Perception & Psychophysics
1993, 54 (2), 145-156

Perception of local shape from shading

RODERIK G. F. ERENS, ASTRID M. L. KAPPERS, and JAN J. KOENDERINK
Utrecht Biophysics Research Institute, Utrecht, The Netherlands

Theoretically, metric solid shape is not determined uniquely by shading. Consequently, human vision has difficulty in categorizing shape when shading is the only cue. In the present research, subjects were required to categorize shaded quadric surfaces. We found that they were rather poor at this task; they confused hyperbolic and elliptic (both convex and concave) shapes easily. When a cast shadow visually indicated the direction of the illuminant, they were able to notice the concavity or convexity of elliptic shapes. However, they still confused elliptic and hyperbolic ones. Finally, when an animated sequence of eight intensity patterns belonging to one quadric shape had been displayed, the subjects were able to categorize the quadrics. However, the results are still quite moderate. Our experiments indicate that local shading structure is only a weak shape cue when presented in the absence of other visual cues.

With the present generation of rendering algorithms, it is apparently relatively easy to generate very "natural" images on a computer screen. The three-dimensional impression created by these images is very compelling. However, such an informal observation has to be quantified. Very little research has been done on the perception of solid shape on the basis of shading.

Computer images provide several visual cues that potentially contribute to the perception of solid shape (Gibson, 1950). The role of shading in these images is complicated because extracting shape from shading is not a trivial task. The shading is determined by the surface orientation, the illuminant direction, and the surface properties. This means that the same luminance distribution can be generated by several different surfaces. For instance, the luminance distribution of concave and convex spheres is exactly the same. Thus, in order to find the local surface structure on the basis of shading, one has to make assumptions about the illuminant direction and the surface properties (Horn, 1975, 1977; Pentland, 1984, 1989). Another method to estimate the local surface structure from shading is to combine the information given by several intensity patterns of one shape (Woodham, 1980). Most computational theories on shape from shading calculate the surface shape locally (Horn, 1975, 1977; Pentland, 1984, 1989; Woodham, 1980). It is not clear whether human observers are able to judge the local shape of an object on the basis of shading.

Several psychophysical experiments report observers' ability to estimate local surface structure from shading (Bülthoff & Mallot, 1988; Mingolla & Todd, 1986; Todd & Mingolla, 1983; Todd & Reichel, 1991). These experiments indicated that human observers are very poor at this. Other experiments have shown that observers cannot distinguish local shapes such as concave and convex spheres when they do not know the illuminant direction (Berbaum, Bever, & Sup Chung, 1984; Ramachandran, 1988). In most of these studies, a very restricted set of shapes such as spheres (Berbaum et al., 1984), cylinders (Todd & Mingolla, 1983), or ellipsoids (Bülthoff & Mallot, 1988; Mingolla & Todd, 1986) have been used. Here, we investigate whether human observers are able to categorize a wide range of local shapes (elliptic, hyperbolic, and cylindric) on the basis of shading and the possible effect of information about the illuminant direction on the responses. Furthermore, we investigate whether human observers can estimate local shape from a sequence of intensity patterns belonging to the same shape.

This research was supported by the SPIN project "3D-Computer Vision" of the Dutch Ministry of Economic Affairs. Address correspondence to R. Erens, University of Utrecht, Princetonplein 5, Utrecht 3584 CC, The Netherlands.

Describing Local Shape

Any solid shape can be approximated on a point-by-point basis by a collection of local surface patches. The shape of such a local surface patch can be described by a Taylor expansion:

$$F(x,y) = A_1 + (B_1x + B_2y) + \tfrac{1}{2}(C_1x^2 + C_2xy + C_3y^2) + \text{higher order terms.}$$

The zeroth position and first-order attitude do not contribute to the local surface shape, because they are dependent on the observer-object geometry only. They can be made to vanish by taking the point of interest at the origin and the z direction parallel to the surface normal. When the surface patch is taken small enough, the terms higher than second order may be neglected. The only relevant term to describe local shape in the Taylor expansion is the second-order (quadric) term. Thus, every solid shape can be approximated locally by a collection of quadric surface patches. The expression of a quadric surface is:

$$F(u,v) = \tfrac{1}{2}(\varkappa_1u^2 + \varkappa_2v^2), \tag{1}$$

where $\varkappa_1$ and $\varkappa_2$ are the so-called principal curvatures, which are the minimum and maximum curvatures of the surface patch, respectively (Aleksandrov, Kolmogorov, & Lavrent'ev, 1963; Spivak, 1975). The directions of the principal curvatures are called the *principal directions* and are denoted by u, v in Equation 1. The principal directions are always perpendicular to each other.

For visual psychophysics of shape, one needs a convenient description of local shape that complies with our intuitive notion of *shape*. For example, spheres of different sizes look the same and thus should be indicated by the same "shape measure." The principal curvatures give complete control of the local surface geometry, but are perceptually nonintuitive, because they will be different for shapes that "look alike." It is obvious that one quantity should define the actual *shape* scale independently and another one should define the *scale* of the shape. Koenderink (1990) proposed such quantities (see also Koenderink & van Doorn, 1992), the shape index (S) and the curvedness (C).

The shape index describes the local surface geometry scale independently and is given by:

$$S = -\frac{2}{\pi} \arctan \frac{\varkappa_1 + \varkappa_2}{\varkappa_1 - \varkappa_2} \qquad (\varkappa_1 \geq \varkappa_2). \qquad (2)$$

The shape index gives a one-dimensional continuous scale on which local shape can be represented (see Figure 1). On the shape index scale, the shapes are classified roughly into three classes—the convex and concave elliptic shapes and hyperbolic shapes. These classes are separated by the cylinders (Figure 1). The shape index is between -1 and -0.5 for concave elliptic surfaces and between 0.5 and 1 for the convex ones. Hyperbolic surfaces (saddles) have a shape index between -0.5 and 0.5. The cylinders ($S = \pm 0.5$) will be on the boundaries between elliptic and hyperbolic shapes. Whenever a cylinder undergoes a deformation, however slight, it will immediately become elliptic or hyperbolic.

The amount of curvature of the local surface shape is given by the curvedness:

$$C = \sqrt{\frac{(\varkappa_1^2 + \varkappa_2^2)}{2}} \qquad (m^{-1}). \qquad (3)$$

The curvedness varies from zero for a flat surface up to infinity for an extremely curved surface.

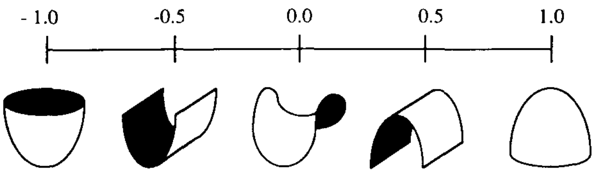


Figure 1. Local shapes presented on the shape index scale. The shape index is between -1 and -0.5 for concave elliptic surfaces, between 0.5 and 1 for convex, and between -0.5 and 0.5 for hyperbolic shapes. The cylinders will be on the boundaries between elliptic and hyperbolic shapes and the shape index is -0.5 or 0.5.

