Categorical Local Shape Perception

Pascal Mamassian, Daniel Kersten, and David C. Knill

Department of Psychology
University of Minnesota
Minneapolis, MN 55455

submitted January 30, 1995
revised July 6, 1995 [1]

[1]Correspondence should be sent to: Daniel Kersten, N218 Elliott Hall, 75 East River Rd, Minneapolis, MN 55455, USA. Email: kersten@eye.psych.umn.edu, pascal@cns.nyu.edu, knill@cattell.psych.upenn.edu. Manuscript submitted to *Perception*.

Abstract

How well do observers perceive the local shape of an object from its shaded image? We address this complex problem by first deriving a potential representation of local solid shape, and by presenting the results of a simple psychophysical experiment. Our descriptor of local solid shape, called shape characteristic, provides a viewpoint independent continuum between hyperbolic (saddle-shaped) and elliptic (egg-shaped) points. We then study the ability of human observers to make categorical judgments of local solid shape. We investigated this question using a smooth "croissant", a simple object made of two connected regions of elliptic and hyperbolic points. Observers decided whether the surface was locally elliptic or hyperbolic at various points on the object. The task was natural, and the observers could reliably partition the shaded image of the object into one elliptic and one hyperbolic region. The ability of observers to perform this partition shows that they can, at least implicitly, localize the parabolic curves on a surface. This ability to locate the parabolic curve can in turn be exploited for other purposes, in particular to segment an object into its parts.

A central issue in object perception is how the shape of an object is represented by the visual system. Shape may be represented in a variety of ways that depend on the visual task and the stages of processing in a given task. However, several factors might be common to all these representations. In particular, we take as a fundamental property of any shape representation that it should be viewpoint invariant. This property alone precludes traditional representations in terms of depth values or local surface orientation (Marr, 1982). We suggest here that a natural local representation of solid shape is given by the dichotomy between hyperbolic and elliptic patches.

Hyperbolic points are the points on the surface where the shape can locally be described as a saddle, while elliptic points correspond to egg-shaped regions. The boundaries between hyperbolic and elliptic regions are the parabolic curves, at which loci the surface is locally cylindrical. Parabolic curves are associated with important events occuring on a surface, such as the formation of a new occluding contour segment, a new shadow or a new specular point (Koenderink, 1990; Mamassian, 1995; Longuet-Higgins, 1960).

In this paper, we take the first steps towards characterizing human ability to categorize local solid shape in terms of view-independent descriptions. In the first section, we derive a local representation of solid shape based on the hyperbolic-elliptic dichotomy. We then describe a psychophysical experiment whose purpose was to see how reliably human observers could partition a simple surface into hyperbolic and elliptic regions. Encouraged by the results of the experiment, we discuss the advantages to have a local solid shape representation such as the one built from our viewpoint invariant descriptor.

1 Local Solid Shape Space

In this section, we seek a local shape descriptor which can partition hyperbolic and elliptic points on a surface. We consider the merits of three such descriptors, namely, the Gaussian curvature, the shape index of Koenderink (1990), and a new descriptor that we have called *shape characteristic* (Mamassian, 1993).

1.1 Local Shape Descriptors

Gaussian Curvature

A reasonable starting point for a local shape descriptor is that it should not depend on the position and orientation of the surface. In other words, a shape descriptor should be defined only from the surface curvatures. Since surfaces are two-dimensional manifolds, we need to know how the surface is curved in exactly two independent directions. For instance, we can take the two principal curvatures of the surface, which are the maximum and minimum changes in surface orientation at one point (cf. Hilbert and Cohn-Vossen, 1932/1952). The Gaussian curvature K is then simply the product of the principal curvatures.

The fundamental property associated with the Gaussian curvature is that it is a bending invariant. In other words, if a surface patch is bent (but not teared), its Gaussian curvature will remain the same for each of its points. This powerful property of differential geometry is also the main weakness of the Gaussian curvature for a psychological descriptor of shape. Assuredly, to squeeze an hemisphere into an half-ellipsoidal patch is to transform its perceived shape, so our psychological descriptor should not be invariant in this situation.

Shape Index

A potential shape descriptor which is not invariant under bending was proposed by Koenderink (1990). The shape index S was defined as the polar angle in the principal curvature space, that is, as a function of the two principal curvatures κ_1 and κ_2:

$$S = \frac{2}{\pi} \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right) \qquad \text{with } \kappa_1 \leq \kappa_2. \tag{1}$$

Since its definition involves only the principal curvatures, the shape index does not depend on the position or orientation of the object (just like the Gaussian curvature). One additional and important characteristic of the shape index is that it is also invariant against uniform scalings, so that seeing an object at different distances will not affect its shape. Together, these two properties establish the viewpoint independence of the shape index.

Nevertheless, our concern with the shape index descriptor is that it mingles two different surface qualities, namely the hyperbolic-elliptic with the concave-convex. Along the shape

index dimension, the local shape varies continuously between hyperbolic and elliptic, but also abruptly from concave to convex. For this reason, the minimal hyperbolic patches become rather arbitrarily the "null" shapes, while cylindrical patches are associated with a shape index of plus or minus one half. However, it is worth noting that while an hyperbolic patch can become elliptic only through a shape transformation, a single surface can be at the same time concave and convex (for instance, an egg-shell seen either from the inside or the outside). It thus appears that the concave-convex dichotomy is an artefactual property of the shape index definition which should be avoided for a psychological descriptor of local shape.

Shape Characteristic

For psychological purposes, it seems that a suitable shape descriptor should share with the Gaussian curvature the property that the local shape varies monotonically from hyperbolic to elliptic, and share with the shape index the property of viewpoint independence. Such an entity was introduced by Mamassian (1993). The shape characteristic χ is defined as the ratio of the principal curvatures at one point of the surface:

$$\chi = \frac{\kappa_1}{\kappa_2} \qquad \text{with} \quad |\kappa_1| \leq |\kappa_2|. \tag{2}$$

The shape characteristic varies from -1 to $+1$, from hyperbolic to elliptic points. The null shape characteristic is the parabolic point. The sign of the shape characteristic is therefore identical to the sign of the Gaussian curvature. In addition, the shape characteristic is a unitless entity, thereby borrowing the viewpoint independence property from the shape index. Finally, it can be noticed that the shape characteristic varies continuously on a smooth surface, with rare first order discontinuities at umbilical and minimal lines (where χ equals $+1$ and -1 respectively).

