Correspondence in pictorial space

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

and

FRANK E. POLLICK and MITSUO KAWATO
ATR Human Information Processing Research Laboratories, Kyoto, Japan

We have investigated psychophysically determined image correspondences between pairs of photographs of a single three-dimensional (3-D) object in various poses. These correspondences were obtained by presenting the pictures simultaneously, side by side, and letting the subject match a marker in one picture with a marker (under manual control) in the other picture. Between poses, the object was rotated about a fixed vertical axis; thus, the shifts of the veridical correspondences (with respect to the surface of the object) were very nearly horizontal. In fact, the subjects produced appreciable scatter in both horizontal and vertical directions. The scatter in repeated sessions and between data depends on the local (landmarks) and global (interpolation) structure of the pictures. Since the object was fairly smooth (white semigloss finish) and nontextured, the only way to establish the correspondence is by way of the "pictorial relief." The relief is some largely unknown function of the image structure and the observer. Apparently, more immediate entities (e.g., the shading or the contour) cannot be used as such, since they vary with the pose. We compare these data with results obtained with a surface attitude probe on a single picture. We studied various measures of consistency both within a single method and between methods. We found that subjects were confident in establishing correspondences, but results scattered appreciably in a way that depended on both global and local image structure. Correspondence results for various pose angles were mutually very consistent, but only to a minor extent with results of attitude measurements. The main finding was that subjects could establish correspondence on the basis of their 3-D interpretation (pictorial relief), even if the 2-D graytone distributions are quite different.

Pictorial space is, at first glance, very similar to the space we move in, at least in some aspects. For instance, it appears to be extended in three independent directions. Pictorial space often contains opaque pictorial surfaces that obstruct our vision into the depth dimension of pictorial space, much like the surfaces of (typically) opaque objects in our field of view. Such pictorial surfaces have a two-dimensional (2-D) structure that can be probed with simple psychophysical methods and that appears as well defined and consistent (Koenderink, van Doorn, & Kappers, 1992, 1994). However, unlike the space we move in (which is isotropic, homogeneous, and Euclidean for most practical purposes and appears to possess a structure that is largely independent of ourselves), pictorial space is anisotropic, and unhomogeneous, and its structure is highly volatile in the sense that it depends critically upon the observer and the observer's relation to the picture (Hildebrand, 1893/1945; Koenderink et al., 1994). Viewing conditions, such as monocular, binocular, distance,

and attitude to the picture plane, critically affect the depth dimension, which is, in many respects, distinct from the dimensions that span the visual field or the frontoparallel plane. Interindividual concordance is more variable for configurations in pictorial space than for configurations in our mutual physical environment. For instance, large differences in the depth scale are routinely encountered.

Whereas a physical object remains essentially unchanged when we alter its position and/or attitude in physical space, such invariance becomes questionable and an issue of empirical study for pictorial objects in pictorial space. Of course, "the space we move in" exists perceptually only if we actually *move* in it (Mach, 1886/1959)— that is, for example, we can walk around the object or manipulate it in our hands. If we sit down and generally enjoy the scene in front of us from a single vantage point, the environment appears much like pictorial space (Gibson, 1950). Any movement relative to the scene induces dynamic changes in perspective that are rich carriers of information concerning the three-dimensional structure of the scene— that is, if you know how to use this optically specified structure. Human observers use this information; to humans, this type of optical structure is information concerning spatiality. This pertains both to the active, locomoting observer and to the passive observer watching TV or movies.

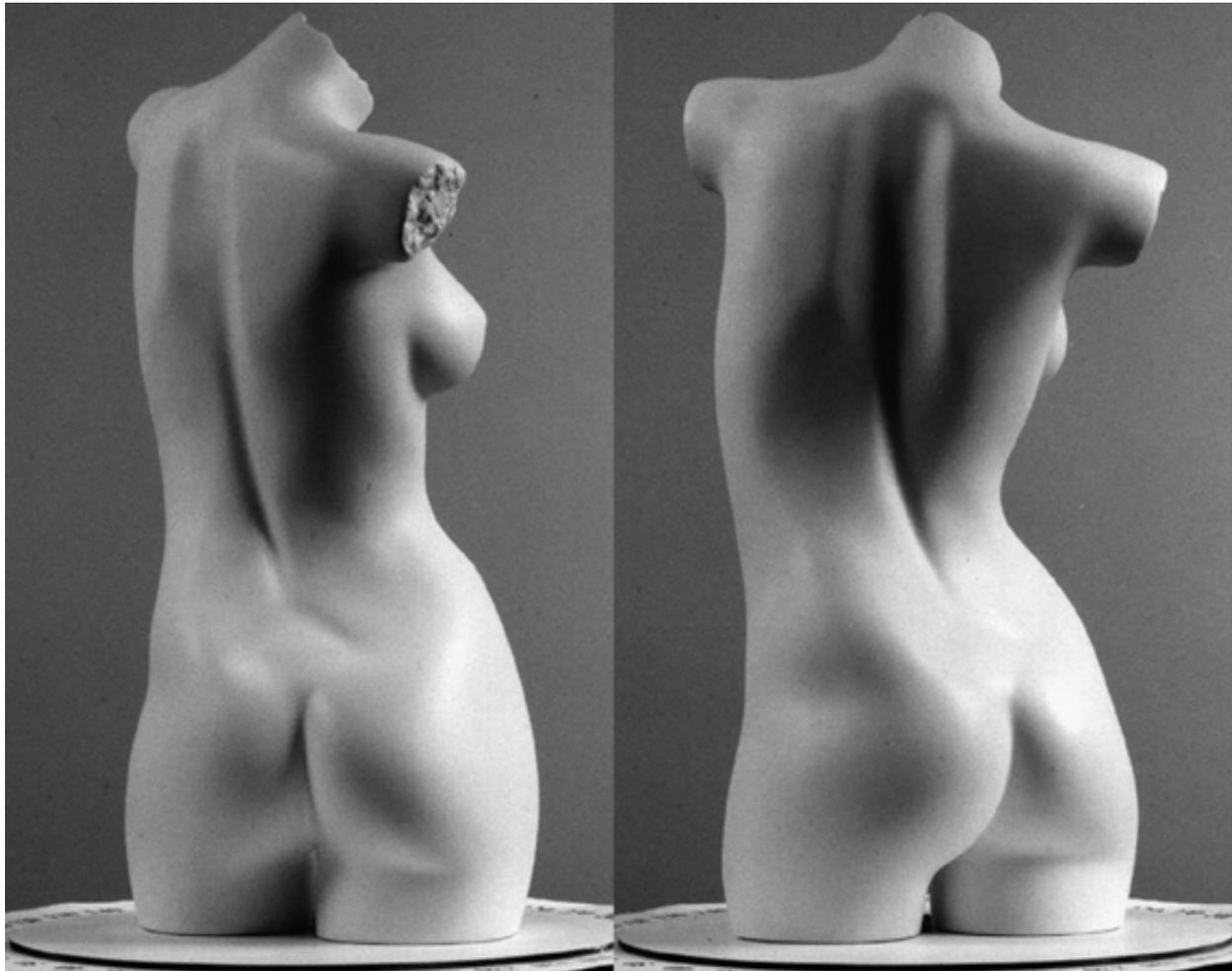One important issue that distinguishes such continuous temporal changes from mere multiple perspectives is

Figure 1. Examples of stimuli. Left, the fiducial (0º) pose. Right, the 45º pose.

that, in the former case, the "image correspondences" are much more apparent than in the latter case. A correspondence between two pictures is defined as a mapping of the picture elements from one picture to those of the other in such a way that the elements mapped on each other are perceived as the images of identical structural elements in pictorial space. If the picture changes continuously, the correspondence changes continuously (for small temporal differences, the correspondence regresses to the identity), and the problem is much easier—though by no means trivial—than in the case of discrete changes. When two photographs of a face are presented in different poses, it is clear enough that a correspondence exists (e.g., the nose, eyes, and mouth can be immediately pointed out in the two pictures), but it is much harder to establish the correspondence on a point-by-point basis.

It should be noticed that there exists no unique way to define correspondences between pictures; a correspondence is necessarily operationally defined. One may object that, in cases where the history of the process by which the pictures were produced is available, we may possess a "veridical" correspondence. However, this type of corre-

spondence is defined by both the pictures and the history of their production and not by the set of pictures as such. Thus, even in this case, the pictures cannot be said to possess a unique correspondence. This is indeed intuitively evident in extreme cases. For example, think of two pictures of a featureless white planar object: The pictures are all uniformly gray, and the veridical correspondences are defined only by the history of production and are not inherent in the pictures at all. In our case, the correspondence is defined via the matches made by a human observer.

When a rigid object moves in Euclidean space, the problem of correspondence is known to be reduced enormously: There is no need to establish the correspondence on a point-to-point basis. The correspondence is simply a Euclidean isometry that can be characterized by a mere 6 degrees of freedom (a rotation about some axis and a translation according to Hamilton-Rodriguez's theorem; Whittaker, 1944). This is quite different in pictorial space, because, in pictorial space, "isometries" are meaningless.

