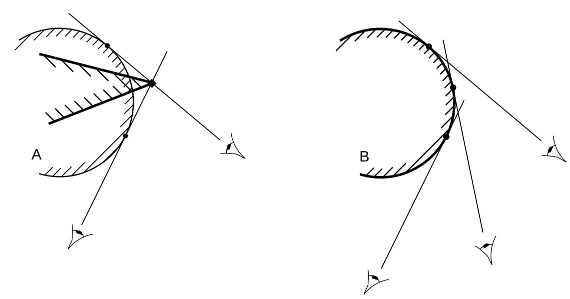

Figure 5: This figure illustrates that stereopsis and motion parallax are conflicting cues to estimate the nature of an edge. From only two views, a smooth wedge can always be interpreted as a sharp one (a). ¿From three or more views, the correct radial curvature of the surface can be estimated (b).

If the human visual system uses both stereopsis and motion parallax to determine the radial curvature of an occluding contour, we might then expect a drop in performance when both cues are simultaneously used. We have studied this question by presenting the stimulus either under egomotion alone, stereopsis alone, or both together.

3 Experiment I: Real Objects

We now investigate the ability of human observers to perceive the surface curvature from the occluding contour motion. We describe here a first experiment where the stimuli were real objects; in the next section, a virtual environment will be used instead. Observers had to discriminate between sharp and smooth edges in depth of a vertical wedge placed in front of them. Four viewing conditions were explored, the observer being either passive or active, and the object being seen either monocularly or binocularly.

3.1 Methods

Stimuli

The stimuli consisted of a set of five triangular wedges oriented vertically (Figure 6). The wedges differed in that one of their three edges was more or less smooth. The smoothness of this particular edge was manipulated by varying its radial curvature, from 0 to 4 mm in steps of 1 mm.

The wedges were otherwise matched for their width and depth extents to avoid to provide an alternative cue to distinguish them. Care was also taken to reduce as much as possible any texture or shading cue on the object. For this purpose, the wedges were built out of finely polished aluminium, and covered with two layers of black matte paint to obtain a uniform reflectance.

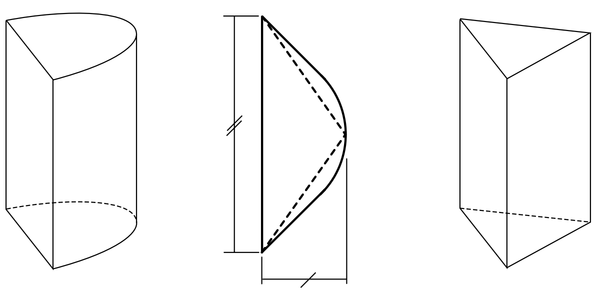

Figure 6: The left figure shows an example of a "smooth" wedge, and the right figure the "sharp" wedge. As indicated by the middle figure, the objects were matched for their width and depth extents (10 and 20 mm respectively). Relative to this latter figure, the observer was looking from the lower left.

Apparatus

To reduce the shading on the surface, the experimental room was painted in matte black, and all surfaces in view of the object were covered with black cardboard to attenuate the effects of mutual illumination. The wedges were presented in front of a computer monitor in an otherwise dark room. The monitor displayed a random texture background to enhance the discrimination of the edge motion. Finally, a reduction screen was placed in front of the object, so that neither the top nor the bottom of the wedge could be seen (Figure 7).

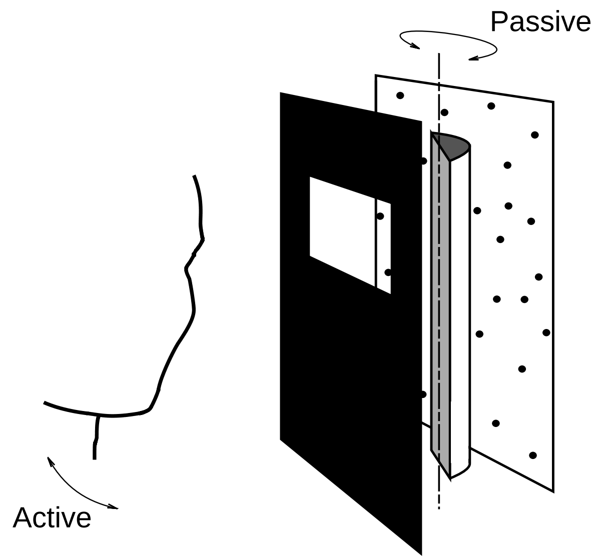

Figure 7: The observer was looking at the object from behind a reduction screen. In the passive conditions, the wedge was rotating about an axis located on its base. In the active conditions, the wedge was stationary and observers was asked to move from left to right.

