



# Perception of Three-dimensional Shape from Ego- and Object-motion: Comparison Between Small- and Large-field Stimuli

T. M. H. DIJKSTRA,\* V. CORNILLEAU-PÉRÈS,† C. C. A. M. GIELEN,\* J. DROULEZ†

Received 20 December 1993; in revised form 20 June 1994

**We compare the performance in the detection of the shape of concave, planar and convex surfaces for small-field (8 deg) and large-field (90 deg) stimuli. Shape is perceived from head translations, object translations and object rotations. We find large differences between small-field and large-field stimulation. For small-field stimulation performance is best for object rotation, intermediate for self-motion and worst for object translation. For large-field stimulation performance is similar across conditions. Few errors on the sign of the curvature are found for self-motion for both field sizes, indicating that self-motion information disambiguates the curvature sign. For object rotation with small-field stimulation, the concave-convex ambiguity is strong with many apparent deformations. In contrast, large-field curvature signs are always accurately reported, suggesting that the weight of the rigidity hypothesis depends on field size.**

Active vision Structure-from-motion Three-dimensional shape Human

## INTRODUCTION

A new field of research in the computer graphics and computer vision communities is devoted to active vision, i.e. vision by an actively moving observer (Azarbayejani, Starner, Horowitz & Pentland, 1993). Within computer vision active vision is seen as a means for a robot to extract three-dimensional (3D) information from the environment by using ego-motion information from nonvisual sources in the evaluation of visual information. However, it is unclear whether and to what extent human observers use nonvisual information in this direct way. So far only a few psycho-physical studies have been performed in this field and very few comparisons between the perceptual effects of active and passive vision have been made.

The relationship between active movements and 3D shape perception has been pioneered by Rogers and Graham (1979) who simulated a corrugated surface on an oscilloscope screen. The motion of the dots on the screen was linked to the movements of the observer (subject movement): some horizontal lines of dots moved with the observer and some other lines moved in opposite direction, creating a compelling view of a surface

with horizontal corrugations. They also did the experiment with movement of the oscilloscope and linked the motion of the dots to the oscilloscope movement (object-translation). They found that the perceived depth of a surface is about 15% higher when the motion parallax is generated by active movements of the observer rather than by movement of the stimulus presented to a stationary observer.

A good interpretation of the decreased performance in the object-translation condition is difficult for several reasons. First, head movements in the active condition were not stored. Therefore, the movement of the stimulus relative to the head might have been different in the subject-movement and object translation conditions. Second, it is known that the fixation of a point of the stimulus is better in the subject-movement condition than in the object-translation condition. For a study comparing fixation of a target by vergence movements in active and passive conditions see Erkelens, van der Steen, Steinman & Collewijn (1989). Also, the otolith-ocular reflex might contribute to a better retinal image stabilisation during ego-motion (Buizza, Leger, Droulez, Berthoz & Schmid, 1980; Bronstein & Gresty, 1988; Baloh, Beykirch, Honrubia & Yee, 1988). Finally, there was no fixation point in the stimuli of Rogers and Graham. This is of importance since it has been demonstrated recently that performance in detection of the sign of curvature is dependent on fixation (Hayashibe, 1991).

Therefore, Cornilleau-Pérès and Droulez (1994) compared the sensitivity for the detection of curvature of a

\*Laboratory of Medical Physics and Biophysics, University of Nijmegen, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands [Email tjeerd@mbfys.kun.nl].

†Laboratoire de Physiologie de la Perception et de l'Action, CNRS, 15 rue de l'École de Médecine, F-75270 Paris Cedex 06, France [Email vcp@ccr.jussieu.fr].

moving surface for the conditions of subject-movement and object-movement. They constructed the experiment so that the relative translation between the object and the observer was identical in all conditions. They tested both conditions from the Rogers and Graham experiment and added a third condition (object-rotation) where the movement of the object was a rotation in depth around the fixation point. The stimuli, with a diameter of 8 deg, were either plane or convex with a fixed curvature and the observer's task was a forced choice between plane and convex. The results show that curvature sensitivity is far higher in the subject-movement condition than in the object-translation condition, and that the object-rotation condition yields the best performance. To explain their findings, the authors invoked the global image motion which results from different oculomotor behaviour in the three conditions, as the main factor which determines the performance. This explanation is based on the fact that global image motion is known to impair the visual sensitivity to differential velocity (Nakayama, 1981), and on the fact that the detection of surface curvature from motion is likely to be mediated by the processing of spatial variations in image velocity (Cornilleau-Pérès & Droulez, 1989).

The optic flow fields plays a double functional role in visual perception: it provides an observer with exteroceptive information about the structure (distance, slant and curvature) and motion (velocity and rotation rate) of objects in the environment as well as with proprioceptive information about the movements of the observer in the environment (velocity and rotation rate). From theoretical studies on optic flow processing it is known that the parameters of the relative motion between observer and object cannot be separated from the recovery of the structure of the object (Waxman, Kamgar-Parsi & Subbarao, 1987; Droulez & Cornilleau-Pérès, 1990; Koenderink & van Doorn, 1992). Hence, the visual system could take advantage of self-motion to improve its ability to solve the problem of structure from motion.

The first goal of the present paper is to investigate whether proprioceptive ego-movement information (knowing where and how fast you are moving) is used directly in the perception of shape. Self-motion is processed both from non-visual information, such as efference copies and vestibular signals, and from visual information. Although different variables interact in the perception of self-motion, the size of the stimulus is one of the major factors which influences vection or the control of stance (for a review see Warren & Kurtz, 1992). In particular, when a lamellar flow field due to the frontal translation of a plane is presented in central vision, Stoffregen (1985, p. 561) found that compensatory body sway is very small for a stimulus width of 20 deg, and increased much as this width reaches 40 deg. Similarly, Post (1988) showed a large reduction of circular vection when the stimulus size was 30 deg wide, rather than full-field. Therefore, the small stimuli (8 deg diameter) used by Cornilleau-Pérès and Droulez (1994) are poor in terms of visual information about self-

motion. In order to create a stronger impression of self-motion we extend the experiment of Cornilleau-Pérès and Droulez to large field stimuli (90 deg visual angle).