The problem becomes conceptually even more challenging when we consider two photographs of geometrically congruent objects (i.e., two identical copies of a 3-D
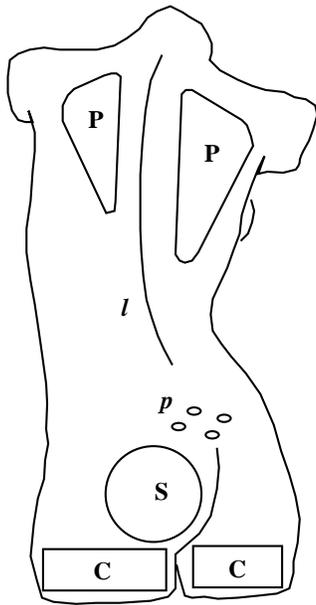
Figure 2. Typical regions with qualitatively different landmark-like features: P, planar polygonal facets; S, spherical areas; C, cylindrical with vertical axis; *l*, linear features; *p*, punctate features.

object) that have been painted in different textures, that are differently illuminated, that assume different attitudes with respect to the environment, or that are photographed from different angles, with different focal lengths, and so forth. In such cases, the correspondence has to be based on local structures of pictorial shape referenced to the global pictorial spatial structure. Yet there is no question that human observers can perform such tasks as a matter of routine, though we know little about the tolerances. This is a very powerful handle on pictorial shape. We have previously investigated such possibilities on small pieces of surface generated by computer graphics (Phillips, Todd, Koenderink, & Kappers, 1995). In this paper, we investigate correspondence in pictorial space for photographs of complicated, real objects.

Please note that such a task may turn out to be either trivial or impossible in simple (nongeneric from a mathematical point of view) cases. For instance, if the photographs are of a smooth sphere, the task is impossible (or totally ambiguous), because any point on one picture corresponds to any point in the other picture in the sense that there exists a movement in Euclidean space that would have yielded the result. Suppose that one puts three pockmarks on the sphere in general position. The task can then be done via interpolation (or extrapolation) from these local marks. The task has become trivial. In the latter case, one expects the reproducibility to depend on location. Near the marks, it will be near perfect; far from the marks, it will be sloppy. Both triviality and complete ambiguity (impossibility) can be avoided by using objects that are sufficiently smooth on a fine scale (no pockmarks) but also sufficiently structured within the region of interest (to avoid the barren landscape of the

sphere). Evidently, scale is the key issue here. For a given object, the task may be simultaneously ambiguous on a microlevel and trivial on a macrolevel. This suggests that the correspondence will be defined only within some tolerance region that depends on the image structure (which again depends on the structure of the object that is rendered in the picture). This yields a handle for psychophysical investigation: The scatter in repeated settings will reflect the regions of ambiguity.

The method of correspondences allows us to study the nature of the depth dimension in pictorial space in an entirely novel way. The depth dimension of one picture may lie (in the extreme case) in the frontoparallel plane of another picture (consider portraits en face and en profil). Looking for correspondences then effectively becomes a method of probing the depth dimension for one picture through indication of a frontoparallel location in another picture. In this paper, we compare the result of such a correspondence task with that of surface attitude estimations obtained by an independent psychophysical method.

It is hard to say on which type of (no doubt higher order in the sense of differential geometry) optical structure the subjects base their correspondences. We have set up the task such that it would almost certainly involve the recognition of 3-D shape features (distribution of curvature). However, we know that human observers can (and routinely do) establish correspondences between different objects (e.g., faces and bodies of different people; Thompson, 1942). It is hard to see how one might survive without constantly establishing such correspondences. The science of correspondences is the theory of "homology" in biology, perhaps starting with Goethe's (1971) insightful writings on the Urpflanze (generic plant). Historically, homology has always been treated with suspicion (contempt may be closer to the facts; see Remane, 1971, and Riedl, 1990). However, modern DNA analysis has basically confirmed all conclusions from homology (Riedl, 1990); thus, there cannot be anything wrong with it apart from the fact that we do not fully understand its principles. In a way, perceiving homologies captures much of the creative part of perception. Such considerations add another dimension to this type of study. The present report only scratches the surface.

METHOD

Stimuli

The stimuli were photographs of rigid objects. The objects were mannequins sold for the purpose of displaying fashion items in shop windows. Such objects may be obtained in a variety of poses for both genders and in any number of identical copies. For this experiment, we used a single item that was painted in a semigloss, even-white finish. This precluded the use of textural elements that might be used as landmarks and thus would (at least conceptually) trivialize the correspondence task. The object was mounted on a turntable with a vertical axis. Positions could be set at quarter-degree accuracy. The turntable was mounted on a column in a photographic studio where we had full control over the lights. The background luminance was carefully controlled to provide a small (factor of −0.4 log unit) contrast with the object. The camera and the light sources were fixed, whereas the object was turned in different poses. This presents certain
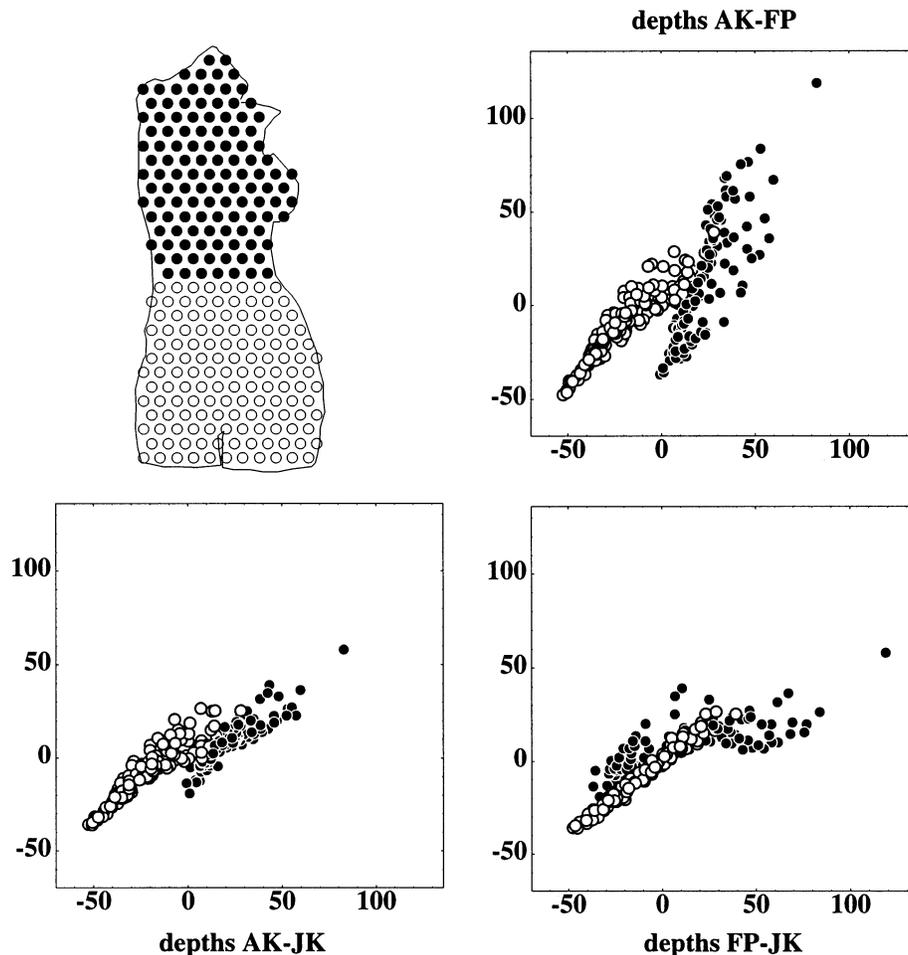
Figure 3. Scatter diagrams of depths from attitude settings for pairs of subjects. Also shown (upper left) is a segmentation of the stimulus in two parts, with the corresponding sets of data points marked in the scatterplot.

difficulties, because a good illumination scheme for one pose is quite likely disadvantageous for another (Hattersley, 1979; Hunter & Fuqua, 1990; Nurnberg, 1948). The object was illuminated by two broad sources or flood lights, placed a little in front of and displaced laterally to the left and right of the object. This yields a shading pattern that brings out the surface relief quite well at most parts of the object for all orientations of the object. This was important for the experiment, and it is difficult to achieve with more "natural" lighting (in the sense that we are exposed to it often), such as the standard high broad source from frontal left (the "Rembrandt illumination," often used by portrait photographers; Nurnberg, 1948). With the illumination we used, there is a dark "nucleus" in the center of the shape but only minor cast or body shadows. Though this shading pattern had advantages for our experiment, it presents unsurmountable problems to current "shape-from-shading" algorithms (Horn & Brooks, 1989). The typical shape-from-shading algorithms fail (i.e., produce a grossly nonveridical result) because the direction of illumination is ill defined, because the surface of the object is far from being Lambertian but rather shiny, and because, due to the nonconvex nature of the object, vignetting and interreflections violate prior assumptions. The human observer apparently easily reads such photographs, although it is unknown with what result.

Photographs were made from a fixed camera position and various orientations (we refer to these as *poses*) of the object. A pose is characterized by the orientation of the turntable with respect to a fiducial pose. The optical axis of the camera was horizontal, and it met the object at midheight. The elevation of the object subtended about ±5° of visual angle. We used orientations of 0°, 11.25°, 22.50°, 45°, 67.50°, and 90°. We used Polaroid monochrome transparency 35-mm film. The slides were scanned to 8 bit per pixel at 1,850 lines per inch. They were trimmed and subsampled in Photoshop. Final images were 600 pixels high at 72 pixels/inch. The grayscale fitted the 8-bit monochrome of our monitor. In Figure 1, we present examples of the stimuli.

In all photographs, there was the image of a mark on the (fixed) base of the turntable. This allowed us to use coordinates in the picture planes that have origins that correspond to the same point in the scene. This fact was of use in the processing of the data.

