Perception & Psychophysics
1992, 52 (5), 487-496

Surface perception in pictures

JAN J. KOENDERINK, ANDREA J. VAN DOORN,
and ASTRID M. L. KAPPERS
Utrecht Biophysics Research Institute, Utrecht, The Netherlands

Subjects adjusted a local gauge figure such as to perceptually "fit" the apparent surfaces of objects depicted in photographs. We obtained a few hundred data points per session, covering the picture according to a uniform lattice. Settings were repeated 3 times for each of 3 subjects. Almost all of the variability resided in the slant; the relative spread in the slant was about 25% (Weber fraction). The tilt was reproduced with a typical spread of about 10°. The rank correlation of the slant settings of different observers was high, thus the slant settings of different subjects were monotonically related. The variability could be predicted from the scatter in repeated settings by the individual observers. Although repeated settings by a single observer agreed within 5%, observers did not agree on the value of the slant, even on the average. Scaling factors of a doubling in the depth dimension were encountered between different subjects. The data conformed quite well to some hypothetical fiducial global surface, the orientation of which was "probed" by the subject's local settings. The variability was completely accounted for by single-observer scatter. These conclusions are based upon an analysis of the internal structure of the local settings. We did not address the problem of veridicality, that is, conformity to some "real object."

We addressed the problem of the *internal consistency* of data structures generated via the visual inspection of pictures, in the presence of these pictures. We used photographs of existing objects, but made no further references to these objects.

Many previous investigators have used local probes to measure the internal representation of three-dimensional (3-D) surfaces from looking at pictures. For instance, Bülthoff and Mallott (1992) have measured depth maps, whereas Stevens (1983a, 1983b), Stevens and Brookes (1987), and Todd and Akerstrom (1987) have developed methods to probe the slant and tilt distribution. These authors also address the important problem of conformity to the real (3-D) object, on the basis of shape from shading, texture gradients, disparity, and so forth.

In the present study, we used a method related to that proposed by Mingolla and Todd (1986). The aim here was to test, by quantitative means, whether observers sample some coherent surface. In order to do so, it was necessary to obtain somewhat more extensive data than usual.

METHOD

Stimuli

The stimuli were photographs of rigid objects, displayed on a CRT tube measuring 640 × 480 pixels (or less), displayed at 8 bits of greytone. The objects were pieces of sculpture with clear-cut smooth shapes, as viewed under general room illumination. Presentation was on a screen (8-bit RGB monitor) attached to a Macintosh IIfx. Another screen was used for user interactions and cuing messages. The subject responded by using a mouse.

The picture used for this study depicts a marble piece ("The Bird," finished in 1912 by Constantin Brancusi, 1876–1957), which is kept in the Philadelphia Museum of Art (see Figure 1). The picture was chosen from Wittkower's (1977) book on sculpture.

Subjects

The subjects were the authors, who performed the settings without a preliminary training period. All had normal, or corrected-to-normal, binocular vision (A.K. was emmetropic, A.D. slightly myopic, and J.K. slightly presbyopic). All had had extensive experience with psychophysical experiments.

Procedure

The viewing distance was 500 mm. The full screen subtended 26° × 20° of visual angle. The picture itself was only part of the screen. (The bounding box of "The Bird" measures 140 × 356 pixels.) The pixels were spaced at 2.5′ of arc intervals. Viewing was with the right eye only, from a centered and frontal position. Maximum luminance of the screen was 10 cd/m^2. For the experiment, the room was darkened and the subject was constrained with a head- and chinrest.

In a preliminary session, the subject was told that only the smooth, polished surface above the base ("The Bird" is feetless) and the area up to the fractured-looking top surface (it has an unfinished or fragmentary look) were relevant to the experiment. The subjects were asked to move a cursor over the outline of this part and to repeat this 3 times. The subjects agreed rather precisely (rms deviation of 0.6 pixels). We then predefined a grid consisting of 225 points in a regular hexagonal lattice on the area defined by the common outline. All the subjects performed settings based upon this same lattice, thus enabling intersubject consistency checks.

The subject was to adjust the shape of a "gauge figure," used as a probe, which appeared overlaid in red over the greytone picture. The gauge figure was the orthographic projection of a thin

The authors are indebted to the SPIN 3-D Computer Vision Program of the Dutch Ministry of Economic Affairs and to the InSight project of the Basic Research Actions of the ESPRIT program of the EC. Address correspondence to J. J. Koenderink, Utrecht Biophysics Research Institute, Buys Ballot Laboratory, Princetonplein 5, 3584 CC Utrecht, The Netherlands.

Figure 1. Monochrome halftone photograph of the piece of sculpture "The Bird," finished in 1912 by Constantin Brancusi (1876–1957), from the Philadelphia Museum of Art. Only part of this picture is used in the experiment, showing all of the smooth continuous surface of "The Bird," but only part of the "feet." The coordinate origin is taken at the lower left, with the *x*-axis running horizontally toward the right, the *y*-axis vertically upward. We associate a *z*-axis ("depth dimension") with the *x*- and *y*-axes: It is orthogonal to both the *x*- and *y*-axes, whereas the *xyz* system is lefthandedly oriented.

circular disk, pierced orthogonally through the center with a straight-line-shaped axle. The axle protruded a distance equal to one disk radius from both sides of the disk. The radius of the disk measured 75′ of arc (see Figure 2). The subject controlled the slant and tilt of the projection of the gauge figure with the mouse. The instructions were to adjust the gauge figure in such a way that the disk *looked tangent to the surface* of the depicted object. The question asked was: "Could this be a red circle *painted upon the surface*?" If the configuration "looked right," the mouse button was pressed by the subject and the gauge figure disappeared. The task was an easy one: Many naive subjects have tried it and, even on first try, did as well as our experienced subjects. The subject was allowed unrestricted time for the setting, but typically performed the task well within 10 sec. The gauge figure was centered on one of the fiducial locations. In the course of the experiment, all lattice points were visited once in random order. In a number of different sessions, this routine was repeated.