Procedure

One of the five wedges was presented randomly, and the task of the subject was to decide whether the right-most edge of the object was *sharp* or *smooth*. The sharp object (i.e. the one with a radial curvature equal to 0 mm) was presented 12 times, while the smooth objects (i.e. the ones with a radial curvature between 1 and 4 mm) were presented 6 times each. The subjects were told before starting the experiment that sharp and smooth objects would not be presented equally often.

Four viewing conditions were run as a blocked factor in a randomized block design. Viewing was either passive or active, and either monocular or binocular. In the passive condition, the observer was instructed to stay still, and the object was rotating about a vertical axis passing through its base (cf. Figure 7). The motion was sinusoidal (frequency 1 Hz, amplitude 23 deg) and controlled by an oscilloscope connected to a step-motor. In the active condition, the object was static, and the observer was asked to move his upper body to the left and to the right. Viewing distance was approximately 50 cm. No time limit was imposed, but each trial was about 4 sec long. The subjects did not receive any feedback during the experiment.

Subjects

Seven observers participated in this experiment (between 24 and 34 years old). All subjects were naive relative to the purpose of the experiment, and had normal or corrected-to-normal visual acuity.

3.2 Results and Discussion

The objects differed by the radius of radial curvature of their right-most edge. As expected, as the radial curvature radius increased, the object was perceived more often smooth by the seven subjects (Figure 8). However, the sharp wedges were not always judged to be sharp, but instead were perceived smooth half of the times. In other words, there was a large bias to assume an object to be smooth since the data points rarely go below chance level (50%).

We fitted the data with a cumulative Gaussian with two degrees of freedom. Because of the smoothness bias, we imposed that the cumulative Gaussian varied in the interval $[0.5, 1]$ instead of $[0, 1]$. The best fits for the four conditions are shown superimposed on Figure 8. A threshold between smooth and sharp decisions can then be obtained as the radius value which corresponds to an object perceived 75% of the times smooth (75% is the point of subjective equality when the bias is taken into account). These threshold radii varied between 3.0 and 3.8 mm for the four viewing conditions (for a viewing distance of about 50 cm). Even though there was a trend that binocular viewing led to a smaller threshold radius, the four thresholds were not significantly different from each other.

The finding that performance was not the worst in the passive monocular condition is rather disturbing. Following the analysis of Cipolla and Blake (1992) outlined in the introduction, it should be quasi-impossible to differentiate sharp from smooth wedges when the observer's motion is unknown. Before calling the analysis of Cipolla and Blake into question, one should wonder whether the observers could use some other cues available in the stimuli and whether the simulated observer's motion was really ambiguous. It is indeed possible that the observers used the fact that the object's motion was identical in all trials of the passive viewing condition, and identified the sharp wedge as being the one for which the occluding contour moved the slowest. Since the subjects were naive and did not receive any feedback on their performance, it is however unlikely that the subjects managed to discover and use this strategy.

Alternatively, it is also possible that in spite of our precautions some residual shading was visible on the object, or that observers were able to use a change in accommodation to the edge moving in depth. In order to control for the potential effects of these undesired properties of our stimuli, we decided to replicate the experiment using virtual objects instead of real ones.

4 Experiment II: Virtual Objects

Because of the complexity of the task of controlling precisely the appearance of a real object, the use of computer graphics is appealing. In the previous experiment where real objects were observed, the estimate of the radial curvature of an edge of the object could have indeed been facilitated by some cues other than the motion of the edge. The psychophysical experiment described in this section was instead performed by simulating the presence of a virtual object in front of the observer. All four viewing conditions of the previous experiment have been replicated. In particular, the active and binocular viewing conditions required the use of some special equipments detailed below.