In our study, the local shapes are defined in terms of the shape index and curvedness and the responses of the subjects are categorized on the shape index scale. Similar types of experiments on human visual perception of shape (shape from stereo: de Vries, Kappers, & Koenderink, 1993; shape from motion: van Damme & van de Grind, 1991) have shown that the shape index and curvedness are very suitable for psychophysical research on shape perception.

GENERAL METHOD

Subjects

Four subjects (3 male and 1 female) participated in the experiment. All had normal or corrected-to-normal vision and viewed the screen with their dominant eye. All the subjects knew the purpose of the investigation.

Apparatus

The shaded images were generated on an Apollo DN590 computer using a graphical rendering package. The images were displayed in 8-bit gray tone on a high-resolution monitor screen, 39.0 × 29.0 cm (22.1° × 16.5°), with 1,280 × 1,024 pixels. The mean luminance of the screen was 26 cd/m². The subjects viewed the monitor screen monocularly with a natural pupil at a distance of 100 cm. The experiments were done in a totally dark room. A chinrest was used to restrict head movements.

Stimulus

The stimuli were images of lambertian shaded quadric shapes, which is defined in Equation 4. The principal curvatures $\varkappa_1$ and $\varkappa_2$ can be calculated from the shape index and curvedness by Equations 2 and 3. Here we stress again that the shape index and curvedness are purely local shape measures; usually, they will change over the surface of an arbitrary object. This holds also for quadric shapes. We define the quadric shape by its shape index and curvedness at the origin. Farther away from the center, both quantities will be slightly different (one can calculate from Equation 4 that at the stimulus boundary a quadric shape will be more cylinder-like and less curved).

The normal distance from the tangent plane in the origin of the quadric shape is given in Cartesian coordinates:

$$z = \frac{1}{2}(\varkappa_1 x^2 + \varkappa_2 y^2). \qquad (4)$$

A set of 501 quadric shapes (generated according to Equation 4) with varying shape indexes (from -1.000, -0.996, ..., 0.996, 1.000) was used in the study. For all shapes, the curvedness was 30 m^{-1}. The positive z-axis pointed in the direction of the viewer; thus, the view of the quadric shape was frontoparallel. The images were lambertian shaded and projected perspectively for a viewing distance of 100 cm. The images had a diameter of 7.5 cm (4.3°).

The perspective projection gives a specific outline (this is not an occluding contour!) to the quadric shape (which is defined on a square Cartesian coordinate patch), by which the displayed shape can be easily recognized. Therefore, a randomly frayed gray mask was superimposed on the stimulus to cover the outline (see Figures 2 and 3). The mask was identical during each series of measurements and had a mean luminance of 26 cd/m². Another cue that subjects may use to recognize the quadrics is the fixed directions of the principal curvatures, which are aligned with the x-y frame (see Figures 2 and 3). Thus, the quadrics would be oriented horizontally. To prevent the subjects from using this information, the quadric shape was rotated around the z-axis over a random angle for each trial.

Procedure

Images from the set of 501 quadrics were selected randomly and displayed on the monitor screen and the subjects were required to categorize the quadric shapes. The subjects' responses are given on the shape index scale, which, for simplicity, is divided into eight equal parts. This gives eight "shape" categories (Table 1).

The stimulus and the subject's response are expressed in terms of one of these categories on the shape index scale. All the subjects were thoroughly familiar with the scale of the shape index and curvedness. Three-dimensional sculpted examples of quadric surfaces were used to verify whether the subjects were able to handle the eight categories on the shape index scale. Before the measurement started, several training sessions were held in which the subjects received feedback after each trial about the displayed and responded category.

The stimuli were displayed during a 4-sec interval, then eight menu buttons appeared on the screen representing the shape categories. The subjects could choose a category by moving the cursor with the computer mouse to the appropriate menu button and confirmed their choice by pressing the mouse button. The response time was not limited. During the measurements, no feedback concerning the validity of the answers was given but after finishing a session, the subjects were allowed to review their results. For each experiment, three series of 200 measurements were performed—each series on a different day.

a

b

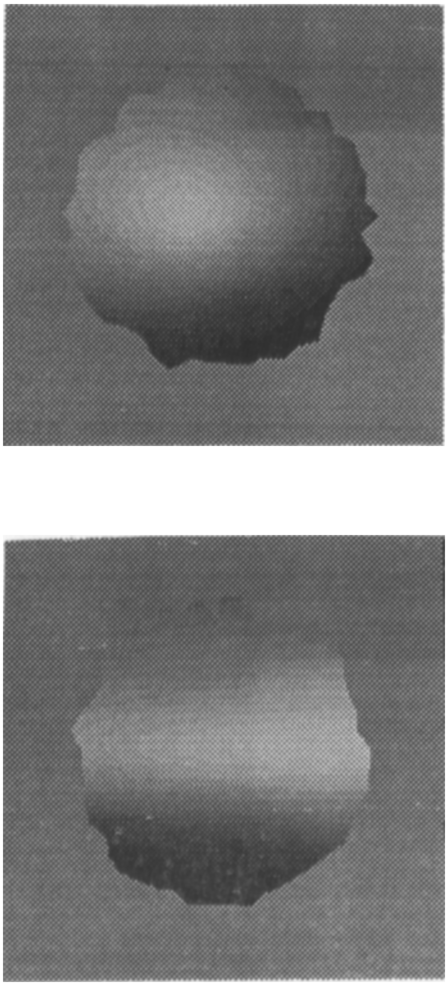


Figure 2. Illustration of the stimuli used in Experiment 1. The quadric shapes are diffusely shaded and masked with a random frayed aperture. The quadric shape (a) is convex elliptic, with shape index 0.9 and curvedness 30 m^{-1}. This shape belongs to the "symmetric" group. An example of an "elongated" shape (b) shows a cylinder-like saddle with shape index 0.4. The illuminant direction in both images is from the upper left corner and both shapes are orientated horizontally (the largest principal curvature is in the direction of the positive x-axis). The figure clearly shows that elongated and symmetric shapes are distinguishable. In the experiment, the orientation of the quadrics was randomized.