RESULTS

Experimental Paradigm

The settings were converted to slant and tilt angle. The slant measured the degree of turnout of the frontoparallel position and was specified as either an angle in the range of 0° (frontoparallel orientation) to 90° (seen "edge on"), or as the tangent of this angle. It will be clear from the context which interpretation is being used at any time. The tilt measured the orientation of the slant (compass direction) and was specified as an angle in the range of 0° to 360° from a reference direction (e.g., left to right direction). Alternatively, the setting was converted into a *depth gradient*, which is a vectorial quantity. The modulus of the gradient equals the tangent of the slant, whereas the direction of the gradient specifies the tilt in a coordinate independent manner.

The results of a typical session are depicted in Figure 3. For reasons of clarity, this figure depicts the results of a pilot run with fewer settings than those used in the actual experiment, in which a triangulation with 225 vertices was used.

In the paradigm, the settings were interpreted as the orientation of hypothetical surface elements of some fiducial surface "perceived" by the subject on the basis of optical data provided by the picture. The data structure produced by the subject is a sampled depth-gradient field, which is described by 225 two-dimensional vectors. Thus, the data structure has 3 × 450 degrees of freedom and is completely specified by 1,350 numbers.

The internal-consistency check was simply a check of whether the data structure represented a possible sampled depth-gradient field.

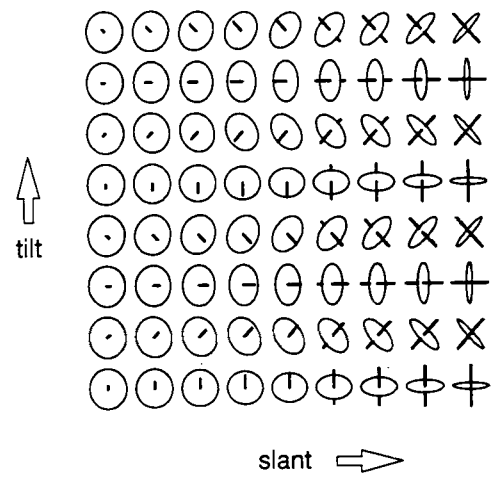

Figure 2. A collection of gauge figures used in the experiment. (In reality, the apparent orientation is continuously variable.) Rows depict constant tilt, columns depict constant slant. This figure roughly illustrates the response parameter space.

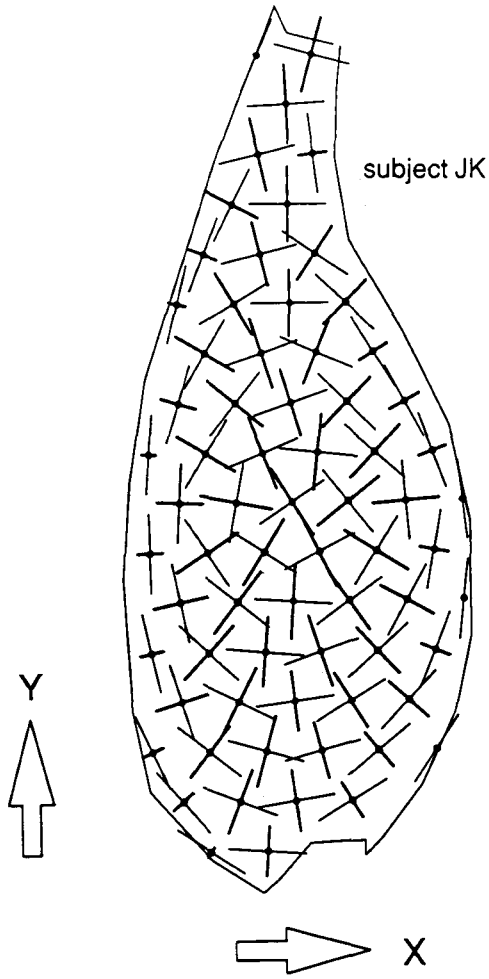

Figure 3. Result of a session (Subject J.K.). For the sake of clarity, this figure applies to a triangulation with less than one third of the number of vertices used in the sessions reported in this paper. The tilt direction is indicated by the thin lines of fixed length, the slant by the thick lines of varying length. The latter lines are projections of the axle of the gauge figure, thus they become shorter for larger slants. The outline may be compared to the photograph in Figure 1. It encloses all of the visible smooth surface, but excludes the "feet," the pedestal, and the roughened top area.

Internal Consistency

A local depth gradient specifies the spatial orientation of a local surface element (its slant and tilt), a so-called "contact element" (Burke, 1985). We introduce Cartesian coordinates (x,y) in the picture plane. Let z denote a third dimension ("depth"), then a function $z(x,y)$ specifies a "surface," at least in a formal sense. At any point (x,y) of the picture, we can imagine many different contact elements $(\partial z/\partial x, \partial z/\partial y)$. All of these contact elements exist in the "contact bundle" (Burke, 1985), which is the Cartesian product of picture space and orientation space. The contact bundle is an abstract four-dimensional space with coordinates $(x,y,\partial z/\partial x,\partial z/\partial y)$.

