



Influence of Surface Attitude and Curvature Scaling on Discrimination of Binocularly Presented Curved Surfaces

SJOERD C. DE VRIES,* ASTRID M. L. KAPPERS,* JAN J. KOENDERINK*

Received 8 March 1993; in revised form 13 January 1994

We report on the ability of human observers to discriminate local second-order shape of quadratic stereo-defined surfaces. Local second-order shape can be specified by two parameters: the curvedness (a scale-dependent quantity describing overall curvature of a shape) and the shape index (a scale-independent quantity describing the shape's appearance). We measured shape index discrimination thresholds of shapes subject to attitude and curvedness transformations. Results show that neither slant nor curvedness manipulations affect or bias observers' judgements of surface shape. Furthermore, ideal detector simulations show that observers do not perform optimally: they do not exploit all available information.

Stereopsis Shape discrimination Surface curvature Disparity curvature Slant

INTRODUCTION

In some aspects, the human stereo vision system is remarkably robust and accurate. An example of robustness is the fact that it can handle a 15% isotropic scaling of the image of one eye (Julesz, 1971). The accuracy of stereo is illustrated by the fact that at the fixation point, performance of depth discrimination is comparable to that of two-dimensional vernier tasks (Berry, 1948).

An intriguing question is whether the extraction of higher-order depth structure is also robust and accurate. In the field of three-dimensional shape-from-stereo research, Rogers and Cagenello (1989) show that curvature of cylinders situated at different distances from the observer could be compared accurately. This result, however, seems to contradict the findings of Johnston (1991), who reports a dependence of the perceived curvature on the observation distance. Actually, in Johnston's case, the task was to compare depth and height of the presented elliptical cylinders and not to judge curvature. The ratio of depth and height needed to perceive the cylinder as circular varies at different viewing distances. At far distances, Johnston's results suggests an apparent flattening of the shape, while at close range they indicate an elongation in the radial direction. This is explained by assuming that observers use a *scaling distance* that differs from the real viewing distance.

In our research we primarily explore the robustness of

three-dimensional shape perceptions with respect to a family of quadratic shapes subject to curvature scaling and attitude changes. As a byproduct of this research we are able to discuss the effects of misestimation of distance at our (fixed) viewing distance.

1.1. Shape definitions

We use shape measures which were proposed by Koenderink (1990) and which have been exploited in a number of recent papers (de Vries, Kappers & Koenderink, 1991, 1993; Erens, Kappers & Koenderink, 1991; van Damme & van de Grind, 1991, 1993).

These shape definitions are based on two quantities that play an important role in the description of local shape: the principal curvatures. It is known from differential geometry that of the infinitely large family of curves generated by taking normal sections through a point on a surface, the curves that have minimum and maximum curvature (the principal curvatures K_{\max} and K_{\min}) lie in perpendicular planes. The curvature of the curves of all other normal sections relate to these principal curvatures in a very simple manner:

$$\kappa(\phi) = K_{\max} \cos^2 \phi + K_{\min} \sin^2 \phi, \quad (1)$$

where ϕ is the angle between K_{\max} and the curve in question (see e.g. Aleksandrov, Kolmogorov & Lavrent'ev, 1963). Equation (1) is called *Euler's formula*. Local shape can be characterized by a point in K_{\max} , K_{\min} space. In this space we find spherical shapes on the $K_{\max} = K_{\min}$ line, cylinders on the $K_{\max} = 0$ and the $K_{\min} = 0$ lines and symmetric saddles on the $K_{\max} = -K_{\min}$ line (see Fig. 1). Similar shapes have the

*Utrecht Biophysics Research Institute, Department of Medical and Physiological Physics, Princetonplein 5, 3584 CC Utrecht, The Netherlands [Email sjoerd@fysae.fy.ruu.nl or sdevries@fys.ruu.nl].

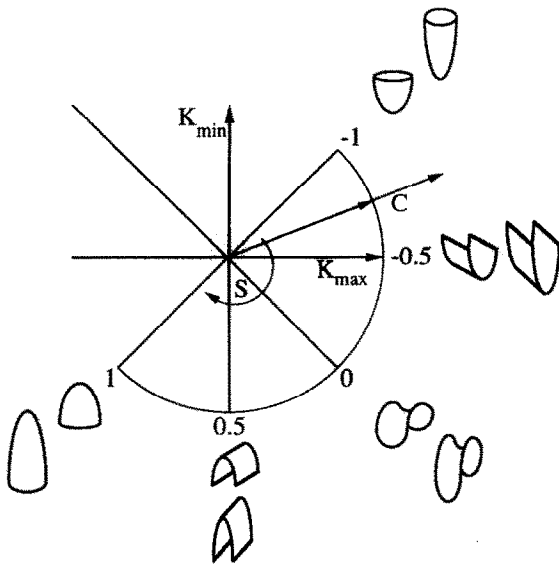


FIGURE 1. Objects in the K_{\max} , K_{\min} space. In this space, objects can be characterized by an angular coordinate, the shape index S , and a radial coordinate, the curvedness C .

same angular coordinate; varying the radial coordinate changes only the overall "amount" of curvature, but not its intuitive shape description.

Pursuing this description of the K_{\max} , K_{\min} space in polar coordinates, Koenderink (1990) defined two quantities, the shape index (S) and the curvedness (C) as

$$S = -\frac{2}{\pi} \arctan \frac{K_{\max} + K_{\min}}{K_{\max} - K_{\min}} \quad (2)$$

$$C = \sqrt{\frac{K_{\max}^2 + K_{\min}^2}{2}}, \quad (3)$$

where S is the angular coordinate specifying shape type or "quality of shape" and C the radial coordinate specifying "quantity of shape" (see Fig. 1). The shape index ranges from -1 to 1 ; it parametrizes a continuous family of shapes. This range can be subdivided into a number of surface types: from -1 to $-\frac{1}{2}$ concave elliptical, from $-\frac{1}{2}$ to $\frac{1}{2}$ hyperbolic and from $\frac{1}{2}$ to 1 convex elliptical. The curvedness varies from 0 cm^{-1} (flat surfaces) to infinity (sharp edges, needles). Unlike the shape index, the curvedness is scale dependent. Changing the curvedness of a patch is equivalent to scaling.

Previous studies (de Vries *et al.*, 1991, 1993) make it clear that these quantities provide a convenient parameterization for psychophysics. Subjects quickly become accustomed to the use of these quantities and are able to assign reasonable accurate shape index values to quadratic shapes.

1.2. Aim of study

From the definition of S and C , it follows that they are mathematically independent. To find out whether or not these quantities can be extracted independently by the human visual system is one of the goals of our psychophysical experiments.

Another quantity that can be independently varied is the slant, which we define as the angle between the surface normal at the point of interest and the line of sight. In the binocular case, the line of sight is commonly defined as the line connecting the point of interest and the cyclopic eye. This cyclopic eye, then, is often defined as lying in between both eyes, in the middle of the straight line connecting them, or as lying behind the eyes on the intersection of the Vieth-Müller circle and the line that bisects the lines of sight of each eye. In most cases (object not too close, and not far from the forward direction) the different cyclopic eye definitions do not make much difference. We will use the second definition. In our experiments, the point of interest is always on the z -axis, which goes through the cyclopic eye (the z direction is the straight-forward direction). Hence, the slant is also the angle between the surface normal and the z -axis.

Slant is not an object property, but a feature of the observer-object geometry and as such can be adjusted independent of shape. Of course, slant does influence the retinal image of an object, as can be seen in Fig. 2, which depicts the shift of disparity levels due to a change in slant.

Errors in the measurement of slant could very well propagate into surface curvature measures. This is evident from the expression for the surface curvature in a certain direction (in this case along the x -axis):

$$\kappa_x = \frac{z_{xx}}{(1 + z_x^2)^{3/2}}. \quad (4)$$

The second derivative which dominates this equation in near fronto-parallel situations ($z_x \approx 0$) might be approximated by the difference between two slant measurements, because slant is related to the first derivatives of the surface. In oblique situations, uncertainty in the slant determination has an additional effect, because of its presence in the denominator of equation (4).

Rogers and Cagenello (1989) point out that disparity gradients, which are connected to slants, vary roughly in inverse proportion to distance. Hence, slant estimation depends crucially on a correct determination of distance. In principle, absolute distances cannot be measured on the basis of retinal information only. One also needs knowledge of interocular distance and particularly the vergence angle (see Appendix). The motor signal from the eye muscles, which could be a measure of vergence, appears to be only crudely available to the visual system (Richards & Miller, 1969; Steinman & Collewijn, 1980; Regan, Erkelens & Collewijn, 1986; Foley, 1991; Johnston, 1991). Hence, a system that calculates curvature from slants by multi-local slant estimates is prone to error.

However, it is also possible that surface curvature is derived more directly from disparity information. Rogers and Cagenello (1989) suggest that curvature disparities would be suited for this task. They also show that the second derivative of the disparity field is roughly independent of the distance from the observed object, whereas by halving the viewing distance absolute

disparities are quadrupled and disparity gradients are squared in magnitude. Subjects perform well in equating curvatures of cylinders at different distances, which indicates that second-order derivatives are being used. This notion is supported by the fact that slanted flat surfaces (without a contrasting background) do not induce a percept of slant well (Mitchison & McKee, 1990; Gillam & Ryan, 1992; Cagenello & Rogers, 1993).

The aim of research reported here is threefold: to determine how competent the visual system is in discriminating shape, how shape discrimination depends on overall curvedness (and indirectly, scale), and how surface attitude influences the discrimination performance.

To achieve these goals we devised two experiments. In the first experiment we measured shape index discrimination thresholds. The objects were presented fronto-parallel and had a fixed curvedness value. For a number of subjects, the influence of another, higher, fixed curvedness value was examined.