a

b

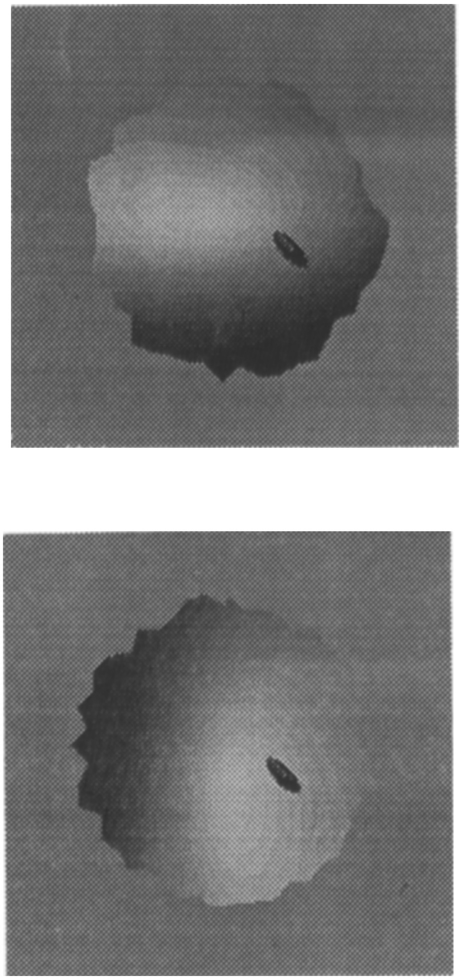


Figure 3. Illustration of the stimuli used in Experiment 2. A cast shadow visually indicates the azimuth of illuminant direction. A convex (a) elliptic (shape index 0.8) and concave (b) elliptic surface (shape index −0.8) are shown. The curvedness of both quadrics is 30 m^{-1}. The illuminant direction is from the upper left corner (as is indicated by the cast shadow) and the shapes are orientated horizontally (the largest principal curvature is in the direction of the positive x-axis). Here we can see that, because of our definition that $\kappa_1 \geq \kappa_2$, the convex "cylinder-like" shapes would be oriented horizontally, whereas the concave ones would be oriented vertically. To prevent the subjects from using this information, the quadric shape was rotated around the z-axis over a random angle for each trial (see text).

Table 1
The Ranges of the Eight Shape Categories

Category Type	Shape Index Range	Mnemonic
1	(−1.00...−0.75)	Concave spherical
2	(−0.75...−0.50)	Concave
3	(−0.50...−0.25)	Saddle rut
4	(−0.25... 0.00)	Shallow saddle rut
5	(0.00... 0.25)	Shallow saddle ridge
6	(0.25... 0.50)	Saddle ridge
7	(0.50... 0.75)	Convex
8	(0.75... 1.00)	Convex spherical

EXPERIMENT 1
Local Shape From a Single Image Without Information About the Illuminant Direction

Here, we investigated the ability of observers to categorize quadric shapes on the shape index scale when no information about the illuminant direction was present in the image.

Predictions

When the illuminant direction is unknown, several quadric shapes will be represented by the same luminance distribution. First of all, we will predict analytically which quadric shapes are indistinguishable on the basis of their luminance distribution. The luminance $I(x,y)$ at position x,y of a diffusely shaded image is given by:

$$I(x,y) = \rho\lambda(\mathbf{N}\cdot\mathbf{L}), \tag{5}$$

where ρ is the surface albedo, and λ is the intensity of the light source. $\mathbf{N}$ is the surface normal and $\mathbf{L}$ is the illuminant direction pointing toward the light source. The surface normal is calculated from Equation 4 by:

$$\mathbf{N} = \frac{(-z_x, -z_y, 1)}{\sqrt{z_x^2+z_y^2+1}},$$

where z_x and z_y denote the first derivative of z to x and y, respectively. Then, for a quadric surface, the normal at point x,y is:

$$\mathbf{N} = \frac{(-\varkappa_1 x, -\varkappa_2 y, 1)}{\sqrt{\varkappa_1^2 x^2+\varkappa_2^2 y^2+1}}. \tag{6}$$

The illuminant direction is given by:

$$\mathbf{L} = (\cos\phi\ \sin\theta,\ \sin\phi\ \sin\theta,\ \cos\theta), \tag{7}$$

where ϕ is the azimuth of the illuminant direction, which is the angle between the positive x-axis and the projection of the illuminant direction on the x-y plane. θ is the inclination of the illuminant direction, which is the angle between the positive z-axis and the illuminant direction. For a quadric surface, the luminance at position x,y is given by Equation 5:

$$I(x,y) = \rho\lambda\frac{-\varkappa_1 x\ \cos\phi\ \sin\theta - \varkappa_2 y\ \sin\phi\ \sin\theta + \cos\theta}{\sqrt{\varkappa_1^2 x^2+\varkappa_2^2 y^2+1}}. \tag{8}$$

This is the expression for the luminance distribution at the surface of the object. The luminance distribution on the computer display is the perspective projection of Equation 8. Due to perspective effects, the isophotes, that is, contours of equal luminance, for concave ellipsoids will be somewhat wider than for convex ones.

Consider another quadric surface $(\varkappa_1', \varkappa_2')$ and another illuminant direction (θ', ϕ'). The luminance at position x,y of the quadric surface is then given by:

$$I'(x,y) = \rho\lambda\frac{-\varkappa_1' x\ \cos\phi'\sin\theta' - \varkappa_2' y\ \sin\phi'\sin\theta' + \cos\theta'}{\sqrt{(\varkappa_1')^2 x^2+(\varkappa_2')^2 y^2+1}}. \tag{9}$$

The different quadric surfaces will yield approximately the same stimulus if $I = I'$ (perspective effects neglected). For this condition, we obtain

$$\varkappa_1' = \varkappa_1\frac{\cos\phi\ \sin\theta\ \cos\theta'}{\cos\phi'\ \sin\theta'\ \cos\theta}$$

and

$$\varkappa_2' = \varkappa_2\frac{\sin\phi\ \sin\theta\ \cos\theta'}{\sin\phi'\ \sin\theta'\ \cos\theta}.$$

In our setup, the inclination of the illuminant direction was always 45° during the experiment. Therefore, both inclinations θ and θ' of the illuminant directions are equal. We find that identical luminance distributions may yield different quadric shapes:

$$\varkappa_1' = \varkappa_1\frac{\cos\phi}{\cos\phi'} \qquad \varkappa_2' = \varkappa_2\frac{\sin\phi}{\sin\phi'}. \tag{10}$$

Equation 10 clearly shows that quadric shapes are indistinguishable on the basis of their luminance distributions when the azimuth of the illuminant direction is unknown. Observers can easily confuse convex and concave if the azimuth of the illuminant direction is reversed over 180°; this is the well-known crater illusion. By reversing the azimuth over 90° or 270°, elliptic and hyperbolic shapes also are indistinguishable. Thus, local shape reconstruction from a single image in the absence of information about the azimuth of the illuminant direction leads (at least) up to a fourfold ambiguity. This is also reported by Blake and Brelstaff (1988). In terms of the shape index scale, this means that Categories 1, 4, 5, and 8 cannot be distinguished from each other nor can Categories 2, 3, 6, and 7 (see Table 1 for the meaning of these categories). However, the two groups can be distinguished. The first group of categories represents the so-called "symmetric" shapes, which are the symmetric, concave and convex elliptic, and hyperbolic shapes. The second group represents the so-called "elongated" shapes, which are the elongated concave and convex cylinder-like shapes. Examples of these shapes are shown in Figure 2.

In this discussion, we have neglected the perspective effects. The perspective projection slightly alters the actual stimulus from the luminance distribution, as given by Equation 8. There is a small difference (less than 5%)

between the intensity distributions of concave, convex elliptic, and hyperbolic shapes. The perspective cue might help subjects to allot ambiguous shapes to the right category.

Method

Procedure. The subjects were required to categorize the shaded quadric shapes on the shape index scale. In order to prevent the subjects from using any information they may have acquired regarding the azimuth of the illuminant direction, the latter was chosen randomly for each trial from four possibilities: 45°, 135°, 225°, or 315°. The inclination of the illuminant direction was always 45°. Examples of the stimuli are given in Figure 2. In this first experiment, the subjects could only categorize the quadric shape on the basis of the shading cue and the perspective cue.