A *surface* $z(x,y)$ naturally induces a field of contact elements, namely, one contact element $(\partial z/\partial x,\partial z/\partial y)$ at every point (x,y) of the picture. Such a field is referred to as a "section of the contact bundle" (Burke, 1985). Our empirical data are a (sampled) section of the contact bundle. However, not every section of the contact bundle specifies a surface: In order for an arbitrary vector field to represent a physically realizable depth-gradient field, the field of contact elements has to be *integrable*.

An intuitive explanation of "integrability" is as follows: If one sews local contact elements together to form a "quilt" (moving them freely in depth as required), then one should be able to piece together a *surface*. That this is not necessarily possible for an arbitrary field of contact elements can be understood from the following reasoning: Consider a chain of contact elements, that is, a string of contact elements located on a closed curve in the projection. If these are pieced together, then one should be able to *close the chain* in 3-D. If it turns out that the initial and final elements are at *different depths*, then the chain cannot be embedded in any surface (see Figure 4). One *must* require that arbitrary chains can be closed; otherwise, the vector field does not allow a consistent interpretation as a depth-gradient field.

In mathematical terms, this can be framed on the local level. The vorticity of a vector field vanishes identically if the vector field allows interpretation as a depth-gradient field; in that case, there can be no "eddies." Nonvanishing curl indicates inconsistency, and the root mean square of the curl is a convenient numerical measure of this inconsistency.

That the curl has to vanish is evident from this simple reasoning: Let the function $z(x,y)$ denote the depth at a location (x,y) in the picture. Then the depth-gradient field is composed of the two-dimensional vectors $[z_x(x,y),\ z_y(x,y)]$, where the subscripted coefficients denote partial derivatives with respect to the coordinates. The curl of this field is given by the expression $\partial z_x(x,y)/\partial y - \partial z_y(x,y)/\partial x$, which again equals $z_{xy}(x,y) - z_{yx}(x,y)$. Vanishing of the curl thus implies equality of mixed partial derivatives. This again guarantees that an integral manifold, that is, a surface $[x,y,z(x,y)]$ exists.

Because the data are specified on a discrete lattice of points, we had to reformulate this a bit. Consider two adjacent fiducial points. The contact elements will be at different depths. The least depth difference that allows us to sew them together occurs when the seam bisects the connecting segment. Then the depth difference equals the scalar product of the *average* depth gradient and the connection vector. In this way, we can easily find the depth differences implied for every pair of adjacent fiducial points.

Consider three pairwise adjacent fiducial points, A, B, and C. Let the depth differences be denoted z_{AB}, z_{BC}, and z_{CA}. Consistency requires that one has $z_{AB}+z_{BC}+z_{CA} = 0$. This condition exactly guarantees that the chain of three links closes. Only if this condition pertains, may one assign three depths z_A, z_B, and z_C (up to a common addi-

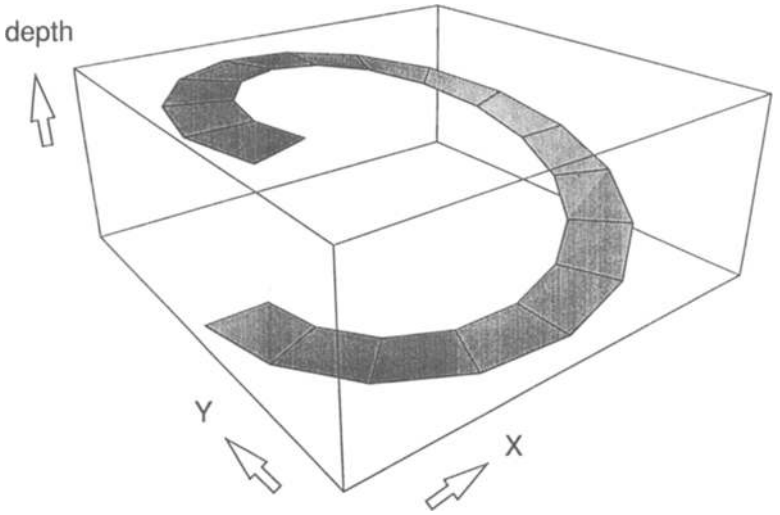

Figure 4. Graphic illustration of the condition of nonclosure of surface elements along a closed curve in the projection: The surface elements at the initial and final points of the curve (identical in the projection!) are at different depths. (This depth difference is a measure of the constraint violation for the loop.) Such a chain of contact elements cannot be found on a smooth surface.

tive constant) such that $z_{AB} = z_A - z_B$ and so on. Otherwise, one can't find a consistent triplet of depths and the specification of the triplet (z_{AB}, z_{BC}, z_{CA}) doesn't permit an interpretation in terms of a configuration of points on a surface.

In general, one cannot expect to be able to assign a set of depth values to the fiducial points in such a way that the depth differences for all adjacent pairs are exactly reproduced. Instead, we proceeded to find the unique set of depth values that reproduced the depth differences in the least squares sense. (Of course, the absolute depth must remain undetermined.) The remaining root-mean square deviation per fiducial point was a convenient measure of the degree of *inconsistency*. The analysis is straightforward. We analyzed the data with the Mathematica (Wolfram, 1988) package. The fit was found immediately via the singular value decomposition of the set of linear equations for the depth differences. (If z_i denotes the ith vertex, and δz_{ij} the estimated depth difference for the vertex pair i,j, then the equations are $z_i - z_j = \partial z_{ij}$.) We arbitrarily set the average depth to zero in order to fix the scale, and thereby resolved the remaining ambiguity.

