

The visual contour in depth

JAN J. KOENDERINK, ANDREA J. VAN DOORN, and ASTRID M. L. KAPPERS
Helmholtz Instituut, Universiteit Utrecht, Utrecht, The Netherlands

and

JAMES T. TODD
Ohio State University, Columbus, Ohio

We asked subjects to match points on the surface of a smooth three-dimensional (3-D) shape with points on the surface of another object that was geometrically identical to the first object but was placed in a different pose, was differently textured, and was differently shaded. In all cases, the fiducial point was on the rim of one of the objects (i.e., the boundary of the visible region of the surface), whereas the matching point was well within the silhouette of the other object. This allowed us to draw (preliminary) conclusions concerning the way monocular human observers are able to handle the neighborhood of the rim, where the local slant assumes arbitrarily high values. All experiments were done in real space with real objects (no computer-simulated scenes), the points being indicated with laser beam illumination. The subject was given control over the direction of the laser beams and was thus able to perform the task by adjustment from the vantage position. We studied both consistency (whether the subject's judgments were invariant against changes of relative pose) and veridicality (whether the depth of the visual contour as calculated from the settings agreed with the true distance as measured by mechanical means). Subjects caught much of the 3-D structure of the contour but did deviate appreciably and apparently idiosyncratically from the true geometry.

One of the most basic constraints in optically guided behavior (or *vision* for short) is due to the fact that typical objects in the ken of a human observer are optically opaque solids. This is true for the typical office or living environment (walls of our houses, furniture, other people, etc.), as well as for outside environments (buildings, cars, trees, hills, etc.). For one thing, this means that we see only the frontal side of most objects, and we are unaware of optical specification of about half of their surfaces. Obviously, the frontier between the visible and invisible parts is of vital interest. This is the "occlusion boundary" or "visual contour." In this paper, we need to distinguish sharply between the visual contour as an entity in the two-dimensional (2-D) visual field and the frontier between visible and invisible parts on the surface of the object in three-dimensional (3-D) space. We denote the former entity as the *contour*, and the latter as the *rim*. Thus, the contour is the central projection of the rim, where the center of projection is the vantage point of the observer. Both the contour and the rim depend on the object, as well as on the relative position of the observer with respect to the object (Blaschke, 1967; Cipolla & Blake, 1990;

da Vinci, 1989; Giblin & Weiss, 1987; Koenderink, 1990; Koenderink & van Doorn, 1976, 1982, 1984).

In many cases, it might be of interest to distinguish also between the *silhouette* and the contour proper, where the silhouette is the boundary of the visual projection of the object. In this paper, there is no need to do so; we deal only with the contour proper.

Despite the vital importance of the rim, most people possess only a rudimentary and most often erroneous understanding of its structure. For instance, it is common enough to hear the opinion that the rim is a planar curve and runs in a frontoparallel (to the observer) plane. This is indeed the case for any sphere that confronts the observer. The spheres also exhaust the class of objects for which the opinion holds true. More generally, the rim is a planar curve (though not necessarily frontoparallel) for all quadrics (e.g., triaxial ellipsoids) but for no other smooth solid objects (Blaschke, 1967). Thus, the common opinion holds true in certain degenerate cases that indeed often apply within the confines of the laboratory, but in few other places. Yet even Marr (1982), writing expressly on the geometry of vision, makes this common mistake.

A moment's reflection reveals that the rim is a general space curve: Just imagine the rim of a long copper wire, arbitrarily bent. Clearly, you can produce any space curve in this thought experiment. If the rim is a general space curve, then it has now to recede from and then to advance toward the vantage point as the eye traces the contour. Thus, we can indicate alternating near and far points on the contour. In a small neighborhood of a near point, the points on the contour are more distant than the near point

This collaboration was supported by NATO Exchange Grant CRG 920675. J.J.K. and A.M.L.K. were supported by the Human Frontiers Science Program. A.J.v.D. was supported by the Netherlands Organization for Scientific Research (NWO). J.T.T. was supported by Air Force Office of Scientific Research Grant F49620-93-1-0116. Correspondence should be addressed to J. J. Koenderink, Helmholtz Instituut, Universiteit Utrecht, Buys Ballot Laboratorium, P. O. Box 80 000, NL-3508 TA Utrecht, The Netherlands (e-mail: j.j.koenderink@fys.ruu.nl).

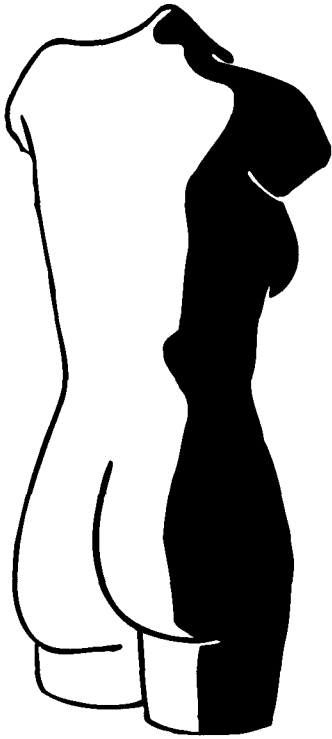


Figure 1. The mannequin in 90° pose with a body shadow revealing the rim for the fiducial pose. The light direction is from left to right, horizontal, and in the plane of the picture. This picture was traced from a photograph made from a direction perpendicular to the primary viewing direction with the mannequin in the fiducial pose and illuminated with a point source at the vantage point. The photograph was processed in Adobe Photoshop, so as to emphasize the contour and major isophotes (in order to reveal the body shadow boundary), printed and enlarged, and then traced by hand.

itself, whereas many of the points on the visible part of the surface in a small neighborhood of the near point must be closer than the near point (because the surface curves into depth toward the contour): Thus, the near point appears as a special type of “saddle.” In a small neighborhood of a far point, the points on the contour are closer than the far point itself, whereas the points on the visible part of the surface in a small neighborhood of the far point must also be closer than the far point: Thus, the far point appears as a special type of “extremum” (a depth maximum; see Koenderink, 1990.) At the near and far points, the tangent direction of the rim is always orthogonal to the visual direction. Because the tangent direction at the rim and the visual direction are also *conjugate* directions of the second fundamental form of the surface (Blaschke, 1967; Koenderink & van Doorn, 1984), this means that both the visual direction and the tangent at the near and far points of the rim are *principal* directions of the surface (Koenderink, 1990). In other words, you have a near or far point at the contour when the visual direction happens to coincide with one of the principal directions of the surface, which is why the sphere is a degenerated case:

All directions are principal directions; thus, every point of the contour is a near or far (no distinction here) point.