EXPERIMENT 2
Local Shape From a Single Image With Information About the Illuminant Direction

Information about illuminant direction is needed in order to remove some of the inherent ambiguity. In most pictures and real scenes, there may be numerous cues that indicate this direction, such as interreflections between objects, shadows, or the luminance profile near occluding contours. In Experiment 2, we placed a small stick on a surface, which cast a shadow on that surface (see Figure 3). The azimuth (ϕ) of the illuminant direction in Equation 8 is visually indicated by the shadow. The influence that a cast shadow has on observers' responses to convex and concave ellipsoids has often been investigated (Berbaum et al., 1984; Ramachandran, 1988). However, the influence of a cast shadow on the discrimination between elliptic and hyperbolic shapes is less clear and has never been investigated.

Predictions

The cast shadow indicates the azimuth of the illuminant direction, so this information is available in the image. From Equation 10, one could conclude that quadric shapes will be distinguishable. However, this is not true; even when the illuminant direction is known, some ambiguities in shape estimation remain. This can best be shown by a series of graphs of the luminance distribution of the stimuli. In Figures 4 and 5, the luminance distributions of a shaded elliptic (shape index 0.8) and a hyperbolic surface (shape index -0.2) are plotted according to Equation 8 (the κ_1, κ_2 in Equation 8 are calculated from the shape index and curvedness by Equations 2 and 3). In each graph, the azimuth of the illuminant direction is known and identical (from the upper left corner), but the quadric surface is orientated differently. The orientation of the quadric surface differs 45° for each graph.

First, we will discuss the luminance distribution on an elliptical surface ($S = 0.8$) for different orientations of the surface and fixed illuminant direction. Figure 3 shows that the luminance distribution is more or less identical for different orientations of an elliptic patch. On a *convex* elliptic patch, as is shown in Figure 4, a luminance gradient will always point *away* from the illuminant direction, but for a *concave* one, the luminance gradient will always point *toward* the illuminant direction. Thus, a cast shadow will directly solve the concave-convex ambiguity in elliptic patches if one knows that they are elliptic.

On a hyperbolic surface, the luminance distribution differs enormously for various orientations of the patch, as can be seen in Figure 5. The luminance gradient may have any possible direction with respect to the illuminant direction. Therefore, luminance patterns like those on elliptic patches can be obtained just as easily on hyperbolic patches. Thus, even with information about the illuminant direction in the image, elliptic and hyperbolic shapes cannot be distinguished.

In terms of the shape index scale, this means that subjects can be expected to respond in all eight categories when a hyperbolic surface is shown. Also, an elliptic shape can be interpreted as a hyperbolic one. However, it is not possible to interpret the convex elliptic shapes as being concave. On the shape index scale, this means that concave elliptic shapes can end up in Categories 1-4 and the convex elliptic ones can end up in Categories 5-8 (see Table 1 for descriptions of these categories).

Method

Procedure. The subjects were required to allot quadric shapes to one of the eight response categories. In the images, a cast shadow provided information about the azimuth of the illuminant direction. On each trial, the illuminant direction was chosen randomly from the same range of possibilities as in Experiment 1. Examples of a concave and convex elliptical shape are given in Figure 3. Again, due to the perspective projection, the stimulus contained a perspective cue that might help the subjects to allot ambiguous shapes to the right category.

EXPERIMENT 3
Local Shape From an Animated Sequence of Intensity Patterns

By varying the illuminant direction, one can obtain several different luminance distributions belonging to the same surface. From these luminance distributions, it is theoretically possible to estimate the local surface structure uniquely (Woodham, 1980).

Predictions

Woodham (1980) showed theoretically that the local surface structure can be reconstructed uniquely from at least three different luminance distributions. The luminance distributions that belong to the same shape will be obtained by varying the illuminant direction. Actually, one simply gets three expressions for the luminance value at positions x,y (see Equation 8). From these three expressions, one can calculate relatively easily the principal curvatures. Thereafter, by Equation 2, the shape index at each position (x,y) can be found.

In terms of the shape index scale, this means that every quadric surface can be allotted to the right shape category.