Surfaces Defined in the Experiment

The surfaces computed from the empirical data are the best fits in the sense of the least squares deviation from the depth differences based upon pairs of observed orientations. The procedure defines surface perception for a certain class of pictures in an operational sense.

The surfaces obtained in this manner had to be interpreted with due caution. For instance, were the computed depth values to be referred to a polar coordinate system centered on the subject's entrance pupil (or first nodal point), or to a Cartesian system attached to the screen? What about the focal length of the camera that took the photograph in the first place? It isn't easy to resolve such questions, nor is it a priori clear that such a resolution is at all possible. We note that the problem is *irrelevant* as long as the question of veridicality is not presented. In this paper, we (arbitrarily) refer the depth values to a Cartesian coordinate system attached to the screen.

An initial study should investigate the nature of the empirically determined surfaces per se. The outcome might be trivial (e.g., planar, frontoparallel patches, or spherical shells centered on the entrance pupil of the subject's eye). Subsequent studies need to investigate these surfaces in terms of dependence on the structure of the pictures and, finally, in terms of the depicted objects. In this paper, we restrict the discussion to the issue of internal consistency.

Intrasubject Variability

Repeated observations by a single observer correlated well. Figure 5 shows a scatterplot that contains data from all points for two repeated experiments by Subject J.K. The gradient magnitudes are plotted against each other. The data are proportional with factors differing a few percent from unity. These scale differences were small, but sometimes statistically significant.

In order to quantify these scale differences, we studied the distribution of the logarithms of the ratios of the magnitude of the gradient to the average value for all sessions. This variable turned out to be almost normally distributed. A t interval for the mean at 95% confidence did not always include the value zero. For the 3 subjects and three

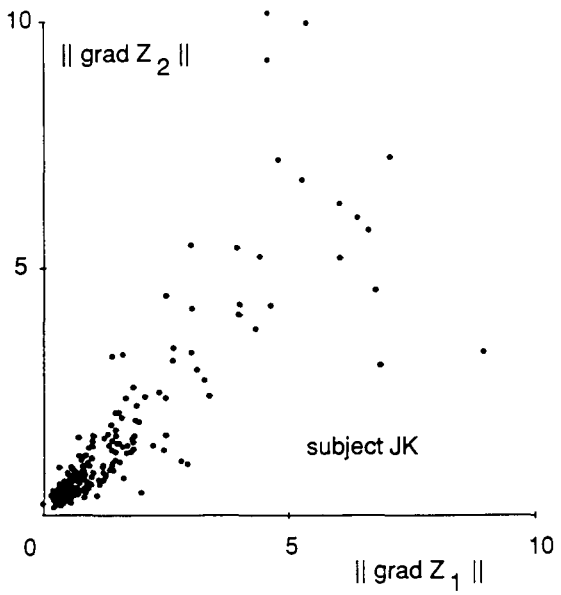

Figure 5. Scatterplot of the gradient magnitudes (that is, $\|\nabla z(x,y)\|$) for the settings of two different sessions for a single subject (J.K.). Such a plot offers a realistic insight into the variability of repeated settings.

runs, we obtained a nonzero value four times out of nine. Corresponding deviations of the depth scaling amounted to zero (five times), 4% (twice) and 10% (twice) from unity. We concluded from this that apparently random deviations of the order of 5% were the rule. Apart from these scale differences, the reproducibility was satisfactory.

In Figure 6, the estimated variance as a function of the magnitude of the gradient for binned data (gradient magnitude range divided into bins with an equal number of cases per bin) is shown for Subject J.K. A Weber law fits the data (linear regression yields a fraction of 24%). The same is true for the other observers, except for the fact that Subject A.K., especially, had outliers that corresponded to vertices at the contour (a few percent of the vertices). Such outliers spoil a linear regression analysis. A robust estimate for the relative spread is the median of the logarithms of relative spreads per vertex. (For Subjects A.D. and J.K., the difference with a linear regression was slight; for Subject A.K., the robust estimate was similar to a linear regression after discarding the vertices with slant in excess of 4 depth pixels per picture-plane pixel.) We obtained the following estimates: Weber fractions were 29% (Subject A.D.), 26% (Subject A.K.), and 17% (Subject J.K.). We concluded that Weber fractions of about 1/4 were the rule.

In order to study the nature of the deviations, we computed the average gradient for all points and then plotted the deviations in single settings *parallel* to the average gradient direction against the components *orthogonal* to it. (see Figure 7, Subject J.K.). The resulting cloud was close to normally distributed with principal axes along the coordinate directions. A principal components analysis revealed that the ratio of the eigenvalues were 15 (Subject A.D.), 19 (Subject A.K.), and 31 (Subject J.K.). Clearly, the variance in the slant dominates; the tilt is rather more precise if the slant is not too small. (When the slant vanishes, the tilt is obviously undetermined.) The spread in the slant alone took care of 94%-97% of the total variance.

We found that the scatter in the tilt direction was about 10° for average slant values (e.g., unit slant). Thus, the tilt was comparatively well defined, whereas the slant was rather uncertain.

Consistency of the Data

In order to judge the consistency of the data, we performed the following analysis for every face (triangle) of the triangulation:

1. For every edge, we found the average gradient for the data on the vertices, then converted it into the depth difference over the edge by taking the scalar product of the average gradient with the (directed) edge vector. This depth difference is expressed in terms of pixels.
2. The sum of the (signed) depth differences over the boundary of the face should add up to zero. Instead, we may expect a finite *mismatch*, or violation of surface consistency. The constraint violation is expressed in terms of pixels.
3. We computed the average gradient of the three vertices of the face, then multiplied the modulus of this gradient with the edge length in order to obtain a measure of the depth variation over the face. This depth variation is again expressed in terms of pixels.