In special cases (Koenderink & van Doorn, 1982), the tangent to the rim may actually run straight into depth. In such cases, the projection of the tangent to the rim degenerates to a point. This is the case of the *ending contour*, where the visual direction happens to coincide with an asymptotic direction of the surface. Such cases are not at all rare in realistic circumstances. For instance, the contour of a human body seen from a reasonable angle typically has several of such “cusps.”

The structure of the visual contour with the alternation of near and far points and the additional inflections was clearly envisaged by Leonardo da Vinci (1989). The inflections indicate the intersections of the parabolic curves of the surface with the rim (Koenderink & van Doorn, 1976). Leonardo tells the painter to eyeball the pattern of to-and-fro movements of the rim in depth for real objects and express it in paint. We do not know whether Leonardo discovered these facts through his (uncanny) powers of observation or whether he first reasoned it out and confirmed the conclusion visually. In any case, given the importance of the contour for optically guided behavior and the widespread misconceptions concerning its true nature, it seems of some interest to address the issue of the visual perception of the surface of objects near their rims.

The contour has been studied in mathematics (differential geometry and topology), computer vision, and human psychophysics. Whereas early geometers settled issues of planarity and frontoparallelity (Blaschke, 1967), more recent work has investigated the apparent curvature of the contour and its generic singularities (Arnold, 1984; Cipolla & Blake, 1990; Faugeras, 1993; Giblin & Weiss, 1987; Koenderink, 1990; Koenderink & van Doorn, 1976, 1982, 1984). The study of singularities already considers the evolving contour in dynamic situations. Euclidean motions that affect the pose of the object with respect to the vantage point are especially relevant since they cover the case of the moving observer. Almost all work in computer vision and human psychophysics has been focused on the dynamic case because single images are necessarily ambiguous, whereas time sequences of images or continuously evolving ones specify 3-D properties in an objective sense (Cipolla & Blake, 1990). Early observations by Mach (1959) indicated that rigidly moving ovoid objects are often perceived as continuously deforming, jelly-like ones. This is typically the case when the objects contain few or no surface markings. More recent research has departed from these informal observations (Cortese & Andersen, 1991; Norman & Todd, 1994; Norman, Todd, & Phillips, 1995; Pollick, Giblin, Rycroft, & Wilson, 1992). Human perception tends to become more veridical the more generic (complicated shapes, general movements) the situation (e.g., Norman & Todd, 1994). Research in computer vision has clarified many possible reasons for this. Perhaps not surprisingly, shape from contour is least ambiguous and most robust in generic situations. In this study, we consider rigid, stationary, generic objects viewed

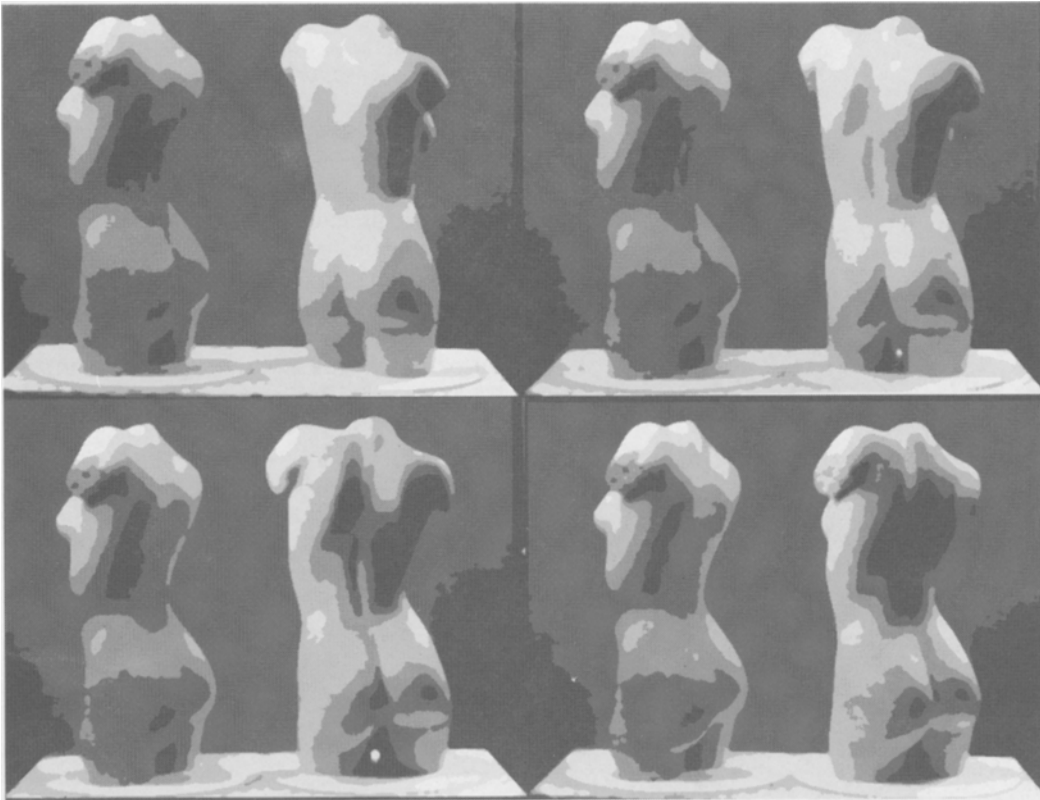


Figure 2. Posterized photographs of the field of view of the subject from the vantage position. The posterization brings out the differences in shading. The four panels show the pose differences used in the experiment: top left, 90°; top right, 70°; bottom left, 45°; bottom right, 20°. In each panel, the fiducial object is on the left. The apparent height of the mannequins is about 13° of visual angle. (Apparent small differences in the left-hand object—these should all be equal—are due to slight inaccuracies in the photographic process, subsequent digitalization, and posterization.)

from a stationary vantage point. Thus, most of the analyses and psychophysical data do not immediately apply to the present study.

When an object is illuminated by a point source of light at the vantage point, the boundary of the body shadow coincides with the rim (Koenderink, 1990), because we have simply reversed the direction of the light rays. We used this simple fact to “photograph” the rim as a true space curve: Figure 1 shows our stimulus seen in profile view (distance from the vantage point increases toward the right) with the body shadow marked. If the rim were a planar, frontoparallel curve, the boundary of the body shadow would appear as a vertical straight edge in this picture. This is clearly far from being the case. The rim has a rather complicated depth structure, dependent on the surface geometry of the object.