Since our results suggest that proprioceptive information is not used in a direct way in the perception of shape, the question arises whether the ego-movement information is used at all. A possible use for this information may be to assist in fixating a particular point on the object. Cornilleau-Pérès and Droulez (1994) explained their findings by a different amount of retinal slip in the different movement conditions. This explanation can be tested by extending the experiment of Cornilleau-Pérès and Droulez to large-field stimuli. Fixation is better for a large field of view as shown by van den Berg and Collewijn (1986). They showed an increase in the gain of pursuit eye movements when a large grating is superimposed on the target to be pursued.

In the object-rotation condition, we noticed an ambiguity between concave and convex spheres. This ambiguity was already reported by Hayashibe (1991) and Rogers and Rogers (1992). Rogers and Rogers found that both perspective and non-visual information about self-motion contribute to disambiguation. Another goal of this paper was thus to compare the efficiency of self-motion and perspective information in disambiguating the sign of curvature. Hence, instead of asking the subject to report only the presence or absence of surface curvature, we also require that he reports the sign of curvature.

## METHODS

Wide-field and small-field experiments were performed in different laboratories. We therefore start with a separate description of each of the set-ups.

### *Experimental set-up for large field stimulation*

Stimuli with a resolution of  $1152 \times 900$  pixels were generated with a SUN4/260 CXP workstation. The video images were projected on a large translucent screen (dimensions  $2.5 \times 2$  m) by a Barco Graphics 400 video projector. The stimuli were green (phosphor P53) and had a luminance of  $0.5 \text{ cd/m}^2$ . The frame rate was 66 Hz and the stimuli were also presented at this rate. The translucent screen was homogeneously white without any visible texture.

The subject stood in front of the screen at a distance of 50 cm wearing an eye patch to cover one eye and a helmet on which six infra-red light-emitting diodes (IREDS) were mounted. The positions of these IREDS were measured with a WATSMART system at a rate of 400 Hz. The 2D coordinates from the two cameras of this system were converted in real-time to 3D coordinates. The 3D coordinates were sent to the SUN4 by direct memory access. The SUN4 was programmed to generate an image of a 3D shape, viewed from the current position of the eye. The algorithm to achieve this is explained below. Part of the output of this algorithm

is the position of the eye in 3D space and the orientation of the head. The orientation of the eyes, i.e. the direction of gaze was not measured. It should be noted that in the context of this article eye position means the *position* in 3D space, not the *orientation* of the eyes. The delay in the feedback loop between eye translation and position change of a pixel in the middle of the screen was measured using a turntable and was found to be  $43 \pm 3$  msec. The small variability was probably caused by the fact that SunOS is not a real-time operating system.

The position of each eye was calculated using a quaternion algorithm due to Horn (1987). Each session started with a calibration procedure in which the subject faced the cameras and held two additional IREDs in front of the eyes. The position of these two IREDs and of those on the helmet was sampled for 200 msec at 400 Hz in this calibration configuration. From these data the position of each eye relative to each IRED on the helmet was calculated. The rotation of the helmet relative to its orientation in the calibration configuration was calculated using a real-time implementation of Horn's algorithm for a planar figure (Section 5 of Horn, 1987). Thus the position of each eye could be calculated during the experiment using the known positions of the eyes relative to the IREDs on the helmet. Assuming the accuracy of the WATSMART system to scale linearly with the calibrated volume, which in our case was a cube with sides of length 0.6 m, the accuracy of the position of one IRED is estimated to be 3 mm (Ball & Pierowsky, 1987) and the systematic error of eye position relative to the simulated scene is estimated to be about 5 mm. The dynamic noise in the eye position was approximately white and had a SD of 0.5 mm.

#### *Experimental set-up for small-field stimulation*

The stimuli were presented on the monitor of a Silicon Graphics workstation (resolution  $1280 \times 1024$  pixels, frame rate 60 Hz). The stimuli were white (phosphor P22), had a luminance of  $1.4 \text{ cd/m}^2$  and were presented at a rate of 30 Hz (each frame is displayed twice).

The subject was sitting at a distance of 72 cm from the monitor with one eye covered. He had a light-weighted helmet on his head on top of which was fixed a mobile bar. The weight of the bar was sufficiently small so as not to hamper head movements. It was mobile in a pulley with very low friction, and could therefore translate along itself. The pulley could rotate around the vertical and horizontal axes passing through its centre. Three potentiometers delivered analog signals linearly or sinusoidally related to each of the translations of the head (up-down, left-right and backwards-forwards). These signals were converted to digital by a microcomputer and were then sent to the workstation through an RS232 bus at a rate of 9600 baud. The workstation was programmed to generate a video image of a 3D shape, viewed from the current position of the eye. The delay in the feedback-loop was 55 msec.

The microcomputer was used to calibrate the three head translation signals. Repeated calibrations

performed on 105 points lying within a parallelepiped centred on the median subject's head position (30 cm in horizontal, 20 cm in vertical, 6 cm in depth) showed that the mean error on head position is 1.7 mm, with a maximum of  $< 5$  mm. A restricted calibration was performed prior to each experiment, in order to estimate the potentiometer offsets and gains that could vary in time.

#### *Stimuli*

Because of different technical constraints, the parameters of the large-field stimuli (hereafter LF) and small-field stimuli (hereafter SF) were not precisely the same. However, as shown in the section on control experiments, they were generally sufficiently similar so that the two experiments were comparable.

Stimuli were curved or flat surfaces covered with 300 (LF) or 400 (SF) random dots, each of diameter 0.2 deg (LF) or 0.02 deg (SF). The distribution on the surface was such that the density of dots was uniform per solid angle. This was done to minimize the possibility to use the local density of dots as a feature to estimate the curvature of the surface. The large-field stimulus covered a range between 2 and 45 deg of visual eccentricity (field of view 90 deg), and had a fixation cross of  $2 \times 2$  deg at the centre. The small-field stimulus covered a range between 0 and 4 deg of visual eccentricity (field of view 8 deg), and had a bright fixation dot of diameter 0.05 deg at the centre. The shape of the large-field stimulus was a section of a sphere which could have a curvature of  $-0.67, -0.33, -0.17, 0, 0.17, 0.33$  or  $0.67 \text{ m}^{-1}$ . The shape of the small-field stimulus was a section of a sphere which could have a curvature of  $-5, -4, -2.85, 0, 2.85, 4$  or  $5 \text{ m}^{-1}$ . Negative curvatures denote concave sphere segments, curvature 0 denotes a plane and positive curvatures denote convex sphere segments. It should be noted that the rim of the stimulus is a planar curve which had the same projection for all curvatures and hence could not be used as an artifactual cue.