This series of pictures was well suited for the present experiment because ambiguous (spherical, cylindrical with vertical axis) regions and (partly) trivial (polygonal planar facets, linear features, punctate features) regions alternated. Moreover, because the pose was a rather twisted one (an emulation of the baroque "figura serpentinata"; Lomazzo, 1958), even minor pose variations caused large changes in the contour, thus minimizing the value of the contour as a landmark.

The fact that subjects must be expected to be familiar with the general structure of the human body aids them in finding nameable parts (arm, shoulder, scalula, etc.), thus speeding up the task of finding a rough correspondence, but it does not provide enough of a handle to find the precise correspondence. The poses are such that
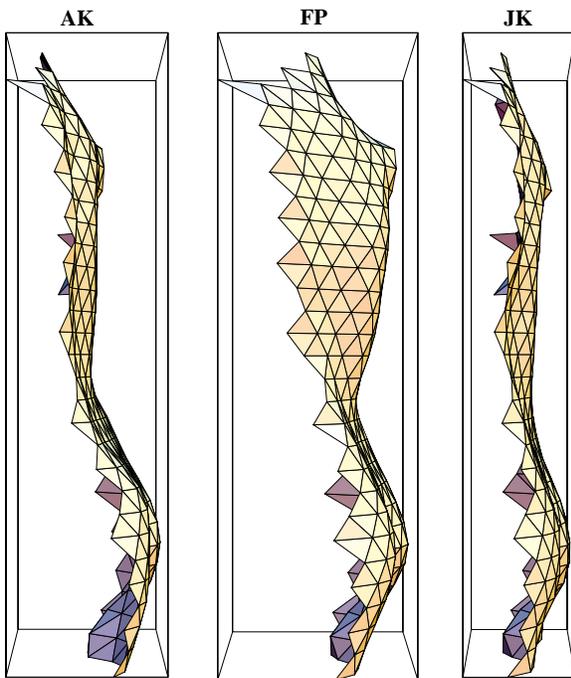
Figure 4. Profile views predicted from the attitude settings.

the bilateral symmetry of the body does not play a role in the precise correspondence, but, again, it may help to quickly zoom in on the approximate location.

The image for the fiducial pose (0º) was triangulated. First, we traced the silhouette; then, we fitted 256 vertices in a regular hexagonal lattice within the silhouette. This defined a basic triangulation; after slight hand editing (removing or slightly displacing vertices on the outline), we ended up with a triangulation of 430 faces, 681 edges, and 252 vertices. Data structures were calculated that allowed us to do various geometrical calculations based on this triangulation with ease. The triangulation was used only to generate the stimulus presentation and in the final data evaluation software; the subjects needed not be aware of its existence.

The images were presented on an RGB monitor in 8-bit monochrome. They were about 17 cm high and were viewed from 1 m in exactly the right perspective. Viewing was binocular. No fixation conditions were imposed, but head position was restricted via seating instructions. The lights in the laboratory were dimmed, though it was not entirely dark. Thus, the monitor and the frame of the picture on the computer's desktop screen were visible, and the subjects were fully aware that they were looking at pictures.

In the presentations, the left image was always the fiducial one of Pose 0º. The right image could be any of the stimuli. All but the Pose 45º stimulus were run in a single session in which each match was set three times over (in random order). For the Pose 45º stimulus, this was repeated three times so that the total number of settings amounted to nine.

In most conditions, there were a number of locations for which the subject could not find a match because the different poses did not reveal exactly the same surface area of the object. Such cases were indicated by the subject's putting the "match" in the extreme corner at bottom left. These cases were sorted out later and processed in a special way.

Subjects

The subjects were authors A.K. (emmetropic), F.P. (2 diopters myopic, corrected to normal), and J.K. (presbyopic, needs no cor-

rection for this viewing distance). All subjects were aware of the goal of the experiment. Subjects A.K. and J.K. had seen the actual object; Subject F.P. had no prior visual or haptic experience with it.

Experiments

As mentioned in the introduction, we performed two types of experiment. In one experiment, we emulated the surface attitude measurements as reported earlier (Koenderink et al., 1992). In another experiment, we determined correspondences for various combinations of poses. This allowed us to compare different methods of probing pictorial space. We also had internal consistency checks for both methods individually. Please note that we do not have the ground truth (i.e., the physical configuration, or the Cartesian coordinates of all the vertices of the triangulation in the fiducial pose). Thus, we are in no position to confront either method with physical reality and assess the veridicality of the results. Comparison of methods partly makes up for this. It is also the case that the expected relations between the correspondences for the various pose pairs allow a check closely related to the veridicality issue: We may check whether the results are consistent with the projections of a rigid configuration rotated in 3-D Euclidean space. However, the problem of internal consistency for a given method and the concordance/discordance of results from different methods is an issue that comes prior to the veridicality issue when the focus (as is true for our research) is on the structure of visual perception. Of course, the veridicality issue is of immediate importance in most applications.

Surface attitude probing. In the surface attitude task, the subjects adjusted a "gauge figure" to appear to "fit" the pictorial surface (Koenderink et al., 1992). In this case, the gauge figure was an ellipse, and a fit was defined as the appearance of a circle "painted on the pictorial surface." The gauge figure also included the projection of an outward normal vector with a (3-D) length equal to the radius of the circle. The normal helped disambiguate the inherent 180º tilt ambiguity. The subjects adjusted the shape and orientation of the gauge figure to appear as a circle painted upon the pictorial surface (this defines the fit). The actual ellipse is then interpreted as the projection of a circle in nonfrontoparallel attitude, and the attitude of this circle defines the pictorial surface attitude. This attitude either is parameterized by its slant and tilt or is described as a pictorial depth gradient in the visual field. The former parameters are perhaps the more familiar ones to vision researchers. However, the latter choice has some advantages from a formal point of view: For instance, at vanishing slant, the tilt becomes ill defined—a problem that does not appear in the depth gradient description. We have previously described this method in more detail (Koenderink et al., 1992).

All subjects visited each vertex three times in a single session lasting about 1 h. Only the fiducial pose was measured by this method.

Our aim in this experiment was to use an independent method to obtain a measure of the subject's "pictorial relief." This relief is a surface in 3-D pictorial space, and it can be compared with the results of the matching task. It seems more natural to use the subject's individual pictorial relief for this than to use the actual shape of the physical object. For instance, if the data from the attitude probing would contain local depth inversions due to ambiguous shading, then we also expect to find traces of this in the subject's matching data. The veridical surface would be the same for all subjects and would preclude such correlations. The comparison of the surface attitude task with the matching task is also interesting because the former probes the first-order depth structure (attitude is the derivative of depth), whereas the latter probes the (zeroth-order) depth structure more directly. We have previously addressed such comparisons (Koenderink, van Doorn, & Kappers, 1996), and it appears that much can be learned from them.

Image correspondences. In the experiment on simultaneous matching of location, two images were presented simultaneously, side by side (Figure 1). A marker was superimposed on the left image, and the subject was instructed to move another marker to the corresponding location on the right image. Correspondence was ex-
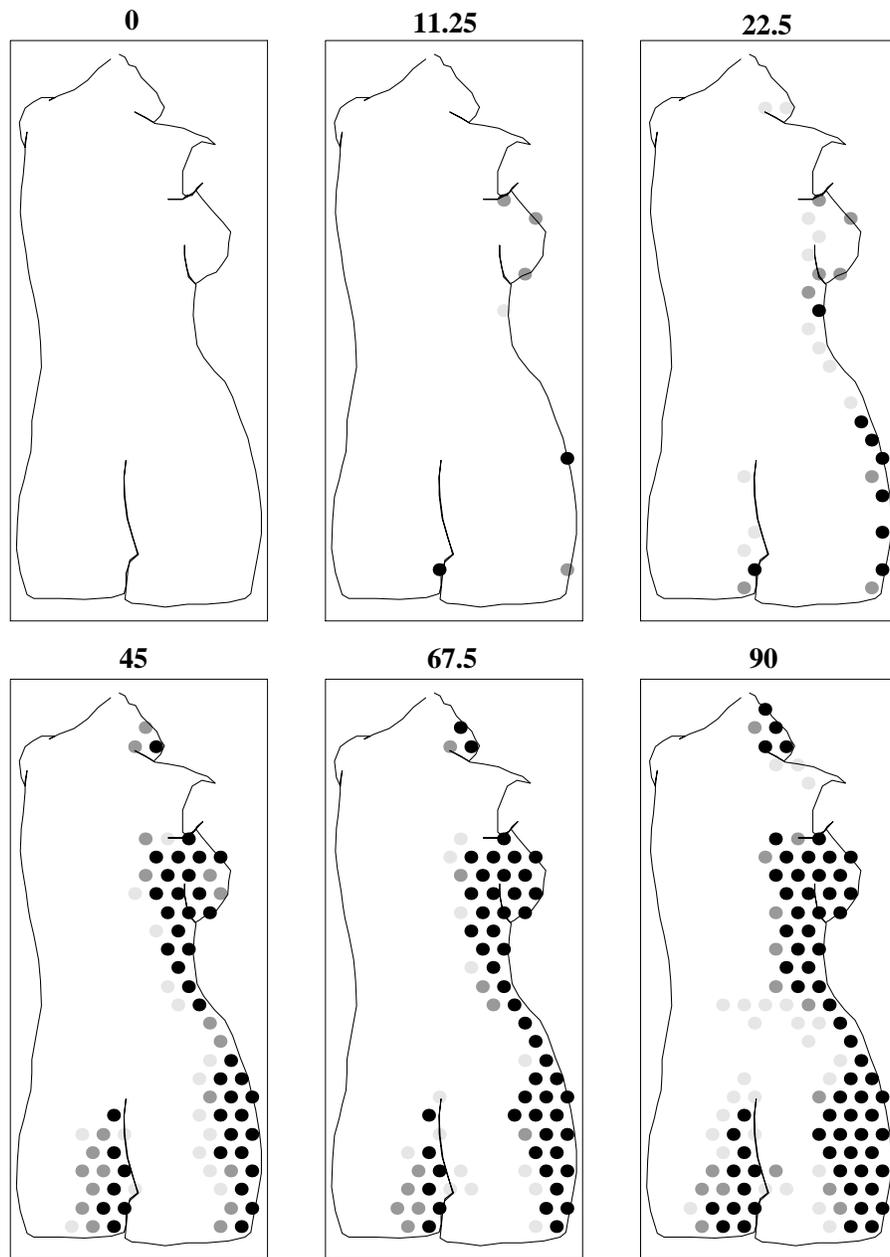
Figure 5. Vertices that were unmatched in one or more sessions are shown in graytone. White (invisible) signifies that the vertex was always matched by all subjects; black signifies that none of the subjects ever matched it. The fair amount of in-between graytones indicates a surprisingly high degree of uncertainty.