There exist many reasons of a geometrical and physical nature why it might be difficult to explore the neighborhood of the rim visually, except for its 2-D character as a contour in the visual field. Near the rim, the surface is almost seen edge on; thus, its details are extremely foreshortened. The depth gradient tends to infinity (i.e., it assumes arbitrarily high values near the contour as the slant angle approaches a right angle). Moreover, near the rim,

small deviations from the overall smooth surface (e.g., tiny hairs sticking out of the surface) become very important. For instance, even shallow unevennesses that go unnoticed when the surface is seen at a reasonable slant may occlude large parts of the surface in the regions near the rim. Many effects of physics also become important near the rim. For instance, extreme deviations of otherwise nearly “Lambertian behavior” are the rule rather than the exception near the contour. Of course, such effects often specify important material properties, and they can be crucial in effective optically guided behavior.

These simple facts provide one explanation why the visual contour is never (at least in our experience) satisfactorily rendered in computer-generated (simulated) images: The renderers do not capture the relevant physics. Painters take great pains to render the nature of the visual contours in their abundant variety with the utmost care, because the nature of these entities is of decisive importance for naturalistic renderings (Jacobs, 1988). This is possibly one important reason why computer graphics so often looks less realistic than even mediocre naturalistic paintings. This is also the reason why we went to great pains to do our experiments on real shapes rather than on computer graphics, convenient as the latter may be.

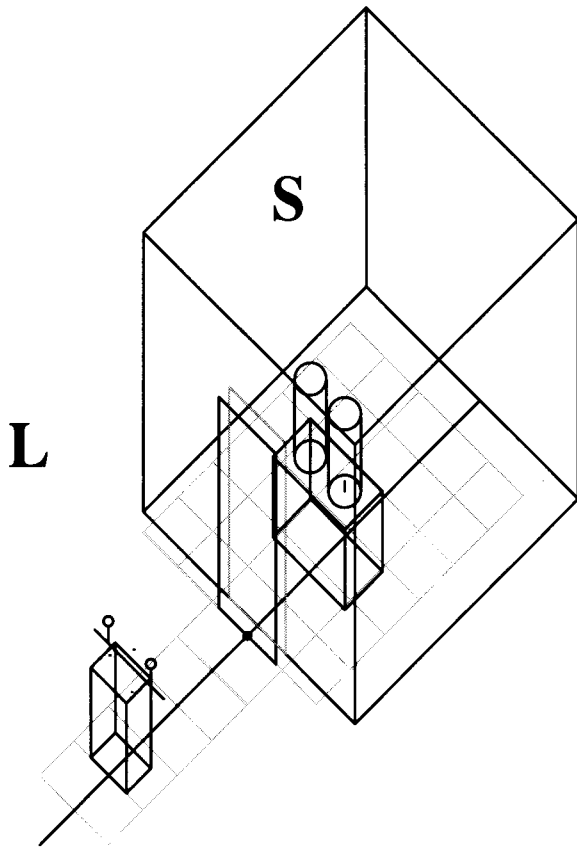


Figure 3. Axonometric rendering of the studio (S) and the laboratory space (L). The gridlines on the floor are spaced by 50 cm in both orthogonal directions. Indicated are the subject's table with optical bench and the two aluminum spheres, the door opening, and the object table with the two objects. The mannequins are shown as cylinders of the correct (maximum) width and height.

In this paper, we address a very simple question: Do observers see *where* on the surface the rim actually is? In other words, can they identify points on the rim in terms of landmarks on the objects as seen in a more favorable pose? This question can easily be addressed experimentally when we confront the observer with two geometrically identical objects in different poses. We specify a point on the rim of one of the objects (the "fiducial point") and ask the observer to indicate the corresponding point on the other object where it lies well within the area delineated by the contour (see Figure 2). Observers can perform such a task with great precision when both the fiducial point and its corresponding point are well within the areas delineated by the two contours (Koenderink, Kappers, Pollick, & Kawato, 1997), even when the two (geometrically identical) objects are painted with very different textures (Phillips, Todd, Koenderink, & Kappers, 1995): Subjects typically do not solve the task by attending to local textural detail; rather, they refer locations to the curvature landscape over some wider region. In the present task, we expected large deviations between the actual corresponding point and the point indicated by

the subject in such a way that the indicated points may correspond to large depth shifts of the rim.

METHOD

Objects and Apparatus

For objects, we used mannequins expressly sold for the display of fashion items. These mannequins can be obtained in a variety of postures of both genders and in any number of identical copies. Many have complicated shapes, yielding rims that have a rich pattern of near, far, and inflection points. We mounted the objects on turntables, which were mounted on a single tabletop; the axes of rotation were 55 cm apart. Both turntables were horizontal and coplanar; the axes of rotation were thus vertical and in the frontoparallel plane. One object (the left one for the subject) was put in a fixed pose, whereas the other object (the right one) could be set in any of a number of poses (Figure 2) parameterized by the angle subtended by a reference direction (in identical coordinate frames attached to the objects) at the vantage point. Thus, we corrected the difference of the poses for the apparent rotation of objects in the visual field when they have been translated in the frontoparallel plane.

The table with the mannequins was placed in a studio that could be viewed by the subjects from the laboratory through a door opening (an axonometric rendering of the studio and laboratory spaces is shown in Figure 3). This configuration enabled us to place light sources almost ad lib but out of sight of the subject. We applied a studio illumination that revealed surface detail to good advantage (sources frontal, to the side, somewhat elevated) and illuminated both objects similarly (but, of course, not identically) without undue interactions, such as shadows cast by one object onto the other. The background was uniform (a large studio roll of background paper) and illuminated such as to avoid undue contrast at the contour but, at the same time, revealing the contour clearly. The mannequins were geometrically identical. They were covered with a fine "granite" texture, rendered as dark on light-gray speckling. Each object had another instance of this texture, thus correspondences could not be established on the basis of texture elements.

The subject was seated at a table with a chinrest, the eye at about half the height of the mannequins and in their median plane. The subject table also carried an optical bench, with two large (10 cm diameter) aluminum spheres placed on spherical cups mounted on riders on the optical bench. The spheres were held in place through

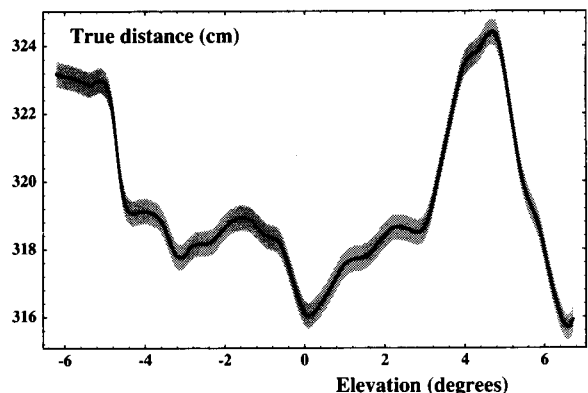


Figure 4. The true distance (in centimeters) of the rim as a function of the elevation (in degrees of visual angle) in the visual field. The ribbon indicates the envelope of three independent measurements and, thus, the most pessimistic estimate of the tolerance. The drawn curve indicates the mean. Note that the rim indeed moves back and forth in depth, often quite markedly for even small changes of elevation.