The stimuli were shown for 6 sec in a dark room and on a dark background. The fixation point was a small cross or a bright dot in the centre of the simulated surface and was straight in front of the subject at the beginning of a trial. The distance from the eye to the fixation point was chosen randomly between 40 and 60 cm (LF) or between 75 and 85 cm (SF). This made it difficult for the subject to use the mean retinal velocity as a cue for the shape (see subsection about control experiments). At the start of each trial the tangent plane at the fixation point was fronto-parallel. Due to the head-movements the viewing distance and the orientation of the tangent plane changed in the course of a trial.

We compared thresholds of curvature detection in three conditions: a subject-movement condition, an object-translation condition and an object-rotation condition. In the subject-movement condition subjects moved in left/right direction at a frequency of 0.33 Hz (LF, SF) or 0.5 Hz (SF) and with an amplitude of 10 cm. Pilot results and a control experiment on subject VCP showed the effect of frequency to be very small. A

metronome helped the subjects to maintain a constant frequency. This frequency and amplitude of movement were trained at the beginning of each session by giving the subject feedback about his movement. Subjects could readily perform this with a relative standard deviation in amplitude of movement of about 10% (Table 1). We stored a time series of the translation of the eye together with the positions of the random dots relative to the eye on disk. This information was used later in the two object-movement conditions to generate the same projections on the optic array.

In the object-translation condition the head of the subject was fixed using a chinrest and the stimulus translated with the translation of the head previously recorded in the subject-movement condition. Thus the subject had to make tracking eye movements in order to fixate the fixation point.

In the object-rotation condition the head was also fixed but the motion of the stimulus was a pure rotation in depth. This rotation was calculated from the previous translation by adding a simulated eye rotation so that the stimulus on the optic array of the subject was the same as in the subject-movement condition. The torsion component of rotation was set to zero: the torsion of the head was negligible and the torsion of the eye was not measured but is known to be small for eye orientations of up to 10 deg (Mok, Ro, Cadera, Crawford & Vilis, 1992).

### Protocol

The stimulus was either a concave, planar or convex surface with equal probability. The subject's task was a forced choice between concave, planar or convex. No feedback on performance was given.

Each experiment consisted of five sessions and lasted approx. 45 min each. Each session consisted of two subsessions. Each subsession consisted of six blocks in fixed order: for the left eye subject-movement, object-rotation, object-translation, then for the right eye subject-movement, object-translation and object-rotation. Each block consisted of 18 stimuli: two repetitions for each of the six curvatures and six repetitions for the plane in random order. So in all sessions together there were 40 repetitions per movement condition and per curvature, 20 for the left eye and 20 for the right eye. For the plane these numbers are 3 times as high.

TABLE 1. Comparison of experimental parameters and movement characteristics for large-field and small-field stimulation

Field of view (deg)	Subject	Movement frequency (Hz)	Peak-to-peak amplitude (cm)	SD peak-to-peak amplitude (cm)
90	TD	0.33	21.8	2.1
90	MG	0.33	23.1	2.4
90	PS	0.33	19.8	2.2
8	TD	0.33	26.4	3.3
8	VCP	0.33	22.9	1.3
8	OV	0.5	19.5	2.8

### Subjects

Three subjects participated for each field of view, one of the authors (TD) was tested for both SF and LF stimulation. Three subjects were naive as to the purpose of the experiment (MG, PS and OV). All subjects had normal vision or corrected to normal vision wearing contact lenses.

## RESULTS

In Fig. 1 we show the results for large-field stimulation for the three subjects. In general the stimuli with the largest absolute curvature (both convex and concave) are perceived with almost 100% accuracy. For smaller absolute curvatures, the probability of a correct response decreased gradually. The percentage of planar responses generally peaks at zero curvature and decreases when absolute curvature becomes higher. The general features of these curves are roughly the same for each movement condition. This is also the message from Table 2 where we compare the performance in detection of absolute curvature for the different movement conditions. The percentage correct responses (PCR) does not differ significantly between the three movement-conditions. Subjects always perceive a rigid shape and find the object-translation condition more difficult than the other two, although their performance is not significantly worse.

In Fig. 2 we show the results for small-field stimulation for the three subjects. One subject (TD) was also tested for large-field stimulation (see Fig. 1). The results for the three movement conditions are very different from one another. The curves for the subject-movement condition are qualitatively the same as for large-field stimulation, albeit that performance is somewhat lower for subject TD. The curves for the object-translation condition are close to chance level, which is very different from the large-field result. In this condition only the percentage of concave responses at a curvature of  $-5 \text{ m}^{-1}$  for subject OV is clearly different from chance level. For object-rotation the percentage of planar responses shows the normal profile with a peak at zero curvature. The width of this curve, which is a measure for the performance of detection of absolute curvature, is smaller in the object-rotation condition than in the subject-movement condition. The other two curves do not converge to 100% for the extreme curvatures in contradistinction with large-field stimulation. Since the number of false positives for flat surfaces is very low for the larger curvatures, this indicates that subjects had no problem in distinguishing a flat or curved surface but had difficulties in detecting the sign of curvature. Table 2 shows that performance in the detection of absolute curvature is best in the object-rotation condition, and that the subject-movement condition yields slightly worse performance than the object-rotation condition. The differences between the three conditions are significant to the level  $P < 0.05$ , and performance in the