4.1 Methods

Stimuli

The objects simulated in this experiment were very similar to the real ones used in the previous experiment. The stimuli consisted of a vertically oriented wedge presented in front of a textured background (Figure 9). When the stimulus was seen binocularly, the background was placed in the plane of the monitor (zero disparity), and the wedge was simulated to lie 10 cm in front of the background. The wedge was oriented such that its tip was on the right side of the object. The tip of the wedges differ in their radius of curvature (cf. again Figure 6). For the "sharp" wedge whose radius of curvature was zero, the angle at the tip was 90 degrees. The wedges were matched for their width and depth extents, uniformly shaded and not textured.

Depending on the viewpoint, two or three edges of the objects were visible. The right-most edge was the one that subjects had to decide whether it was sharp or smooth in depth. The other edge always visible was the front edge of the base of the wedge (the middle line in Figure 9). This edge provided the possibility to the observers to use the motion parallax strategy discussed in the introduction. When the observer was located on the left of the object, the back edge of the base of the wedge was also visible (the left line in Figure 9). All these edges were sub-pixel positioned using a line anti-aliasing technique in which each line was given a thickness of one pixel (cf. Foley et al., 1990).

Apparatus

The images displayed on a high-resolution (1280×1024 pixels) 19in. monitor were controlled by an Indigo (Silicon Graphics, Inc., Mountain View, CA) with an Elan graphics board. Some experimental conditions required the display to be updated according to the observer position in front of the monitor. The three-dimensional position of