plained to the subjects as congruence in 3-D pictorial space (an explanation that seemed self-evident to the subjects). The subjects used the computer mouse device to move the marker, watching it move over the picture. No time constraint was imposed. On the average, the subjects used about 15 sec per setting. When the subject was satisfied with the marker position, the mouse button was clicked, and the next trial was initiated. The presentation program selected locations from the triangulation vertex list in randomized order. During one session, all vertices were visited three times. A single session lasted between 1 and 2 h; 1 subject did the task within the hour, 1 subject took up to 2 h, and the 3rd subject was intermediate.

All sessions (matching and attitude probing together) were done in random order, different for each subject. The subjects were denied the opportunity to see their data until after the conclusion of all sessions. The subjects were also denied knowledge of the actual pose angle differences for the various sessions. Subject F.P. did not know the overall list of the pose angles used, whereas Subjects A.K. and J.K. did.

The correspondence task was novel, and we had little previous experience with it (Phillips et al., 1995). For instance, it was unclear how well subjects perform the (trivial) task on two identical pictures, and it was not clear whether the task would be possible at all in the case of different poses. After all, it is not easy to see how
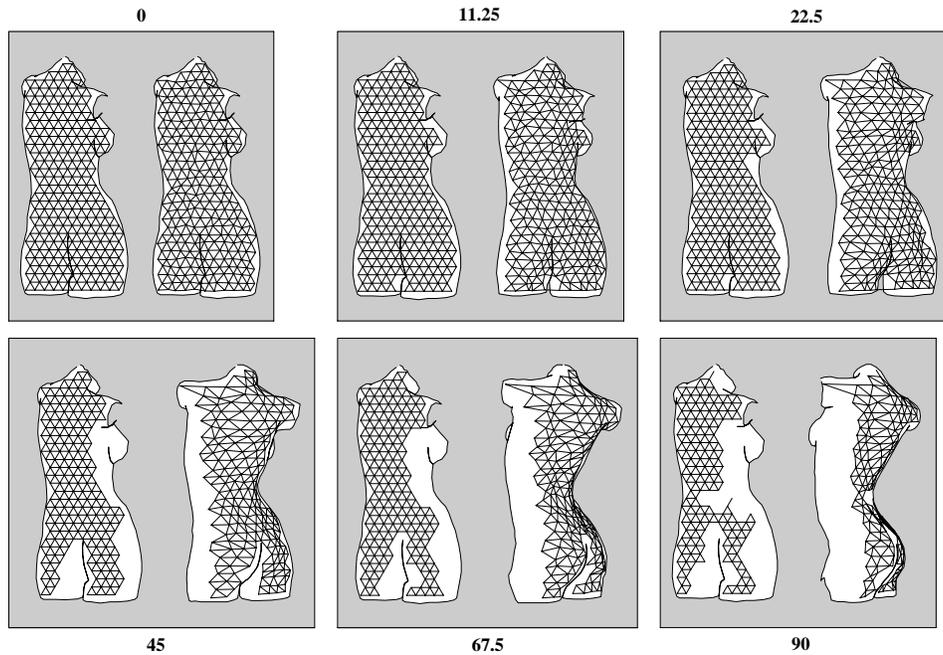
Figure 6. Raw matches of Subject F.P. (In each panel, the left figure is the 0º pose.)

current computer vision methods might establish correspondences in these cases. As it turned out, all subjects performed the task in all combinations of poses with confidence. The task seemed natural to them, in the sense that they did not doubt the existence of a correspondence, though some correspondences are more obvious than others when precision is required (e.g., in more or less uniform areas, the exact location is doubtful).

There exist conceptual differences between the matching of a pair of identical pictures and a pair of pictures corresponding to different poses. In the former case, the task is, in principle, trivial because it may be done through a mere correlation of a 2-D graytone distribution. A correlation of the global picture would, no doubt, allow us to reach subpixel accuracy. There is no reason to expect differences between the horizontal and vertical directions. Moreover, we did not expect the subjects' results to be correlated, since any inaccuracy would have been due to a random cause. In the latter case, the 2-D structure of the pictures is less useful, and it might often lead us astray. The larger the pose angle difference, the more different the two pictures will be as 2-D graytone distributions. Correspondences are expected to be shifted in the horizontal direction, whereas the vertical coordinates will be identical. The subjects' results were expected to be correlated, since they all base their settings on the same pictorial differences—that is, the same deterministic cause.

It might be argued that the self-matching task is spurious because it can be done without any 3-D interpretation at all. This is true, but it is certainly an interesting limiting case since one does not expect that the behavior of the subject will be much different when the pose angle difference is very small. We expect that, in all cases, there is some mixture of 2-D and 3-D features being treated as "landmarks" by the subject.

The subject could not easily use local landmark type features in this task. The possibility of establishing a correspondence depends on the structure of the "curvature landscape" because the contour, the shading, and the specularities do not offer a handle, and textural features cannot be used. For spheres and vertical cylinders, the task is even completely ambiguous. The upper leg regions are examples of vertical cylinders, whereas the buttocks are examples of nearly

spherical parts (see Figure 2). For planar patches, the task can be solved only by interpolating from the edges. Examples are the shoulder blades, which are nearly planar triangular regions. There, the task is not unlike what it would be for a cube. For linear vertical elements, the task is well defined—an example is the spine. It is indeed most unlikely that the subjects would ever make errors that cross the spine. Well-localized features are the vertices of the shoulder blade regions, various bifurcations of ridges, and the dimples at the sacral region. At such points, the task becomes trivial, and we would expect local covariance ellipses to be very small.

Due to perspective effects, we expected some vertical deviations. These are maximal for the 90º pose angle difference and positions in either the top or the bottom part of the picture. Maximum excursions were estimated at 8 pixels. For the large majority of cases (moderate pose angle differences, central part of the picture), these excursions are of pixel or even subpixel dimensions.

RESULTS

Surface Attitude Probing

The results from the surface attitude probing are, in all respects, similar to results of previous experiments with this method. We found that the scatter in repeated settings resided primarily in the slant component of the depth gradient, whereas the standard error in the tilt component was several times (typically three to four times) less. The standard error in the slant component amounted to about 6%–7% of the magnitude of the depth gradient, with a minimum value of about 0.025 ("pixels in depth" per pixel). The data were consistent with the existence of an integral surface within the tolerance given by the scatter in repeated data. The integral surface (we refer to it as *the pictorial relief due to attitude probing*) was roughly similar for the 3 subjects, though the differences were sig-
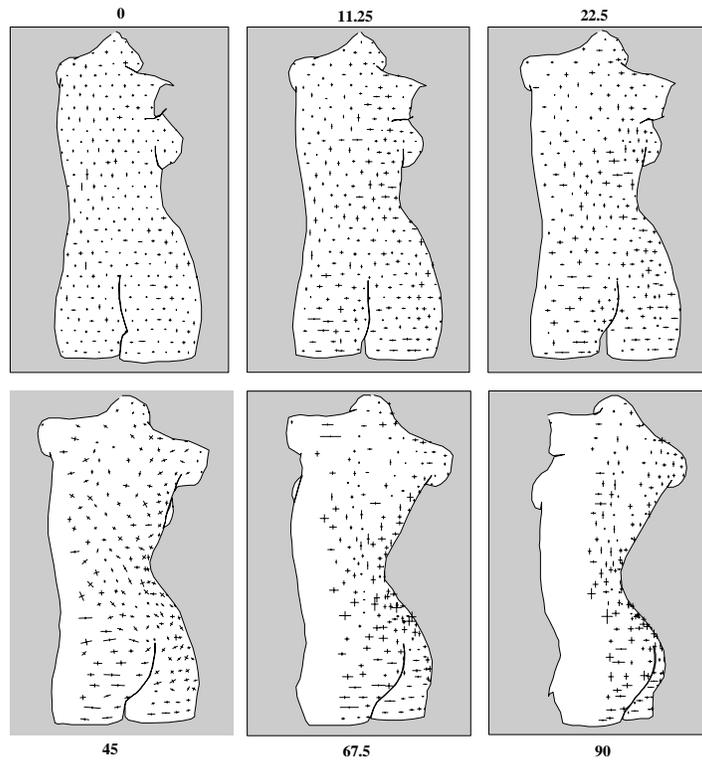
Figure 7. Subject F. P.'s standard errors in the horizontal and vertical directions for all pose angle differences.

nificant (see below). The pictorial relief is a true 3-D object, and we may compute a view of it from aside. Such a view yields a prediction of the profile and, thus, of the correspondences for a 90º pose angle difference.