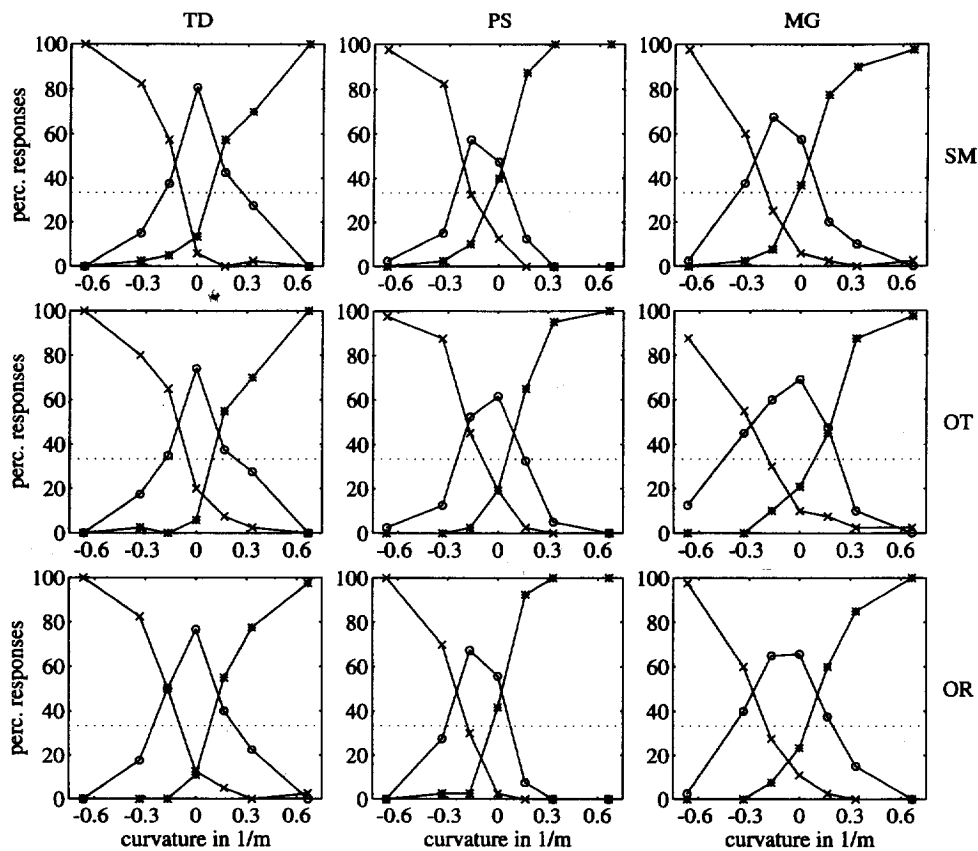


FIGURE 1. Response curves for subjects TD, PS and MG for large-field stimulation in each of the three movement conditions.  $\times$  percentage of concave responses;  $\circ$  percentage of planar responses;  $*$  percentage of convex responses. Left column, subject TD; middle column, subject PS; right column, subject MG. Upper panels, subject-movement (SM); middle panels, object-translation (OT); lower panels, object-rotation (OR). The horizontal dotted line indicates chance level.

object-translation condition does not, in general, differ from chance to the level of  $P < 0.1$ .

In order to compare the performance in large-field stimulation with the performance in small-field stimulation we calculated that the response curves of large-field stimulation need to be scaled in curvature by somewhat more than a factor 10 in order to be comparable with the response curves of small-field stimulation. For subject TD, the only subject tested in both set-ups, the factor is 17 for the percentage of planar responses in the object-rotation condition. We determined this scaling factor by an *ad hoc* procedure, where we fitted both response curves with a gaussian and compared the widths of the gaussians.

Subjects perceive the shape always as rigid in the subject-movement condition with small-field stimulation. They find the object-translation condition with small-field stimulation very difficult, which is reflected in their performance being near chance level. Subjects sometimes perceive the shapes as being hyperbolic, i.e. of negative gaussian curvature. In that case they report the sign of the largest absolute curvature. Subjects frequently report apparent deformations of the stimulus for the object-rotation condition with a small field of view. This is reflected in the large number of errors in the sign of the curvature for object-rotation. The percentages of curvature-inversions for this condition are: TD 40.0%, VCP 24.2% and OV 29.2%. From the response curves

TABLE 2. Comparison of the performance of curvature detection expressed in percentage correct responses (PCR) of the three movement conditions for large-field and small-field stimulation.

Field of view (deg)	Subject	Diff in PCR		Diff in PCR		Diff in PCR	
		OR - SM	Significance	SM - OT	Significance	OT - ch	Significance
90	TD	-1.9	$t = -0.3$ , ns	1.9	$t = 0.3$ , ns	43.6	$t = 8.4$ , $P < 0.005$
90	MG	0.8	$t = 0.3$ , ns	1.1	$t = 0.4$ , ns	34.4	$t = 20.0$ , $P < 0.005$
90	PS	1.9	$t = 0.9$ , ns	-3.6	$t = -1.3$ , ns	41.7	$t = 19.8$ , $P < 0.005$
8	TD	8.6	$t = 2.1$ , $P < 0.1$	13.9	$t = 2.9$ , $P < 0.05$	-2.8	$t = 0.9$ , ns
8	VCP	5.6	$t = 3.2$ , $P < 0.05$	27.2	$t = 8.0$ , $P < 0.01$	4.7	$t = 3.7$ , $P < 0.05$
8	OV	10.8	$t = 5.6$ , $P < 0.01$	14.2	$t = 4.7$ , $P < 0.01$	2.2	$t = 0.9$ , ns

SM, subject-movement; OT, object-translation; OR, object-rotation; ch, chance level.