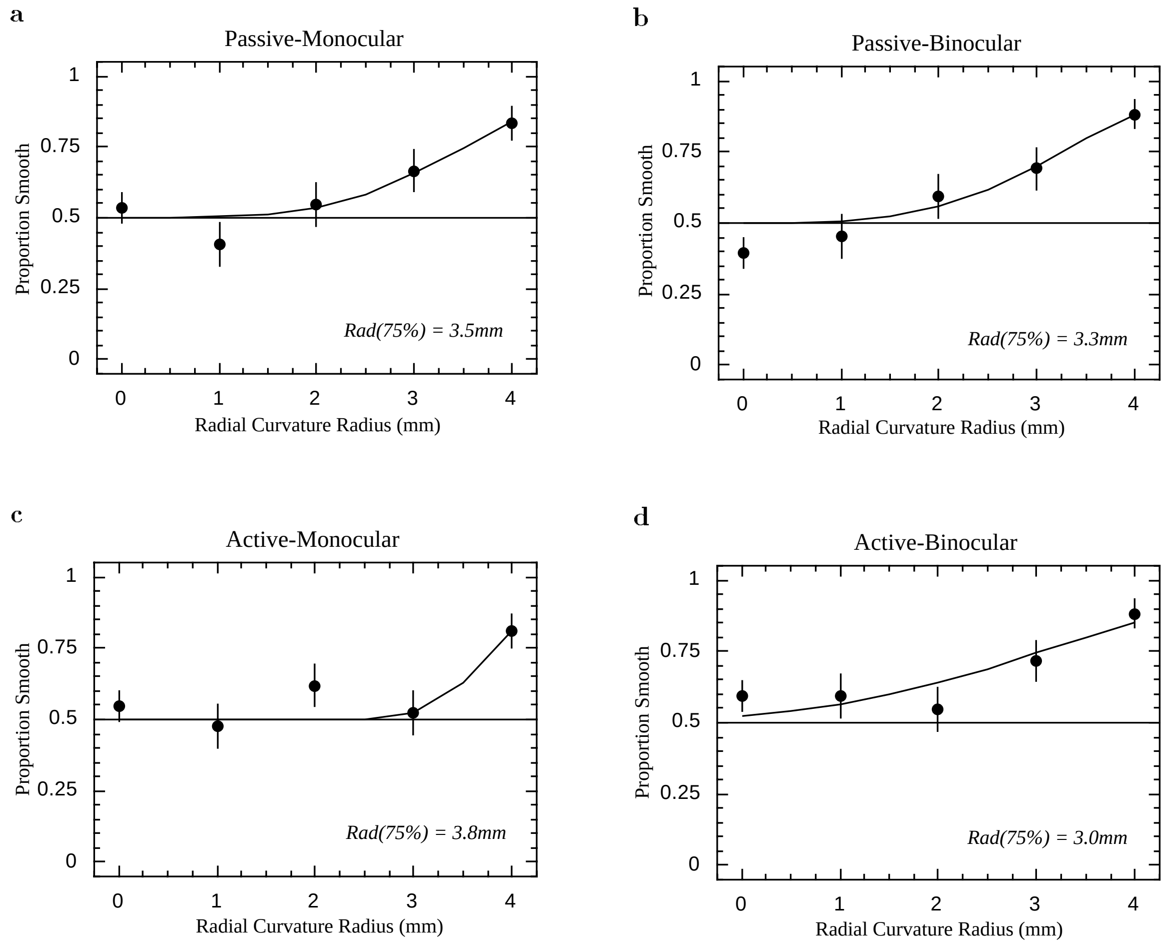

Figure 8: Experimental results with real objects and for the four viewing conditions. Each figure shows the proportion of times a wedge was judged to be smooth. Error bars reflect the standard deviations of the corresponding binomial distribution. The psychometric curve is the best fit of the data by a cumulative Gaussian with two degrees of freedom and a proportion-smooth bias taken to be 50% in all four viewing conditions. The radius threshold was found by taking the 75% correct from this psychometric curve.

the head together with its yaw were sensed using an ultrasonic system (Logitech Inc., Fremont, CA). The yaw of the head was necessary to locate the position of the two eyes. The sensing system emitted ultrasonic signals at 50Hz from each of three transmitters arranged within a triangle (base 275 mm, height 195 mm). These ultrasonic signals were received by three microphones inserted into a pair of CrystalEyes stereo glasses (StereoGraphics Corp., San Raphael, CA).

The precision of the sensing system was estimated to be about 2 mm (standard deviation 1 mm) within the used 1 m tracking range. The maximum tracking speed was 750 mm/s, well above the range of head speeds used during active viewing. There was an estimated time lag of a third of a second between the time the head moved and the stimulus was updated on the monitor. This time lag took into account the computation of the head position and orientation, its transmission over the serial port of the computer, and the wait for the next vertical synchronization of the monitor. The time lag was constant (independent of the head position and velocity), which enabled us to maintain a steady refresh rate of the display at 60Hz.

The stereo glasses consisted of shutter lenses driven by an infrared signal at 120Hz. The dynamic range along the line of sight was 1000:1, so that what was displayed for one eye was almost invisible to the other eye. On the other hand, the decay rate of the monitor phosphors was such that the image displayed for one eye could not be cleared fast enough before displaying the image for the other eye. This effect was found to be quite disturbing, since some "shadow" lines were detectable and disrupted the binocular fusion process. We annihilated it by subtracting from each image a weighted contrast image of the other eye image. The weights

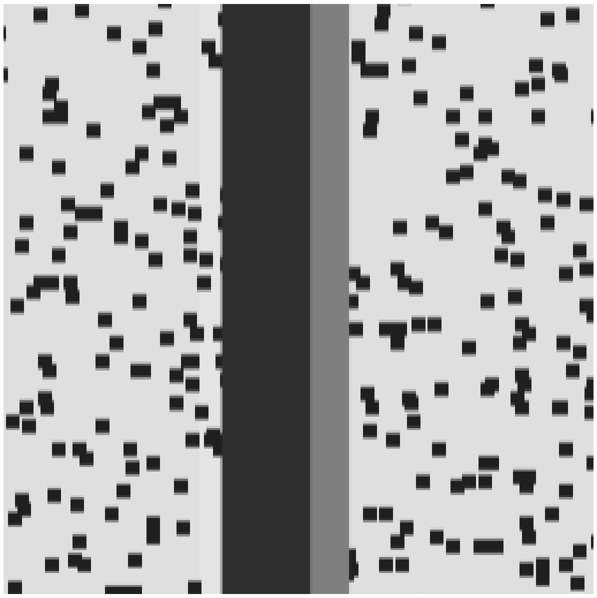

Figure 9: Snapshot of the monitor in the second experiment. The two vertical rectangles represent two faces of the wedge, the dark left one being the base of the wedge. Through lateral displacement, the observer had to decide whether the right-most vertical edge was smooth or sharp. The dots belong to the background.

were a function of the luminance of the displayed image, the luminance of the other eye image, and the phosphor color (the green being the most sluggish). These weights were stored in a lookup table to enable real-time display.