In the second experiment, we used three conditions: a randomized curvedness condition, a randomized slant condition and a combination of both. Again we measured discrimination thresholds.

2. GENERAL METHOD

2.1. Stimuli

The images we presented were random-dot stereograms (Julesz, 1960, 1971), so most monocular cues were eliminated. The small density cue was diminished by selecting half of the number of points from a uniform random distribution in the right eye image, projecting them on the surface of the object to be depicted and then projecting them back to the left eye image and vice versa. In this way the density cue became virtually invisible and tests showed that no shape information whatsoever could be extracted, monocularly.

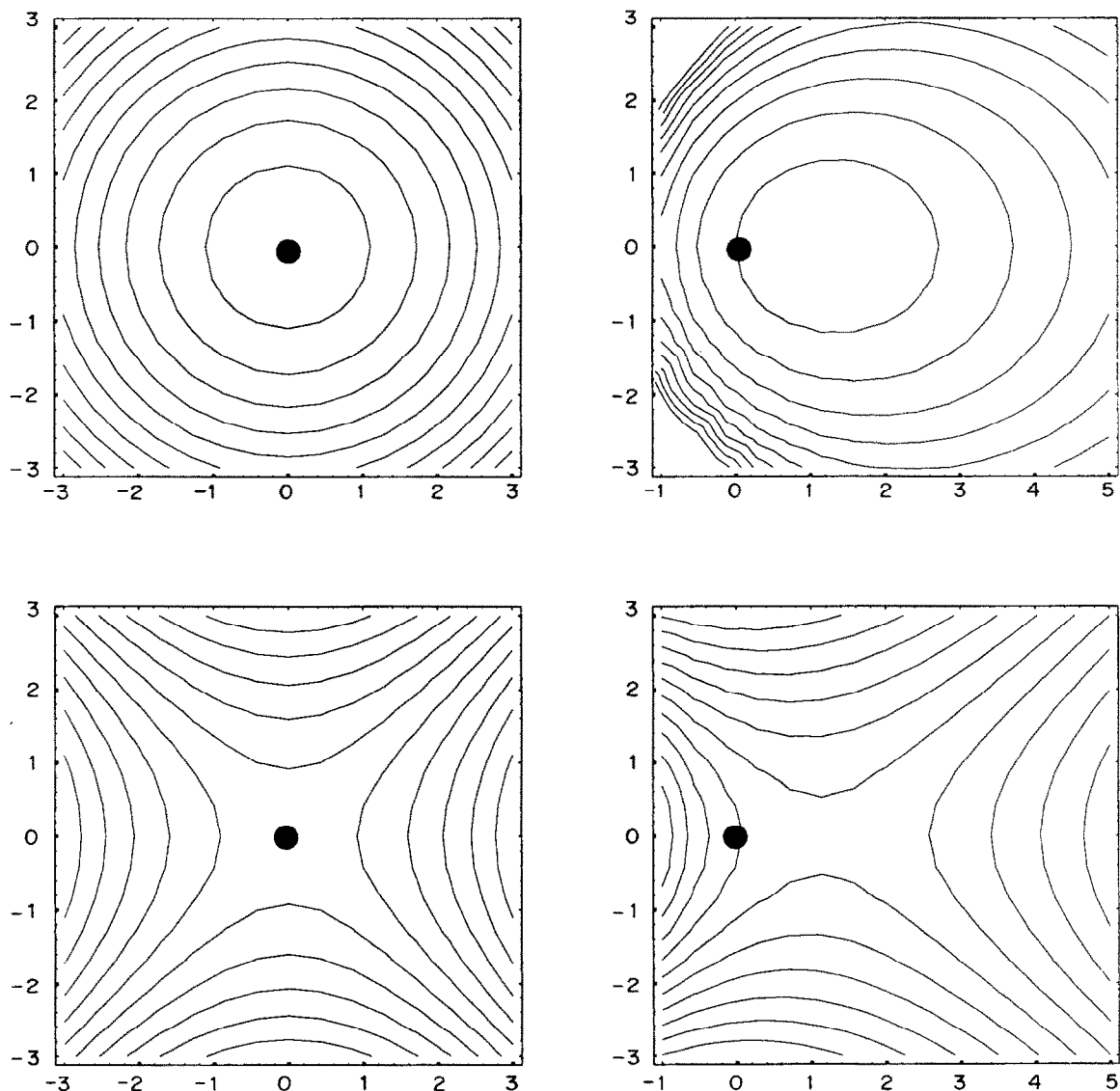


FIGURE 2. Disparity levels for two shapes. Upper panels, $S = 1$; lower panels, $S = 0$; $C = 0.5 \text{ cm}^{-1}$; scales are in cm. The right-hand panels show the effects on these disparity levels of slanting over an angle of 30 deg. The position indicated by the dot is always in the centre of the screen.

The probability that a pixel would be set was 15%. As a test, we generated several hundreds of random-dot patterns for each of a number of surfaces. The pixel values for each position were added together and divided by the number of images generated. In this way we obtained an average density for each position. The density distribution was virtually identical to that of a binomial distribution and centred around 15%. Only in very degenerate situations (like viewing a plane which slants in such a way that it almost points to one eye) does the density distribution deviate noticeably from a uniform random distribution.

The random-dot stereograms depicted quadratic shapes given by:

$$Z = z_0 + \frac{1}{2}(K_1x^2 + K_2y^2), \quad (5)$$

where the z -direction is towards the observer, the y -direction is vertical and the x -direction is horizontal; the origin is the centre of the screen. The form of equation (5) was chosen so that the principal curvatures K_{\max} , K_{\min} are given by K_1 and K_2 , as can easily be checked. K_1 and K_2 can be calculated from equations (2) and (3). In this way we are able to generate shapes with a specific shape index and curvedness.

We stress here, that shape index and curvedness are purely local measures. Every point on a surface has its own S and C value. Generally, the shape index and curvedness change all over the surface. An example of an exception is the sphere, which has the same S and C value everywhere. A paraboloid of rotation with $C = 0.5 \text{ cm}^{-1}$ has at its top a shape index value of 1, while at a horizontal distance of about 3 cm from this

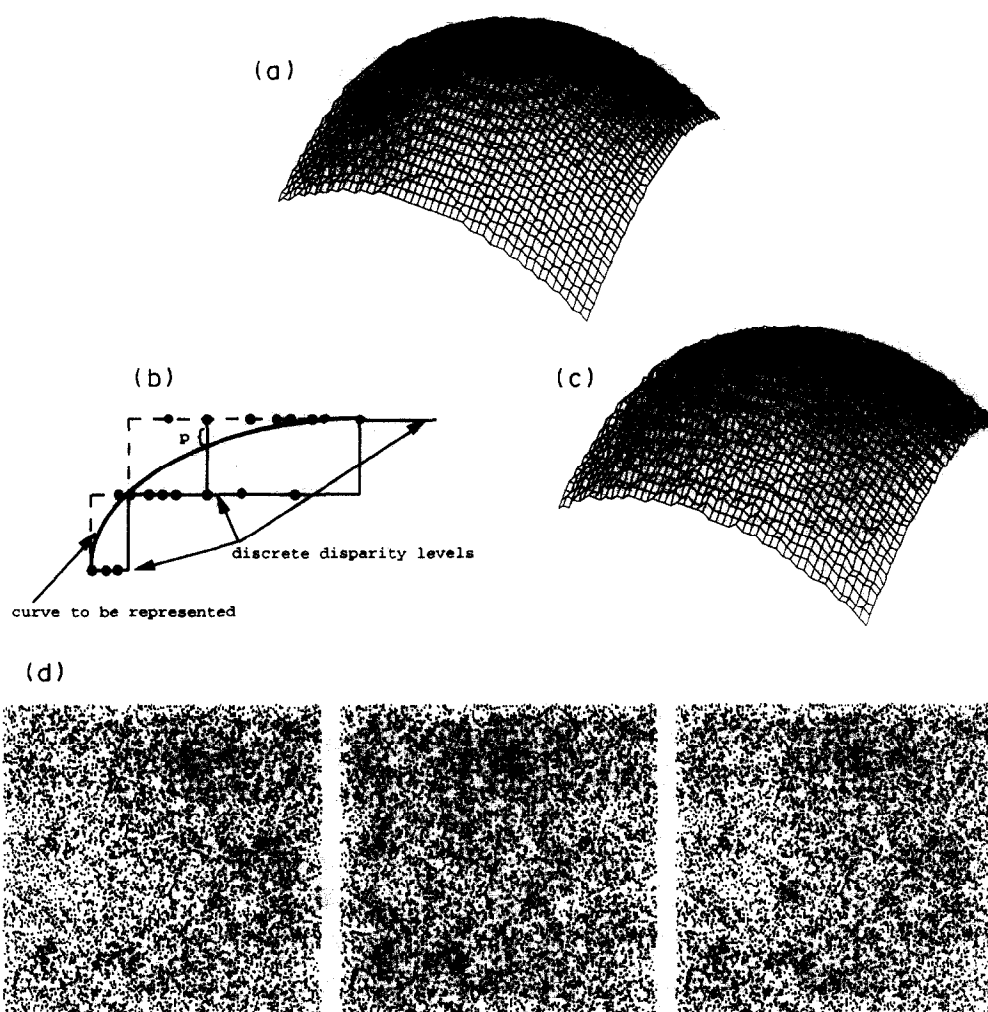


FIGURE 3. The effects of discretization of the disparity levels on a shape with shape index 0 and curvedness 0.5 cm^{-1} . A shape presented in stereo on a raster graphics display necessarily shows terrace-like features (a). Note: the dot density in this figure differs from the dot density in our set up; this is for illustrative purposes. In (b) we show how the terraces can be diminished. Setting the exact height of the dots becomes a probability process. The smaller the distance (p) between the exact height and the next higher disparity value, the higher the probability the dot will be set at this height. The results of this process can be seen in (c). The shape looks smoother and the terraces have almost gone. A random-dot stereogram generated with this method can be seen in (d), where the right-hand pair is for cross viewers and the left-hand pair for parallel viewers. To provide a better illustration, we used some settings that differ from our set-up. The actual boundary was blurred extensively, the pixel density higher and the dimensions larger.

point the shape index value has dropped to 0.65; the local shape is approaching that of a cylinder. When talking about *the* shape index of an object like the one defined in equation (5) we always mean the value at $(0, 0, z_0)$. From the generating equation it follows that the tangent plane to the surface in this point is fronto-parallel.