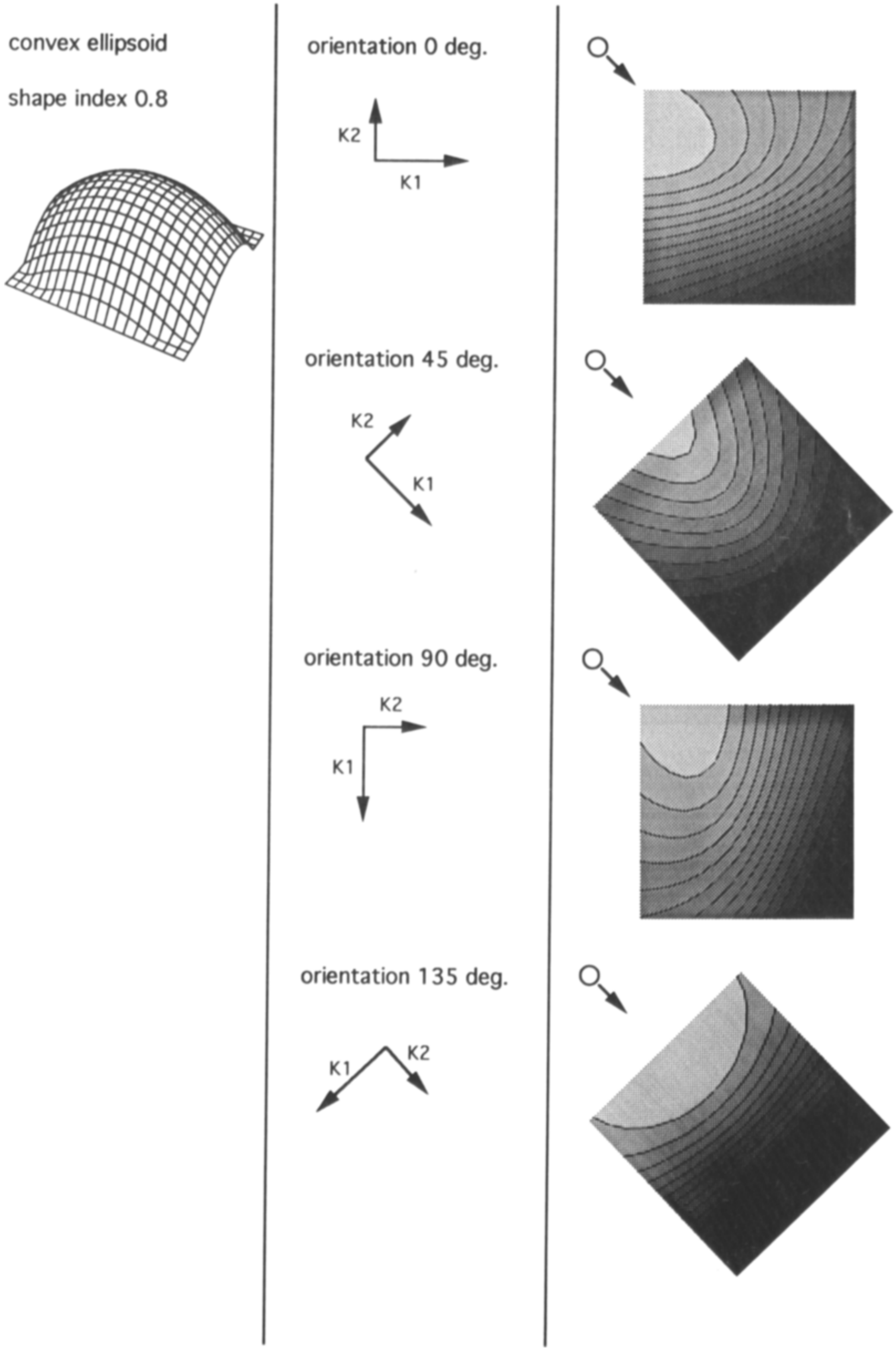


Figure 4. The variation of the luminance distribution on a convex elliptic patch due to different orientations of the surface. The isophotes are drawn for an elliptic shape with shape index 0.8, and the illuminant direction is from the upper left corner. The series of figures in the far right column show several orientations of the surface (shown in the middle column). K1 and K2 denote the maximum and minimum principal curvatures, respectively. The luminance distribution does not change dramatically for different orientations of the convex elliptic shape. The maximum intensity value is always on the same side as the light source and the luminance gradient is always more or less in the same direction, pointing away from the light source.

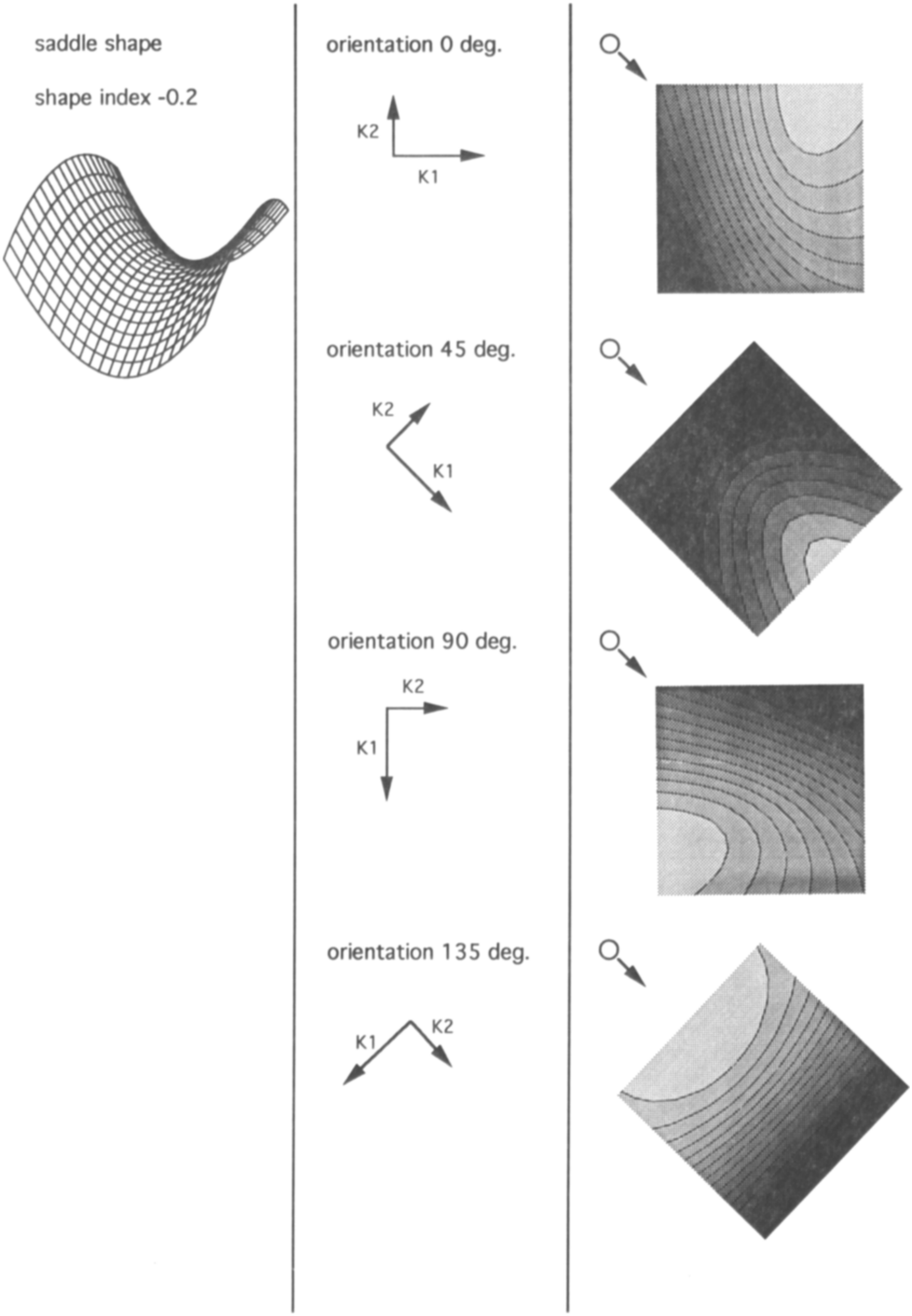


Figure 5. The variation of the luminance distribution on a concave hyperbolic patch, due to different orientations of the surface. The isophotes are drawn for a hyperbolic shape with shape index -0.2, and the illuminant direction is from the upper left corner. The series of figures in the far right column show several orientations of the surface (shown in the middle column). K1 and K2 denote the maximum and minimum principal curvatures, respectively. The luminance changes dramatically for different orientations of the hyperbolic shape. The maximum intensity can be anywhere on the surface and the direction of the luminance gradient also varies for different orientations of the hyperbolic surface.

Method

Procedure. In this experiment, the subjects were shown an animated sequence of eight different luminance distributions belonging to a certain quadric shape and were required to categorize the quadric into one of the eight "shape" categories. The orientation of the quadric surface was fixed and the azimuth of the illuminant direction changed over the surface from 45° to 135° in eight steps. The different intensity patterns were presented for 0.5 sec, so the total exposure time of the quadric shape was the same as in the previous experiments. The illuminant direction was again indicated by the shadow cast by a small stick (as in Experiment 2). So, by comparing the different shading patterns, it would be theoretically possible to unambiguously categorize the shape. The question was whether the human observers could do this.

RESULTS

For all the experiments, we show the results of Subject R.E., who is one of the authors, and Subject S.dV., who was not familiar with shape-from-shading experiments, but was experienced in using the shape index scale. The results of the other subjects are quite similar qualitatively, and are shown in Table 1.

The responses of Subjects R.E. and S.dV. for the three experiments are shown in Figures 6, 7, and 8, which show a 3-D plot of the stimulus–response matrix. The axis on the left gives the stimulus categories and the one on the right gives the response categories. The z-axis shows the responses as a percentage of the total number of stimuli in a particular category. We call this the score (%) for each stimulus–response combination. If the subjects categorized the quadric surfaces perfectly, the scores would be 100% and on a diagonal for which the stimulus and response categories were identical. The results for each series of measurements were reproduced within 10% accuracy.