Procedure

The subjects were asked to discriminate between two wedges that differ in the radial curvature of their right-most edge. One of each pair of wedges was always sharp, that is the radial curvature radius was zero (cf. again Figure 6). Within one trial, subjects saw sequentially a sharp and a smooth wedge, and had to report the order of presentation. The order of presentation was randomized from trial to trial. The choice of such a 2-Alternative-Forced-Choice paradigm enabled us to discard the apparent bias for smoothness found in the first experiment. For each trial, the curvature of the "smooth" wedge to be displayed was automatically selected by a Quest adaptive procedure (Watson and Pelli, 1983); the maximum radius permitted to be selected was arbitrarily set to 8 mm. The "sharp" and the "smooth" wedges were presented 5 seconds each. The 76% correct ($d' = 1$) was taken to be the threshold to discriminate sharp from smooth wedges.

Four conditions were intermixed between blocks of trials: i) the object was either viewed monocularly or binocularly; ii) the observer was either active or passive. In active viewing, the observer was instructed to move around the object at his own convenience, from left to right. In passive viewing, the observer stayed still while the display was updated by a virtual observer translating from left to right at a constant speed of 250mm/sec (the simulated translation amplitude was 40 cm). 72 trials were recorded per condition.

The observers were sitting in front of the monitor at a distance of approximately 50 cm. No feedback was given to the subject during the experiment.

Subjects

Five subjects (between 27 and 34 years old) participated in this study. The observers were naive regarding the purposes of the experiment, apart from the first author who was also the only experienced psychophysical subject. None of these subjects participated also in the previous experiment. All subjects had normal or corrected-to-normal visual acuity.

4.2 Results and Discussion

For each of the four viewing conditions, the data were pooled over all five subjects. Figure 10 shows the proportion of times the subjects reported the correct presentation order for a particular radius of curvature of the "smooth" wedge. Because an adaptive procedure was chosen to present the stimuli, some radii of curvature were seen less often than others (this is reflected by some large error bars computed from the appropriate binomial distribution). These data were fitted with a cumulative Gaussian with two degrees of freedom (the mean and the standard deviation of the Gaussian). We took the 76% correct on this psychometric function to be the radius threshold to discriminate "sharp" from "smooth" wedges.

The threshold radii varied between 3.7 and 11 mm for a viewing distance of about 50 cm. In contrast to the first experiment, the active viewing conditions led to significantly lower thresholds than the passive conditions. This result is consistent with the theoretical analysis outlined in the introduction, and suggests also that the subjects in the first experiment might have been using an alternative cue.

When the object was seen passively, the radius threshold came close to the maximum radial curvature radius of the range of objects displayed. It is therefore reasonable to conclude that observers were incapable to discriminate "sharp" from "smooth" wedges in these conditions.

Viewing the virtual scene monocularly and actively led to a radius threshold of 3.7 mm. In comparison, when the object was seen binocularly and actively, subjects' performance did not significantly