Figure 8 shows the results obtained in this manner for Subject A.D. The average amount of violation was zero (95% confidence t interval for the mean was $(-0.297, 0.342)$), whereas the spread increased with increasing depth variation. A linear regression without constant term on the absolute value of the violation versus the depth vari-

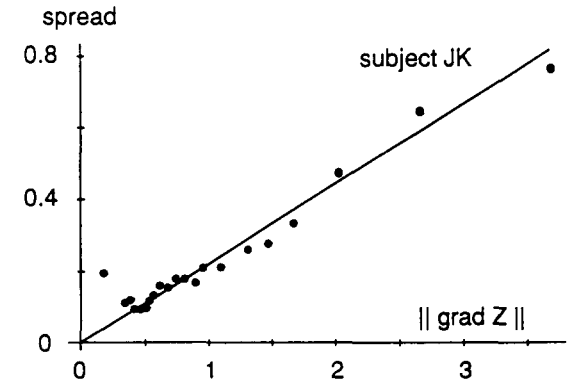

Figure 6. The spread of repeated settings by a single subject (J.K.) as a function of the average gradient magnitude. For the sake of clarity, the data points have been coarse-grained through binning of the magnitude scale. The line is the linear regression without constant term.

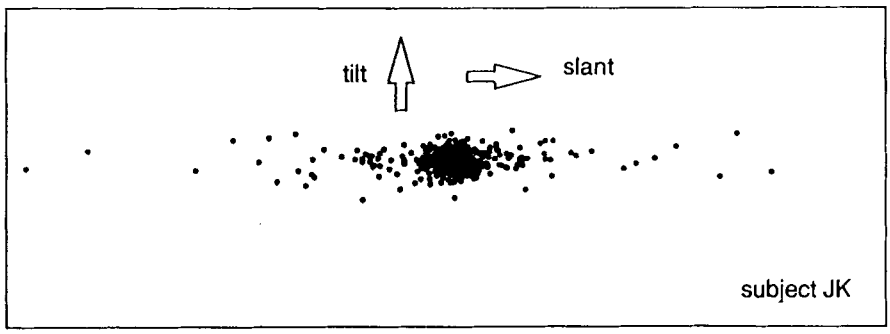

Figure 7. Scatterplot of the components of the deviation from the mean gradient in repeated settings by a single subject (J.K.) in the average gradient direction (horizontal axis) and orthogonally to it (vertical axis). The scales are identical for both axes. Thus, the elongated impression of the point cloud indicates the anisotropy of the settings realistically. The slant (gradient magnitude) is much less precisely reproduced than the tilt (gradient direction).

ation yielded a Weber fraction of 10.0% ± 0.5%. A more stable measure, the median of the relative modulus of the violation, yielded a Weber fraction of 9.0%. The same observation applies to the data of the 2 other subjects. The data are well described with a Weber's law, with Weber fractions of 9.0% (Subject A.D.), 12.5% (Subject A.K.), and 8.2% (Subject J.K.).

We did not derive the expected value of the violation on the basis of the variances in repeated settings, taking the anisotropic distribution into account analytically. Instead, we found this value through a Monte Carlo simulation. This has the advantage that the sensitivity to variation in the assumptions can easily be studied. The value used here was based on the following assumptions: (1) the gradient is uniform over a face, (2) the corresponding surface normal is distributed uniformly over all directions, and (3) the variance is due to a relative spread of 25% in the slant direction. The amount of constraint violation is conveniently expressed as a Weber fraction. We obtained an estimate of this Weber fraction of 7%; this value rises to about 9% if the variance is assumed to be isotropic, and rises only insignificantly if the gradient field is assumed to vary over the face.

The default hypothesis was that the constraint violation is explained through the spread in repeated settings alone. Then the empirically determined fields of contact elements are integrable within the experimental tolerance. We found that for none of the 3 subjects can we reject this hypothesis at the 95% level.

We conclude that the constraint violation for a face of the triangulation is about 10% of the total depth variation over the face. This inconsistency is primarily due to uncertainties in the slant settings and is expected from the variability of repeated settings by a single observer (see above).

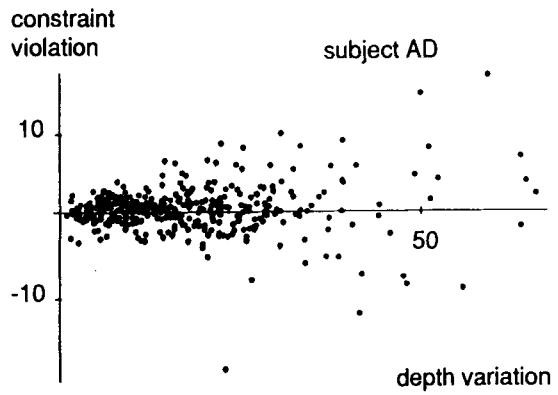

Figure 8. Scatterplot of the closure violations as a function of the depth variation over a face of the triangulation for a single setting for Subject A.D.

Nature of the Best-Fitting Surface

The surfaces produced by the subjects were smooth ones; elliptical (in case of "The Bird," only convex) and hyperbolical patches can be discerned. The surfaces were articulated in *depth*, that is, the curvatures, for instance, in the sagittal plane are comparable to the curvatures (as apparent from the contour) in the frontoparallel plane. There appears to be a parabolic curve at the edge of the "neck" of "The Bird," neatly meeting the contour at its inflection points. Indeed, the data allow a variety of differential geometrical properties of the surfaces to be computed.