The differences between the pictorial reliefs of the 3 subjects were significant (thus, at least not all three pictorial reliefs can be veridical). This was evident from scatter diagrams of the depths for pairs of subjects (Figure 3), as well as from the predictions of the profile (Figure 4). The scatterplots show a curious multimodal structure. We found (by trial and error) that the modes corresponded closely to a segmentation of the stimulus in an upper and
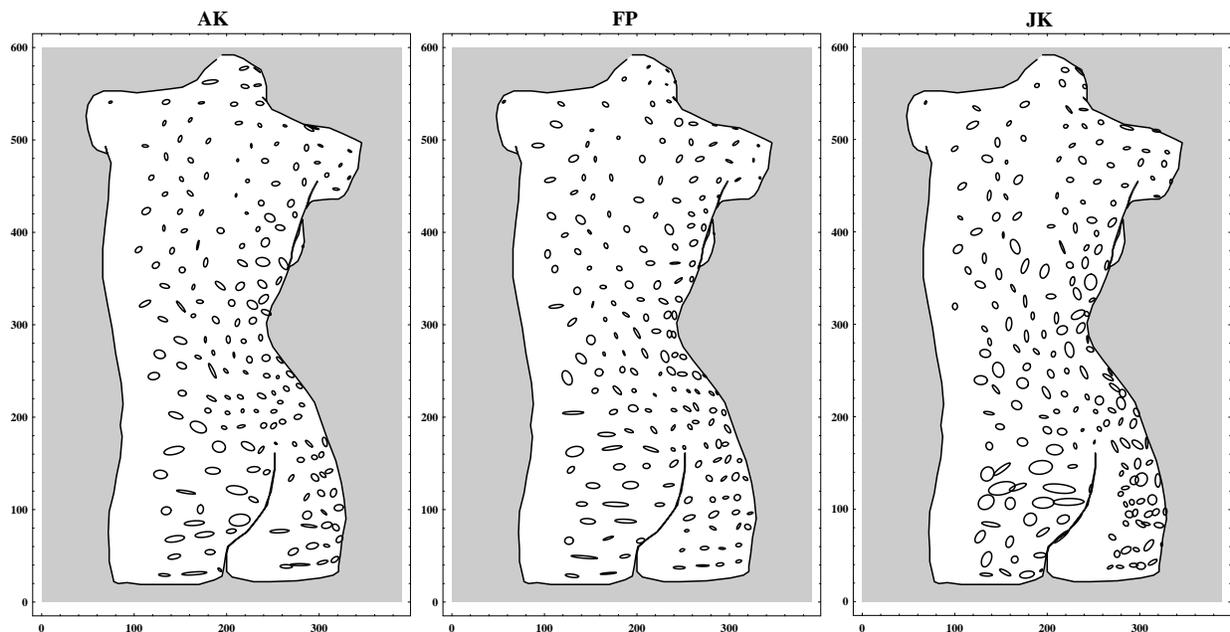

Figure 8. Covariance ellipses for the 45º pose difference.

Table 1
Results of a Regression of the Parameters of the
Covariance Ellipses for Pairs of Subjects Over All Positions

Subject Pairs	Magnitude	Orientation
A.K.–F.P.	.26	.64
A.K.–J.K.	.18	.60
F.P.–J.K.	.16	.59

Note—Magnitude is the area of the covariance ellipse; presented here are the R^2 values. The Orientation column contains correlations of the orientation of the covariance ellipses; this measure is explained in more detail in the Appendix.

lower part. Indeed, if we mark the data points in the scatterplot with their membership of such a segmentation (as has been done in Figure 3), we see that the two sets correlate much better than the data set as a whole.

From the predictions of the profile, we see that, for Subject F.P., relative to Subjects A.K. or J.K., the shoulder girdle–ribcage region was twisted with respect to the pelvic region. This was apparently the cause of the multimodal structure of the scatterplots. Such a difference is more complicated than a mere linear transformation (affinity in pictorial space). We have found such differences even in single objects when we change the relative salience of different monocular depth cues (Koenderink et al., 1994). The difference here may tentatively be interpreted as resulting from the individual subjects' putting different weights on various monocular depth cues.

Simultaneous Matching of Location

For the matching task, a match does not necessarily exist. We monitored when the subjects could not find a matching location (Figure 5). In Figure 5, the points that were left unmatched have been marked. Although there was some overlap, the subjects clearly disagreed. The decision changed from session to session and varied between subjects. However, on the whole, we found the expected pattern. It is perhaps surprising that even for such a basic task as deciding whether or not a match exists, there can be considerable disagreement.

The matching induces a deformation of the fiducial triangulation into a warped version of it, due to the equivalent of an "optical flow" (Figure 6).

For all pose angle differences, we had at least three repeated settings. This allowed us to map the regional distribution of both the horizontal and the vertical standard error (Figure 7). This makes sense because we have a priori reason to expect differences in these dimensions. We found that, in many regions, the horizontal scatter dominated. However, the vertical scatter was, by no means, small, and it appeared to follow some deterministic pattern.

We found that the correlation of the settings between the subjects increased as the pose angle difference increased. Values of R^2 up to .3 were found in the horizontal, and values of up to .15 were found in the vertical. For the identical picture pair, the subjects' results are not significantly correlated at all. For a large pose angle difference, the scatter in repeated data shows up a clear topographical structure, which repeats in repeated sessions of a single object

and tends to be similar for different subjects. This pattern evidently depends on the structure of the pictures.

For the pose angle difference of 45º, we have nine repeated settings, allowing us to estimate covariance ellipses and study the anisotropy of the scatter in somewhat more detail (Figure 8). We found that the orientations (see Appendix, Correlation of Orientations section) of the covariance ellipses correlated well between subjects, but that the magnitudes correlated only barely significantly (Table 1). The R^2 values for the magnitude were about .2, whereas the orientation correlation was about .6, corresponding to an angular spread of about 50º (the correlation of orientation far exceeded the noise level of the correlation, which amounted to .05).

We indeed spot indications of most of the expected effects of topographic distribution (Figures 7 and 8) of landmarks. Clearest are perhaps the large horizontal ambi-

Consistent scatter extremes   (all subjects)
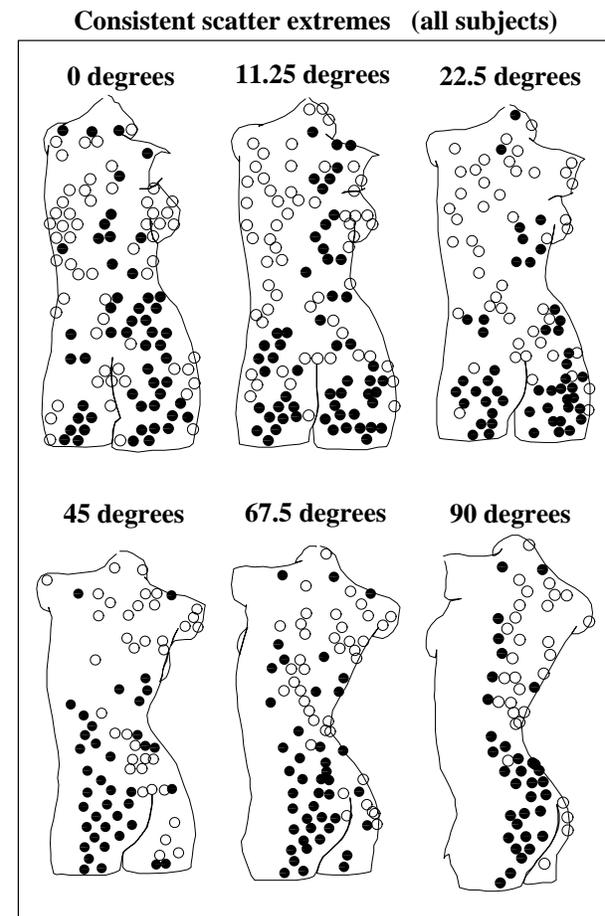
Figure 9. Maps of the distribution of high, low, and conflicting values of the scatter in repeated data. Data are over all subjects. We have only marked "high" (filled circles) and "low" (open circles) values. A "high" value was when at least 2 of the 3 subjects had a score above the 75% quartile range; a "low" value was when at least 2 of the 3 subjects had a score beneath the 25% quartile range. A "conflict" was when a score was not in the high or the low category and when the value was in the upper 75% quartile range for 1 subject and in the lower 25% quartile range for another subject.

Table 2
Frequency of Outliers of the Scatter
in Repeated Correspondences

Angle (degrees)	Score		
	High	Low	Conflict
0	2.1	3.7	−0.9
11.25	2.9	3.7	−1.5
22.50	2.8	3.7	−1.8
45	2.4	1.9	−3.7
67.50	3.0	2.5	−3.4
90	2.9	1.8	−1.8

Note—Data are over all subjects, specified per pose angle difference. A "high" score was when at least 2 of the 3 subjects had a value in the above 75% quartile range. A "low" score was when at least 2 of the 3 subjects had a value beneath the 25% quartile range. A "conflict" score was when a value was not in the "high" or "low" category and when the value was in the upper 75% quartile range for 1 subject and in the lower 25% quartile range for another subject. The frequencies have been normalized by subtraction of the expected value and division by the standard deviation for a pure chance situation.

guity in the regions that are approximately cylindrical with vertical axis (upper legs) and the small overall ambiguity in the sacral region. It is also evident that no correspondence ever crossed the furrow of the spinal column. There seems to be a tendency of the covariance ellipses to decrease in magnitude toward the contour.