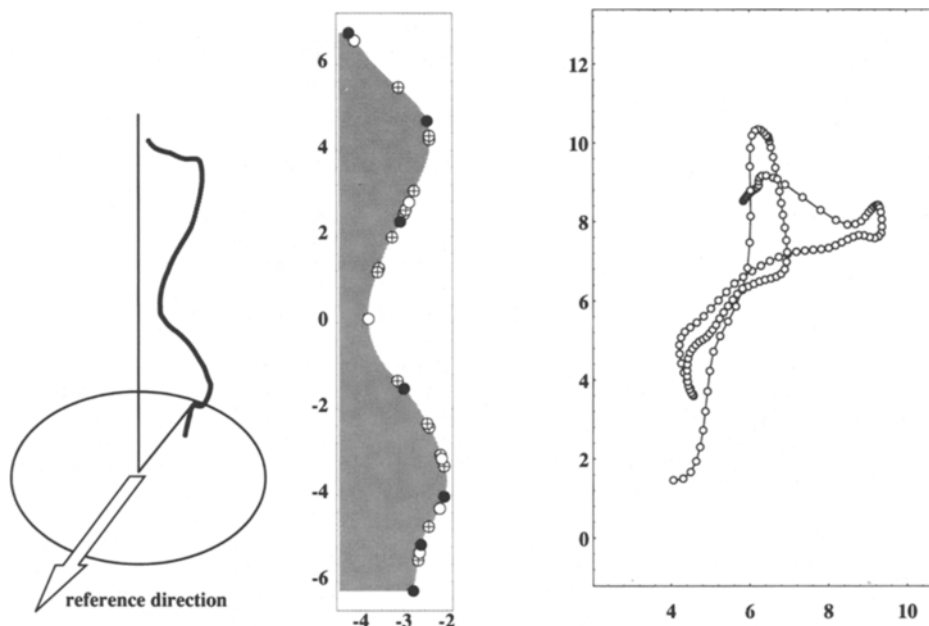


Figure 5. The true rim and contour. On the left, the rim as a space curve drawn in true proportion over the turntable top and with the axis of rotation of the mounted mannequin. The reference direction indicated on the turntable top is the straight-ahead direction for the subject. In the middle, the silhouette with characteristic points of the contour (filled circles, far points; open circles, near points; circles with pluses, inflection points). Units are degrees of visual angle (azimuth horizontal, elevation vertical). On the right, the orthographic projection of the rim on the horizontal plane. Units are centimeters. Distance away from the subject is plotted upward; horizontal is rightward.

gravity; the cups were slightly greased. As a result, the spheres could easily be rotated, but they remained motionless when left untouched. These spheres were pierced and took commercially available (red diode) laser pointers. The subject could easily manipulate the aluminum spheres, thus casting laser beams into the studio in a precisely controlled way. One beam was used to indicate the fiducial point on the rim of the first object; the other beam was manipulated by the subject so as to indicate the corresponding point on the second object. The laser dots were not unduly bright but were easily visible and were perceived as being on the surface of the objects.

Both the studio and the laboratory were carefully calibrated and carried a coordinate grid on the floor (not visible to the subject). We could find the position of any point in these rooms (about 100 m³) with millimeter accuracy. The vantage point, the centers of the laser spheres, and the positions of the turntables were known with precision. The door between the studio and the laboratory had an additional significance in this experiment: When the door was closed, we obtained a coordinate grid in the plane of the door (itself accurately known in the laboratory coordinate system) that allowed us to read the positions where the laser beams pierced the plane of the door. Since we knew two points of the laser beams with great precision (the centers of the aluminum spheres and the points in the doorplane), we knew the absolute position of the laser beams in space. Thus, we could use the lasers as a position-measuring device by finding the point of intersection of the laser beams. We define the point of intersection as the point of closest approach—that is, the midpoint of the common perpendicular of the laser beams.

To determine the rim with even greater precision, we also attached a pulley to the subject table (position carefully calibrated), through which we ran a thin steel wire, several meters in length. The wire was kept taut by a weight on the end near the table. The other side of the wire held a pencil-shaped pointer used to touch points in the scene. The height of the weight (the distance from the ground)

could be read on a scale. By pointing at a reference point in the studio, we first calibrated the wire length. The distance to any point on the rim could be determined with millimeter accuracy. The position measurements via laser triangulation and steel-wire pointing were carefully checked and crosschecked with encouraging results: We could pinpoint arbitrary points within the studio and the laboratory with millimeter accuracy and perfect reproducibility over time spans of weeks.

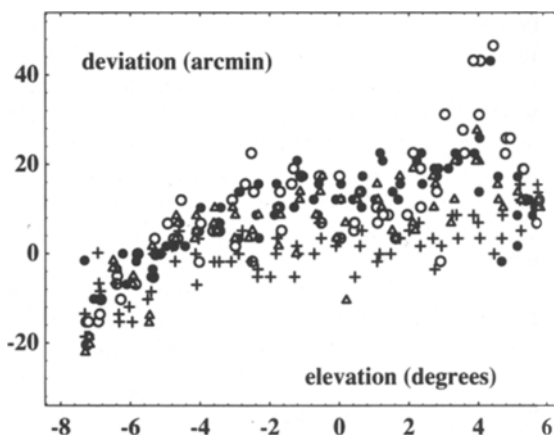


Figure 6. The mismatches in the elevation (in minutes of arc of visual angle) as a function of the elevation (in degrees of visual angle) in the visual field for all 4 subjects. The pose difference was 90°. Subject A.D., open circles; Subject A.K., filled circles; Subject J.K., plus signs; Subject J.T., open triangles.