The quadratic shapes subtended a circular domain with a radius of 3.5 cm (3.3 deg). The simulated quadratic surfaces were translated 7.5 cm in front of the screen ($z_0 = 7.5$ cm) to ensure that no part of the shape would ever lie behind the screen, since this could yield an unwanted cue.

Starting at 1.5 cm (1.4 deg) from the centre the shape was blurred by background pixels. The ratio of background to object pixels increased as a Gaussian function of distance from the unblurred region in such a way that at the boundary about 90% of the pixels were background. Again, this was done to remove an unwanted shape cue, in this case the tell-tale boundary contour. This trick works very well. The irregular fringe that is the result of this process, if visible at all, is very hard to use. A drawback is that untrained observers experience more problems perceiving blurred-boundary objects than unblurred ones. This is another indication that depth contrast plays an important role in stereopsis (Brookes & Stevens, 1989a, b; Stevens & Brookes, 1987, 1988; Rogers & Graham, 1983).

Due to the discrete character of the screen, disparities on the screen could only change stepwise, as is shown in Fig. 3(a).

In order to smooth the surface, the disparity of a point on the screen could be increased by 1 pixel, with a probability proportional to the rounding error that was made by truncating the calculated disparity of that point [see Fig. 3(b)]. In this way the perceived average height (z) of a small neighbourhood of a point was about the same as it would be in the continuous case. There is some evidence that this averaging does indeed take place (Westheimer, 1986; Westheimer & Levi, 1987; Parker & Yang, 1989). A disadvantage of this method is that the location of the surface in the z -dimension appears slightly blurred. A typical result is shown in Fig. 3(c, d).

To nullify possible effects of the orientation of the objects, each object was rotated around the normal at the surface in the point of interest, with rotation values chosen randomly between 0 and 360 deg.

2.2. Apparatus

To depict the stereo images we used an ATARI 1040 ST computer connected to a pair of LCD shutter spectacles. The images were calculated on an HP/Apollo DN 10000 workstation, because the ATARI could not generate them fast enough for them to be used for practical purposes.

The frame rate of the ATARI monitor (P4 white phosphor) is 70 Hz. Screen dimensions were 20.75×13.0 cm (18.7×12.2 deg), with 640×400 pixels. The LCD spectacles were toggled in synchrony with the vertical retrace of the monitor in such a way that after

every even vertical retrace the monitor showed the left eye image and after every odd retrace the right eye image. We used a chin rest to restrict head movements. Viewing distance was 60 cm. Experiments were done in a dark room; average luminance of the screen was 45 cd/m^2 .

2.3. Psychophysical procedure

We used a two-alternative forced-choice task combined with a probabilistic staircase paradigm. A reference and a test object were shown in random order for 3 sec each, separated by a 0.5 sec interval during which a white screen was shown (except in the case of subject CVG who had problems maintaining the correct vergence angle during and after this interval; in his case, we used a random-dot stereogram depicting a circular object at the same place as the other objects). The reference object had a fixed shape index; the shape index of the test object was varied depending on the answers of the subjects. The staircase procedure was designed to converge to a 76% correct score. At this point, the distance between test and reference object on the shape index scale defined the threshold value.

Subjects knew the shape index value of the reference object (S_0 ; the value was entered by the subjects at the start of each run) and knew that the test object (S) was either to the left or to the right (on the shape index scale) of the reference object. During a staircase procedure the sign of $S - S_0$ was constant. This sign was reversed each time we measured the discrimination threshold of a particular reference shape, except for the scale endings, which could be approached unidirectionally only. The subjects' task was to indicate which of the two objects had the higher/lower shape index. Alternatively, because the subjects knew the shape index value of the reference shape, the task could be rephrased as "which of the two objects is the reference object?" Responses were given by means of a joystick. No feedback was given.

The staircase procedure consisted of two phases. The first phase was used to approach the threshold shape index. The answers during this phase were not recorded. At least 10 responses were required before the subjects could enter the next phase. The starting distance of the zoom phase was normally set at the threshold of the previous measurement of the same reference shape index. There was no fixed criterion for entering the measuring phase; usually 10 presentations were enough to adapt to the task. Unusually high thresholds at the end of the zoom phase usually indicated that the subjects had used the wrong approach direction, i.e. they mistook the reference for the test object. In cases like this, the zoom phase was repeated. The measuring phase consisted of the presentation of 40 reference-test object pairs. The threshold distance was the average shape index distance between these reference-test pairs. Thresholds were determined for each reference value three or four times on different days. The average value of these measurements yielded the final threshold.

Due to the high sampling of the shape index scale (test objects were separated by 0.002 of the shape index scale), individual images were not repeated very often during the staircase process. To prevent recognition of the reference object on the basis of its texture, we generated 10 different versions of each reference object. That is to say, we made a different random-dot texture; all other characteristics (except rotation) of the shape remained the same. In this way, each reference object was shown on average five times. The fact that repetition (at a low rate) occurs in the case of both reference and test objects prevents repetition from being used as cue.

2.4. Subjects

Eight undergraduate students took part in some or all of the experiments as part of their course requirements. Also, a member of our department performed most of the experiments, as did one of the authors. Subjects were all knowledgeable about the purposes of the experiments. They had normal or corrected-to-normal vision and were tested with partially decorrelated random-dot test images (Julesz, 1971). All subjects were able to perceive the quadratic shapes within 3 sec (after a few trials). In calculating the stereograms we used the subjects' own interocular distances, which ranged from 5.7 to 6.6 cm.

3. EXPERIMENT 1: BASIC SHAPE DISCRIMINATION TASK

3.1. Procedure

In this experiment we used quadratic shapes with a fixed value of curvedness $C = 0.5 \text{ cm}^{-1}$; these were presented fronto-parallel. To be more precise, the surface normal at the point on the surface we were interested in (the intersection of the symmetry axes) was in the direction of the line of sight. The procedure was as described in the Methods section. All 10 subjects took part in this experiment; they either measured discrimination thresholds for one-half of the shape index scale only (subjects LE, MA, JO, TT, FG and RE) or for the whole scale (subjects SDV, CVG, MVD and LB). In another version of this experiment we used curvedness values of $C = 1.0 \text{ cm}^{-1}$ (subjects MVD, SDV and LB) and $C = 0.75 \text{ cm}^{-1}$ (subject LB). It should be clear that doubling the curvedness doubles all curvatures of the objects used.

3.2. Results and discussion

Graphs of the obtained thresholds as a function of shape index are shown in Fig. 4. The results show that thresholds are highest in umbilical areas (spherical convexities/concavities) and in the symmetric saddle region. We find the lowest thresholds close to the cylindrical regions. (See Fig. 4.)

A W-shape shows up for all subjects (or half a W for those who measured thresholds for one half of the scale only), but there is a strong intersubject variability.

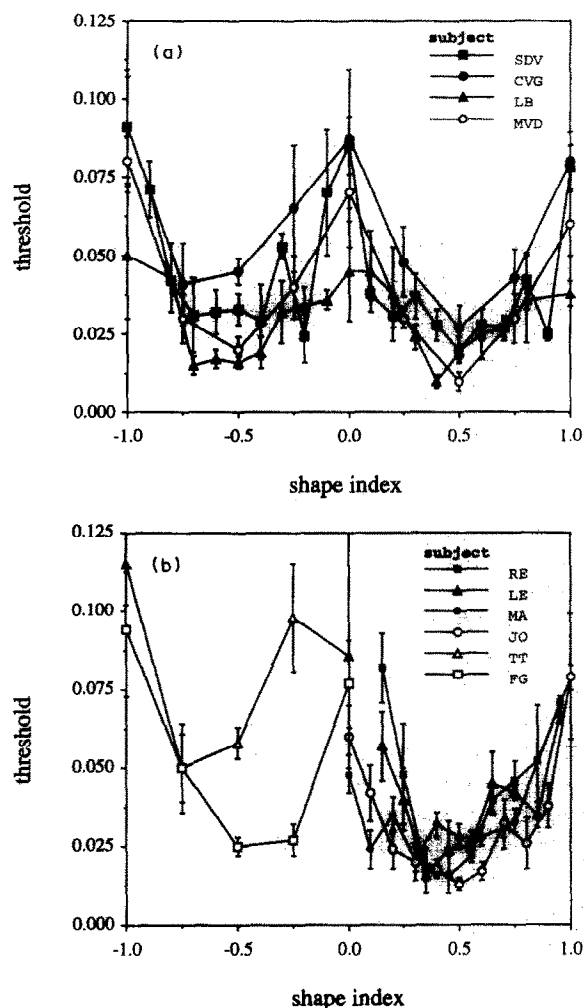


FIGURE 4. Discrimination thresholds as a function of the shape index for all participating subjects. Curvedness was fixed at $C = 0.5 \text{ cm}^{-1}$. Thresholds were defined as the distance along the shape index scale at which subjects scored 76% correct in the two-alternative forced-choice task. (a) Results for subjects who measured the whole shape index range. (b) Results for subjects who measured one half of the scale only.