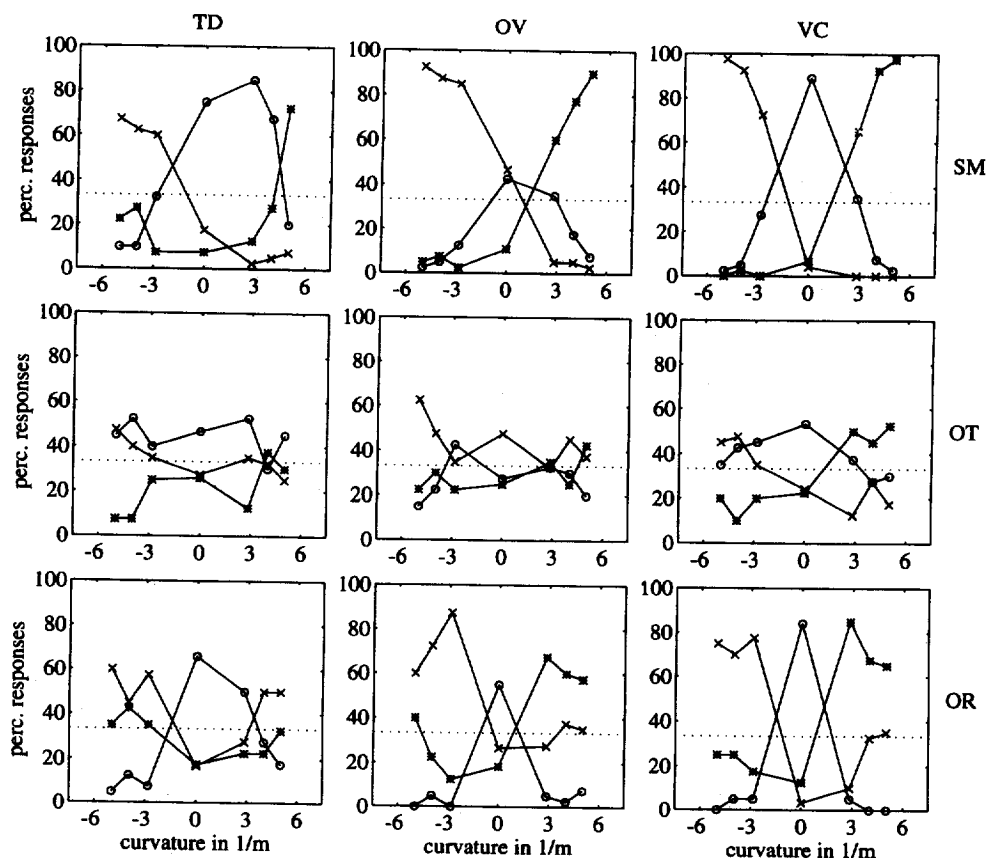


FIGURE 2. Response curves for subjects TD, OV and VC for small-field stimulation in each of the three movement conditions.  $\times$  percentage of concave responses;  $\circ$  percentage of planar responses;  $*$  percentage of convex responses. Left column, subject TD; middle column, subject OV; right column, subject VC. Upper panels, subject-motion (SM); middle panels, object-translation (OT); lower panels, object-rotation (OR). The horizontal dotted line indicates chance level.

in the lower panels of Fig. 2 (also for subject VCP) one can see a more surprising fact: the percentage of reversals does not decrease with increasing absolute curvature. Thus, on one hand performance for detection of absolute curvature becomes better with increasing curvature, because the percentage of planar responses becomes smaller for larger absolute curvatures. On the other hand the percentage of convex responses for concave stimuli is seen not to decrease with decreasing (increasing absolute) curvature. The same holds true for the percentage of concave responses for convex stimuli.

Figure 3 summarizes the results, and shows for each subject and condition the average percentage of errors. The part of this percentage concerning only the sign of curvature is indicated as the hatched part of each bar, while the remaining part of each bar corresponds to the errors in reporting the presence of absolute surface curvature. As noted above the results vary strongly across conditions in small field but not in large field. In particular the sensitivity to absolute surface curvature varies across conditions in small field; it is maximal for object-rotation, intermediate for self-motion, and minimum for object-translation. No such variations are observed in large-field. Further, the number of errors in the curvature sign is low in all conditions in large field but varies across conditions in small field. In this last

case the depth ambiguity is strong for object-rotation (the hatched parts of the bar are close to the chance level of reporting the wrong sign of curvature of 33.3%), but much reduced for subject-motion.

In summary, the main results are that the performance for curvature detection in the three movement conditions is roughly constant for large-field stimuli but shows great differences for small-field stimuli. Small-field stimuli give a reasonably good performance for the subject-motion condition and a performance near chance level for the object-translation condition. Although the distinction between a flat or curved surface is made very well in the object-rotation condition, ambiguity between convex and concave surfaces is present for the small-field stimuli.

#### Control experiments

First we performed a control session with the set-up of the large-field stimulation but using the same field of view as in the small-field stimulation, using the monitor of the SUN4. In particular, we took 300 random dots, a frame rate of 66 Hz, a screen distance of 50 cm and a simulated viewing distance between 40 and 60 cm. Further, we took the curvatures 10 times as high as in the large-field stimulation, thus covering a large range of absolute curvatures than the small-field stimuli. This

gave essentially the same results as we found for small-field stimulation. The only difference was a clear above chance level performance with no depth reversals for the larger curvatures in the object-translation condition. This shows that the different frame rates, number of dots and viewing distance are unimportant in a comparison of the results of large-field and small-field stimuli.

It is unlikely that the dot density accounts for the differences between large-field and small-field stimulation: in pilot experiments where we varied the dot density with a factor 2, we found that all movement conditions were affected equally and we found a small effect on performance (this was reported before, e.g. by Norman & Lappin, 1992). Unfortunately, we could not perform a control experiment where dot density in

large-field and small-field stimulation was equal because the two field sizes differ by a factor 120 in area. Using the dot density of large-field stimulation for small-field stimuli leads to only 3 dots, which does not lead to a clear percept of a solid shape. Using the dot density of small-field stimulation for large-field stimuli leads to about 50k dots, which is beyond the capabilities of modern graphics machines.

Our experiment is based on the assumption that the subjects really report on the 3D structure that they have perceived. The need for verifying that no other cue incidentally related to surface curvature is used by the subjects, has recently been underlined in a series of papers (Braunstein & Todd, 1990; Sperling, Doshier & Landy, 1990). There are two main cues that could be used as artifactual cues by the subjects in experiments on the perception of surface curvature from motion: (1) the spatial variations of dot density that occur in the extreme positions reached by the object relative to the observer (recall that the dot density was uniform per solid angle for the median position); (2) the magnitude of the image velocity of the moving dots relative to the observer.

In the object-rotation condition the position of the dots relative to the eye is exactly the same as in the other two conditions, if a perfect stabilization of the image occurred (if the subject kept fixating the fixation point). Therefore we can assume that if the subject based his responses on any of the above artifacts, then he could perform particularly well in the object-rotation condition for which image stabilization is the easiest. So we restrict the discussion to this latter condition.

For large-field stimulation we did a control experiment in which subject TD was shown the extreme positions of the object for 3 sec each. Otherwise the stimuli were equal to the large-field stimuli of the experiment. The responses did not significantly differ from chance level. We conclude that there is little evidence that the static variations in dot density are used as a cue for curvature.