In Figure 9, we show a frontal view of the triangulated surface *produced* by Subject A.D. This figure is trivial in the sense that the frontal views for all the subjects were *identical* and did not depend on the actual settings. We only include the figure because it allows a good impression of the triangulation used in the experiment. Although the probe only appeared on the vertices, the subjects were never directly *confronted* with the triangulation and only saw the picture.

That the surfaces produced in this way are 3-D entities is brought out in *profile* views. In a side view, we obtain a contour that is not unlike the contours in the frontal view. In Figure 10, we present such a profile view for Sub-

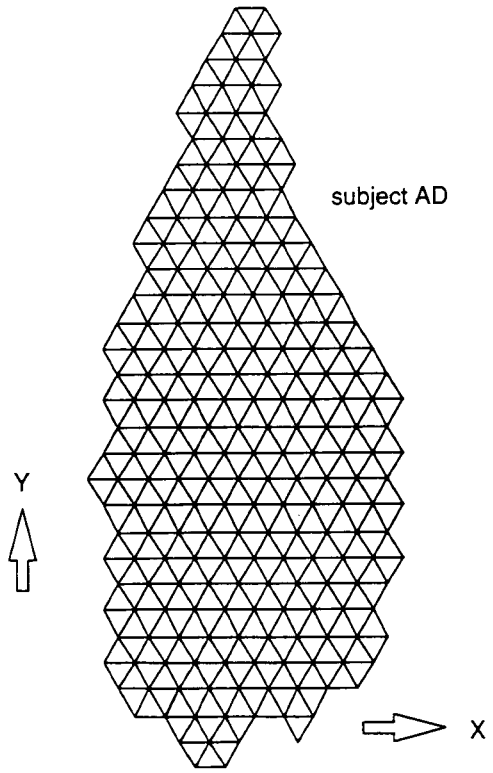

Figure 9. Frontal view of the fiducial surface by Subject A.D. Note that this view was identical for all the subjects and doesn't even depend on the settings, because it is the view that was actually presented. (Of course, the subjects never saw the triangulation, as in this figure, but looked instead at the picture of Brancusi's "The Bird." However, the probe appears only on the vertices of this triangulation.)

ject A.D. In this view, the depth values implied by the data are apparent.

In Figure 11, we present the profile view of the surface produced by Observer A.K. The surfaces produced by Observers A.D. and A.K. are very similar; the main differences occur very close to the contour. At the contour, the gradient magnitude is theoretically infinite, which is why any small difference is strongly magnified. This affects only a few percent of the data. On the interior (almost all points), the surfaces are nearly identical up to a depth scaling.

In horizontal cross section, the "perceived" photograph turned out to be close to a circle for Subject J.K., was somewhat more articulated for Subject A.D., and was somewhat flatter for Subject A.K. Thus, the percepts are close to surfaces of revolution. In Figure 12 (Subject A.D.) and Figure 13 (Subject A.K.), we depict the extreme cases (the data for Subject J.K. were in-between) of a view from below.

It is a priori very likely that the contour had an appreciable influence on the shape of these surfaces. Whether the subjects would have produced the same or similar surfaces on the basis of a bare outline drawing remains to be studied.

All 3 subjects agreed upon the vertices at which the slant vanished (and thus the tilt becomes indeterminate). Such points play an important role in the *ordinal* depth structure (Todd & Reichel, 1989).

The subjects detected two singular points in the field of contact elements on "The Bird"—one at the center of the convex "belly," and the other at the narrow "neck" at the top (see Figure 14). These singularities are of different types (opposite topological index), which becomes apparent when one studies the structure of the field of contact elements in the immediate neighborhood of these points. At one point, the surface is a local depth minimum, at the other, a saddle of the depth map.

Intersubject Agreement

The internal consistency only quantifies how well the data structure produced by the subject represents a *surface*. The issue of veridicality, that is, how well such a surface actually fits the object depicted, is not addressed here. In this section, we examine to what degree the 3 subjects produced the *same* surface. If the various surfaces agree within a narrow tolerance, then intersubject agreement is counted as high, even though such a high degree of agreement need not imply veridicality.

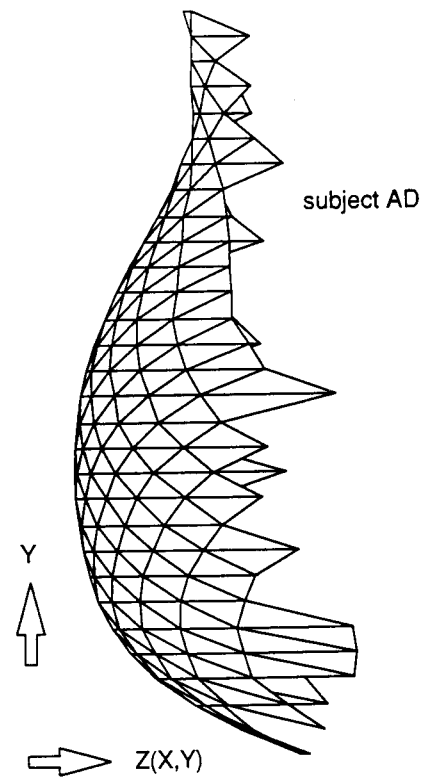

Figure 10. A profile view of the surface for Subject A.D. Note that the depth dimension is immediately evident in this figure.