Lowest thresholds vary between 0.01 and 0.05 (remember, these values are distances along the shape index scale, 2 being the maximum possible distance). Highest thresholds vary between 0.05 and 0.11.

We checked whether the W-shape was a result of the physical constraints inherent in the stimuli. In a later section we report on an ideal detector simulation, which shows that the response is determined by the visual system.

The results of the experiments with higher values of curvedness are shown in Fig. 5.

Although the curvedness is doubled, which means that *all* curvatures are doubled, the thresholds do not change noticeably. An analysis of variance (ANOVA) shows no influence of curvedness on the graph as a whole (subject SDV, $P = 0.49$; subject LB, $P = 0.90$; subject MVD, $P = 0.051$). In the area of the cylindrical shapes an influence of curvedness is discernible in the case of subjects SDV and LB. An ANOVA of the data points of the cylindrical areas does indeed show a

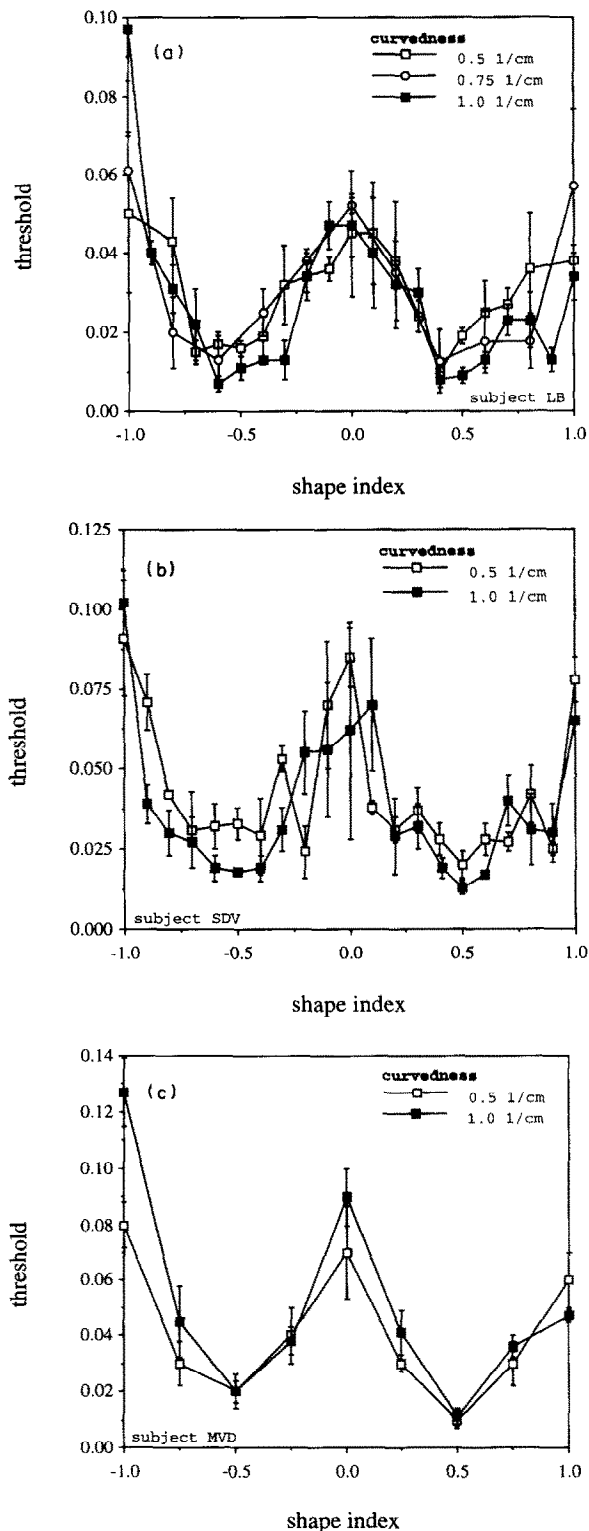


FIGURE 5. Discrimination thresholds measured for a number of fixed curvedness values. (a) The results for subject LB, (b) subject SDV, (c) subject MVD.

significant effect of curvedness for LB ($P = 0.0012$) and SDV ($P = 0.0001$) but not for MVD ($P = 0.18$).

The results of MVD differ in another aspect from those of SDV and LB. While in the latter cases the average thresholds decrease with increasing curvedness, the average thresholds of MVD rise. The mean values are: SDV, 0.042 and 0.038; LB, 0.029 and 0.022;

MVD, 0.042 and 0.054 for $C = 0.5$ and 1.0 cm^{-1} respectively, but, as we said above, these changes are not significant.

3.2.1. Shapes with constant shape index. One of the first proposed explanations for the lower values at the cylindrical regions was based on the observation that the shape index of a cylinder, unlike most other shapes, is the same all over the surface. For instance, the shape index close to the mask bordering a paraboloid of rotation ($S = 1.0$) is about 35% lower than in the centre. A constant shape index may help some kind of integration mechanism in that it increases shape measuring accuracy by lowering the influence of noise.

In order to verify this assumption, we performed a control experiment. Instead of using a paraboloid of rotation as our $S = 1.0$ shape, we used a hemisphere. Just like cylinders, spheres have a constant shape index all over their surface. And just as a cylinder, the sphere is not a generic shape, that is to say: it is no longer a sphere when there is an infinitesimal perturbation of the surface. A family of different shapes close to the sphere was generated by the expression $z = \sqrt{r^2 - \epsilon x^2 - y^2}$. Shape index and curvedness depend on r and ϵ . Hence, by manipulating r and ϵ we were able to keep the curvedness fixed at $C = 0.5 \text{ cm}^{-1}$ and generate the necessary shape index range.

To prevent the diameter of the distorted spheres being used as a cue for the shape index we randomized our mask size. Furthermore, the maximum value of the mask was always less than the diameter of the distorted sphere.

Thresholds were measured for subjects CVG, LB and SDV. These did not deviate significantly from previous results for the paraboloid of rotation ($P = 0.98, 0.93$ and 0.94 respectively).

What does this experiment show? First, the lower scores at $S = \pm 0.5$ are not due to the fact that the shape index is constant over these surfaces. Second, subjects are able to determine second-order shape features (like the shape index) on certain higher-order surfaces. A Taylor series development of a Monge-patch representation [i.e. $z = f(x, y)$] of the spherical surface contains prominent fourth-order terms which are missing in the shapes we used. Of course, the results should not be generalized to other higher order surfaces, because, for instance, the fourth-order terms in this case did not perturb the axes of symmetry of the second-order surface. Since the principal curvatures are aligned along the symmetry axis in this case we may have had a cue here. More research in the area of higher order surfaces is required.

In the General Discussion we will present a model which can account for the W-shaped threshold curve.

4. EXPERIMENT 2: INFLUENCE OF RANDOMIZED SLANT AND CURVEDNESS

In this experiment, we examined the influence of two parameters: curvedness and slant. We did this by randomizing their values. In the case of curvedness, this meant that subjects could not use the absolute curvature

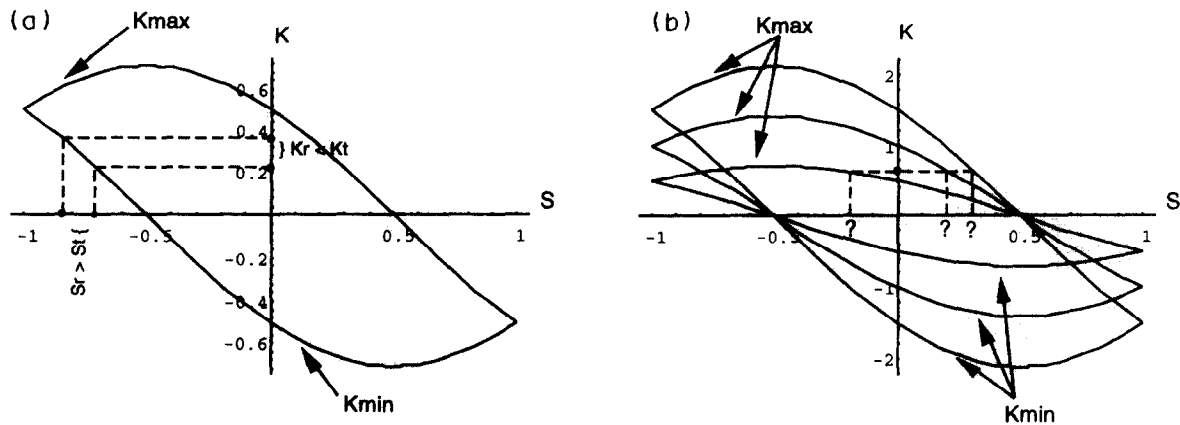


FIGURE 6. The relation between shape index value (x -axis) and principal curvatures (y -axis). The graphs show the need for an experiment with randomized curvature. Comparing one of the principal curvatures of the two objects (for instance the smallest in absolute value as some subjects claimed they did) yields enough information to perform the shape index discrimination task. In the example of (a), finding that one object has a smaller principal curvature than the other results in the conclusion that it must have a higher shape index. Randomizing the curvedness of both objects precludes the use of this possibility, as shown in (b). Curvature–shape index graphs are drawn here for three values of the curvedness, 0.5, 1.0 and 1.5 cm^{-1} . In the current experiment, we used 10 values between 0.3 and 1.25 cm^{-1} . In this set up, measuring one principal curvature no longer gives enough information. An ordering of two principal curvatures among the two objects does not give information about their ordering in the shape index dimension.

comparisons that they could in Expt 1 (see Fig. 6). This was possible because for fixed curvedness there is a direct relation between principal curvature and shape index value [see equations (2) and (3)]. This relation could have been utilized by the subjects to simplify the task.