Since one expects to see a correspondence of the magnitude of the scatter over subjects, we performed analyses pertaining specifically to the magnitudes. Because linear regression appeared to reveal this pattern only barely (as indicated earlier), we designed a more robust statistic. We marked cases where 2 of the 3 subjects scored a magnitude in the 75%–100% quartile range ("high" score), scored a magnitude in the 0%–25% quartile range ("low" score), or where neither of the above was the case but 1 subject scored in the 75%–100% quartile range and another subject in the 0%–25% quartile range ("conflict" score). We then subtracted the expected values for a pure chance occurrence and divided by the standard deviation for the pure chance case. The results are collected in Table 2; a topographic map of the scores is presented in Figure 9. In all cases but that of the equal pictures (pose angle difference zero), the high and low scores were much above chance level, whereas the conflict scores were much below it. This means, of course, that the results depend on local image structure, thus causing the subjects to act in a similar manner.

In the scatterplots of vertical deviations against height in the picture, we fail to detect a trace of the expected excursions (because of the effects of central perspective) in the vertical direction. The results appear to be dominated by either random or perhaps picture-content–induced scatter. This is clear from the fact that the vertical deviations also occur (and are of the same order of magnitude) in those cases in which we can rest assured that the actual values are zero—that is, for the matching of identical pictures and for the matching of the horizontally extended region at midheight in any other case. The vertical scatter must have been due to other effects.

A priori, these might have been random variations and/ or perhaps variations induced by the local structure of the picture. The latter component would be expected to show up in a correlation between repeated settings of a single subject and perhaps in a correlation of the settings of different subjects for the same stimuli. Such correlations were barely significant at the 5% level (see Figure 10.)

A simple measure of overall consistency is obtained if we compute profile predictions (see Appendix, Depth Structure From Single Correspondence section). We may compute a profile from matchings at any pose angle difference, and the profiles are a "common currency" in which the data at different pose angle differences can immediately be compared (Figure 11). All profiles are very similar, indicating excellent consistency over all pose angle differences. The main differences are in the amount of scatter. The predicted profiles are very noisy at small pose angle differences and much smoother at large ones. This is exactly as expected from the standard shape-from-motion reconstruction algorithms.

From the correspondences of all pose differences we may find a more direct measure of internal consistency. This is the case because a small number of such correspondences suffices to compute the 3-D shape. Once the shape is known, all other correspondences can be predicted. Thus, any large number of correspondences must contain dependencies and will almost certainly be inconsistent (see Appendix, Relations Between Multiple Correspondences section).

If we assume that the pose angles are known to the experimenter (not to the subject), we may find the full 3-D structure from just a single correspondence. Indeed, we have (see Figure 12): $x_\theta = \cos\theta x_0 + \sin\theta x_{90}$. Here x_0, x_{90}, and x_θ denote the horizontal position of corresponding pixels for the 0°, 90°, and θ poses. Here, we use the fact that x_{90} is simply the depth for the 0° pose. If we measure only x_θ, we can solve for the depth (x_{90}) since all the other variables are known (since x_0 is the position in the fidu-
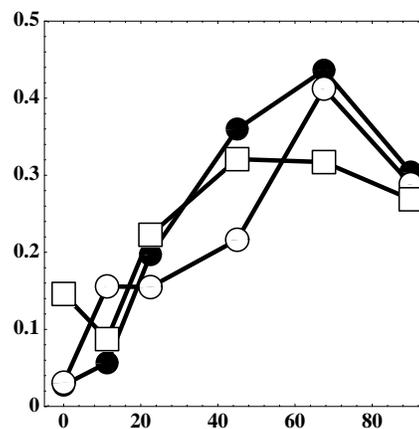
Figure 10. R^2 values for a regression of vertical scatter between pairs of subjects, as a function of the pose difference in degrees. Filled circles, A.K.–F.P.; open circles, A.K.–J.K.; squares, F.P.–J.K.
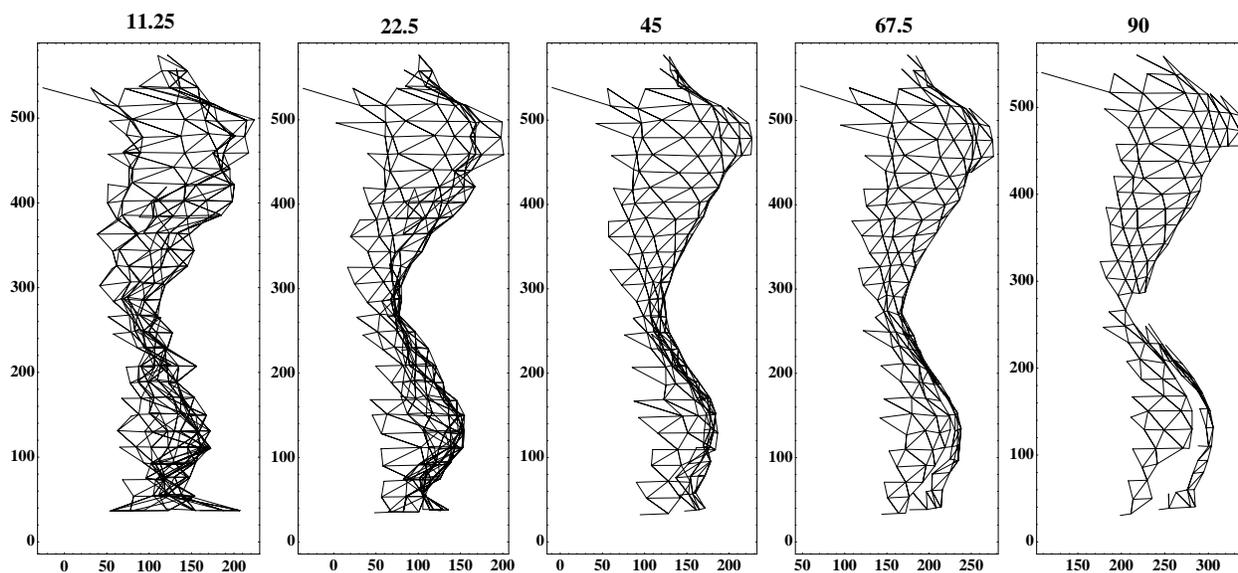
Figure 11. Profiles computed from matches at pose angle differences of 11.25º to 90º for Subject F.P.

cial image, $\cos\theta$ and $\sin\theta$ are assumed to be known). We then can use the relation $x_\varphi = \cos\varphi\, x_0 + \sin\varphi x_{90}$ to predict the correspondence x_φ for any desired pose angle φ.

An overall check of consistency reduces simply to a check on the rank of the complete data matrix (see Appendix, Relations Between Multiple Correspondences section). We found that the first two singular values indeed dominated. Together, they accounted for more than 99.5% of the total sum of squares of the singular values (ratio of the second to third largest singular values was about 15 for all subjects). The data matrix was very close to a consistent one for each of the subjects: Residuals amount to root mean square pixel shifts of about 3 pixels, which is roughly to be expected from the scatter in repeated sessions. From the singular values decomposition, we immediately obtain estimates of the pose angle differences (Figure 13 shows that the data conform exceptionally well to the model of a rotation about a fixed axis): We obtain errors of the order of 1º to 2º of turn angle (R^2 of .998). This might seem remarkable in view of the fact that the subjects found it very hard to estimate the pose angles and often confused the stimuli. However, the subjects saw only pairs of stimuli at any time, and, for a single pair, this type of analysis fails: Depth of relief and pose angle difference can be traded for each other and are both individually ambiguous.

The result means that the data are consistent with an interpretation in terms of orthographic projections of a rigid object rotated by various amounts about a fixed axis in Euclidean space. Thus, this type of consistency also indicates a high degree of veridicality.

The residuals from the consistency analysis show different patterns for the 3 subjects. They do not appear to have been completely random; however, since similar deviations occurred in rather large spatial clusters, apparently there was some nonrandom topographical structure.

These patterns were different for all subjects, and they apparently indicate some idiosyncratic component. Straightforward correlation of the residuals between subjects reveals only very weak—though just significant at 95% confidence level—correlations (R^2s of .07–.1). In order to make sense of this, much larger (and varied) data sets would be required.