To account for the two degrees of freedom spanned by the two principal curvatures, we need a complementary quantity. While the shape characteristic describes *how* the surface is curved, we shall now seek to describe *how much* the surface is curved. To describe the amount of curvature in a local patch, we shall draw our inspiration from Attneave (1954),

who suggested that points of maximum curvature in a line drawing were highly informative about shape. Extending Attneave's argument to surfaces, we define the *curvature magnitude* η as a descriptor of the maximum curvature at one point of the surface, that is as a monotonic function of the largest principal curvature:

$$\eta = -\kappa_2 \qquad \text{with} \quad |\kappa_1| \leq |\kappa_2|. \tag{3}$$

The local solid shape space parameterized by the shape characteristic and the curvature magnitude is illustrated in Figure 1. As we have already noticed, the shape characteristic divides hyperbolic from elliptic points, while the curvature magnitude now separates concave from convex patches (Figure 2). At the boundary between hyperbolic and elliptic half-spaces lie the parabolic points, and at the boundary between concave and convex half-spaces lie the planar patches.

Figure 1: about here.

Figure 2: about here.

We now investigate the ability of human observers to classify the local shape on a shaded smooth object into hyperbolic and elliptic patches. It is important to note that since parabolic curves separate hyperbolic from elliptic regions, this categorical discrimination task corresponds to an implicit localization of the parabolic curve on the surface.

2 Experiment

2.1 Methods

Subjects

Among the four subjects who participated in this experiment, two were the first and third authors ([PM] and [DK]). The other two observers ([JH] and [TK]) were naive relative to the purposes of the experiment, but were graduate students also working in visual perception.

Apparatus

A shaded object was simulated using a graphics computer (a 4D35 Silicon Graphics workstation). The object was displayed on a high-resolution (1280×1024 pixels) 19 in. color monitor. The pixel brightness was quantized to 8 bits, providing a maximum of 256 different grey-levels. The screen was gamma-corrected in order to have a linear relationship between the grey-level values from the colormap and the displayed pixel brightness. After correction, the brightness varied from 0.26 cd/m^2 to 110 cd/m^2. A reduction screen was placed between the screen and the observer, so that the display could be seen only through a small aperture which was out of focus for the observer. Subjects sat in an otherwise dark room, with their head resting on a chin-rest. Viewing was monocular (the other eye covered by an eye-patch), and the viewing distance was 50 centimeters.

Stimulus

The object chosen was "croissant"-shaped, obtained from bending the long axis of an ellipsoid of revolution along a circular arc. This object has the advantage of being the simplest smooth object with just one elliptic region and one hyperbolic region. Any other bounded object, which contains these two regions (an "egg" is purely elliptical, and is therefore excluded), has either a more complex topology (for instance a "donut" whose genus is different from zero), or contains more than one elliptic or hyperbolic regions (a "peanut", for example). Figure 3 shows the iso-shape and iso-curvature lines on the stimulus.

Figure 3: (a) and (b) about here.

The shading on the croissant was computed directly from Lambert's law. The intensity of the light source and the albedo of the surface were chosen so that the brightest points on the surface (facing the light source) were rendered using the highest value of the colormap (110 cd/m^2). Since a simple Lambertian reflectance model was used, there were no specular highlights and no mutual illumination (the attached shadows were black, i.e. the lowest entry of the colormap). The light sources and the object pose were so chosen as to avoid any visible cast shadow. The shaded croissant was displayed on a dark grey background

(7.1 cd/m^2).

The croissant was always seen from the same viewpoint, under orthographic projection. It extended 11 by 13 deg of visual angle. Before each trial, the whole croissant was translated in the image plane, in order to always display the point at which the measurement was done at the center of the screen.

Design

The image of the croissant shape was uniformly sampled by ninety-two points. Four possible illumination conditions were used (Figure 4). The light source was located at infinity, either at the viewpoint (i.e. at zero slant), above the croissant ($(slant, tilt) = (35°, 120°)$), below it ($(slant, tilt) = (65°, -25°)$), or behind it ($(slant, tilt) = (105°, 60°)$).

Figure 4: (a), (b), (c), and (d) about here.

The task of the subjects was to identify a point highlighted in red on the surface as being either elliptic or hyperbolic. Whenever the local surface shape appeared more hyperbolic, the observer had to press the left button of a three-button computer mouse, and inversely when the surface appeared more elliptic to press the right button. To help the observers remember which mouse button corresponded to which shape, two small 3-D wire plots were displayed on top of the screen which represented an hyperbolic and an elliptic patch respectively. After each trial, the shape disappeared for one second until the beginning of the next trial.

The experiment followed a randomized blocked design with 8 repeated measures for the 92 measurement points. The blocked factor was the illumination condition so that, within a single block of trials, the illumination condition was kept constant and the subject visited all 92 points in a random order. Each subject ran 32 such blocks.

2.2 Results

The observers ability to categorize the local surface shape is illustrated in Figure 5. Filled circles represent measurement points which were judged more often to be hyperbolic over

repeated trials, and unfilled circles those which were judged more often elliptic. The radius of each circle indicates the consistency with which this point was judged either hyperbolic or elliptic. This figure shows that the observer succeeded to partition the object into two regions which correspond roughly to the hyperbolic and elliptic regions of the displayed shape.

Figure 5: about here.

One convenient way to summarize the results is to use the formalism of signal detection theory (Green and Swets, 1966). To the physical dimension provided by the shape characteristic, we can associate a psychological dimension along which the observer has to decide whether the point he is looking at is hyperbolic or elliptic. In this decision space, the sensations produced by the hyperbolic and elliptic shapes are represented by probability density functions. If we assume that these distributions are normal, we can thus compute the sensitivity d' of the observer to discriminate hyperbolic from elliptic shapes (the larger d' is, the more distinguishable are the two shapes). We can also determine the value of a criterion c which characterizes an overall preference of the observer for one shape over the other (the more negative c is, the larger is the bias of the observer to respond "elliptic" when the point was actually hyperbolic). The computed d' and c for each observer are presented in Table 1. We can see that for the particular stimulus used in our experiment, the sensitivity is fairly high (d' larger than 1.7), and that three out of four subjects exhibited a bias to favor the elliptic response (negative c).

Table 1: about here.

We shall now analyze further each observer's performance by focusing successively on the shape characteristic, the curvature magnitude, and the illumination condition.

Shape Characteristic

Since the task of the observer was to partition the surface along the shape characteristic dimension, it is reasonable to start our analysis with this attribute. Let us call the *elliptic*

score the proportion of times the observer decided that the surface was elliptic rather than hyperbolic. Figure 6 plots the elliptic score as a function of the shape characteristic for two observers. The sharpness of the transition between the hyperbolic and elliptic regions characterizes the observer's sensitivity. It is interesting to note that the elliptic score is a monotonic function of the shape characteristic and that all the inaccuracy of the subjects was concentrated around the parabolic curve of our surface.

Figure 6: (a) and (b) about here.

To analyze the observers' performance in more details, we ran an analysis of variance on the elliptic score as a function of the shape characteristic, the curvature magnitude, and the illumination condition. For this purpose, the 92 measurement points were binned along the shape characteristic into four groups of 23 points each, of means -0.35, 0.07, 0.27, and 0.32 respectively. The measurement points were also binned along the curvature magnitude into four groups of means 0.36, 0.39, 0.43, and $0.64\,\mathrm{deg}^{-1}$.