In the randomized case, subjects had to compare the shapes using more global features than just one curvature. A calculation of the ratio of the principal curvatures, for instance, would get rid of the influence of randomizing curvedness, but this computation would yield another source of errors and thresholds would probably rise as a result. On the other hand, if subjects had not used absolute curvature measures previously, we would not expect thresholds to change.

4.1. Procedure

The procedure remained basically the same. Three slightly different versions were used. In the first version (referred to as “randc”), the curvedness values of the objects were selected randomly from a range between 0.3 and 1.25 cm^{-1} . These values were chosen so that the curvedness was neither too low (i.e. there were enough height steps) nor too high (i.e. a depth range that is small relative to the viewing distance).

In the second version (“rands”), we kept curvedness constant at $C = 0.5 \text{ cm}^{-1}$ while varying the slant of the object. The slant values were selected from the range 0–30 deg (SDV) and 0–27 deg (CVG, TT and FG), in discrete steps of 3 deg. Tilt (the direction of slant) was kept fixed at 0 deg, i.e. horizontal.

In the third version (“allrand”), the two random factors were combined. We noticed during the measurements made by subject SDV that a combination of high curvedness and slant values yielded situations where the centre of the quadratic shape is occluded by its rim. Consequently, we lowered the maximum slant

that we used with subjects CVG, TT and FG from 30 to 27 deg.

In these three versions, all other conditions were the same as those in Expt 1. Subjects CVG and SDV sampled the whole shape index scale and subjects TT and FG sampled the negative side (the results for the two sides do not appreciably differ anyway, see Fig. 4).

4.2. Results

The results of the three versions of Expt 2 and the basic experiment (“allcon”, Expt 1) are shown in Fig. 7. The intuitive prediction that adding features—that were expected to increase the difficulty of the task—would result in higher thresholds was not fulfilled. The lowest thresholds were generally not obtained in the simplest case, the allcon condition. Clearly, no common pattern of influence of experimental conditions was found among the subjects. Ranking the conditions by their mean threshold we get (from low to high means): TT: randc, allrand, rands, allcon; FG: allcon, randc, rands, allrand; CVG: allrand, allcon, rands, randc; SDV: rands, allcon, randc, allrand.

Performing an ANOVA on the data, we find significant effects of shape index ($P < 0.00001$) and of subjects ($P = 0.002$). There is an insignificant effect of experiment type ($P = 0.11$), and of session number ($P = 0.77$) showing there was no learning effect.

The individual data for each subject show an insignificant influence of experiment type for subject FG ($P = 0.14$) and subject CVG ($P = 0.07$). In the case of subjects TT and SDV, the influence is significant ($P = 0.0005$ and $P = 0.0037$ respectively). In subject TT’s case, this is due to the fact that all randomized conditions give rise to a noticeable lower threshold than the allcon condition. The significant influence of experiment type for subject SDV is due to the allrand

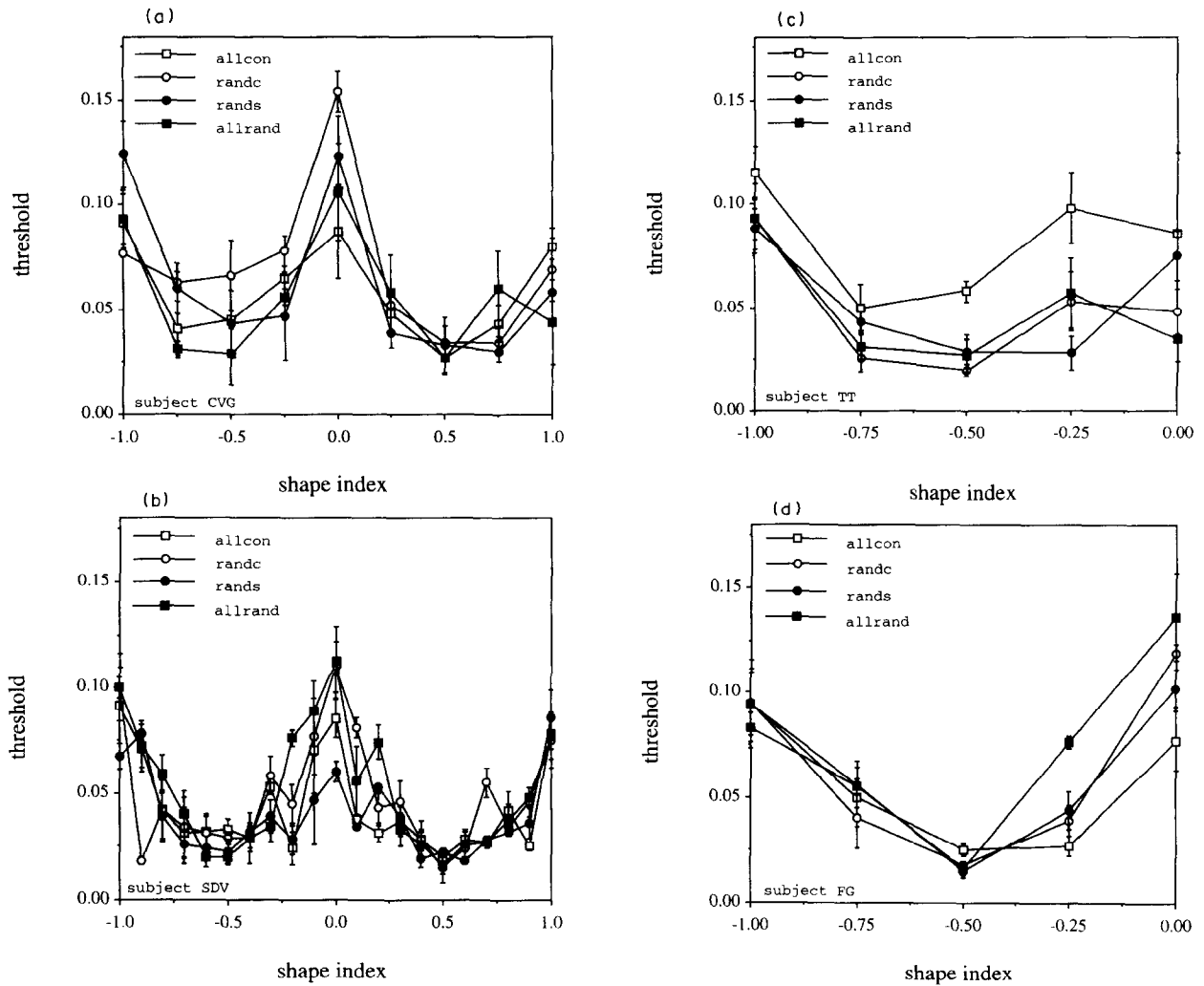


FIGURE 7. The results of the four versions of the discrimination experiment. Whole-scale measurements were done by subjects CVG and SDV (a, b) and half-scale measurements by subjects TT and FG (c, d).

condition. A comparison, by means of t -tests, of the mean threshold of the allcon condition with that of the other conditions, reveals a significant difference for the allrand condition only ($P = 0.06$, 0.19 and 0.006 for randc, rands and allrand respectively).

4.2.1. A closer look at the influence of the slant. In the section above, we compared two conditions of slant: no slant and randomized slant. There is sufficient data for us to analyse the influence of slant more precisely.

Each step in the staircase procedure consisted of two objects with randomly selected slant values. During each measurement, 40 pairs were shown and this was repeated three or four times. When this was done for different values we obtained a large database of reference slant, test slant and answer triples. This database can be used to calculate a score for each combination of slants. However, we did not calculate percentage correct scores, because in each pair, in addition to the two slants there was another parameter: the shape index distance between the two objects. Obviously, a correct answer with respect to two objects that differ considerably in shape index should be valued less than such an answer

in the case of objects that differ less in shape index. The reverse holds for incorrect answers.

Hence, we calculated a merit value which incorporated the distance between the shape index of the test object and that of the group's average (\bar{S}) (by "group" we mean the set of 40 responses during one staircase measurement) scaled by the group's standard deviation (σ). The score for a correct response was:

$$\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{|\bar{S} - S_{\text{ref}}| - |S_{\text{test}} - S_{\text{ref}}|}{\sigma} \right). \quad (6)$$

An analogue correction was applied to the incorrect answers.

We examined two aspects of the resulting matrix (scores indexed by slant values). First, there is the difference between the diagonal elements (reference and test slant equal) and the off-diagonal elements (unequal slants). Second, a calculation of the correlation of scores and slant values could tell us whether a fronto-parallel view yields better performance than a more oblique view.

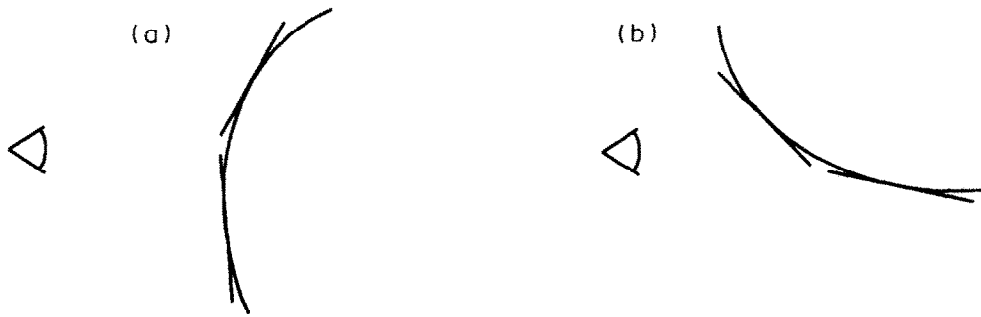


FIGURE 8. The same curve viewed from different angles. In (a), the point on the curve directly in front of the observer is almost fronto-parallel. In (b), this point has a much larger slant. One way of measuring the curvature in this point would be to take the difference between the slants of two nearby points. When the errors in the slant measurements follow a Weber-law, the errors in the curvature measurement should vary linearly with the amount of slant (assuming that no additional errors are made in the last stage of curvature calculation).