Experiment 1

Figure 6 shows that shading information alone is not enough to categorize the quadric shape correctly. The average percentage of correct responses (scores on the diagonal) is 13.4% and 16.0% for Subjects R.E. and S.dV., respectively. This is not significantly above the 12.5% that would be obtained if the observers simply guessed the category. The perspective cue did not seem to help the subjects categorize the quadrics. They categorized most quadrics to the response Categories 4–8, which indicates that most of the shapes were interpreted as being convex. This effect would possibly have been more pronounced if the subjects had not known a priori that all stimulus categories were equally likely.

We predicted theoretically that "elongated" and "symmetric" shapes were distinguishable on the basis of their luminance distribution. The elongated and symmetric stimulus categories are presented in Figure 6 by the rows of blank and shaded score bars, respectively. Note that the scores of Subject R.E. for the elongated (blank score bars) shapes are primarily grouped around Category 6 (about 65% of the scores), whereas the symmetric (shaded score bars) shapes are found mainly in Categories 7 and 8 (about 53% of the scores). The distribution of the scores

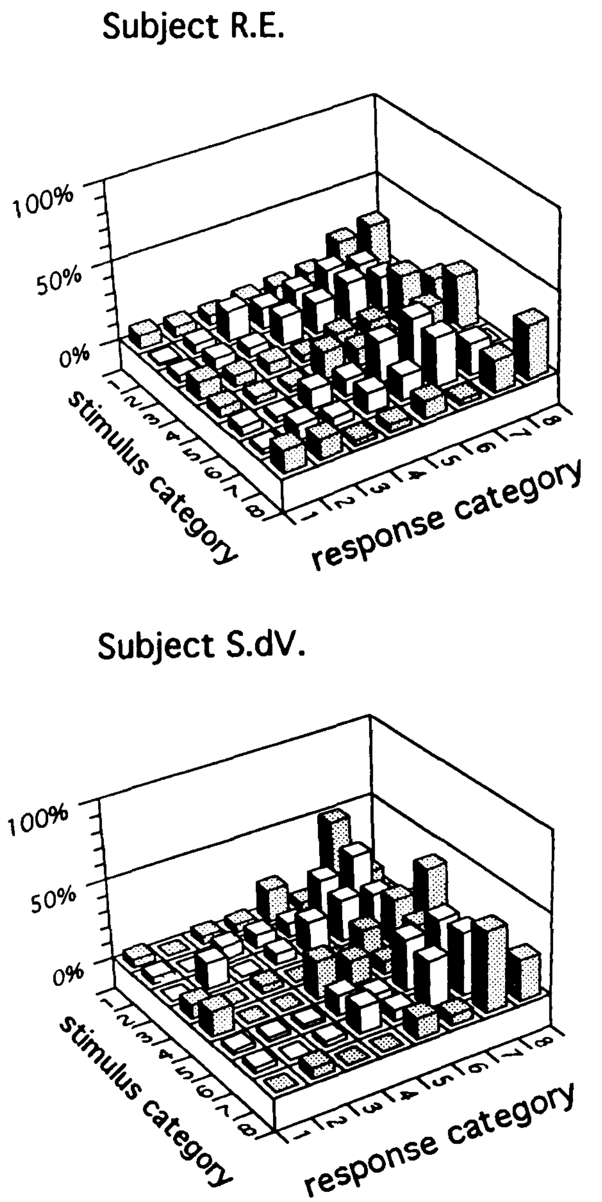


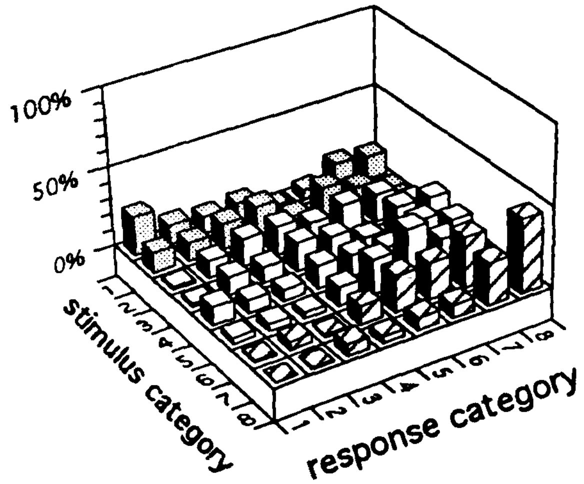
Figure 6. The responses of Subjects R.E. and S.dV. for Experiment 1 with no information about the illuminant direction available. The stimulus categories and response categories are given on the left and right axes, respectively. The z-axis shows the scores of the subject at each stimulus–response combination. The percentages of correct responses are given by the diagonal, for which the category number of the stimulus and response are identical. The shaded score bars indicate the results for the "symmetric" shapes and the blank score bars refer to the scores for the "elongated" shapes.

along the response categories is also quite specific for each group of elongated and symmetric shapes, as can be seen in the figure. Analysis of variance (ANOVA) shows that the effect for Subject R.E. is significant (p = .015). The results for Subject S.dV. are less clear, because he tended to allot most of the shapes to Category 7. However, his scores for the elongated shapes are also grouped around

Category 6 (about 73%) and can be found for the symmetric shapes in Categories 7 and 8 (about 60%). The distribution of the scores along the response categories is rather identical for the elongated shape group (blank score bars), but not for the symmetric shape group (p = .25). In conclusion, the results indicate that the subjects were able to distinguish between elongated and symmetric shapes purely on the basis of shading.

It might be that quadric shape is better categorized for certain values of the azimuth of the illuminant direction.

Subject R.E.



Subject S.dV.

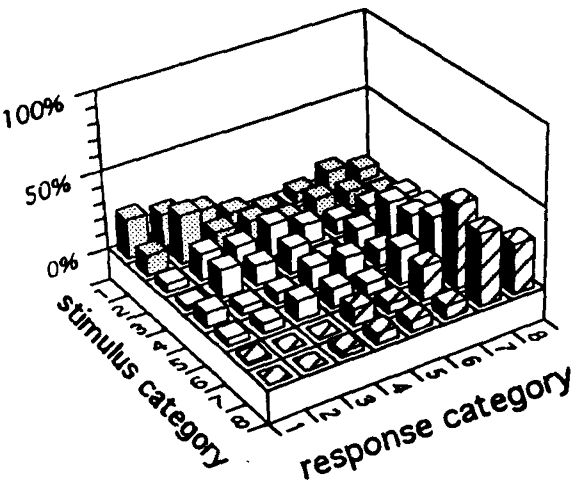
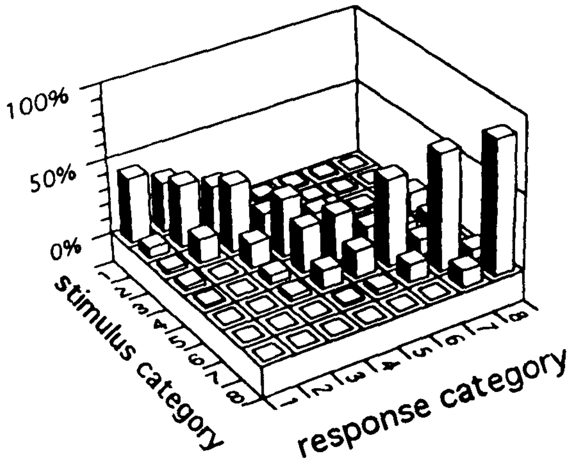


Figure 7. The responses of Subjects R.E. and S.dV. for Experiment 2, in which a cast shadow indicated the illuminant direction. The 3-D bar plot presents the scores; correct responses are given by the diagonal for which the category number of the stimulus and response are identical. The shaded score bars indicate the results for the concave elliptic shapes, the blank score bars give the scores for the hyperbolic shapes, and the hatched score bars refer to the scores for the convex elliptic shapes.