As could have been expected from the plots of figure 6, the analysis of variance revealed a highly significant main effect of the shape characteristic on the elliptic score (for [PM]: $F(3, 331) = 220.7, p < 0.01$; for [DK]: $F(3, 331) = 244.8, p < 0.01$; for [JH]: $F(3, 331) = 181.7, p < 0.01$; for [TK]: $F(3, 331) = 99.9, p < 0.01$).

Curvature Magnitude

Our local solid shape has two dimensions, indexed by the shape characteristic and the curvature magnitude. As the curvature magnitude approaches zero, the surface flattens out. At the limit when the curvature magnitude is null, the surface is locally planar, and no categorization into hyperbolic or elliptic is possible. It is therefore of interest to test how the surface curvature affects the performance of our observers. The analysis of variance revealed no effect of the curvature magnitude for any of our subjects (for [PM]: $F(3, 331) = 1.93, n.s.$; for [DK]: $F(3, 331) = 1.06, n.s.$; for [JH]: $F(3, 331) = 2.28, n.s.$; for [TK]: $F(3, 331) = 0.15, n.s.$), and, for three out of four subjects, there was no interaction between the curvature magnitude and the shape characteristic (the exception was observer

[JH]: $F(9,331) = 2.58, p < 0.01$). This negative effect of the curvature magnitude is puzzling, for we would have expected a drop of performance as the surface gets locally planar. It is however consistent with a similar negative effect of surface curvature on perceived local shape from binocular disparities (de Vries, Kappers and Koenderink, 1993).

Illumination condition

From the same analysis of variance described in the previous sections, we now examine the influence of the light direction on the observer's performance. Only the two naive subjects showed a significant main effect of the illumination condition (for [PM]: $F(3,331) = 2.04, n.s.$; for [DK]: $F(3,331) = 0.26, n.s.$; for [JH]: $F(3,331) = 9.06, p < 0.01$; for [TK]: $F(3,331) = 6.48, p < 0.01$), and one of these subjects also showed a significant interaction of the illumination condition with both the shape characteristic and the curvature magnitude (for [JH]: $F(9,331) = 4.21, p < 0.01$, and $F(9,331) = 5.04, p < 0.01$, respectively). Although significant for the two naive observers, the effect of the illumination condition was small and difficult to interpret.

3 Discussion

Our psychophysical experiment has demonstrated that human observers could reliably partition the image of an object into hyperbolic and elliptic regions. It should be emphasized that the task of our experiment was very natural and did not require any training (even for the two naive observers who had no interest in differential geometry). Our results should thus be appreciated relative to the apparent inability of human observers to discriminate quadratic patches from shading alone (Erens, Kappers and Koenderink, 1993b). We shall therefore discuss now what information in the image the observers might have been using to perform the task. We then conclude with some general remarks about the need of a local solid shape representation for the visual system.

3.1 Information in the Image

We shall examine here the relationships between the shape characteristic and the shading at a regular point of the surface, around a peak of irradiance, next to the attached shadow boundary, and also at the occluding contour.

Shading at a Regular Point

There exists a simple relationship between the shape characteristic χ and the pattern of shading on the surface. For a Lambertian surface, the direction of the shading gradient is given by:

$$\tan\theta \, \tan\alpha = -\chi, \tag{4}$$

where θ is the angle that the isophote makes with the first principal direction, and α is the azimuth of the light source (cf. Appendix). In other words, *the orientation of the isophote is entirely specified by the shape characteristic*, and does not depend on the curvature magnitude or some higher order surface derivatives.

One interesting correlate of the above equation is the behavior of an isophote as the light source is displaced. As the light rotates about the normal at one point of the surface, the isophote will rotate in the same direction if the point is elliptic, but in the opposite direction if the point is hyperbolic. If the point is parabolic, the orientation of the isophote will not change as the light moves. This property would therefore enable us to determine the sign of the shape characteristic, although it could not have been used by our observers since the measurements were done on a static image.

Peaks of Irradiance

One class of singularities of the field of shading are the peaks of intensity on the surface, or equivalently, the peaks of irradiance in the projected image of the surface (Koenderink and van Doorn, 1980). For a Lambertian surface, the peaks of irradiance correspond to points on the surface which face the light source.

One interesting property of the peaks of irradiance is that the shading is again completely

determined by the shape characteristic. As we show in the Appendix, *any isophote in the neighborhood of a peak of irradiance is an ellipse whose aspect ratio is the absolute value of the shape characteristic.* However, this property will be again of little use for our observers, since the isophote shape is determined by the absolute value of the shape characteristic, not its sign. An elliptic and an hyperbolic point can therefore induce an identical shading pattern in the neighborhood of a peak of irradiance. Similar ambiguities between elliptic and hyperbolic surfaces from their shading have also been reported by Bruss (1982) and Erens et al. (1993b).

Attached Shadow Boundary

The attached shadow boundary is the curve formed by the points where the light rays graze the surface. The shape of the attached shadow boundary is linked to the structure of the underlying surface (Knill, Mamassian, and Kersten, 1993). Since the shading gradient is larger in the vicinity of the attached shadow (cf. Appendix), we might expect the local solid shape to be more accurately determined at the attached shadow boundary. Erens (Erens, Kappers and Koenderink, 1993a; Erens and de Haan, submitted) showed that indeed, in regions of high contrast, observers perceive better the gradient direction and the curvature of the field of isophotes.

Occluding Contour

Koenderink (1984) showed the remarkable result that the occluding contour of a smooth object unambiguously determines the local solid shape of the object. More precisely, convex parts of the occluding contour map to elliptic regions on the surface, concave to hyperbolic, and inflections of the occluding contour correspond to parabolic points. In other words, *for a given point on the occluding contour, the sign of the contour's curvature is identical to the sign of the shape characteristic.* In comparison to other sources of information in the image, the occluding contour is the most reliable source to discriminate elliptic from hyperbolic regions.

Performance of the Observers

The analysis of the information in the image triggered our interest for the performance of the observers in the vicinity of these sparse sources of information. For this purpose, we computed the image distance of each measurement point to the closest point on the occluding contour, the closest point on the attached shadow boundary, and the closest peak of irradiance. Since no shadow was visible when the light source was at the viewpoint, this illumination condition was discarded when performance was computed relative to the distance to the shadow boundary (and similarly for the peak of irradiance when the light source was behind the object). Although we showed that the irradiance peaks were not directly informative to discriminate elliptic from hyperbolic, we include them for comparison.

Figure 7: (a) and (b) about here.

Figure 7 show, for two subjects, the percentage of correct responses as a function of the distance to a potential source of information (the measurement points were here binned in segments equally spaced along the logarithm of distance). Although these plots should be interpreted with caution since they were obtained for only one object, we can nevertheless note the following trends. The performance is usually very high in the close neighborhood of the attached shadow and occluding contours, and seems to drop as the measurement point gets farther away from one of these sources, especially beyond approximately one degree of visual angle. No such fall off is observed for the irradiance peak.