The results show that the diagonal values do not differ significantly from the off-diagonal elements ($P = 0.31$, $P = 0.72$, $P = 0.08$, $P = 0.82$, for subject SDV, CVG, FG and TT respectively). Interestingly, in three out of the four cases the off-diagonal scores are higher than the diagonal scores.

Spearman Rank Correlations of slant values vs scores were calculated. No significant correlations were found for either constant reference and varying test slant or constant test and varying reference slant.

4.3. Discussion

The addition of a random curvedness condition means that the subjects have to refrain from using absolute curvature comparisons. This did not increase the thresholds. Hence, it appears that the visual system is able to perform relative-curvature tasks reasonably well. This is close to saying that the visual system performs some kind of scale-independent shape measurement. Of course, this does not mean that it should necessarily be a shape index measurement.

The random slant experiment demonstrates that shapes can be discriminated irrespective of their slant (at least, within the range used). As we show in the Appendix, the perceived curvature is necessarily distorted (even for an ideal detector) if a faulty vergence estimate is applied, the amount of distortion being governed by the amount of slant and the amount of distance misestimation. Deviations are expected to be largest in the $S = \pm 1$ and 0 areas. The fact that the thresholds do not rise noticeably indicates that an overestimation—if present at all—is rather small and is certainly not as high as found by Johnston (1991) for this viewing distance. An explanation for this relatively good distance estimation may be that our viewing distance always remained the same. This might have helped to reduce the positional uncertainty.

If curvature is calculated using bilocal slant measurements instead of direct differentiation we might expect any specific error pattern in these slant measurements to show up in the curvature determination as well. This restricts the way in which errors emerge in the

determination of slant. In slant estimation, for instance, a Weber-law behaviour would yield curvature errors that vary linearly with the slant (see Fig. 8). We did not find this.

5. IDEAL DETECTOR SIMULATIONS

In order to judge human performance in the discrimination task, we performed an ideal detector simulation of this experiment. The detector was ideal in the sense that it received the dots in the stereograms as matched pairs, so that it did not have to deal with the *correspondence problem*. Otherwise, all conditions (discrete disparity levels, distances, number of sampled shape index values, staircase procedure etc.) were the same as for the human observers.

A polynomial surface was fitted to the obtained set of heights by means of a least squares method. From the fitted surface, the detector calculated the shape index. This value was then used in the two-alternative forced-choice task.

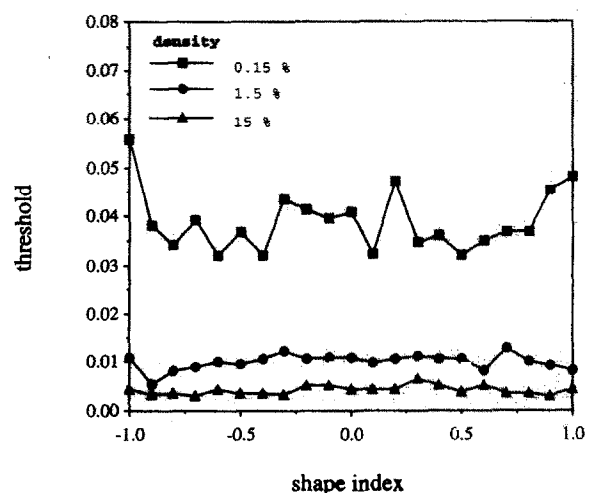


FIGURE 9. Thresholds obtained by an ideal detector which performed the same discrimination task as the human subjects who used a 15% density.

The detector scored thresholds between 0.0045 and 0.0055, which is about 10 times lower than in the human case. Also, the shape of the ideal detector threshold graph does not resemble the human W-graph at all (see Fig. 9).

The product moment correlation between the detector and the human threshold values (subject SDV) is $r_{19} = -0.16$, $r_{0.05,19} = 0.43$ (for one particular run of the detector, other runs with new random seeds yield different, but low correlation values).

If we decrease the dot density, threshold levels of the ideal detector rise. Thresholds were found to vary in inverse proportion to the square root of the number of points, as might be expected. Hence, at a dot density of 0.0015, 1/100th of the standard density, the thresholds of the ideal detector are of the same order as those of the subjects, though still differing in shape index dependency. In this case we find the insignificant correlation value of $r_{19} = 0.31$.

Noticeable is the finding that thresholds at $S = \pm 1$ are now about 20% higher than thresholds at other shape-index values. This might indicate that a small fraction of the performance at these values is due to the discreteness constraints in the stimuli or to their degenerate nature (all normal curvatures are principal curvatures in the umbilical case). The latter possibility seems to be the more probable, because the same principal curvatures that are present in the symmetric convex and concave shapes are also present in the symmetric saddle; had the depiction of a certain curvature value caused any problems, then we would certainly expect these problems with saddles as well.

A control experiment revealed that the human visual system reacts differently to lower dot densities. We measured the thresholds at $S = 1.0$ (subjects CVG, LB and SDV) and at $S = 0.5$ (subjects LB and SDV) using stereograms with a dot density of 0.015 (1/10th of the standard density). We found no significant differences ($P = 0.13$, $P = 0.17$ and $P = 0.38$ respectively in the $S = 1.0$ case and $P = 0.63$ and $P = 0.26$ respectively in the $S = 0.5$ case) between the threshold value at low and high density. This means that, unlike the ideal observer, the human observer does not use all available information.

6. GENERAL DISCUSSION

6.1. Consistency of the result with a model using noise considerations

Can we think of a mechanistic explanation for the W-shaped curves? As a starting point, we assume that subjects utilised quantities closely connected to local surface shape. One could use, for instance, a measurement of the principal curvatures (easy to find in a symmetric case like the quadratics; in general, a few more independent measurements are necessary to determine these quantities) or of Gaussian or mean curvature.

We examined how errors in the above-mentioned geometrical quantities propagate into a shape index discrimination threshold.

Thresholds were calculated as follows:

$$\Delta S^2 = \sum_{i=1}^N \left(\frac{\partial S}{\partial q_i} \right)^2 (\Delta q_i)^2, \quad (7)$$

where the q_i are the primary quantities, the determination of which underlies the shape index measurement. Interactions between these quantities are assumed to be negligible.

Rogers and Cagenello (1989) report that curvature discrimination (which they measured using cylinders) yields a Weber-law behaviour. This behaviour is usually modelled with multiplicative noise. We tried a combination of multiplicative and additive noise:

$$(\Delta q_i)^2 = (\alpha q_i)^2 + \beta^2. \quad (8)$$

Of course, this may turn out to be purely multiplicative if $\beta = 0$.

In the case of a measurement of the two principal curvatures, we end up with:

$$\Delta S = \frac{\sqrt{2}}{\pi} \sqrt{\frac{\beta^2}{C^2} + \alpha^2 \cos^2 \pi S}, \quad (9)$$

which has the desired W-shape.

Analogue calculations for a measurement of the Gaussian curvature also result in a W-shape, but this one behaves rather badly: it has asymptotes at $S = -1, 0$ and 1. The mean curvature yields a U-shaped curve (see Fig. 10).

Although the model describes the qualitative behaviour of the discrimination curve rather well, it accounts for only 40–80% of the variance of the data (see e.g. Fig. 11).

The parameter estimates resulting from the first are: α between 0.093 and 0.15, β between 0.015 and 0.03. We excluded subject TT here, who had exceptionally high values of 0.19 and 0.06, respectively (Fig. 7 shows how different his measurements were). α is equivalent to the Weber fraction for curvature discrimination. For this Weber fraction Rogers and Cagenello (1989) found values that range from 0.04 to 0.06. We can say that these results are close to ours in view of the fact that our object orientations were randomized and theirs were not.

The model takes the curvedness value into account. Hence, the experiments with different fixed curvednesses should yield the same parameter values. The actual values can be found in Table 1. It follows that the fitted parameters are reasonably stable. The model predicts a rather modest dependence of the thresholds on curvedness (just as we have found), except for the area around the cylinders [$S = \pm \frac{1}{2}$, see Fig. 10(c)].

The means for this area (loosely defined as the range from -0.6 to -0.4 and from 0.4 to 0.6) relate as 1:0.8:0.6 and 1:0.6 for LB and SDV respectively, whereas model predictions give 1:0.8:0.7 and 1:0.7 (the figures would have been 1:0.66:0.5 had we used $S = \pm \frac{1}{2}$ only). Again, subject MVD's responses deviate from the others, since his thresholds do not change in this area.