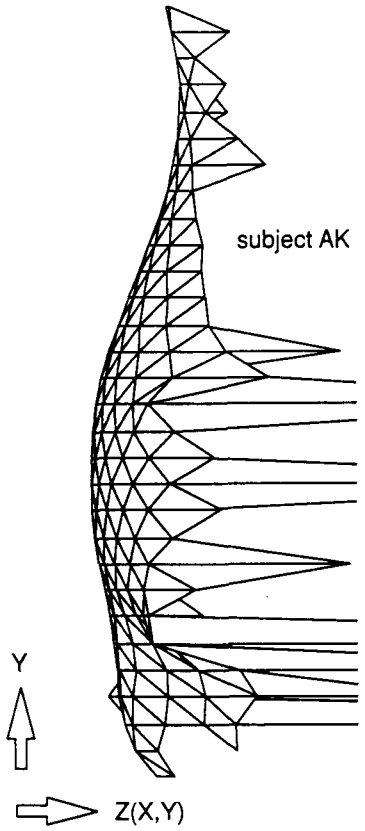

Figure 11. A profile view of the surface by Subject A.K. Note that this surface is much flatter than that for Subject A.D. (Figure 10). Notice also that at the very contour, the settings for Subject A.K. were actually steeper than those for Subject A.D. (Hence the "shoots" sprouting from the boundary, which have been clipped for this figure.)

From scatterplots of gradient magnitudes for all points of the triangulation for pairs of observers, it turns out that there is invariably a nearly linear dependence, except for a few percent of the points, particularly those belonging to the vicinity of the contour. For this analysis, we omit datapoints with a slant in excess of 4 depth pixels per picture-plane pixel (about 7% of the points). Fitting the regression line through the origin reveals a factor of almost 2 between the data of Subjects A.D. and A.K. Figure 15 shows a histogram of the logarithms of the ratios of gradient magnitudes for these observers. Notice that the histogram is offset with respect to zero. Statistical analysis confirms this impression: A t interval at 95% confidence for the mean yields values of the ratio within the range 1.73-1.95. Subject J.K. has a ratio of 0.70-0.79 with respect to Subject A.D., and a ratio of 1.24-1.44 with respect to Subject A.K. (again, t intervals at 95% confidence for the mean).

Apparently, only the depth of relief was different for different observers, whereas the *shapes* were very similar. The scatter is as expected from the scatter already found in repeated settings by a single observer. The correlation between the intersubject settings is high, no matter how it is expressed. The Pearson product-moment correlation and Spearman's rank correlation (rho) for the three intersubject correlations were, respectively: 0.844 and 0.868 for Subject Pair A.D.-A.K., 0.792 and 0.784 for Subject Pair A.K.-J.K., and 0.819 and 0.870 for Subject Pair J.K.-A.D.

One way to state this is to say that the linear regression between Subjects A.D. and A.K. explains 71% (Pearson's product-moment correlation squared) of the variability in each. The outliers in the data (due to points on the boundary) are less troublesome in Spearman's ρ, which is simply the correlation of the *ranks*. Clearly, the concordance in the *order* of the data is very good. Thus, the monotonic relationship between the settings of different subjects is firmly established.

The 80% difference in depth for the surfaces produced by Observers A.D. and A.K. is clearly visible in the profile views.

In order to investigate whether the deviations are isotropic or whether slant and tilt dimensions show up differently, we repeated the analysis in terms of devia-

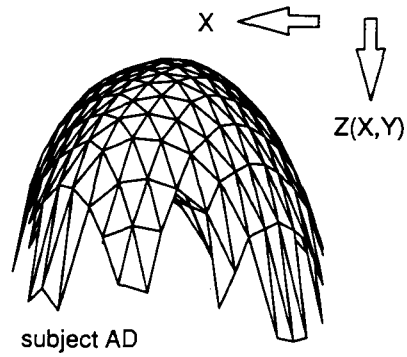

Figure 12. Bottom view of the surface for Subject A.D. This allows a fair appraisal of the fact that the surface is quite close to a surface of revolution. (Of course, only one side is visible.)

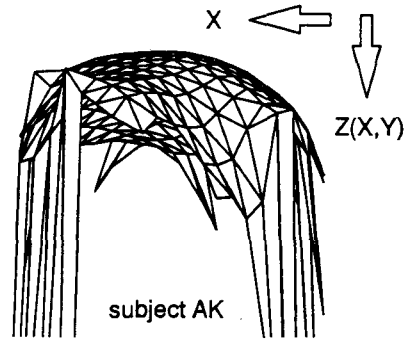

Figure 13. Bottom view of the fiducial surface for Subject A.K. This allows a fair appraisal of the fact that the surface is close to a surface of revolution. (Of course, only one side is visible.) Note that the surface is flatter than that for Subject A.D.

tion components in the direction of the average gradient (this time an average over observers) and orthogonally to it. Again, the scatterplots revealed a very anisotropic distribution—almost all of the variability was due to slant differences. Of course, this is to be expected from the variability already found in single-observer settings.

CONCLUSION

As mentioned, in this study we used a method related to that proposed by Mingolla and Todd (1986). These authors had their subjects judge slant and tilt angles for 13 points on pictures of surfaces (sequentially), indicated by a small superimposed cross. Because they used very constrained surfaces (triaxial ellipsoids), they were able to reconstruct surfaces from the responses. The minor difference used in our study is that we required a judgment of perceptual conformity, not an absolute judgment of angles. Absolute judgments are very difficult to make: Mingolla and Todd report that trained subjects find the task difficult and often take over 1 min to respond. In contrast, our subjects decided almost momentarily whether the gauge figure "fits" or not. Usually, half a dozen settings were tried in about 10 sec before the final decision was reached, and the subjects found this task to be an easy one. Naive subjects new to the task don't score differently or take a longer time than our subjects did. An advantage here was that we could sample somewhat more finely.