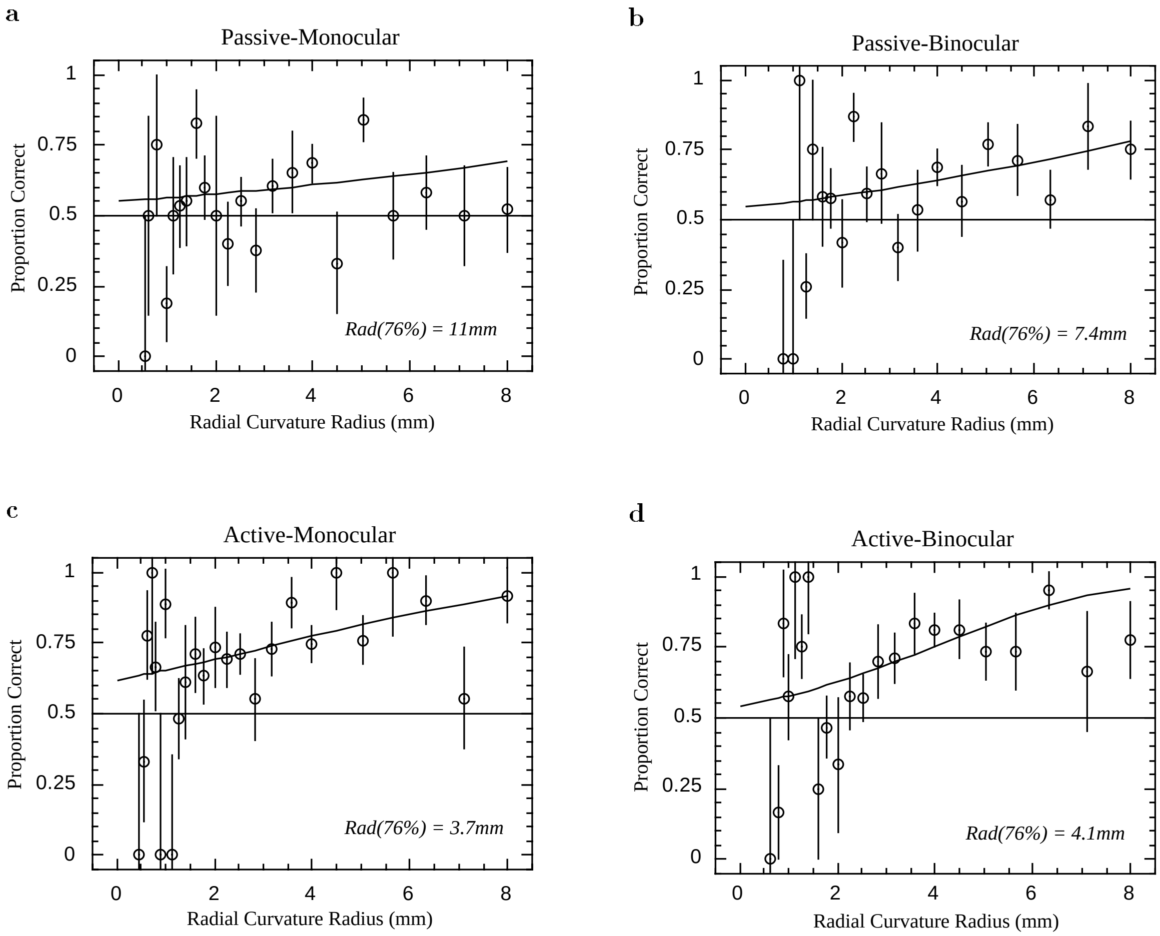

Figure 10: Experimental results for virtual objects. Each figure shows the proportion of correct answers in discriminating sharp from smooth wedges for one particular viewing condition. Error bars reflect the standard deviations of the corresponding binomial distribution. The psychometric curve is the best fit of the data by a cumulative Gaussian with two degrees of freedom. The radius threshold is taken to be the 76% correct from this psychometric curve.

improve or degrade. On one hand, one might have expected a better performance since the edge location was more readily available to the observer, and this information could have been used to improve the estimation of the radial curvature. On the other hand, we have discussed in the introduction why stereopsis and motion parallax were inconsistent cues for the computation of the radial curvature. In addition, this inconsistency between motion parallax and stereopsis might have been enhanced because of the slight time lag in the display refreshment which delayed the stereo information relative to the motion parallax. It therefore appears that stereopsis played only a minor role in our experiment, but it is premature to conclude whether this is due to the nature of the task or to our particular apparatus.

5 Summary and Conclusions

The purpose of our study was to investigate whether human observers could distinguish a sharp from a smooth edge in depth from the motion of this edge. Theoretical considerations suggest that the observer should know his egomotion to perform this task, and that the presence of another sharp edge in the vicinity of the first edge should greatly improve the robustness of the decision.

The psychophysical evidence was gathered from two experiments, one with real objects, the other with computer simulated objects. In the first experiment, we found that the observers had a strong bias to label the moving edge "smooth" in depth. No change in performance was detected when the scene was observed passively, or binocularly.

Since we were concerned with some potential alternative cues available in the stimuli, we replicated the experiment with virtual objects. In this latter experiment, observers could not distinguish sharp from smooth edges in depth if the object was moving by itself. In contrast, if the observers were given the possibility to explore the object themselves, they could discriminate these two kinds of edges when the smooth edge had a curvature radius of about 4 mm at 50 cm viewing distance.