In order to test whether the mean velocity magnitude of dots on the retina of the subject is used as an artifactual cue, we calculated the mean velocity for each trial for large-field stimulation. The mean retinal flow, i.e. the spatial and temporal mean of the retinal velocity is plotted in the upper panels of Fig. 4 for two subjects. As one can see there is a large overlap in mean retinal flow for each of the curvatures. There is still a noticeable trend of mean flow with curvature, especially the sphere with curvature  $0.67 \text{ m}^{-1}$  leads to a higher retinal flow. Also notice that the flow is relatively small for the concave spheres, which are on average closer to the subject, and is relatively high for the convex spheres which are on average farther away. This is because retinal flow depends both on distance from the subject and on the 3D velocity of a dot on the surface. The 3D velocity of a dot is higher at a given eccentricity for a convex sphere than for a concave sphere, because the dot is farther removed from the axis of rotation. From the data in Fig. 4 we calculated response curves based on the

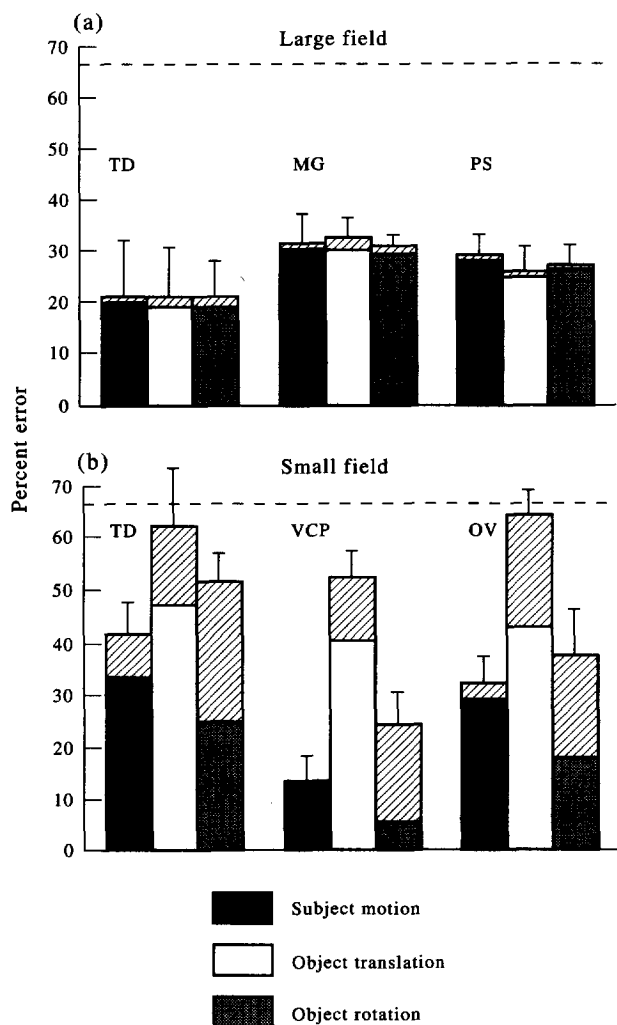


FIGURE 3. Comparison of the results of large-field (a) and small-field (b) experiments. Each bar indicates the percentage of errors which occurred in the concave/plane/convex discrimination task, as averaged over five trials (with the SDs in brackets). The hatched part of each bar shows the percentage of errors restricted to spheres (answer concave when the stimulus is convex, and vice versa). The rest of the bar corresponds to the errors in sphere/plane discrimination. The initials of each subject are indicated above each group of three bars (solid, condition subject-motion; open, condition object-translation; shaded, condition object-rotation). The dotted line indicates chance level for the total percentage of errors.

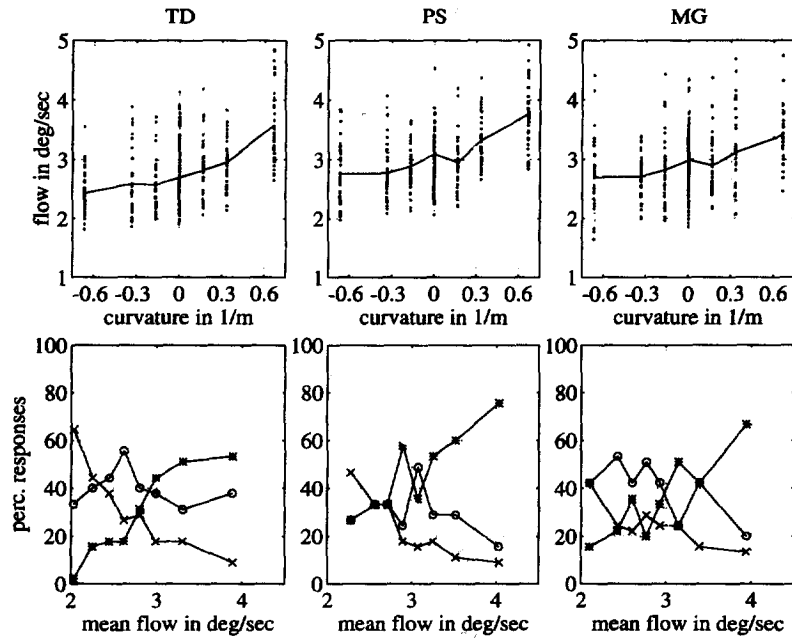


FIGURE 4. Results of calculation to show that subjects do not use mean retinal flow as an artificial cue for curvature detection.  $\times$  percentage of concave responses;  $\circ$  percentage of planar responses;  $*$  percentage of convex responses. Left column, subject TD; middle column, subject PS; right column, subject MG. Upper panels, mean retinal flow vs curvature for every trial. The solid line indicates the average mean flow for each curvature. Lower panels, response curve if the subject would use mean flow for detection.

assumption that the subject used mean retinal flow for curvature detection. To that purpose we subdivided the mean retinal flow from each trial in eight categories with the smallest 45 mean flow values in the first category, etc. We counted the percentage of concave, planar and convex responses in each of these categories. In the lower panels of Fig. 4 we plotted these percentages vs the mean of the 45 flows in each of the categories. Clearly, the results are different from the results reported in Fig. 1. Although there is a tendency of the percentage of concave (convex) responses to decrease (increase) with mean retinal flow, the effects are much smaller than in Fig. 1. We also did the above analysis with the maximum retinal flow and with a different number of categories and found essentially the same. We conclude that there is little evidence that subjects use retinal flow as an artificial cue for curvature detection.