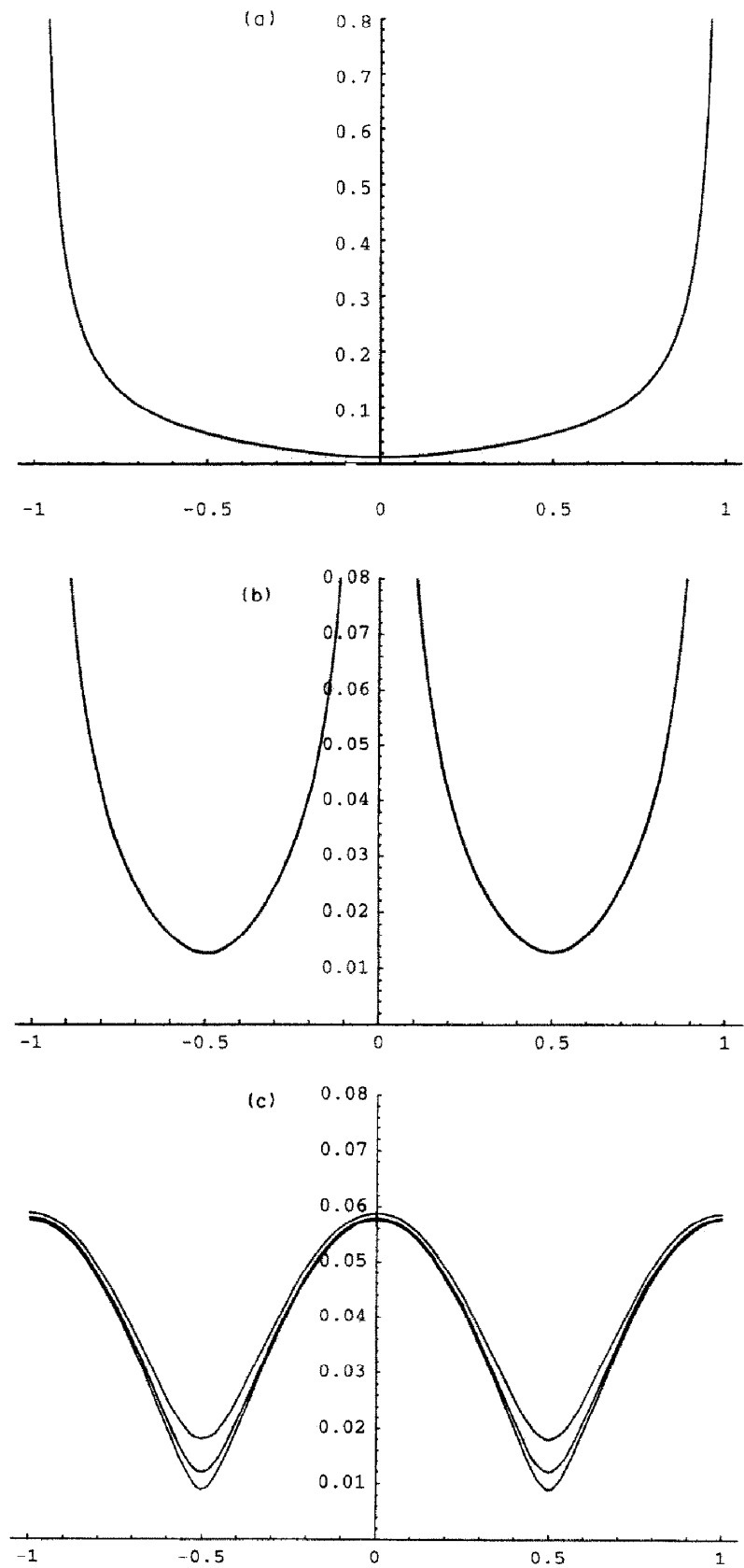


FIGURE 10. Thresholds according to three models. (a) The threshold calculated under the assumption that the Gaussian curvature is the underlying quantity of shape perception. (b) Using the mean curvature and (c) using the principal curvatures. This panel shows the curves for three value of the curvedness, from top to bottom 0.5, 0.75 and 1.0 cm⁻¹ respectively.

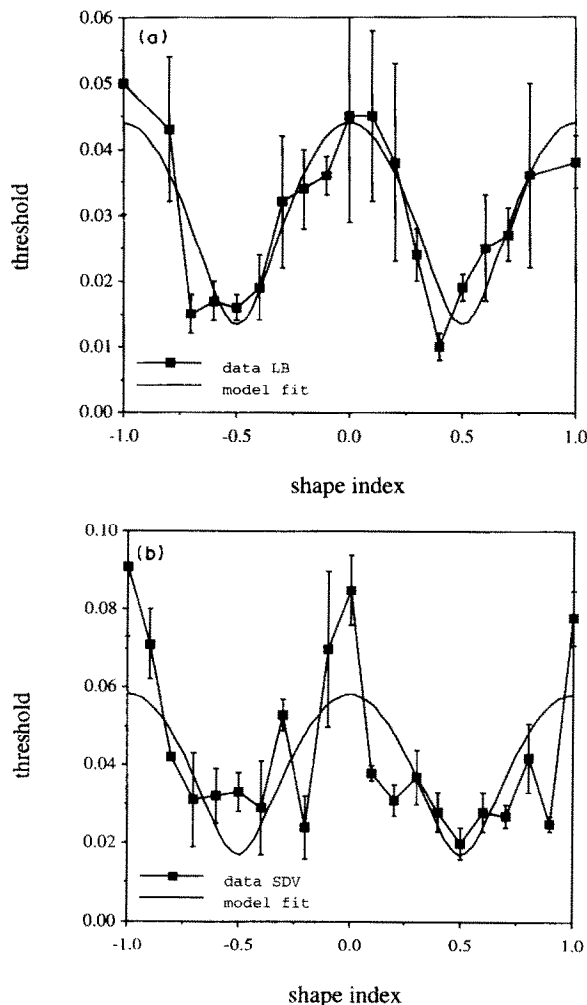


FIGURE 11. Model fits to the data of subjects LB (a) and subject SDV (b). The figure shows one of the best fits and one of the worst fits, respectively.

TABLE 1. Values of the fitted parameters

Subject	α			β		
	$C \text{ (cm}^{-1}\text{)}$			$C \text{ (cm}^{-1}\text{)}$		
	0.5	0.75	1.0	0.5	0.75	1.0
SDV	0.124		0.127	0.019		0.021
LB	0.093	0.095	0.10	0.015	0.015	0.001
MVD	0.13		0.16	0.015		0.027

α is the multiplicative factor, corresponding to a Weber fraction, β is a constant term.

7. CONCLUSIONS

The significance of this model is that it shows that the W-shape itself is not the essence of the results of this research. Probably, this is merely a consequence of the shape index definition and the way subjects arrive at a shape index value. The thresholds themselves are meaningful only when it can be shown that something like a shape index scale comes close to an internal representation of shape. Experiments using multi-dimensional-scaling analysis would presumably be needed to provide evidence for this.

An important finding is that when absolute curvatures cannot be used to perform the tasks, and relative curvature measures are used instead, the thresholds are not increased. The calculation of relative curvature measures appears to be rather accurate and could be a feature of normal visual operations. Equally important is our finding that the attitude of a shape does not interfere with shape judgement itself.

REFERENCES

- Aleksandrov, A. D., Kolmogorov, A. D. & Lavrent'ev, M. A. (1963). *Mathematics, its content, methods and meaning*. Providence, R.I.: American Mathematical Society.
- Berry, R. N. (1948). Quantitative relations among vernier, real depth and stereoscopic depth acuities. *Journal of Experimental Psychology* 38, 708-721.
- Brookes, A. & Stevens, K. A. (1989a). The analogy between stereodepth and brightness. *Perception* 18, 601-614.
- Brookes, A. & Stevens, K. A. (1989b). Binocular depth from surfaces versus volumes. *Journal of Experimental Psychology: Human Perception and Performance* 15, 479-484.
- Cagenello, R. & Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces: The role of orientation disparity. *Vision Research* 16, 2189-2202.
- van Damme, W. J. M. & van de Grind, W. A. (1991). Identification of 3D shapes using motion parallax. In Beek, P. J., Bootsman, R. J. & van Wieringen, P. C. W. (Eds), *Studies in perception and action* (pp. 72-75). Amsterdam: Rodopi.
- van Damme, W. J. M. & van de Grind, W. A. (1993). Active vision and the identification of three-dimensional shape. *Vision Research* 33, 1581-1587.
- Erens, R., Kappers, A. M. L. & Koenderink, J. J. (1991). Limits on the perception of local shape from shading. In Beek, P. J., Bootsma, R. J. & van Wieringen, P. C. W. (Eds), *Studies in perception and action* (pp. 72-75). Amsterdam: Rodopi.
- Foley, J. M. (1991). Binocular space perception. In Regan, D. (Ed.), *Binocular vision, vision and visual dysfunction* (Vol. 9, pp. 75-92). London: Macmillan Press.
- Gillam, B. & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception* 21, 427-439.
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research* 31, 1351-1360.
- Julesz, B. (1960). Binocular depth perception of computer-generated patterns. *Bell Systems Technical Journal* 39, 1125-1162.
- Julesz, B. (1971). *Foundations of cyclopean perception*. London: University of Chicago Press.
- Koenderink, J. J. (1990). *Solid shape*. Cambridge, Mass.: MIT Press.
- Mitchison, G. J. & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research* 30, 1781-1791.
- Parker, A. J. & Yang, Y. (1989). Spatial properties of disparity pooling in human stereo vision. *Vision Research* 29, 1525-1538.
- Regan, D., Erkelens, C. J. & Collewijn, H. (1968). Necessary conditions for perception of motion in depth. *Investigative Ophthalmology and Visual Science* 27, 584-597.
- Richards, W. & Miller, J. F. Jr (1969). Convergence as a cue to depth. *Perception & Psychophysics* 5, 317-320.
- Rogers, B. J. & Cagenello, R. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature* 339, 135-137.
- Rogers, B. J. & Graham, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science* 221, 1409-1411.
- Steinman, R. M. & Collewijn, H. (1980). Binocular retinal image motion during active head rotation. *Vision Research* 20, 415-429.
- Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics* 46, 183-195.
- Stevens, K. A. & Brookes, A. (1987). Depth reconstruction in stereopsis. *Proceedings of the First IEEE International Conference on Computer Vision* (pp. 682-686). Washington, D.C.: IEEE Computer Society.

- Stevens, K. A. & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research*, 28, 371–386.
- de Vries, S. C., Kappers, A. M. L. & Koenderink, J. J. (1991). 3D shape recognition with stereo. In Beek, P. J., Bootsma, R. J. & van Wieringen, P. C. W. (Eds), *Studies in perception and action* (pp. 58–64). Amsterdam: Rodopi.
- de Vries, S. C., Kappers, A. M. L. & Koenderink, J. J. (1993). Shape-from-stereo, a systematic approach using quadratic surfaces. *Perception & Psychophysics* 53, 71–80.
- Westheimer, G. (1986). Spatial interaction in the domain of disparity signals in human stereoscopic vision. *Journal of Physiology* 370, 619–629.
- Westheimer, G. & Levi, D. M. (1987). Depth attraction and repulsion of disparate foveal stimuli. *Vision Research* 27, 1361–1368.