Subject R.E.



Subject S.dV.

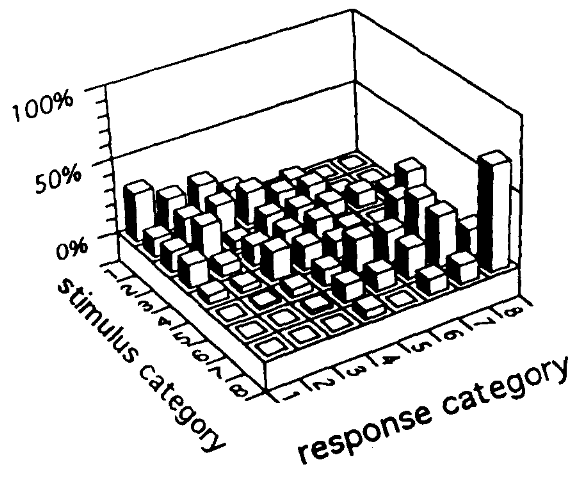


Figure 8. The responses of Subjects R.E. and S.dV. for Experiment 3, in which an animated sequence of eight luminance patterns of one shape was displayed. The scores of the subject at each stimulus–response combination are given on the z-axis. The percentages of correct responses are given by the diagonal, for which the category number of the stimulus and response are identical.

For Subject R.E., we find that the average correct scores (stimulus category = response category) depends significantly on the azimuth of the illuminant direction (45° = 19.9%, 135° = 21.5%, 225° = 6.3%, and 315° = 5.7%). He correctly categorized more shapes when the illuminant shone from above on the shape. For Subject S.dV., we do not observe the influence of the azimuth of the illuminant direction on the percentage of correct responses (45° = 12.7%, 135° = 17.0%, 225° = 17.0%, and 315° = 17.7%).

Experiment 2

The scores for both subjects are presented in Figure 7: the two rows of shaded score bars represent the responses to concave elliptic stimuli, the four rows of blank score bars show the responses to hyperbolic stimuli, and the two rows of hatched score bars show the responses to convex elliptic stimuli. First, we concentrate on the results for the concave and convex elliptic shapes (the shaded and hatched score bars, respectively). The scores show that when information about the illuminant direction was present in the image, the subjects categorized concave elliptic shapes to the concave side of the shape index scale (Categories 1-4; $p = .012$ for R.E. and $p = .0001$ for S.dV.) and the convex elliptic shapes were reported as convex (Categories 5-8; $p = .0006$ for R.E. and $p = .022$ for S.dV.). However, the subjects still seemed to be unable to distinguish between elliptic and hyperbolic shapes. The percentage of correct scores for the concave elliptic shapes are significantly above chance level only for Subject S.dV. ($p = .11$ for R.E. and $p = .0003$ for S.dV.). The scores for the convex elliptic are much higher ($p = .003$ for R.E. and $p = 0.024$ for S.dV.). In Table 2, the percentages of correct responses are shown for all the subjects, grouped into concave elliptic, hyperbolic, and convex elliptic shape categories. Here we can see that only the scores for the convex elliptic shapes are significantly above chance level.

For the hyperbolic stimuli, we found that all the subjects chose any shape category when a "symmetric saddle" (Categories 4 and 5) was presented. The elongated saddles (Categories 3 and 6) are more or less grouped around the correct category. However, for these saddles, we also find responses in the other shape categories (see Figure 7). An ANOVA indicated no significant increase in the scores ($p = .42$ for R.E. and $p = .12$ for S.dV.). For the hyperbolic shapes, the percentage of correct responses was, for all the subjects, not significantly above chance (see Table 2).

When comparing the results with the experimental predictions, we indeed find that the subjects reported concave elliptic shapes as concave, and convex elliptic shapes as convex. Hyperbolic shapes evoked responses on the entire shape index scale, especially the symmetric saddles. These findings imply that, when information about illuminant direction is present in an image, human observers can distinguish between concave and convex elliptic shapes, but still will confuse hyperbolic shapes with elliptic.

Table 2
Percentages of Correct Responses in Experiments 2 and 3

Shape	S.dV	A.K.	W.vD.	R.E.
	Experiment 2			
Concave elliptic	28	20	32	18
Hyperbolic	16	21	17	21
Convex elliptic	42	28	52	39
	Experiment 3			
Concave elliptic	26	40	39	42
Hyperbolic	14	26	38	43
Convex elliptic	53	64	61	83

Experiment 3

Figure 8 shows that the categorization of quadrics is improved when a sequence of eight intensity patterns of the same shape is displayed. The scores of Subject R.E. especially show a major improvement; his results are always grouped around the correct responses on the diagonal. The scores of Subject S.dV. are less drastically increased in this experiment. The results for the convex elliptic shapes were slightly better than in Experiment 2, whereas the correct scores for hyperbolic and concave elliptic shapes remained identical (see Table 2).

In this experiment, the subjects reported that the sequence of intensity patterns was sometimes confusing. The first image gave some impression of the shape, which could be confirmed or not by the rest of the sequence. For elliptic patches, the following patterns mainly confirmed the first impression of the shape, but for hyperbolic shapes the subjects had the impression that the sequence was internally inconsistent. Table 2 shows that the convex elliptic shapes especially were categorized quite well. The performance of all the subjects on categorizing the hyperbolic and concave elliptic shapes was much poorer. Furthermore, the subjects reported that they obtained a response category by reasoning out the information given by the animated sequence. This probably explains the differences in performance between the subjects in Experiment 3. Subject R.E. shows the highest scores, but he was also the experimenter and was very familiar with the experimental conditions.

GENERAL DISCUSSION

First of all, we summarize the major experimental findings.

1. On the basis of a single shaded image and in the absence of information about illuminant direction, the categorization of local shape was rather poor. The only local shapes that the subjects seemed to be able to distinguish were elongated (cylinder-like) shapes from symmetric (sphere-like) shapes.

2. When the illuminant direction was made visually available in the image (e.g., by a cast shadow), the subjects could distinguish between concave and convex elliptic surfaces. However, elliptic and hyperbolic shapes were still confused.