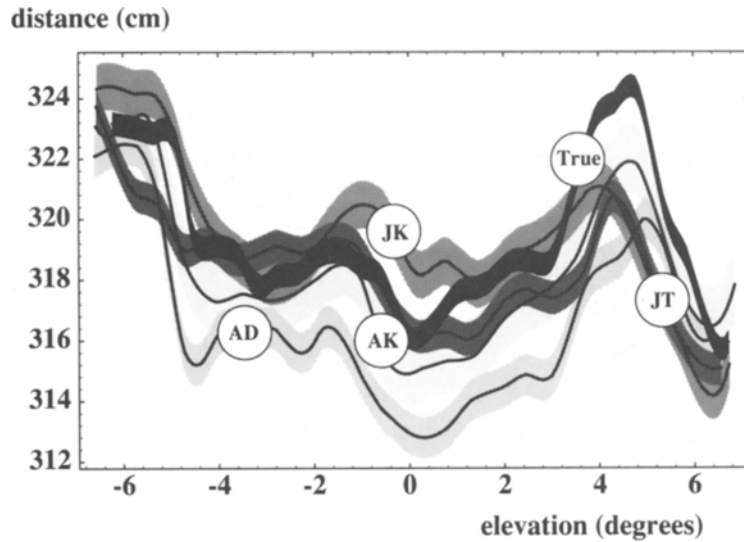


Figure 7. The distance to the rim (in centimeters) as a function of the elevation (in degrees of visual angle) for all subjects (gray ribbons) and the true distance to the rim (black ribbon). The ribbons represent the envelopes of three independent measurements. Pose difference 90°.

Initial Calibration

The left mannequin was placed in a pose that yielded a smooth contour (no cusps or self-intersections) and strong in-depth modulation. At one point, the pose led to an “almost cusp” (a small turn would indeed produce an ending contour); thus, the rim ran almost straight into depth at that place (this is clearly visible in Figure 1). This mannequin was then fixed in this pose.

The fiducial rim was measured by pointing at it (looking from the subject’s vantage point) with the right laser pointer (since the left one could not reach the rim) and then finding the distance with the

steel-wire probe. These observations were then converted to absolute room coordinates (height and horizontal distances from the room coordinate origin) and to polar coordinates in the subject’s frame (distance, azimuth, and elevation). We also plotted the contour in absolute coordinates on the plane of the door.

In Figure 4, we show the true distance of the contour as a function of the elevation. In this range graph, we have indicated the envelope of several independent measurements (different days). The uncertainty is mainly due to the fact that the laser dot on the mannequin is rather elongated in the direction of the beam, since the

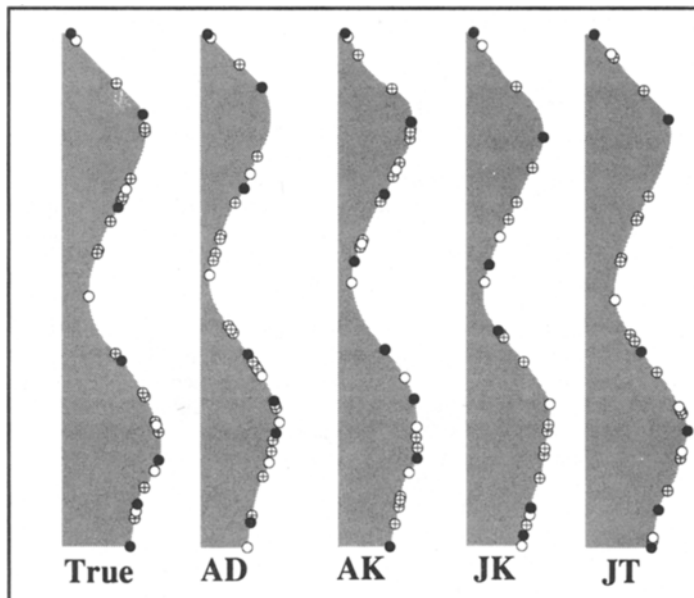


Figure 8. The silhouette with characteristic points of the contour (filled circles, far points; open circles, near points; circles with pluses, inflection points) for all 4 subjects. The true silhouette with its characteristic points is also shown in Figure 5. Pose difference 90°.

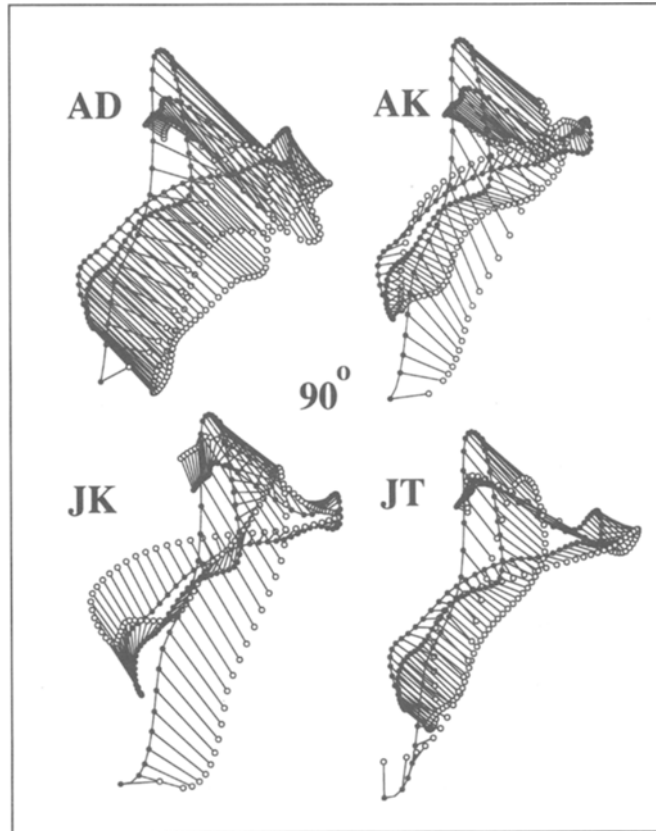


Figure 9. The projections of the rim on the horizontal plane for all 4 subjects for the pose difference of 90° . The projection of the actual rim is shown in Figure 5. For each subject, we show the true projection with that calculated from the subject's settings, corresponding points being joined by straight line segments. Thus, we obtain a field of deviations from veridicality. Note the marked differences between the subjects, but also note the equally obvious global similarities.

beam skims the surface. From the horizontal projection, we see that the rim is indeed a general space curve with a strong movement in depth (see Figure 5, right panel).

In the left panel of Figure 5, we show the rim. In the center panel of Figure 5, we show the contour as plotted in the visual field of the subject, as well as the characteristic points of the silhouette. These characteristic points are of two types: The *inflection* points mark inflections of the contour as a curve in the visual field (such a curve is a "spherical curve," i.e., a 1-D trace of visual directions). The *near* and *far* points mark the minima and maxima of the distance function along the contour. In the right panel of Figure 5, we show the projection of the rim on the horizontal plane.