Acknowledgements—This research is supported by the SPIN project “3D Computer Vision” of the Dutch Ministry of Economic Affairs. The authors wish to thank André Noest and Peter Werkhoven for their comments on the manuscript in its draft form and Lon Boonen for his contributions to the research.

APPENDIX

To calculate the effects of misestimation of the vergence angle we use bicentric coordinates $\gamma, \lambda, \epsilon$ (see e.g. Koenderink, 1992). γ corresponds to half the vergence angle, λ is the direction of the cyclopic line of sight and ϵ is the elevation out of the horizontal plane (see Fig. A1). In terms of these bicentric coordinates the Cartesian coordinates are written as:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{e}{\sin 2\gamma} \begin{pmatrix} \sin 2\gamma \\ \cos \epsilon (\cos 2\lambda + \cos 2\gamma) \\ \sin \epsilon (\cos 2\lambda + \cos 2\gamma) \end{pmatrix} \quad (\text{A1})$$

where e is half the interocular distance.

With bicentric coordinates we have a convenient tool to describe positions in terms of gaze directions. It is simple to calculate the perceived position of a point when a certain erroneous γ is assumed by the visual system. This can be done by first calculating the correct bicentric coordinates using the inverse of equation (A1), then changing one bicentric coordinate and finally finding the distorted Cartesian coordinates by means of equation (A1).

The same procedure can be followed for other objects like three-dimensional curves. To every assumed gaze direction there belongs an object that is distorted with respect to the original. When the visual

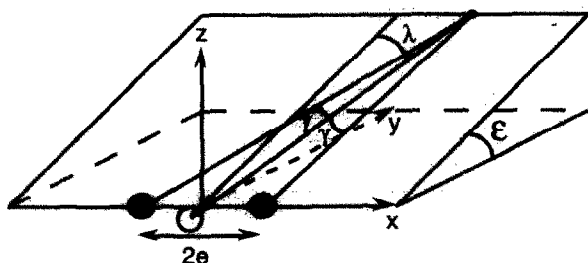


FIGURE A1. The definition of the bicentric coordinates $\gamma, \lambda, \epsilon$. γ equals half the vergence angle, λ is the angle between the straightforward direction and the cyclopic line of sight in the epipolar plane and ϵ is the elevation out of the horizontal plane.

system does not use the motoric information from the eye muscles it is unable to distinguish any of the objects from this set.

We used planar curves to calculate the effects of incorrect gaze estimates on the shape index of slanted objects. We assumed that the principal curvatures of the unslanted object lie in the $x-y$ and $y-z$ plane and that the direction of the applied slant is also in one of these planes. By taking into account the symmetry of these situations it is evident that the principal curvatures of the original object are mapped to those of the distorted object. Hence, the shape index of the distorted figure can be calculated from the curvature of the two transformed principal curvatures.

A general three-dimensional curve is obtained by parametrizing the coordinates of equation (A1) with one parameter. Slant and curvature of this curve can be calculated by taking derivatives with respect to this parameter. The resulting expressions are very complicated and many pages long; we used a symbolic algebra package to perform this laborious task. For the special cases such as those mentioned above, namely planar curves in the straightforward direction, the complexity reduces somewhat, but the expressions are still quite complicated. Therefore we will show the results only graphically (Fig. A2).

In all the graphs, the correct distance is 52.5 cm, the same as in our set-up. The first graph [Fig. A2(a)] shows how a mistaken vergence influences perceived slant; the tilt direction is horizontal. The results for a vertical tilt are almost the same. There are only small differences [Fig. A2(b)]. However, these are too small—generally lower than 1/10th of a degree—to explain the anisotropies that are known to exist in the perception of horizontal and vertical disparity gradients (Cagenello & Rogers, 1993; Gillam & Ryan, 1992; Mitchison & McKee, 1990).

Figure A2(c) shows the perceived curvature of a horizontal planar curve with a set of horizontal slants, ranging from 0 to 30 deg in steps of 6 deg. The curvature is distorted quite considerably, especially for high slant values.

In Fig. A2(d) we have a vertical curve with slants in the vertical direction. For a 0 deg slant we have the same situation as that of Rogers and Cagenello (1989): a horizontally oriented cylinder (which is curved in the vertical direction). Indeed, as these authors point out, in this case the curvature is almost an invariant. Note that for vertically oriented cylinders and small viewing distances this invariance no longer holds [see Fig. A2(c)]. Rogers and Cagenello had their subjects compare cylinders at 57 and 114 cm, distances which are firmly in the “invariant” range. Results for shorter distances would probably be different.

Rogers and Cagenello also report an anisotropy in the thresholds for discriminating the direction of curvature of vertically and horizontally oriented cylinders. Thresholds were found to differ by a factor of at least 2:1. If we examine perceived curvature of vertical and horizontal curves with zero curvature (these are straight lines of course; they separate positive and negative curvature) then it becomes clear that misestimation of vergence results in errors of different magnitudes for the two orientations [see Fig. A2(g)]. These magnitudes also differ by about a factor of 2:1. Of course, this should only result in a bias. However, if Rogers and Cagenello measured their thresholds by approaching zero curvature from only one side, this bias could contribute to the thresholds.

Figure A2(e, f) presents perceived curvature in the case that the slant direction is in the plane perpendicular to the plane in which the curve lies. A combination of Fig. A2(c) and (e) must be used to calculate the perceived shape index of a horizontally slanted object and Fig. A2(d) and (f) for a vertically slanted object.

The consequences of a misestimation of vergence on the shape index of a horizontally slanted object (the first combination mentioned above) can be seen in Fig. A2(h). Clearly, at $S = \pm 1$ and 0, deviations are largest. While examining these graphs, note that according to Johnston (1991) the correct distance (52.5 cm) is overestimated by about 10–20 cm.

FIGURE A2. (facing page).

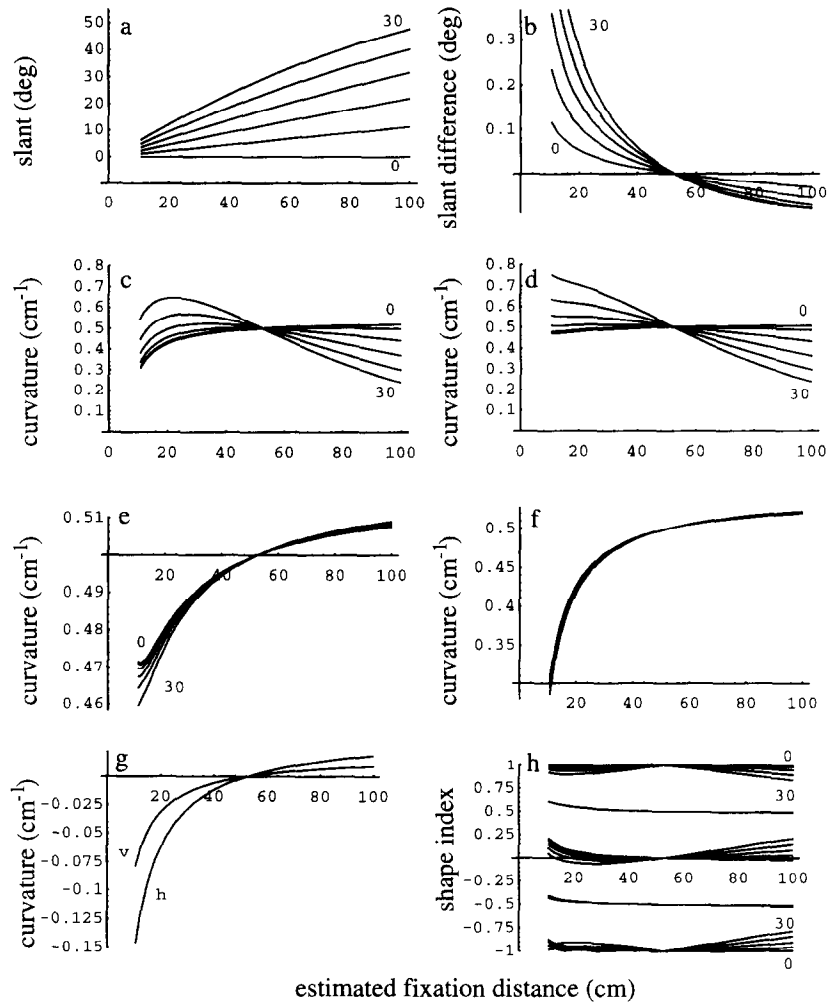


FIGURE A2. Influence of a faulty distance estimation. On the x-axis estimated distance is indicated (where 52.5 cm is veridical). Parameter in all graphs is the slant, ranging from 0 to 30 deg in 6 deg steps. (a) Perceived slant (horizontal). (b) Difference in perceived slant (vertical minus horizontal). (c) Perceived curvature of a horizontal curve with horizontal slant. (d) Perceived curvature of a vertical curve with vertical slant. (e) Perceived curvature of a vertical curve with horizontal slant. (f) Perceived curvature of a horizontal curve with vertical slant. (g) Perceived curvature of a straight line (horizontal and vertical) without slant; the larger deviations are found for the horizontal line. (h) Perceived shape index, for five shape index values (-1 , $-\frac{1}{2}$, 0 , $\frac{1}{2}$ and 1); tilt direction is horizontal. Generally, the zero slant condition has the lowest deviations from the veridical values.