3. From an animated sequence of intensity patterns of the same shape, the subjects could categorize the local shape quite well. However, they reported that they obtained their responses by reasoning out the information given by the sequence.

An arbitrary solid shape can be described locally by hyperbolic and elliptic (convex and concave) patches. For some reason, in most shape-from-shading experiments only the elliptic shapes are actually used. Our research shows that interesting observations can be made when shape-from-shading experiments are extended to all possible local shapes, represented by quadric surfaces.

We found that observers are quite poor at categorizing quadric shapes on the basis of shading alone. Other psychophysical experiments have shown that observers are unable to distinguish between concave and convex elliptic shapes (Berbaum et al., 1984; Ramachandran, 1988). In the present research, we have shown both theoretically and experimentally that only elongated from symmetric shapes could be distinguished. Hyperbolic and elliptic shapes were confused just as easily as were concave and convex ones (see also Blake & Brelstaff, 1988). The hyperbolic-elliptic ambiguity was even more pronounced than the concave-convex ambiguity. If the direction of the illuminant was visually indicated, it was directly possible to distinguish between concave and convex elliptic shapes, but it was still impossible to distinguish between elliptic and hyperbolic ones.

This finding implies that, even when the illuminant direction is known, the local surface structure cannot be estimated uniquely. Therefore, in computational models on shape from shading, as proposed by Horn (1975, 1977) and Pentland (1984, 1989), assumptions about the local surface shape have to be made, even when the illuminant direction is known. Generally, it is assumed in these models that the surface shape is locally convex elliptic. This assumption is true for objects with a convex smooth occluding contour; such objects are at least convex elliptic at the boundary (Koenderink, 1984).

Several notable psychophysical experiments on estimating local shape from shading are reported by Todd and Mingolla (Mingolla & Todd, 1986; Todd & Mingolla 1983; Todd & Reichel, 1991). In these experiments, subjects had to judge the surface orientation of patches on ellipsoids with different eccentricity in the viewing direction. The subjects were rather poor at estimating the local surface orientation. Despite the fact that our experiments were quite distinct, we also found that subjects were very poor at categorizing the local shape on the basis of shading. Only the scores for the animated sequence of images belonging to the same shape look promising. However, our subjects reported that they found the task rather superficial. In addition, the results are only moderate compared with the scores obtained in similar types of categorization experiments done on shape from stereo (de Vries et al., 1993) and shape from motion (van Damme & van de Grind, 1991).

Mingolla and Todd (1986) conclude that human perception of shape from shading is probably not a local process, but involves "global" structures. The results of our research support this point of view. A global description of shape has been proposed by Koenderink and van Doorn (1980, 1982). They have theoretically shown that in the illuminant distribution on a solid shape, several types of singularities (local maxima, minima, and saddle points) will appear (they call these the first, second, and third types of singularity). The first type of singularity appears in the illuminant distribution, in which the surface normal is aligned with the illuminant direction and shows an intensity maximum. The second type of singularity appears only on the boundary curves between hyperbolic and elliptic regions (the so-called parabolic curves). The third type of singularity does not occur on diffusely shaded objects. We know this third type as specular points; they appear on the surface where the surface normal coincides with the bisector of the visual direction and the illuminant direction. The topology of the field of isophotes is determined by the geometry of the solid shape and the illuminant direction. However, the illuminant direction is not necessary for many aspects of a global description of shape. In contradistinction to the local description, in which the surface orientation is calculated from the surface data, the global description is morphologically expressed in terms of the curvature landscape.

The local-global distinction is an important one. Our research clearly shows that subjects cannot distinguish local shapes in isolation. Our subjects also reported that the stimuli looked rather flat. Apparently, they could not judge the surface orientation or curvature correctly. This conforms with the findings of Mingolla and Todd (1986; Todd & Reichel, 1991).

Our research suggests that perception of solid shape from shading is not a local process. However, it is premature to conclude that our findings point to a model of global shape perception. In our experiments, quadric surfaces were used. These surfaces were purely hyperbolic or elliptic, and the luminance distribution did not contain any of the singularities and topological features by which the global shape of an object could be estimated.

Many psychophysical experiments will be needed to resolve the question of local versus global shape perception. In conclusion, the present research shows that human observers are rather poor at categorizing local quadric surfaces on the basis of shading. Even with explicit visual information about the illuminant direction, subjects cannot distinguish elliptic and hyperbolic shapes. It will be of interest to extend the shape-from-shading experiments to more complicated (cubic and higher order) surfaces. Perhaps it will be possible to find some psychophysical indication for global processes in the perception of shape from shading.

REFERENCES

Aleksandrov, A. D., Kolmogorov, A. D., & Lavrent'ev, M. A. (Eds.). (1963). *Mathematics, its content, methods and meaning*. Providence, RI: American Mathematical Society.

Berbaum, K., Bever, T., & Sup Chung, C. (1984). Extending the perception of shape from known to unknown shading. *Perception*, 13, 479-488.

Blake, A., & Brelstaff, G. J. (1988). Geometry from specularities. *Proceedings of the International Conference on Computer Vision* (pp. 394-403). Washington, D.C.: IEEE.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A*, **5**, 1749-1758.

Damme, W. van, Grind, W. A. van de (1991). Identification of 3D shapes using motion parallax. In P. J. Beek, R. J. Bootsma, & P. C. W. van Wieringen (Eds.), *Studies in perception and action* (pp. 72-75). 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.

Gibson, J. J. (1950). *The perception of the visual world*. Cambridge, MA: The Riverside Press.

Horn, B. K. P. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 115-155). New York: McGraw-Hill.

Horn, B. K. P. (1977). Understanding image intensities. *Artificial Intelligence*, **8**, 201-231.

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

Koenderink, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.

Koenderink, J. J., & Doorn, A. J. van (1980). Photometric invariants related to solid shape. *Optica Acta*, **27**, 981-996.

Koenderink, J. J., & Doorn, A. J. van (1982). Perception of solid shape and spatial lay-out through photometric invariants. In R. Trappl (Ed.), *Cybernetics and system research* (pp. 943-948). Amsterdam: North-Holland.

Koenderink, J. J., & Doorn, A. J. van (1992). Surface shape and curvature scales. *Image & Vision Computing*, **10**, 557-565.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics*, **53**, 137-151.

Pentland, A. P. (1982). Finding the illuminant direction. *Journal of the Optical Society of America A*, **72**, 448-455.

Pentland, A. P. (1984). Local shading analysis. *IEEE, Transactions on Analysis & Machine Intelligence*, **6**, 170-187.

Pentland, A. P. (1989). Shape information from shading: A theory about human perception. *Spatial Vision*, **4**, 165-182.

Ramachandran, V. S. (1988). Perceiving shape from shading. *Scientific American*, **259**, 58-65.

Spivak, M. (1975). *A comprehensive introduction to differential geometry: Vol. 3*. Berkeley, CA: Publish or Perish, Inc.

Todd, J. T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 583-595.

Todd, J. T., & Reichel, F. D. (1991). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review*, **96**, 643-657.

Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical Engineering*, **19**, 139-144.

(Manuscript received February 5, 1992;
revision accepted for publication December 10, 1992.)