Experimental Procedure

The right mannequin was placed in some pose (e.g., turned over 90° with respect to the left one as seen by the subject) and fixed. The image of the contour of the left mannequin was placed in position on the doorplane. The doorplane was covered with a thin iron sheet so that we could conveniently affix large pieces of paper kept in place with little magnets. We divided the height in this contour image into about 25 divisions. The subject then aimed the right laser pointer to each of these points on the contour image in turn. When the door was opened, the laser dot indicated the fiducial point on the rim of the left object. The subject then used the left laser pointer to indicate the corresponding point on the right object. The

door was closed again, and this point was marked on the paper. In this way, we could do a full measurement of the contour in about 30 min. The paper is a permanent record that can be evaluated at any convenient time.

Vision was monocular. During the experiment, the subject used a chinrest to fix the head, placing the eye behind a peephole with a diameter of 1 cm. The center of the peephole defined the vantage point.

Subjects

The experiment was performed by 4 subjects (the authors). A.K. and J.T. are emmetropic, J.K. is presbyopic but needs no correction at this viewing distance, and A.D. is slightly myopic and wore her usual correction. All subjects performed the task for pose differences of 90° and 45° . A.D. and J.T. also did the task for pose differ-

Table 1
Correlations of the Depth of All Subjects Mutually and With the True Distance for a Pose Difference of 90°

| | A.D. | A.K. | J.K. | J.T. |
|------|------|------|------|------|
| True | .85 | .74 | .65 | .93 |
| A.D. | | .71 | .53 | .95 |
| A.K. | | | .79 | .81 |
| J.K. | | | | .68 |

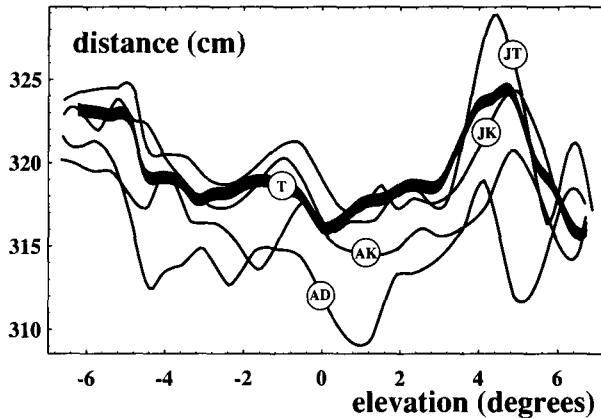


Figure 10. The distance to the rim (in centimeters) as a function of the elevation (in degrees of visual angle) for all subjects (thin curves) and the true distance to the rim (black ribbon). The ribbon represents the envelope of three independent measurements. Pose difference 45°.

ences of 70° and 20°. The measurement at 90° pose difference was repeated three times by all 4 subjects on different days.

Treatment of the Data

The experiment yields sets of pairs of visual directions: The fiducial point on the rim of the left mannequin and the corresponding point (as indicated by the subject) on the right mannequin. We also know the distance to the fiducial point from the initial calibration with the steel wire, and we could (in principle) find the distance to the corresponding point via triangulation with the laser beams.

However, we did not measure the distance to the indicated point on the right mannequin; instead, we transformed the beam toward this point into the coordinate frame affixed to the left mannequin. This makes sense because the objects are geometrically identical. The transformation involves a translation (because the objects are at different places) and a rotation (because the objects are in different poses). Once we have the two laser beams in the frame of the left mannequin, we can find their common intersection. (In practice, the laser beams will be skew; we take the midpoint of the common perpendicular as the best estimate of the point of intersection.) This point can be converted to observer coordinates (distance, azimuth, and elevation). It can thus directly be compared with the spatial position of the fiducial point. The relevant parameter is the distance, since the point is on the fiducial beam and thus has the same azimuth and elevation as the fiducial point (we refer to this distance as calculated from the subject's settings as the *depth* for that subject. In general, this depth differs from the [true] distance). This type of data reduction thus leaves us with a single datum for every fiducial point—namely, the depth of the indicated point that can immediately be compared with the distance of the fiducial point. This method probes the depth structure of the rim.

RESULTS

Pose Difference 90°

Because the mannequins were geometrically identical and were mounted on coplanar turntables with axes at symmetrical positions in a frontoparallel plane, we expected only minor elevation differences between the fiducial and the corresponding point for all pose angle differences. However, we found that these differences were

appreciable for all subjects (see Figure 6). For all subjects, we found nonrandom trends. These mismatches were clearly errors. The trend was similar for all subjects. When we fit regression lines, we found (expressed in minutes of arc) $12' + 2.6\epsilon$ for Subject A.D., $12' + 1.5\epsilon$ for Subject A.K., $2' + 1.4\epsilon$ for Subject J.K., and $8' + 2.0\epsilon$ for Subject J.T., where prime (') denotes minutes of arc and ϵ denotes the elevation in degrees of visual angle. The root mean square (RMS) residual deviations from these regressions were 9' for Subject A.D., 10' for Subject A.K., 13' for Subject J.K., and 10' for Subject J.T.

The reproducibility was quite good for all subjects. The average RMS spread for the depths was 7 mm for Subject A.D., 6 mm for Subject A.K., 8 mm for Subject J.K., and 12 mm for Subject J.T. This spread is quite small with respect to the depth range covered along the contour. We conclude that the subjects did capture appreciable depth variations of the contour with this method.

All subjects showed obvious deviations from veridicality; the RMS depth deviations from the true distance amounted to several centimeters and were indeed very significant (see Figure 7). Having said this, it is equally true that the settings of all subjects did capture at least the qualitative depth structure of the contour. This can be gleaned from the near and far points (Figure 8) and from the projections on the horizontal plane (Figure 9). No subject acted as if the rim were a frontoparallel planar curve; all captured the qualitative depth structure and the general magnitude of the depth variation quite well, though with appreciable quantitative differences. These quantitative differences do not appear to have followed a common pattern but seem to have been idiosyncratic.

Correlating the depths from the various settings of the subjects among each other and with the true distance, we obtain the correlation coefficients compiled in Table 1. The correlation with the true depth structure is a measure of the veridicality and ranges from 0.65 (Subject J.K.) to 0.93 (Subject J.T.). These correlations are very significant, and we may conclude that the subjects indeed managed to capture the true depth structure with their responses. Much of the intersubject correlation must have been due to the correlation with the true distance (the veridicality) common to all subjects. An intersubject correlation of the residuals of each individual subject with the true depth structure corroborates this because it reveals that most combinations (except one) were not significantly different from zero. The one significant (correlation coefficient 0.8) correlation was A.D.–J.T.; these subjects deviated from veridicality in nearly the same way.