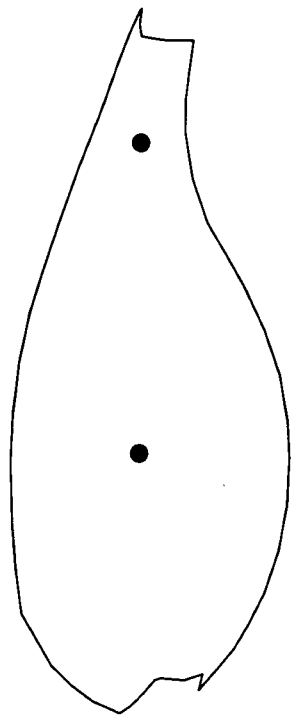

Figure 14. Outline of Brancusi's "The Bird" with the vertices that appear as singularities of the field of contact elements for all 3 subjects.

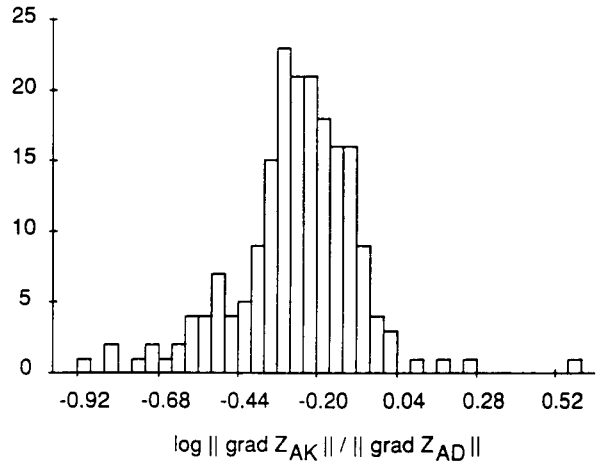

Figure 15. Histogram of the log ratio of gradient magnitudes for Subjects A.D. and A.K. Note that the histogram is offset from zero, indicating a scaling different from the identity. Indeed, a *t* interval for the mean with 95% confidence is −0.29 to −0.23. Thus, the ratio of gradient magnitudes is about 1.8 and is very significantly different from unity.

We have used this method to study the problem of the internal consistency of the data. Mathematically, a gradient field implies that the curl of the field vanishes identically at all points. We found that the empirical violations of this constraint can fully be accounted for through the scatter in repeated settings at single points. With respect to repeatability, the data were consistent with the notion of a sampled smooth surface. (Although this is common knowledge, we know of no prior quantitative checks.)

Stevens (1983a) has general theoretical reasons to expect slant estimates to be less precise than tilt estimates. In our (very specific) case, we found that, although slant and tilt estimates differed appreciably in their variability over sessions, the slant direction was invariably the least precise. The errors we found were smaller than those reported by Mingolla and Todd (1986). Unfortunately, these authors do not report differences in accuracy of tilt and slant judgments. The differences between the paradigms are sufficiently large that such inconsistencies are not surprising.

Different subjects tended to agree in their tilt estimates; the spread is accounted for by the variability already present in single-observer settings. A significant *scale difference* exists in the gradient settings by different observers. Such scaling factors can be surprisingly large; in one case, we found a factor of almost 2. These scale differences are not explained by single-observer scatter. Apparently, the observers agreed only up to some idiosyncratic depth scaling. Such depth scalings also occurred in repeated sessions for single observers, but were much less pronounced (typically 5%). Apparently, people differ

appreciably in their depth scales. It seems a priori likely to us that perceived object contours have an effect that spreads well within their boundaries, an issue that deserves further careful study.

In our present study, the "percept" of the surface is defined in terms of the perceptually best-fitting gauge figure. When different observers disagree (as ours did), this may, of course, be due to an idiosyncratic assessment of the gauge figure, as well as to an idiosyncratic assessment of the surface. The method itself cannot resolve this. (In fact, in the present paradigm, the "problem" was meaningless.) Arguments need to be derived from comparative studies using a spectrum of different response tasks.

Different operational definitions of what will be the perceptual outcome of experiments addressing surface shape (e.g., depth maps, gradient fields, curvature fields, etc.) will, in all likeliness, yield incompatible results (e.g., in the mathematical sense that the gradient should equal the spatial derivative of the distance). Quantitative studies will become possible when detailed depth maps, gradient fields, and so forth, become available. At the moment, we consider it premature to compare our results to those of Bülthoff and Mallott (1992).

REFERENCES

BÜLTHOFF, H. H., & MALLOTT, H. A. (1992). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America*, **A5**, 1749-1758.

BURKE, W. L. (1985). *Applied differential geometry*. Cambridge: Cambridge University Press.

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

STEVENS, K. A. (1983a). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, **46**, 183-195.

STEVENS, K. A. (1983b). Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics*, **33**, 241-250.

STEVENS, K. A., & BROOKES, A. (1987). Probing depth in monocular images. *Biological Cybernetics*, **56**, 355-366.

TODD, J. T., & AKERSTROM, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 242-255.

TODD, J. T., & REICHEL, F. D. (1989). Ordinal structure in the visual perception and cognition of smooth surfaces. *Psychological Review*, **96**, 643-657.

WITTKOWER, R. (1977). *Sculpture, processes and principles*. New York: Harper & Row.

WOLFRAM, W. (1988). Mathematica, a system for doing mathematics by computer. Redwood City, CA: Addison-Wesley.

(Manuscript received August 26, 1991;
revision accepted for publication May 18, 1992.)