For small-field stimulation Cornilleau-Pérès and Droulez (1993, Expt 1C; 1994) performed control experiments which show that the mean retinal velocity and the spatial variations of dot density in individual images are not used as artificial cues in experiments involving the discrimination between convex spherical surfaces and planes. Another argument against the use of mean retinal velocity as a cue for the detection of curvature with small-field stimulation is provided by the high number of curvature inversions (lower panels of Fig. 2) in the object-rotation condition. Since the mean dot velocity is a monotonic function of sphere curvature, the use of this velocity as an artificial cue should result in a better discrimination between convex and concave spheres than between spheres and planes, which is the opposite of what was observed.

## DISCUSSION

We conducted experiments in which we test the performance of detection of curvature of 3D objects with three different movement conditions and two different sizes of the field of view. All movement conditions result in the same motion parallax on the optic array of the observer, hence providing identical information about the 3D shape of the object. The difference between the conditions is the amount of extra-retinal information available. In the subject-movement condition the subject moves his head left/right and is shown a simulated segment of a sphere or plane. In the object-translation condition the head of the subject is stationary but the stimulus translates left/right. In the object-rotation condition the head of the subject is also stationary but the stimulus rotates in depth.

The first conclusion that emerges from our experiments is that the accuracy in discriminating between planar and non-planar surfaces is not increased by proprioceptive ego-motion information. This is borne out by the fact that the performance in the object-rotation condition, where the subject cannot use any proprioceptive ego-motion information, is the same (large-field stimulation) or better (small-field stimulation) than in the subject-movement condition, where the subject can use all available ego-motion information. For small-field stimulation this result was already reported before by Cornilleau-Pérès and Droulez (1994). Both results suggest that proprioceptive ego-motion information is not used directly in the perception of curvature of 3D shapes.

Second, our results are compatible with the



interpretation that proprioceptive ego-motion information is used in the stabilization of the retinal image of the object, thereby enhancing the sensitivity to detection of curvature. Stabilization of the retinal image is important because it affects the sensitivity to detection of differential motion (Nakayama, 1981) and this is thought to be used in the detection of 3D shape (Cornilleau-Pérès & Droulez, 1989; Koenderink & van Doorn, 1992). For small-field stimulation we find performance in the detection of absolute curvature to be highest in the object-rotation condition. We find intermediate performance in the subject-movement condition and the worst performance in the object-translation condition. It is known (Ferman, Collewijn, Jansen & van den Berg, 1987) that stabilization of gaze is better with a stationary observer than with an observer rotating his head. With a small target of 0.2 deg and a stationary head Ferman *et al.* (1987) found a mean retinal slip velocity of 0.4 deg/sec, whereas they found a mean retinal slip velocity of 0.7 deg/sec for subjects rotating their head at 0.33 Hz. In a condition comparable to our object-translation condition, van den Berg and Collewijn (1986, p. 1216, Table 1, discarding subject CE) found a retinal slip velocity of 1.25 deg/sec for a small translating target of 0.15 deg which is poor relative to the previously mentioned slip velocities. Based on these studies, we expect that retinal slip is smallest in the object-rotation condition, intermediate in the subject-movement condition and largest in the object-translation condition. For large-field stimulation we find performance in all conditions to be roughly equal. We could not find any studies comparing retinal slip for a large field of view in all our different movement conditions. For stimulus characteristics comparable to our large-field stimulation, van den Berg and Collewijn (1986, *ibidem*) found a retinal slip velocity of 0.55 deg/sec for a target moving together with a large field. This shows that retinal slip is probably quite small in our object-translation condition, compared to the retinal slip for small-field stimulation. The retinal slip is probably not much smaller in the other two movement conditions, which is consistent with our finding that performance is equal in the different movement conditions. A final point which might influence performance is the fact that subjects always had a part of the stimulus in central vision in large-field stimulation, whereas they had more trouble in fixating the stimulus for the object-translation condition with small-field stimulation. In this last case the stimulation was not always in central vision.

Further, we have evidence that ego-motion information is helpful in disambiguating the sign of curvature, at least when perspective information is not strong enough to disambiguate the curvature of the surfaces. Rogers and Rogers (1992) found that both ego-motion information and perspective information can disambiguate the curvature sign. They used stimuli that were far above the threshold of curvature detection with a depth extent of 3 cm and a field of view of 17 deg. For large-field stimulation all movement conditions in our study lead to unambiguous percepts of the sign of

curvature. The depth range present in our stimuli, for a curvature of  $0.4 \text{ m}^{-1}$ , is 6.5 cm. For small-field stimuli, the subject-movement condition leads to unambiguous percepts of the sign of curvature and the object-rotation condition to many reversals of the sign of curvature. Thus full proprioceptive ego-motion information was enough to disambiguate the sign of curvature whereas the small amount of perspective information (0.52 cm for a curvature of  $4 \text{ m}^{-1}$ ) in the object-rotation condition was not enough. Whether the information from tracking eye movements alone is sufficient to disambiguate the sign of curvature, is not clear from our results but the control session (large-field set-up with small-field stimuli) seems suggestive that they are.

As far as object-rotation is concerned, the finding that many ambiguities are found for a small field of view but not for a large field of view needs to be explained. Assessing the respective roles of *perspective information* (the difference between images from concave and convex surfaces increases with the view angle in perspective projection) and of the *field size* is difficult because these two variables covary in our experiment. However, two findings suggest is mainly field size which influences the ambiguity. First, apparent deformations of the stimulus are frequently seen with small-field stimulation but never with large-field stimulation. These deformations indicate that the amount of perspective information with small-field stimulation is large enough to make the images of concave and convex surfaces visibly different. Second, the amount of perspective information increases with surface curvature and we find that the number of errors in detecting the sign of curvature sometimes increases with increasing curvature. Hence, we speculate that rigidity of the object plays a more important role for peripheral vision, where most motion is caused by ego-motion and thus where objects are moving rigidly, than for central vision where we often observe nonrigid motion in natural situations.

The performance in curvature detection in the small-field object-rotation condition can be compared to the results of Norman and Lappin (1992). They found a percentage of correct responses of 96.5% for the discrimination between a plane and a concave sphere with a curvature of  $4 \text{ m}^{-1}$ . Their field of view was smaller than ours (2 vs 8 deg) whereas their amplitudes (35 vs 10 deg) and movement frequencies (1.2 vs 0.33 Hz) were higher than ours. Further they gave the subjects feedback about their performance. Despite these differences we find a comparable performance in detection of absolute curvature of 93% (average over our three subjects).