3.2 Need for a Local Solid Shape Representation

To conclude, we discuss the need for a local solid shape representation for visual processing. We consider two aspects of this issue, namely the advantages to have a viewpoint independent representation, and the need for a local in addition to a global shape representation.

Viewpoint Dependency

In the introduction of this paper, we have taken for granted that a local shape representation was viewpoint invariant. The question of viewpoint dependency has arisen in the

context of models of object recognition (Biederman, 1987; Tarr and Bülthoff, in press), but has remained a relatively unexplored issue in the perception of local shape. Several studies have focused on local viewpoint dependent descriptors, such as absolute depth (Bülthoff and Mallot, 1988), relative depth (Todd and Reichel, 1989), or slant and tilt (Mingolla and Todd, 1986). However, view-dependent descriptors have disadvantages for some tasks. For instance, the apparent global orientation of a plane is inconsistent with the variation of slant over its surface, because local slant is measured relative to the visual direction. The decomposition of an object into its parts would also seem to benefit from view-independent descriptors. As a final example, the manual prehension of an object requires one to locate stable grasp points on the surface, a task which again only makes sense in an object-centered frame of reference. Nevertheless, the visual information is of course firstly described in a viewer-centered frame of reference, and the fundamental issue then becomes the transformation of viewpoint dependent into viewpoint independent representation (e.g. Andersen, 1987).

Local Representation

It is reasonable to ask: Why have a local representation at all? One possibility is that visual processing has stages in which its hypotheses are verified by reconstructing the input (Mumford, 1992). Verification of early-level visual data would then require some explicit representation of the local contribution to shape. A second reason is that local shape representations could support processing of global shape. Specifically, estimates of local shape may be extracted early on in the visual system to facilitate the parsing of complex objects into parts. It has been suggested that a smooth object could be partitioned, either along the contours of negative minima of a principal curvature (Hoffman and Richards, 1984; Beusmans, Hoffman and Bennett, 1987), or along the parabolic curves which separate hyperbolic from elliptic regions (Koenderink and van Doorn, 1982; Brady, Ponce, Yuille and Asada, 1985; Vaina and Zlaveta, 1990). The results of our experiment suggests that such a segmentation of an object along its parabolic curves is a reasonable possibility.

Acknowledgments

This work was supported by NSF BNS-9109514. A preliminary report of this study was presented at the European Conference on Visual Perception in Edinburgh, Scotland, in August 1993. The authors would like to thank Jan Koenderink and two anonymous reviewers for their critical comments on an earlier version of the paper.

References

Andersen, R. A. (1987). The role of the inferior parietal lobule in spatial perception and visuo-motor integration. In F. Plum, V. B. Mountcastle, & S. R. Geiger (Eds.), *The Handbook of Physiology, Section 1: The Nervous System Volume V, Higher Functions of the Brain Part 2* (pp. 483-518). Bethesda, MD: American Physiological Society.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183-193.

Beusmans, J. M. H., Hoffman, D. D. & Bennett, B. M. (1987). Description of solid shape and its inference from occluding contours. *Journal of the Optical Society of America*, *A 4*, 1155-1167.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115-147.

Brady, M., Ponce, J., Yuille, A. & Asada, H. (1985). Describing surfaces. In H. Hanafusa & H. Inoue (Eds.), *Proceedings of the Second International Symposium of Robotics Research* (pp. 5-16). Cambridge, MA: MIT Press.

Bruss, A. R. (1982). The ikonal equation: Some results applicable to computer vision. *Journal of Mathematical Physics*, *23*, 890-896. Also in Horn, B. K. P. & Brooks, M. J. (1989). *Shape from Shading.* Cambridge, MA: MIT Press.

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.

Erens, R. G. F. & de Haan, E. (submitted). Detection of second order luminance structure. Also as Chapter 3 in: Erens, R. G. F. (1993). *Visual Perception of Shape from Shading*, Unpublished Doctoral Dissertation, Utrechts Biofysica Instituut, Universiteit Utrecht, Utrecht, Netherlands.

Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993a). Estimating the gradient direction of a luminance ramp. *Vision Research*, *33*, 1639-1643.

Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993b). Perception of local shape from shading. *Perception & Psychophysics*, *54*, 145-157.

Ferraro, M. (1994). Local geometry of surfaces from shading analysis. *Journal of the Optical Society of America A*, *11*, 1575-1579.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Hilbert, D., & Cohn-Vossen, S. (1932). *Anschauliche Geometrie*. Berlin: Springer. (English translation: Geometry and the Imagination, New York: Chelsea, 1952 ed.)

Hoffman, D. D. & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65-96.

Knill, D. C., Mamassian, P., & Kersten, D. (1993). *The Geometry of Shadows* (TR 93-47). Department of Computer Science, University of Minnesota.

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. & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta*, *27*(7), 981-996.

Koenderink, J. J. & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception*, *11*, 129-137.

Koenderink, J. J. & van Doorn, A. J. (1993). Illuminance critical points on generic smooth surfaces. *Journal of the Optical Society of America, A*, *10*, 844-854.

Longuet-Higgins, M. S. (1960). Reflection and refraction at a random moving surface. I. Pattern and paths of specular points. *Journal of the Optical Society of America*, *50*, 838-844.

Mamassian, P. (1993). Isophotes on a smooth surface related to scene geometry. In B.C. Vemuri (Ed.), *Geometric Methods in Computer Vision II*, Proceedings of SPIE, Vol. *2031*, pp. 124-133.

Mamassian, P. (1995). *Surface Representations for Visual Perception and Action*, Unpublished Doctoral Dissertation, University of Minnesota, Minneapolis.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W. H. Freeman and Co.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological*

Cybernetics, *53*, 137-151.

Mumford, D. (1992). On the computational architecture of the neo-cortex: II. The role of the cortico-cortical loops. *Biological Cybernetics*, *66*, 241-251.

Tarr, M. J., & Bülthoff, H. H. (in press). Is human object recognition better described by geon-structural-descriptions or by multiple-views? *Journal of Experimental Psychology: Human Perception and Performance.*

Todd, J. T. & Reichel, F. D. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review*, *96*, 643-657.

Vaina, L. M., & Zlaveta, S. D. (1990). The largest convex patches: A boundary-based method for obtaining object parts. *Biological Cybernetics*, *62*, 225-236.

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.

Yuille, A. L. (1989). Zero crossings on lines of curvature. *Computer Vision, Graphics, and Image Processing*, *45*, 68-87.

APPENDIX

Local Solid Shape and Shading

We derive here a number of properties which link the local solid shape with the patterns of shading on a smooth surface. We pay particular attention to the information carried by the shape characteristic.

Figure 8: about here.