Correlations Between the Two Tasks

It is possible to predict the correspondences from the attitude data in a straightforward manner (see Appendix, Correspondence From Attitudes section). We performed linear regressions between these predicted values of the correspondences and the actually measured correspondences for all subjects and all nonvanishing (because then
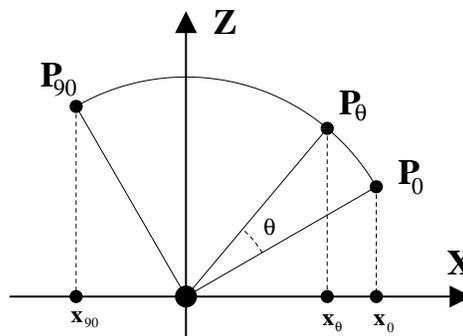

Figure 12. The X-axis points to the right in the frontoparallel plane; the Z-axis points away from the observer, straight into depth. For example, the point P_0 has projection x_0 on the X-axis and projection z_0 on the Z-axis. Then, the projection on the X-axis of this point after a 90º rotation (the point P_{90}) will be $x_{90} = -z_0$. (Thus, the frontoparallel position of the point after rotation over 90º is determined by its depth value in the original pose.) After a rotation over θ, the X-coordinate of the point (now P_θ) will be $x_\theta = \cos\theta, x_0 + \sin\theta, x_{90}$.
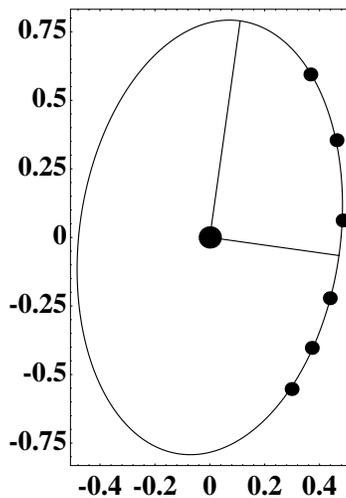
Figure 13. The columns of the *U* matrix from straight singular values decomposition of Subject A.K.'s data matrix and anulling all but the two largest singular values. A best fitting ellipse has been constructed and is seen to fit the points well. This indicates that the data conform to a model that describes the orthographic projection of a rigid object rotating about a fixed axis. Using the metric induced by this ellipse, we obtain the pose angles up to an unknown additive constant. Correlation with the true pose angles yields an R^2 of .9981. Standard error of the regression is 1.42°; slope is between 0.9542 and 1.079 (at 95% confidence level).

the predictions from the attitude task are irrelevant) pose angle differences (Table 3). There appear to be differences between Subject F.P. and Subjects A.K. and J.K. with respect to the intertask correlation. For all 3 subjects, the correlation was very high for the small pose angle differences. For Subjects A.K. and J.K., the correlation dropped rather steeply as the pose angle difference increased, for Subject F.P., this effect was much less (though it is indeed apparent that the correlation also decreased). Yet the effects were at least qualitatively similar.

Since the reconstruction from the correspondences must be nearly veridical (as is evident from the consistency analysis), any differences between the pictorial reliefs of the two tasks must have been due to a deviation from veridicality of the results from the surface attitude task. This is, of course, to be expected, since, in the surface attitude task, the subject was required to make an absolute judgment on the basis of a single picture, whereas, in the correspondence task, the subject was required only to establish a relation between a pair of pictures and needed not make any absolute judgment concerning depth or attitude.

We found differences between the subjects with regard to the degree of veridicality in the result from surface attitude probing. For Subject F.P., a comparison of the (nearly veridical) reconstruction from correspondences and the attitude results reveals a close approximation to veridicality; for Subjects A.K. and J.K., we see a large difference in the twist of the posture (relative rotation of upper with respect to lower body). We have often observed

similar differences between subjects in similar tasks (Todd, Koenderink, van Doorn, & Kappers, 1996).

CONCLUSION

We have studied image correspondences in two simultaneously presented photographs of a single object in different poses. We found that, although no pictorial elements can be used immediately to define such correspondences, subjects nevertheless are quite able to produce such image correspondences. Moreover, the correspondences established in this manner are fully consistent with the multiple perspectives of a rigid object in 3-D Euclidean space. Because simple landmarks in the visual field cannot be used directly, we have to conclude that subjects are apparently able to find the correspondences on the basis of a comparison of the pictorial reliefs induced by the two pictures. This makes the task a relevant one for the psychophysical study of pictorial shape recognition.

The topographical distribution of scatter in repeated trials indicates that both the magnitude and the anisotropy of the scatter is a function of the nature of the pictorial relief. This can be gleaned from a comparison of the covariance ellipses of a single observer with the stimulus and is objectively documented by the correlation of magnitude and orientation of the covariance ellipses between different observers. The scatter can be used as a measure of the scale of the landmarks used by the subject to retrace a point in one picture in the other picture. This scale is fine near most punctate features (e.g., minuscule dimples or pockmarks), coarse and rather isotropic in more or less spherical regions, coarse and rather elongated in cylindrical regions with generators parallel to the axis of rotation, and so on. An analysis of the scatter in correspondences thus allows us to estimate the size of the smallest shape units recognized by the subject. They correspond roughly to the size of marks used by draftsmen to indicate surface relief in academic drawings of the figure (Clifton, 1973; Hogarth, 1981; Jacobs, 1988).

The scatter in the correspondences yields an estimate of the tolerances on the depth that can be recovered from a pair of pictures. For instance, at normal reading distance, depth from binocular correspondence should be at least as good as that obtained here for the smallest pose angle differences. One does indeed obtain a vivid impression of stereo from fusion of the 11.25° pose angle

Table 3
R^2 Values of Linear Regressions of the Measured
Correspondences With the Results of Predictions of
These Correspondences From the Empirical Attitude Data

| | Subject | | |
Angle (degrees)	A.K.	F.P.	J.K.
11.25	.99	.99	.98
22.50	.96	.97	.97
45	.86	.93	.91
67.50	.63	.83	.80
90	.25	.61	.41

difference pair of stimuli, though it is hard to say what the depth tolerances are since the monocular cues probably dominate. If the present estimate may be used to predict the tolerances on the binocular stereo result, then the leftmost profile in Figure 11 gives an impression. This may well be realistic since the stimuli define no disparities in the classical sense (there are no point-by-point correspondences possible) (Bülthoff & Mallot, 1988).

An interesting finding is that the correspondence data are mutually highly consistent and allow the computation of pose angles with high accuracy (about 1.5º). Moreover, these computed pose angles are fully veridical. The finding is intriguing in view of the fact that our subjects often confused the stimuli (i.e., they would spontaneously remark, "I have already done this condition," whereas, in reality, it was a novel pose). Thus, it is not as if they were able to judge the pose angle with great precision. Since we discovered the high degree of consistency in the data only in the final analysis, we were in no position to perform formal experiments to check the accuracy of pose angle judgments. Presumably, these can be off by 10º or more. This is perhaps not surprising since the consistency is over pairs of poses, whereas the "judgments" are on single stimuli. Possibly, the discrimination of pose angle might also be high. This remains an issue for further study.

It would be of considerable interest to compare the quality (and nature) of the correspondences established by the human observer with those obtained by state-of-the-art computer vision algorithms. However, this is not feasible at the moment, since we know of no algorithms that successfully deal with this type of image pairs. The best algorithms correlate textural detail in a patchwise fashion. In our case, texture is not present, whereas (due to the differences in shading of the various poses) correlation of coarse graytone variations is bound to fail.

A quite different task—namely, the probing of surface attitudes in a single image—turns out to yield results that predict the correspondences for not-too-large pose angle differences (up to, say, 45º) surprisingly well. This means that the human observer might, at least in principle, use pictorial information to predict correspondences in binocular disparity or dynamic perspective. Whether this actually plays a causal role in the execution of the matching task is not known.

REFERENCES

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.

CLIFTON, J. (1973). *The eye of the artist*. Westport, CT: North Light.

GIBSON, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

GOETHE, J. W. VON (1971). Die Metamorphose der Pflanzen [The metamorphose of plants]. In *Goethes Werke: Band XIII. Naturwissenschaftliche Schriften I* (6th ed., pp. 64-102). Hamburg: C. Wegner Verlag.

HATTERSLEY, R. (1979). *Photographic lighting: Learning to see*. Englewood Cliffs, NJ: Prentice-Hall.

HILDEBRAND, A. (1945). *The problem of form in painting and sculpture* (M. Meyer & R. M. Ogden, Trans.). New York: G. E. Stechert. (Original work published 1893)

HOGARTH, B. (1981). *Dynamic light and shade*. New York: Watson-Guptill.

HORN, B. K. P., & BROOKS, M. J. (1989). *Shape from shading*. Cambridge, MA: MIT Press.

HUNTER, F., & FUQUA, P. (1990). *Light, science, and magic: An introduction to photographic lighting*. Boston: Focal.

JACOBS, T. S. (1988). *Light for the artist*. New York: Watson-Guptill.

KOENDERINK, J. J., & VAN DOORN, A. J. (in press). The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*.

KOENDERINK, J. J., VAN DOORN, A. J., & KAPPERS, A. M. L. (1992). Surface perception in pictures. *Perception & Psychophysics*, **52**, 487-496.

KOENDERINK, J. J., VAN DOORN, A. J., & KAPPERS, A. M. L. (1994). On so-called paradoxical monocular stereoscopy. *Perception*, **23**, 583-594.

KOENDERINK, J. J., VAN DOORN, A. J., & KAPPERS, A. M. L. (1996). Pictorial surface attitude and local depth comparisons. *Perception & Psychophysics*, **58**, 163-173.

LOMAZZO, P. (1958). Treatise on the art of painting. In E. G. Holt (Ed.), *A documentary history of art* (Vol. 2, pp. 75-82). Garden City, NY: Doubleday.

MACH, E. (1959). *The analysis of sensations and the relation of the physical to the psychical* (C. M. Williams, Trans.; rev. by S. Waterlow). New York: Dover. (Original work published in German in 1886)

NURNBERG, W. (1948). *Lighting for portraiture*. London: Focal.

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.