Finally, binocular viewing of the scene did not help the observers to discriminate the sharp from the smooth edges. At this point, it is not clear whether this result is an artefact of our particular apparatus (a small time lag in the monitor refreshment delayed the stereo information relative to the egomotion), or whether stereo and motion cannot combine because they provide inconsistent information (stereo being always in favor of a "sharp" interpretation). Further studies should clarify this point.

In spite of the negative result of the early work of Wallach and O'Connell (1953), it appears that a moving edge can after all be interpreted as the silhouette of a smooth object. For this to occur however, the observer seems to make use of his own motion. Such a phenomenon, which might be called *active kinetic depth effect*, opens new avenues of research with the recent development of virtual reality technologies.

References

Blake, A., & Bülthoff, H. H. (1991). Shape from specularities: Computation and psychophysics. *Philosophical Transactions of the Royal Society of London, B 331*, 237-252.

Cipolla, R. & Blake, A. (1992). Surface shape from the deformation of apparent contours. *International Journal of Computer Vision, 9*, 83-112.

Cornilleau-Pérès, V., & Droulez, J. (1989). Visual perception of surface curvature: Psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics, 46*, 351-364.

Cornilleau-Pérès, V., & Droulez, J. (1993). Stereo-motion cooperation and the use of motion parallax in the visual perception of 3D structure. *Perception & Psychophysics, 54*, 223-239.

Cortese, J. M. & Andersen, G. J. (1991). Recovery of 3-D shape from deforming contours. *Perception and Psychophysics, 49*, 315-327.

Droulez, J., & Cornilleau-Pérès, V. (1990). Visual perception of surface curvature. The spin variation and its physiological implications. *Biological Cybernetics, 62*, 211-224.

Foley, J.D., van Dam, A., Feiner, S.K., & Hughes, J.F., (1990). *Computer Graphics, Principles and Practice* (2nd ed.). Reading, MA: Addison-Wesley Publishing Company, Inc.

Giblin, P. J. & Weiss, R. (1987). Reconstruction of surfaces from profiles. In *Proceedings of the First International Conference on Computer Vision (London)*, (pp. 136-144).

Goodale, M. A., Ellard, C. G. & Booth, L. (1990). The role of image size and retinal motion in the computation of absolute distance by the mongolian gerbil (Meriones unguiculatus). *Vision Research, 30*, 399-413.

Hildreth, E. C. (1983). *The Measurement of Visual Motion.* Cambridge, MA: MIT Press.

Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape? Perception, **13**, 321-330.

Koenderink, J. J. (1986). Optic Flow. *Vision Research, 26*, 161-180.

Kutulakos, K. N. & Dyer, C. R. (1994). Recovering shape by purposive viewpoint adjustment. *International Journal of Computer Vision, 12*, 113-136.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research, 35*, 389-412.

Lehrer, M., Srinivasan, M. V., Zhang, S. W., & Horridge, G. A. (1988). Motion cues provide the bee's visual world with a third dimension. *Nature, 332*, 356-357.

Pollick, F. E., Giblin, P. J., Rycroft, J., & Wilson, L. L. (1992). Human recovery of shape from profiles. *Behaviormetrika, 19*, 65-79.

Rogers, B. & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8*, 125-134.

Roll, R., Velay, J. L., & Roll, J. P. (1991). Eye and neck proprioceptive messages contribute to the spatial coding of retinal input in visually oriented activities. *Experimental Brain Research, 85*, 423-431.

Sobel, E. C. (1990). The locust's use of motion parallax to measure distance. *Journal of Comparative Physiology A, 167*, 579-588.

Sperling, G., Landy, M. S., Dosher, B. A. & Perkins, M. E. (1989). Kinetic depth effect and identification of shape. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 826-840.

Todd, J. T. (1985). Perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception and Performance, 11*, 689-710.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, MA: MIT Press.

Vaillant, R. & Faugeras, O. (1992). Using extremal boundaries for 3-D object modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*, 157-173.

Wallach, H. & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 205-217.

Watson, A. B. & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113-120.

Waxman, A. M., & Ullman, S. (1985). Surface structure and three-dimensional motion from image flow kinematics. *The International Journal of Robotics Research, 4*, 72-94.