Let us choose one point O on a smooth surface, and let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be its associated referential system such that $\mathbf{e}_1, \mathbf{e}_2$ run along the principal directions (the directions of the principal curvatures), and $\mathbf{e}_3$ is directed along the outward surface normal (Figure 8). In the neighborhood of the origin O, the surface takes the form:

$$\mathbf{P}(u,v) = u\mathbf{e}_1 + v\mathbf{e}_2 + h(u,v)\mathbf{e}_3, \tag{5}$$

where u, v are local parameters, and

$$h(u,v) = \frac{1}{2}(\kappa_1 u^2 + \kappa_2 v^2) + O^3(u,v), \tag{6}$$

with (κ_1, κ_2) denoting the two principal curvatures and the convention that $|\kappa_1| \leq |\kappa_2|$.

In the same coordinate system, the direction towards the light source can be written as:

$$\mathbf{L} = \cos\alpha\ \sin\beta\,\mathbf{e}_1 + \sin\alpha\ \sin\beta\,\mathbf{e}_2 + \cos\beta\,\mathbf{e}_3, \tag{7}$$

where α, β are the azimuth and incidence angles of the light, respectively.

The outward unit surface normal is then:

$$\mathbf{N}(u,v) = \left(1 + \kappa_1^2 u^2 + \kappa_2^2 v^2\right)^{-1/2} (-\kappa_1 u\mathbf{e}_1 - \kappa_2 v\mathbf{e}_2 + \mathbf{e}_3) + O^2(u,v), \tag{8}$$

and, for a Lambertian surface, the irradiance is proportional to the dot product of the direction to the light source and the outward surface normal, that is:

$$\begin{aligned} I(u,v) &= I_0\,\mathbf{L}\cdot\mathbf{N}(u,v) \qquad (9)\\ &= I_0\ \left(1 + \kappa_1^2 u^2 + \kappa_2^2 v^2\right)^{-1/2} (\cos\beta - \sin\beta(\cos\alpha\,\kappa_1 u + \sin\alpha\,\kappa_2 v)) + O^2(u,v). \end{aligned}$$

The gradient of shading at the origin is then (at the first order for a regular point of the irradiance map):

$$\nabla I(0,0) = -I_0 \sin\beta \left(\cos\alpha\,\kappa_1\,\mathbf{e}_1 + \sin\alpha\,\kappa_2\,\mathbf{e}_2\right). \tag{10}$$

Let θ be the angle between the isophote passing through the origin and the first principal direction. Since the isophote is orthogonal to the shading gradient, we immediately obtain:

$$\tan\theta = -\frac{\cos\alpha\ \kappa_1}{\sin\alpha\ \kappa_2},$$

or, defining the shape characteristic as $\chi = \kappa_1/\kappa_2$,

$$\tan\theta\ \tan\alpha = -\chi. \tag{11}$$

This result states the conjugate relationship between the isophote orientation and the component of the light direction in the tangent plane (Mamassian, 1993).

If we denote by $\eta = -\kappa_2$ the curvature magnitude, the magnitude of shading gradient at the origin is given by:

$$\begin{aligned} |\nabla I(0,0)|^2 &= I_0^2\ \sin^2\beta \left(\cos^2\alpha\,\kappa_1^2 + \sin^2\alpha\,\kappa_2^2\right) \\ &= I_0^2\ \sin^2\beta\ \eta^2\ \left(\chi^2\cos^2\alpha + \sin^2\alpha\right). \end{aligned} \tag{12}$$

The magnitude of the shading gradient is dependent on three factors, the elevation of the light, the curvature magnitude, and a function of the shape characteristic and the azimuth of the light. The extrema of irradiance (obtained when $|\nabla I(0,0)| = 0$) therefore occur in only three cases: (i) when the incidence of the light is normal to the surface ($\beta = 0$), (ii) when the surface is locally planar ($\eta = 0$), or (iii) when the point on the surface is parabolic and the azimuth of the light is along the axis of the local cylinder ($\chi = 0$ and $\alpha = 0$) (Yuille, 1989; Koenderink and van Doorn, 1993). Inversely, the shading gradient is large when the surface is highly curved or when the light grazes the surface (in the vicinity of an attached shadow boundary).

From Equations 11 and 12, we can determine the shape characteristic by rotating the light source about the surface normal (i.e. changing α). If the isophote rotates in the same (resp. opposite) direction as the light source, then the origin is an elliptic (resp. hyperbolic)

point; if the isophote does not rotate, it is a parabolic point (in this case, the isophote is always oriented along the axis of the local cylinder). Moreover, the change in amplitude of the shading gradient (ratio of $|\nabla I(0,0)|$ for $\alpha = 0$ and $\alpha = \pi/2$) determines the absolute value of the shape characteristic (if there is no variation, the surface at the origin is either umbilic or minimal).

Finally, we can compute the shape of the isophote in the vicinity of a peak of irradiance (for $\beta = 0$). The isophote which has constant irradiance I_1 is given by (at the first order):

$$I_1 = I_0 \left(1 + \kappa_1^2 u^2 + \kappa_2^2 v^2\right)^{-1/2}$$

or again

$$\kappa_1^2 u^2 + \kappa_2^2 v^2 = \frac{I_0^2}{I_1^2} - 1. \tag{13}$$

This is clearly the equation of an ellipse whose aspect ratio is the shape characteristic χ (the same result can be obtained by computing the principal curvatures of a peak of irradiance on the irradiance map; cf. Ferraro, 1994). To take the example of the sphere for which all points have unit shape characteristic, the above property states that the isophotes are circular near the peak of irradiance. If the sphere is squeezed, then these isophotes are transformed into ellipses whose shapes are determined by the intrinsic shape of the surface.

FIGURE AND TABLE CAPTIONS

Figure 1: In this representation of the local solid shape space, the shape characteristic varies from left to right (from -1 to $+1$), while the curvature magnitude varies from bottom to top (from negative to positive). We consider in this paper that all the patches within one column have the same shape, which is different from any other in another column.

Figure 2: The shape-space is spanned by the shape characteristic χ and the curvature magnitude η. We obtain two distinct partitions, into hyperbolic and elliptic along the χ dimension, and into concave and convex along the η dimension (the sign for η was chosen to place convex patches in the first quadrant).

Figure 3: The stimulus was a croissant-shaped object. (a) The points within a traced line have same shape characteristic χ. Each successive line represents an increment of 0.1 for χ. The thicker line is the parabolic curve. (b) The points within a line have now same curvature magnitude η. The lines are equally spaced on a logarithmic scale between the lowest curvature (in the plane of symmetry) and the highest curvature (at each of the extremities).

Figure 4: The croissant could be illuminated from the viewpoint (a), or from an angle above the object (b), below it (c), or behind it (d). (Due to the photo-reproduction process, these pictures display some isophotes which were not visible on the original stimulus.)