REMANE, A. (1971). *Die Grundlagen des natürlichen Systems der vergleichenden Anatomie und der Phylogenetik* [The basis of the natural system of comparative anatomy and phylogenesis] (2nd ed.). Königstein-Taunus: Koeltz.

RIEDL, R. (1990). *Die Ordnung des Lebendigen* [The taxonomy of life forms]. Munich: Piper.

THOMPSON, D. (1942). *Growth and form*. Cambridge: Cambridge University Press.

TODD, J. T., KOENDERINK, J. J., VAN DOORN, A. J., & KAPPERS, A. M. L. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 695-706.

WHITTAKER, E. T. (1944). *A treatise on the analytical dynamics of particles and rigid bodies with an introduction to the problem of three bodies*. New York: Dover.

APPENDIX

Correspondence From Attitudes

Given the depth structure (depth values at the vertices up to an arbitrary translation into depth) from the surface attitude settings, it is an easy matter to predict correspondences for any pose difference. We denote the position of a vertex in 3-D pictorial space as (x, y, z), where x denotes the frontoparallel horizontal direction to the right, y denotes the frontoparallel vertical direction upwards, and z denotes the depth dimension (i.e., the visual direction away from the subject). Although this is a "left-handed" coordinate system, it is convenient, since the XY-axis represents the picture plane in the conventional manner, whereas the depth increases along the Z-axis (such left-handed systems are quite common in graphics applications).
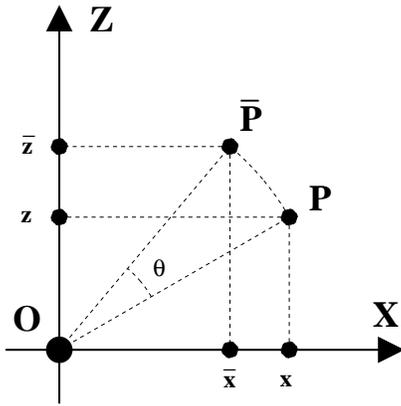
The position of the corresponding point in a view at pose angle θ is given by

$$\bar{x} = x\cos\theta - z\sin\theta + \bar{x}_0,$$

$$\bar{y} = y,$$

where $\bar{x}_0$ is a constant that specifies the coordinate origin in the second view.

Note that one introduces the pose angle θ here, which is known to the experimenter but not to the subject.

Figure A1. The *X*-axis is frontoparallel and toward the right of the observer; the *Z*-axis points into depth. The point *P* is carried into $\overline{P}$ through a pose angle change θ. The coordinates $\{x,z\}$ of *P* are carried into $\{\overline{x},\overline{z}\}$ of $\overline{P}$.

Depth Structure From Single Correspondence

Suppose we have the correspondence $(x,y) \rightarrow (\overline{x},\overline{y})$ for a pose angle difference θ (see Figure A1). We assume orthogonal projection, and $y = \overline{y}$. Then, all of the structure is in the x components.

Because of the rotation over the angle θ, we have the basic relations

$$\overline{x} = x\cos\theta - z\sin\theta,$$

$$\overline{z} = x\sin\theta + z\cos\theta.$$

From the first of these equations, we immediately find an expression for the depth:

$$z = \frac{x\cos\theta - \overline{x}}{\sin\theta} + z_0,$$

where z_0 is an arbitrary constant (the depth of the origin O). Thus, we recover the depth up to an arbitrary depth shift from a single correspondence when the pose angle difference is known. The relation degenerates as the pose angle difference becomes very small: $z - z_0 = (x - \overline{x})/\theta$. When the pose angle difference is a right angle, we simply have $z - z_0 = -\overline{x}$; in other words, the horizontal shift immediately translates into a depth difference. This is also intuitively evident, because, in this case, the depth dimension in one picture becomes the horizontal frontoparallel dimension in the other picture.

Here, we have introduced the pose angle θ, which is unknown to the subject. Thus, one should not assume that the result is also available to the subject. The value of this computation is that it allows one to compare the results obtained for different pose angles in a unitary format. Thus, the method implements a consistency check between sessions with different pose angles.

Relations Between Multiple Correspondences

Since the images are projections of a single object in different orientations in Euclidean 3-D space, there must exist certain relations between them. In our setup, perspective effects can safely be ignored; thus, we perform a simple orthographic analysis. In the orthographic case, the *heights* are invariant since the axis of the turntable is parallel to the focal plane of the camera. We need only consider the horizontal positions in the images. The resulting analysis is straighforward, though perhaps somewhat

unusual for some readers. Additional background is available (Koenderink & van Doorn, in press).

Each image is determined by the relative positions and orientations of camera and object. We capture this by defining a pose as a direction and a position. The direction defines the left-to-right dimension in the focal plane (i.e., it captures the pose), whereas the position defines the coordinate origin in the focal plane. A point on the object can be parameterized by its Cartesian coordinates (x,y,z). The position in the focal plane is (ξ,η), where we have $\eta = y$. Since the height does not depend on the pose, we may safely ignore it. The relevant data is in the ξ coordinate. We have

$$\xi = d_x x + d_z z + \xi_0,$$

where $\mathbf{d} = (d_x, d_z) = (\cos\theta, \sin\theta)$. If all coordinates in the picture planes are referenced to the image of the same landmark in the scene, we may define the axis of rotation to pass through this landmark. We then have identically $\xi_0 = 0$. This will be assumed here.

We define the 2-D vectors $\mathbf{p} = (d_x, d_z)$ and $\mathbf{q} = (x,z)$, allowing us to write $\xi = \mathbf{p} \cdot \mathbf{q}$.

If we consider poses $i = 1, 2, \ldots, n$ and positions $j = 1, 2, \ldots, m$, then we have $\xi_{ij} = \mathbf{p}_i \cdot \mathbf{q}_j$. Here, the ξ_{ij} are *observations* (settings of the subject), whereas $\mathbf{p}_i$ and $\mathbf{q}_j$ are *unknowns*, at least for the moment, because we certainly know d_x and d_z. If we stack all observations to produce the observation matrix Ξ, stack all poses to produce the pose matrix $\mathbf{P}$ and all positions to produce the position matrix $\mathbf{Q}$, we simply have to solve the equation $\Xi = \mathbf{P} \cdot \mathbf{Q}^T$ (here, the T denotes the transpose). Since both $\mathbf{p}_i$ and $\mathbf{q}_j$ are 2-D vectors, the observation matrix Ξ has only rank 2, though its size equals the number of poses (6) times the number of points (about 250). However, measurement errors will spuriously raise the rank to its maximum value (6 in our case).

We solve the equation up to unresolvable ambiguities in a standard manner. First, we perform a singular values decomposition, writing $\Xi = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T$. The diagonal matrix $\mathbf{W}$ contains the singular values. We keep only the two largest ones and set the others to zero, thus obtaining $\mathbf{W}'$. Similarly, we drop corresponding rows and columns from the other matrices, thus obtaining $\Xi' = \mathbf{U}' \cdot \mathbf{W}' \cdot \mathbf{V}'^T$. This effectively projects the problem on the closest consistent (rank 2) problem. We obtain a convenient measure of consistency in the values of the spurious singular values compared with those of the relevant ones. The matrix of residuals $\Xi' - \Xi$ contains more specific information concerning the nature of the inconsistencies.

We estimate the pose matrix as $\overline{\mathbf{P}} = \mathbf{U}' \cdot \mathbf{W}'^{1/2} \cdot \mathbf{A}$, and the position matrix as $\overline{\mathbf{Q}} = \mathbf{V}' \cdot \mathbf{W}'^{1/2} \cdot \mathbf{A}^{T-1}$, where the matrix $\mathbf{A}$ is arbitrary except for the existence of an inverse. The matrix $\mathbf{A}$ represents the inherent ambiguity in the problem. This ambiguity can be resolved by bringing our prior knowledge to bear: $d_x = \cos\theta$, and $d_z = \sin\theta$. Indeed, this completely resolves the ambiguity. Again, when we do this, we assume knowledge of the pose angles—a knowledge that is not available to the subjects. If we leave this ambiguity unresolved, we still find the 3-D structure, but up to an affinity. In resolving the ambiguity using the pose angles, we effectively implement a consistency check over the complete data set.

Correlation of Orientations

Suppose we have two sets of corresponding orientations that we want to compare. Oriented line elements are different from directed line elements in that they are carried over into themselves through a rotation by 180° instead of 360°. We need a measure of correlation or concordance between the two ordered sets.

Intuitively, two oriented line elements are fully correlated when they are lined up (zero angular difference) and anticorrelated when they are orthogonal (90º angular difference) to each other. A number that captures this intuitive notion is the cosine of twice the angle subtended by the oriented line elements: This number reaches a maximum value of $+1$ when the elements coincide and a value of -1 when they are orthogonal.

Consequently, the measure

$$R = \frac{1}{N} \sum_{i=1}^{N} \cos 2(\phi_i - \theta_i),$$

where ϕ_i and θ_i denote the absolute orientations of the two sequences of orientations to be compared, and N denotes the number of items, will assume values between $+1$ and -1. It is zero for large sets of mutually uncorrelated random orientations, unity for equal sets, and minus unity for sets of pairwise orthogonal (or anticorrelated) oriented elements.

For uncorrelated sets of orientations, distributed uniformly over <u>all</u> orientations, the standard deviation of the result will be $1/\sqrt{2N}$. This suffices to judge the significance of the results in our application because the correlations encountered are many times larger than this.

We use this measure to compare the mutual directionality of sets of orientations of covariance ellipses in the plane.