Table 2
Correlations of the Depth of All Subjects Mutually
and With the True Distance for a Pose Difference of 45°

| | A.D. | A.K. | J.K. | J.T. |
|------|------|------|------|------|
| True | .53 | .72 | .82 | .85 |
| A.D. | | .36 | .45 | .56 |
| A.K. | | | .84 | .66 |
| J.K. | | | | .70 |

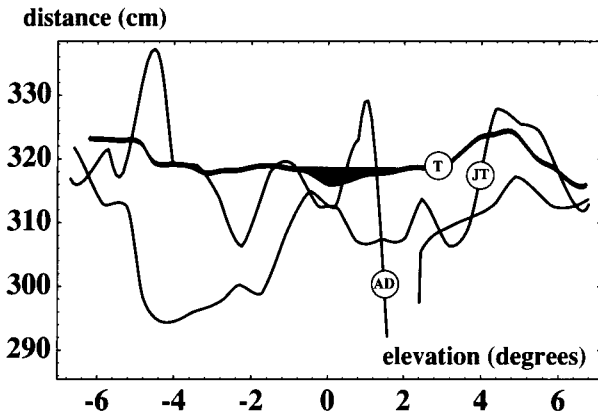


Figure 11. The distance to the rim (in centimeters) as a function of the elevation (in degrees of visual angle) for Subjects A.D. and J.T. (thin curves) and the true distance to the rim (black ribbon). Pose difference 20°. The ribbon represents the envelope of three independent measurements.

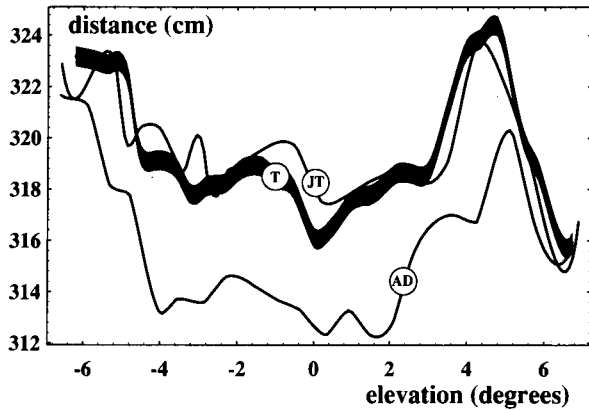


Figure 12. The distance to the rim (in centimeters) as a function of the elevation (in degrees of visual angle) for Subjects A.D. and J.T. (thin curves) and the true distance to the rim (black ribbon). Pose difference 70°. The ribbon represents the envelope of three independent measurements.

Pose Differences Smaller Than 90°

For smaller pose differences, we found basically the same pattern. One difference was that the computation of the spatial position of the indicated point became more uncertain, because the two laser beams (in the frame of the left mannequin) were more nearly parallel. (For a given tolerance in the beam direction—that is, the setting—the tolerance in the depth was inversely proportional to the angle subtended by the beams.) On the other hand, the poses were more similar, and one might perhaps expect the subjects to have done better in an apparently easier task. These effects oppose each other; thus, it is not easy to predict what the results might be. In the extreme case of almost identical poses, one would need extremely high precision in the settings in order to recover the depth: Since we expect some uncertainty in any case, even for essen-

tially identical poses, we expect very noisy results in this limit. Indeed, we found that the data become more noisy with decreasing pose difference.

For a pose difference of 20°, the poses were very similar indeed, and the data are very noisy with a few obvious outliers. The subjects experienced this task as very difficult.

For the 45° pose difference, the depths are compared with each other and with the true distance in Figure 10. The table of correlations (Table 2) reveals a pattern not much different from that at the pose difference of 90°.

Measurements at the pose differences of 20° and 70° are only available for Subjects A.D. and J.T. The depth structures are depicted in Figures 11 and 12. For pose difference of 20°, we have an awkward outlier in the data. The deviations from veridicality increased when the pose difference became smaller. The table of correlations for the pose difference of 20° is Table 3; the table of correlations for the pose difference of 70° is Table 4. As expected from the figures, the correlations deteriorate, and the responses reflect hardly any veridical structure.

Comparison of Pose Angles

In Figure 13, we show the depth structure for all pose angle differences for Subject J.T. Except for the pose angle difference of 20° (for which the task was very difficult, and the calculation of the depth structure numerically rather unstable), all results were similar. The table of correlations (Table 5) reveals that all correlations were significant, though perhaps pose differences that were more alike tended to correlate more than pose differences that were unlike. This conclusion equally applies to the other subjects (the table of correlations for Subject A.D. is shown in Table 6). Apparently, the subjects were able to indicate what was the rim region of one pose on the area within the silhouette of different poses. They could do this more or less irrespective of the actual pose difference. Note that the contour of the right mannequin depended critically on the pose angle difference, and so did the shading. Since this did not influence the results, we may conclude that the subjects used the structure of the perceived curvature landscape to find the correspondence, and they did well at such a task.

Table 3
Correlations of the Depth of Subjects A.D. and J.T. Mutually and With the True Distance for a Pose Difference of 20°

| | A.D. | J.T. |
|------|------|------|
| True | .16 | .53 |
| A.D. | | .06 |

Table 4
Correlations of the Depth of Subjects A.D. and J.T. Mutually and With the True Distance for a Pose Difference of 70°

| | A.D. | J.T. |
|------|------|------|
| True | .82 | .89 |
| A.D. | | .60 |

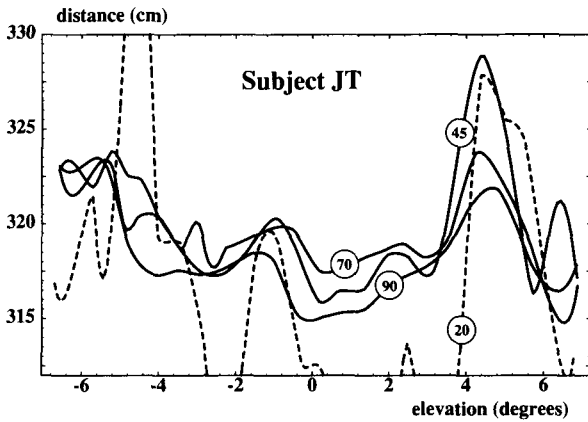


Figure 13. Distance structure for all pose angle differences (drawn curves for 90°, 70°, and 45°; dotted curve for 20°) for Subject J.T.