Figure 5: The performance of observer [PM] is shown here superimposed on the stimulus. At each measurement point, the performance over 32 trials (4 illumination conditions times 8 repeated measures) is represented as a circle. The larger the circle, the more consistent the surface was judged to be hyperbolic (filled circle) or elliptic (unfilled circle). The curve traced on the object is the parabolic line of the displayed object, that is the theoretical boundary between hyperbolic and elliptic regions.

Figure 6: The elliptic scores were monotonic functions of the shape characteristic. The data were well fit by a cumulative normal with two or three degrees of freedom (the third degree of freedom was necessary only for the naive subjects, and corresponded to a vertical compression of the cumulative normal).

Figure 7: The percentage of correct responses is plotted here against the distance to several potential sources of information. Three main sources were considered in the text: the occluding contour, the attached shadow boundary, and the peaks of irradiance. The error bar represents the mean of the standard deviation of the means.

Figure 8: The shading on a smooth surface is mainly a function of the shape characteristic χ. In particular, the orientation of an isophote makes an angle θ with the first principal direction such that $\tan\theta$ is proportional to χ.

Table 1: The sensitivity and response biases for each of the four observers of the local shape discrimination experiment.

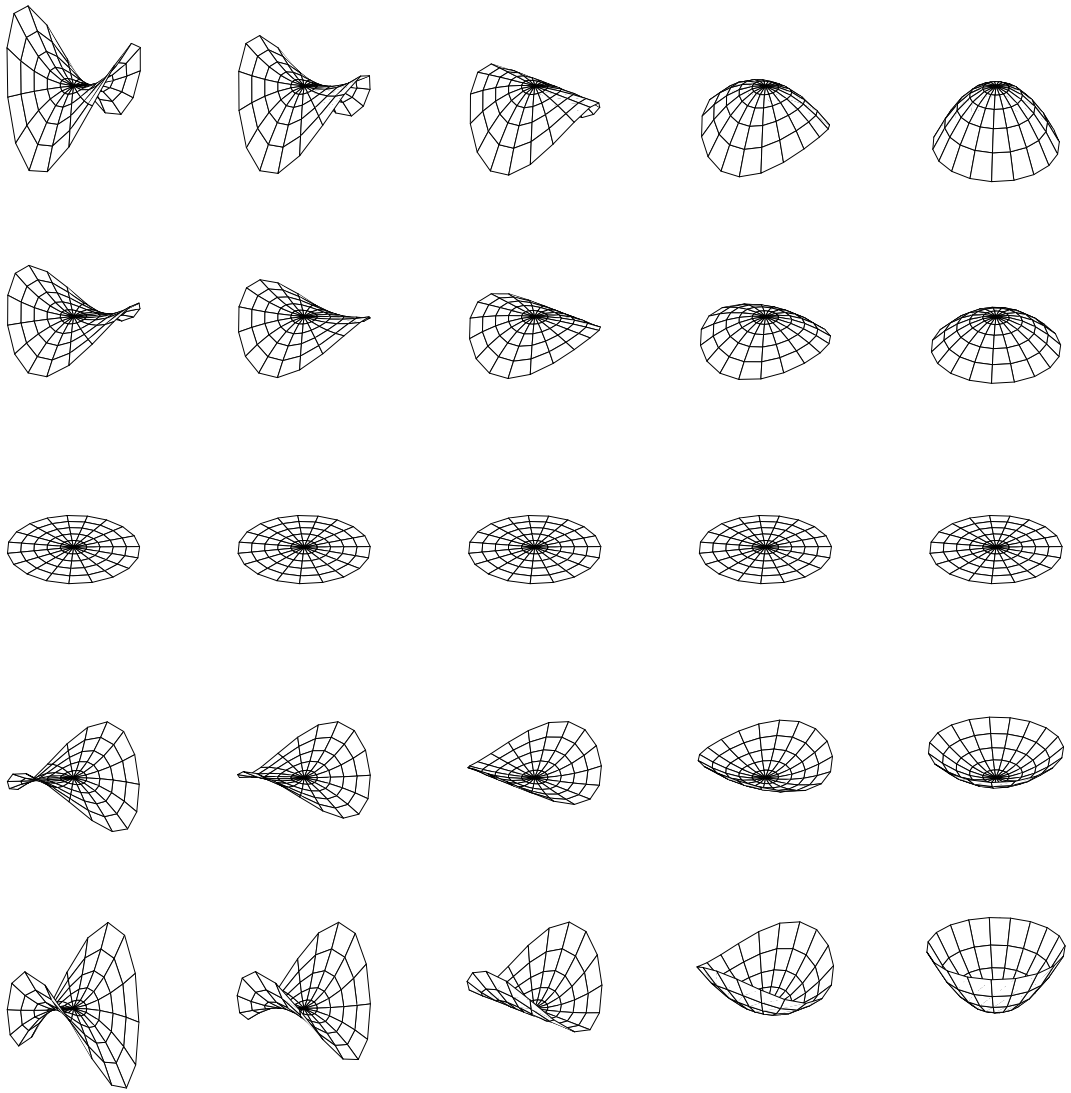

Figure 1