Contrary to Rogers and Graham (1979) we find that curvature detection is best in the object-rotation condition for small-field stimulation. Besides differences in stimulus geometry and field of view there are two important differences. First, in the experiment of Rogers and Graham the subject could see the layout of the experimental room, and this visual information about ego-motion could be responsible for the largest depth being perceived during self-motion. Second, the task of the subjects of Rogers and Graham was to estimate the

amount of depth of the corrugations, whereas our subjects detect the curvature.

Droulez and Cornilleau-Pérès (1990) proposed that the curvature of moving surfaces could be detected through the coding of spin variation—a second-order variation of the optic flow—in the visual pathway. The spin variation is a local operator which is directly proportional to surface curvature when the tangent plane is fronto-parallel. Therefore for similar movements and surface curvature its value in the centre of the median image of our stimuli is identical whatever the size of stimulus. It follows that this model seems incompatible with the tenfold increase in curvature sensitivity which we observed when field size increases. However the model of spin variation can still explain this increase for two reasons. First for a given curvature, the average value of the absolute spin variation over all directions ( $m_{sv}$ ) increases with the slant of the tangent plane. In particular, when the curvature increases from  $0.4 \text{ m}^{-1}$  in large-field to  $4 \text{ m}^{-1}$  in small-field, we calculated that  $m_{sv}$  only doubles in the periphery of our stimuli (while it is multiplied by 10 in the centre). Second, when the diameter of stimulation is multiplied by 10, the area is multiplied by 100. Therefore, within a model of spin variation the number of curvature detectors and the accuracy of the curvature coding would be much higher for our large-field than for our small-field stimuli.

## REFERENCES

- Azarbayejani, A., Starner, T., Horowitz, B. & Pentland, A. (1993). Visually controlled graphics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*, 602–605.
- Ball, K. & Pierowsky, M. (1987). A modified direct linear transformation method (DLT) calibration procedure to improve the accuracy of 3d reconstruction for large volumes. In de Groot, G., Hollander, A. P., Huijting, P. A. & van Ingen Schenau, G. J. (Eds), *Biomechanics XI-B*, Vol. 7-B of the *International series on biomechanics*. Amsterdam: Free University.
- Baloh, R. H., Beykirch, K., Honrubia, V. & Yee, R. D. (1988). Eye movements induced by linear linear acceleration on a parallel swing. *Journal of Neurophysiology*, *60*, 2000–2013.
- van den Berg, A. V. & Collewijn, H. (1986). Human smooth pursuit: Effects of stimulus extent and of spatial and temporal constraints of the pursuit trajectory. *Vision Research*, *26*, 1209–1222.
- Braunstein, M. L. & Todd, J. (1990). On the distinction between artifacts and information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 211–216.
- Bronstein, A. M. & Gresty, M. A. (1988). Short latency compensatory eye movement responses to transient linear head acceleration: A specific function of the otolith-ocular reflex. *Experimental Brain Research*, *71*, 406–410.
- Buizza, A., Leger, A., Droulez, J., Berthoz, A. & Schmid, R. (1980). Influence of otolith stimulation by horizontal acceleration on optokinetic nystagmus and visual motion perception. *Experimental Brain Research*, *39*, 165–176.
- 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. (1990). Three-dimensional motion perception: Sensorimotor interactions and computational models. In Warren, R. & Wertheim, A. (Eds), *The perception and control of self-motion* (pp. 81–100). London: Erlbaum.
- Cornilleau-Pérès, V. & Droulez, J. (1993). Stereo-motion cooperation and the use of motion disparity in the visual perception of 3-D structure. *Perception & Psychophysics*, *54*, 223–239.
- Cornilleau-Pérès, V. & Droulez, J. (1994). The visual perception of 3D shape from self-motion and object-motion. *Vision Research*, *34*, 2331–2336.
- Droulez, J. & Cornilleau-Pérès, V. (1990). Visual perception of surface curvature. The spin variation and its physiological consequences. *Biological Cybernetics*, *62*, 211–224.
- Erkelens, C. J., van der Steen, J., Steinman, R. M. & Collewijn, H. (1989). Ocular vergence under natural conditions. I. Continuous changes of target distance along the median plane. *Proceedings of the Royal Society of London B*, *236*, 417–440.
- Ferman, L., Collewijn, H., Jansen, T. C. & van den Berg, A. V. (1987). Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique. *Vision Research*, *27*, 811–828.
- Hayashibe, K. (1991). Reversals of visual depth caused by motion parallax. *Perception*, *20*, 17–28.
- Horn, B. K. P. (1987). Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, *4*, 629–642.
- Koenderink, J. J. & van Doorn, A. J. (1992). Second-order optic flow. *Journal of the Optical Society of America A*, *9*, 530–538.
- Mok, D., Ro, A., Cadera, W., Crawford, J. D. & Vilis, T. (1992). Rotation of Listing's plane during vergence. *Vision Research*, *32*, 2055–2064.
- Nakayama, K. (1981). Differential motion hyperacuity under conditions of common image motion. *Vision Research*, *21*, 1475–1482.
- Norman, J. F. & Lappin, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, *51*, 386–396.
- Post, R. B. (1988). Circular vection is independent of retinal eccentricity. *Perception*, *17*, 737–744.
- Rogers, B. J. & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.
- Rogers, B. J. & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, *22*, 261–270.
- Rogers, S. & Rogers, B. J. (1992). Visual and nonvisual information diambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, *52*, 446–452.
- Sperling, G., Doshier, B. A. & Landy, M. S. (1990). How to study the kinetic depth effect experimentally. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 445–450.
- Stoffregen, T. A. (1985). Flow structure versus retinal location in the optical control of stance. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 554–565.
- Warren, W. H. & Kurtz K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics*, *51*, 443–454.
- Waxman, A. M., Kamgar-Parsi, B. & Subbarao, M. (1987). Closed-form solutions to image flow equations for 3D structure and motion. *International Journal of Computer Vision*, *1*, 239–258.

---

*Acknowledgements*—This work was partly supported by the Foundation for Biophysics of the Netherlands Organisation for Scientific Research (NWO) and by MUCOM (ESPRIT BRA 3149 and 6615). We would like to thank Peter Werkhoven for stimulating discussions concerning the experimental procedure.