CONCLUSION

We have studied the perception of the depth structure of the visual contour with a simple method based on the indication of corresponding locations on geometrically identical objects seen in different poses (and, thus, differently illuminated in the same light field). The correspondence task is indeed a natural one, and it very likely could be performed by naive subjects if both locations were well within the silhouette (Koenderink et al., 1997). However, it has to be appreciated that this task is much more difficult when the fiducial point is on the contour. Yet subjects are apparently able to handle this task with at least qualitatively (and semi-quantitatively) veridical results. We find no indications that subjects treat the contour as a planar curve running in a frontoparallel plane. The rims reconstructed from their settings are true space curves with depth variations similar to the veridical distances, sometimes overestimating the range variations as they exist.

Note that this method does not purely address the problem of the perception of the depth structure of the visual contour, because the responses of the subject depend critically on the perceptual evaluation of both poses. For a single view of a silhouette, there are many possible interpretations of the rim. However, not all of them will agree with what can be perceived within the silhouette: the shading and texture gradients. However, some leeway (to put it mildly) is left. Fewer interpretations will agree with the relief (from shading, texture gradient, and contour) perceived on the other object. Note that subjects cannot use point correspondences: This would allow them (at least in principle) to solve the photogrammetric problem up to affinities. They also cannot use something akin to "photometric stereo" since the mannequins are not Lambertian, the light field is not homogeneous, and both vignetting and interreflections affect the shading pattern.

We have studied the ability of human observers to establish point correspondences on pictures of 3-D objects in different poses before (Koenderink et al., 1997), be it

only in regions away from the contour. It seems likely that the subjects use combinations of rough photogrammetry, shape from shading, and, perhaps, photometric stereo, filling in from the contour and familiarity cues. The subjects were indeed familiar with the shapes, in that they had been able to view them from various angles at their ease. However, it is easy to overestimate the importance of such familiarity cues. Even our subjects found it difficult to predict changes of contour and/or shading as the objects were slowly turned, and they were often surprised to see what happened in the visual field when they were asked to attend to various details. Whatever the case may be, the responses of the subject will have to depend on a perceived concordance between the surface landscapes on both mannequins. Of course, it is necessary that such a perceived concordance is not trivial (i.e., has topographic structure). Examples of trivial cases would be vertical cylinders or spheres: The task is evidently impossible in such cases. The subjects locate the corresponding position of a fiducial point relative to the curvature landscape in some environment that has to be large enough to include significant shape changes—hence, an appreciable part of the visible object. This is indeed intuitively clear when you image the task on a true sphere: Clearly, there is no way to figure out the correct position of the corresponding point, since it cannot be related to anything. The region needed to locate a point depends on the scale of the surface undulations. For instance, on the mannequins, it would be an easy matter to locate the corresponding point if the fiducial point were on a small significant shape feature, such as the navel, or elongated vertical structures, such as the spinal column. In actual fact, the objects are very globally rounded (much smoother than true human bodies), and significant shape variations take place over distances more like 10 cm or more.

The present task of finding the spatial structure of the rim indirectly via the correspondence task is, of course, different from conceivable more direct tasks. One could, for instance, ask the subject to identify near and far points, to discriminate depths at distinct points of the contour, or to judge the slant of the rim. Such questions appear reasonable in the sense that one expects the subject to be able

Table 5
Correlations of the Depth for Subject J.T.
for All Pose Differences

| | 45° | 70° | 90° |
|-----|-----|-----|-----|
| 20° | .65 | .57 | .49 |
| 45° | | .78 | .79 |
| 70° | | | .82 |

Table 6
Correlations of the Depth for Subject A.D.
for All Pose Differences

| | 45° | 70° | 90° |
|-----|-----|-----|-----|
| 20° | .10 | .39 | .23 |
| 45° | | .67 | .77 |
| 70° | | | .90 |

to handle such questions at least to some degree. However, there does not appear to exist material in the literature that we might draw on.

A priori, it seems likely that such different operationalizations will yield different results. This makes it highly desirable to apply a large battery of methods to this problem and look for patterns. So little is known on the perception of surfaces near the visual contour that a thorough empirical investigation is in order. The present paper has to be regarded as merely a preliminary attempt.

REFERENCES

- ARNOLD, V. I. (1984). *Catastrophe theory*. Berlin: Springer-Verlag.
- BLASCHKE, W. (1967). *Differential geometry*. New York: Chelsea.
- CIPOLLA, R., & BLAKE, A. (1990). The dynamic analysis of apparent contours. In *Proceedings of the IEEE Third International Conference on Computer Vision* (pp. 616-623). Los Alamitos, CA: IEEE Computer Society Press.
- CORTESE, J. M., & ANDERSEN, G. J. (1991). Recovery of 3-D shape from deforming contours. *Perception & Psychophysics*, **49**, 315-327.
- DA VINCI, LEONARDO (1989). *Leonardo on painting* (M. Kemp, Ed.; M. Kemp & M. Walker, Trans.). New Haven, CT: Yale University Press.
- FAUGERAS, O. (1993). *Three-dimensional computer vision: A geometric viewpoint*. Cambridge, MA: MIT Press.
- GIBLIN, P., & WEISS, R. (1987). Reconstruction of surfaces from profiles. In *Proceedings of the IEEE First International Conference on Computer Vision* (pp. 136-144). Los Alamitos, CA: IEEE Computer Society Press.
- JACOBS, T. S. (1988). *Light for the artist*. New York: Watson-Guption.
- KOENDERINK, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.
- KOENDERINK, J. J., KAPPERS, A. M. L., POLLICK, F. E., & KAWATO, M. (1997). Correspondence in pictorial space. *Perception & Psychophysics*, **59**, 813-827.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1976). The singularities of the visual mapping. *Biological Cybernetics*, **24**, 51-59.
- 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. (1984). What does the occluding contour tell us about solid shape? *Perception*, **13**, 321-330.
- MACH, E. (1959). *The analysis of sensations and the relation of the physical to the psychological* (C. M. Williams, Trans.; rev. by S. Waterlow). New York: Dover. (Original work published in German in 1886)
- MARR, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- NORMAN, J. F., & TODD, J. T. (1994). Perception of rigid motion in depth from the optical deformations of shadows and occlusion boundaries. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 343-356.
- NORMAN, J. F., TODD, J. T., & PHILLIPS, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, **57**, 629-636.
- PHILLIPS, F., TODD, J. T., KOENDERINK, J. J., & KAPPERS, A. M. L. (1995). What defines features on smoothly curved surfaces? *Investigative Ophthalmology & Visual Science*, **34**, 847.
- POLLICK, F. E., GIBLIN, P. J., RYCROFT, J., & WILSON, L. L. (1992). Human recovery of shape from profiles. *Behaviormetrika*, **19**, 65-79.

(Manuscript received January 8, 1996;
revision accepted for publication August 8, 1996.)