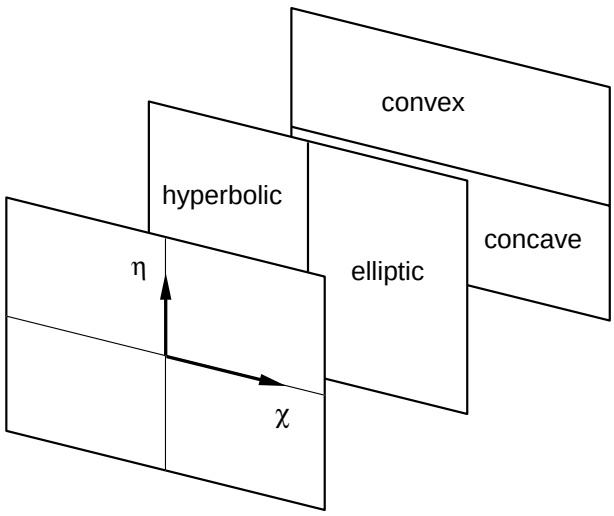

Figure 2

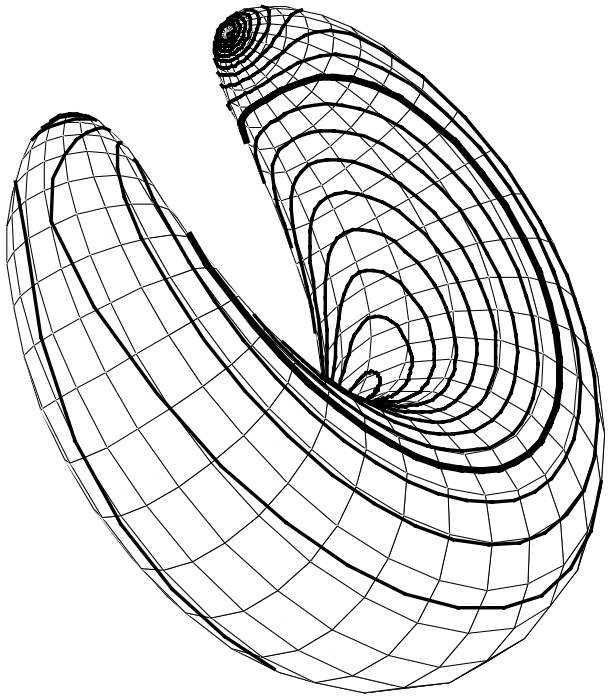

Figure 3a

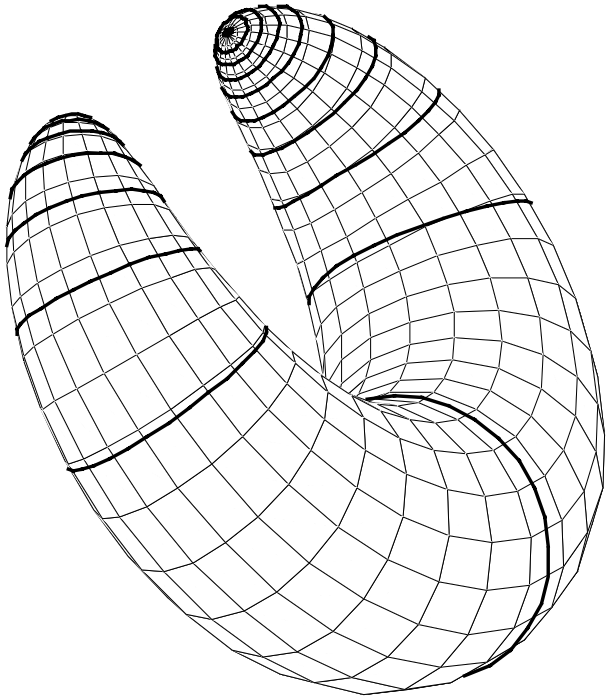

Figure 3b

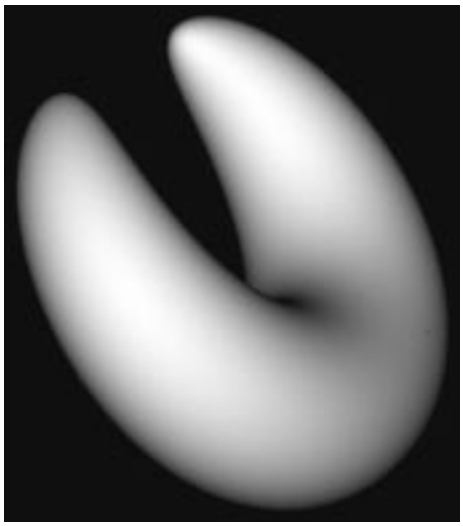

Figure 4a

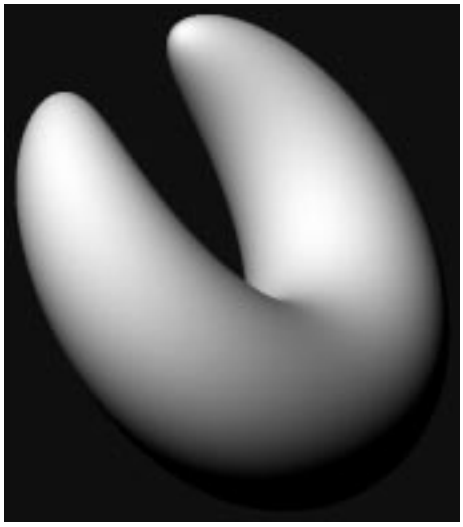

Figure 4b

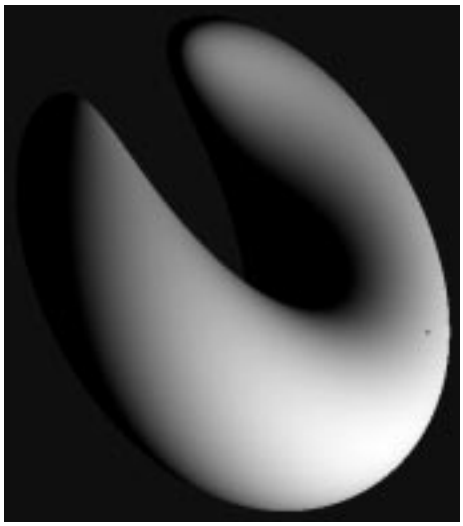

Figure 4c

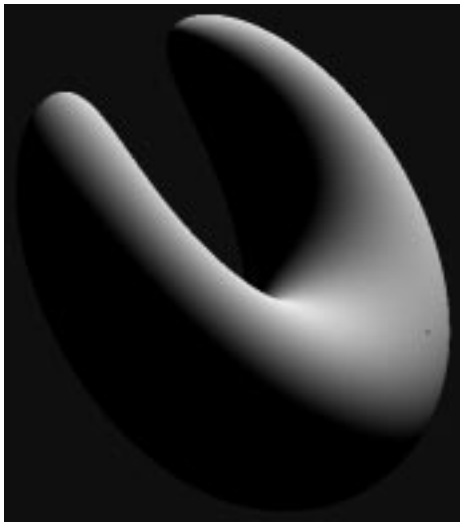

Figure 4d

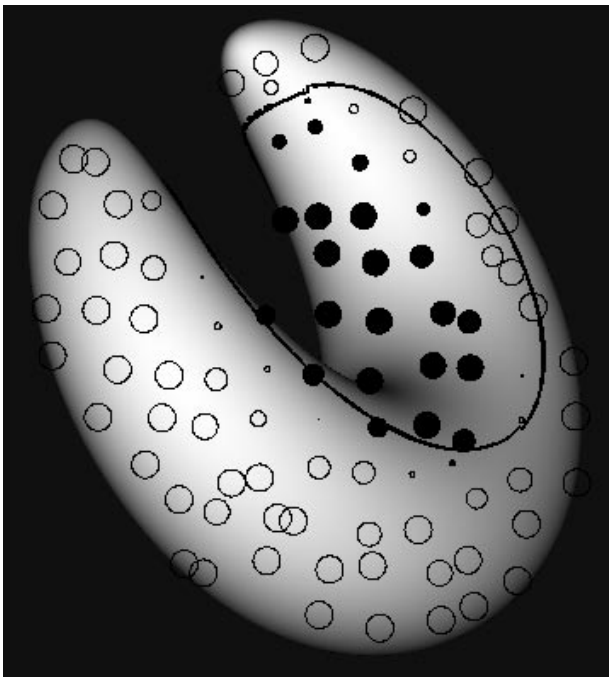

Figure 5

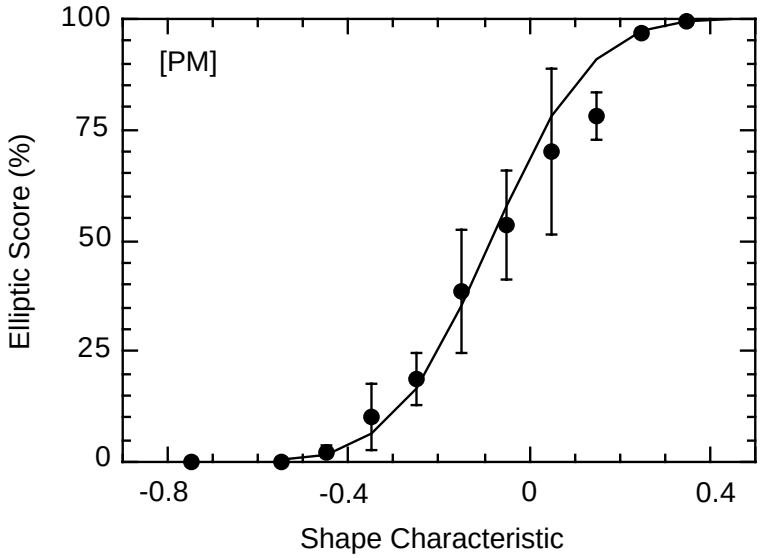

Figure 6a

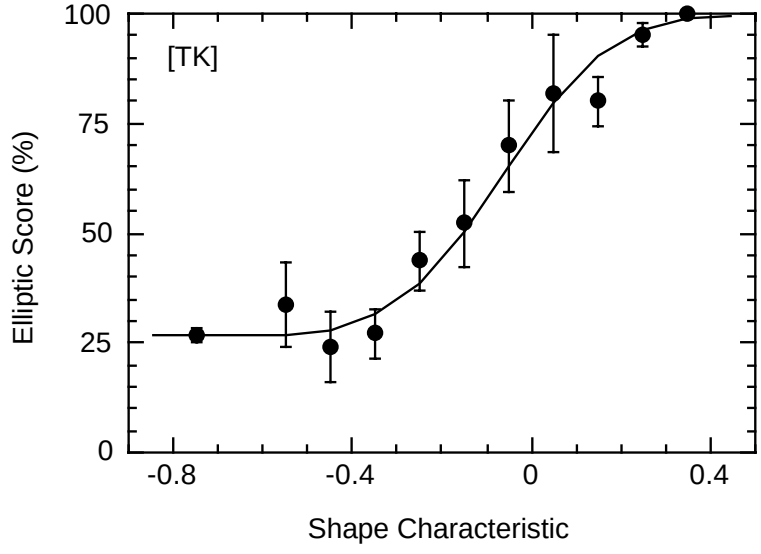

Figure 6b

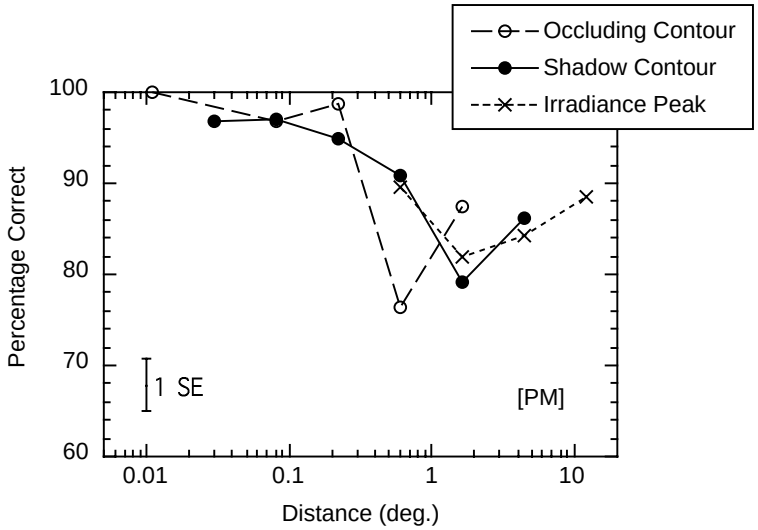

Figure 7a

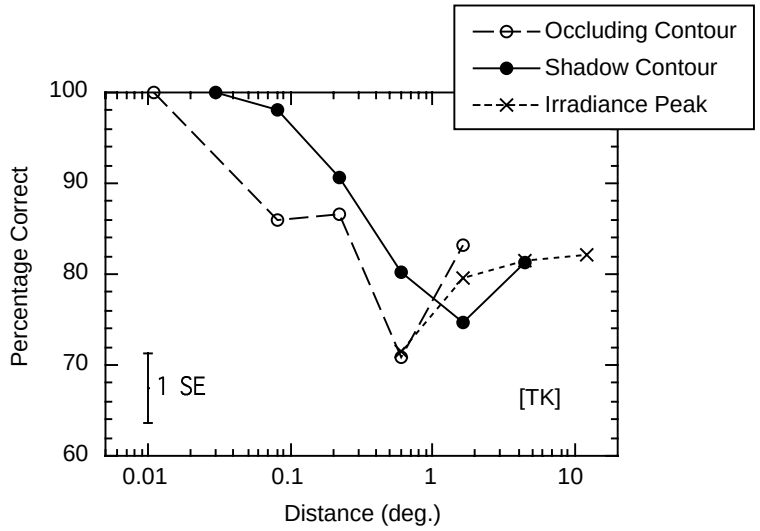

Figure 7b

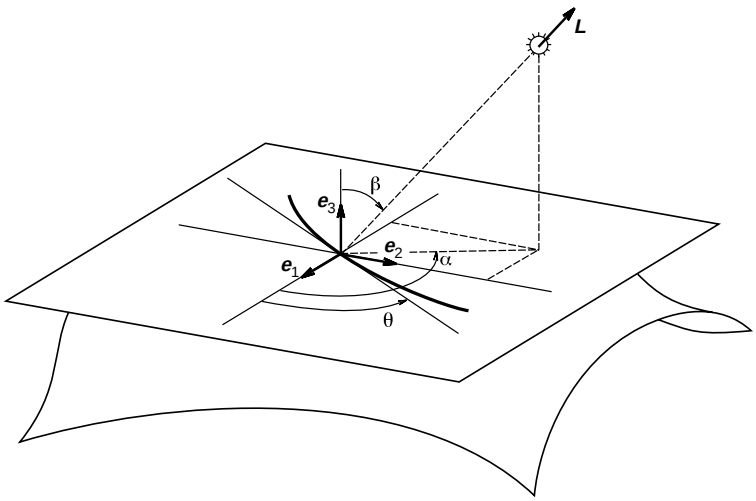

Figure 8

Observer	Sensitivity (d')	Criterion (c)
PM	2.27	-0.32
DK	2.20	0.01
JH	2.01	-0.48
TK	1.73	-0.59

Table 1