Effects of Texture, Illumination and Surface Reflectance on Stereoscopic Shape Perception

James T. Todd[1], J. Farley Norman[2], Jan J. Koenderink[3], Astrid M. L. Kappers[3]

1: The Ohio State University
2: Western Kentucky University
3: Helmholtz Instituut, Universiteit Utrecht

Observers viewed computer generated stereograms of randomly structured smooth surfaces and were required to judge the perceived local orientation at numerous probe points by adjusting a monocular gauge figure. The surfaces were depicted with specular or Lambertian reflectance functions, either with or without identifiable texture elements, and with varying directions of illumination. The results revealed a strong linear correlation between the judged patterns of relief and the actual depicted objects, though there were systematic differences in the magnitude of depth scaling in the different conditions. In general, the accuracy and reliability of observers judgments for the smoothly shaded shiny surfaces was slightly lower than for the textured surfaces and slightly higher than for the smoothly shaded Lambertian surfaces. The direction of illumination had no detectable effect on the observers judgments.

Introduction

Of the many potential sources of optical information about the 3-dimensional structure of the environment, binocular disparity is perhaps the most perceptually compelling. Because human observers have two eyes with overlapping visual fields, each eye receives a slightly different view of the same scene. It is especially interesting to note in this context that binocular overlap reduces the size of the combined visual field relative to what would otherwise be possible if the two eyes faced in opposite directions, as is the case with many other animals. For the ecology of human observers, however, this cost is apparently outweighed by the useful information about relative depth that is provided by the disparities between each eye's view in the region of overlap.

Much of the literature on binocular stereopsis over the past century has been concerned with identifying corresponding points in the projected images of each eye. The modern conceptualization of this is issue can be demonstrated most clearly by the ability of observers to perceive 3-dimensional structure from random dot stereograms, in which each stereoscopic half image contains a dense configuration of small dots (e.g., see Julesz, 1971). For any given dot presented to one eye, the visual system must somehow determine a single corresponding dot with which it should be matched among the many possible targets presented to the other.

One important issue to consider in this context involves the particular image features to be matched. Most theoretical analyses of binocular stereopsis are designed to be employed on changes of image intensity that arise from discontinuities of surface orientation, such as the edges and vertices of a cube, or discontinuities of reflectance, such as small drops of paint that might be spattered on a surface. The desirable characteristic of these particular environmental features is that they produce abrupt changes in image intensity, whose corresponding positions in each eye's view are projectively related to the same physical structure in 3-dimensional space. This property is not shared, however, for changes of image intensity that arise from other types of optical phenomena.

Consider, for example, the occlusion contour of a smoothly curved object such as a sphere. An occlusion contour is formed by the locus of visible points on an object whose surface normals are perpendicular to the observer's line of sight. Because each eye has a different direction of view in binocular vision, the locus of surface points that defines the occlusion for one eye can be quite different from the locus of points that defines the occlusion for the other. Thus, if smooth occlusion contours were matched stereoscopically as if they were
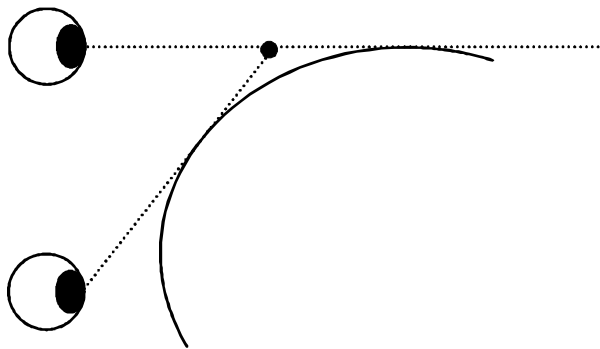
Figure 1 -- Because an occlusion point appears at different positions for each eye, its virtual binocular image can be spatially displaced from the observed surface.

orientation or reflectance discontinuities, then they should appear perceptually at an inappropriate position in 3-dimensional space (see Figure 1).

Smooth surface shading is another common aspect of optical structure that poses theoretical difficulties for the process of stereoscopic matching. For many surfaces encountered in nature the pattern of reflected light in each local region can be modeled as a linear combination of two components, which are often referred to as specular and Lambertian. For the specular component of surface shading, the light reflected off each local region is concentrated around a single primary direction, whereas it is scattered equally in all directions for the Lambertian component (see Figure 2).

Because of the diffuse pattern of reflectance in Lambertian shading, the luminance at any given point will be the same for all possible viewing positions. Thus, any identifiable feature of the luminance pattern, such as local maxima or minima, will always correspond to the same physical point on the surface when it is localized in the two stereoscopic views. That is not the case, however, for the specular components of shading. When observing a shiny surface, highlights appear at points where the surface normal bisects the angle between the line of sight and the direction of illumination. Because each eye has a different direction of view, the highlight for one can appear at a different surface location than the highlight for the other (see Figure 3; Blake & Bülthoff, 1990, 1992), and their binocular disparities may therefore provide misleading information about the 3-dimensional structure of the observed object.

To what extent do these factors influence stereoscopic form perception in human observers? In an early attempt to address this question, Bülthoff and Mallot (1988) presented displays of smooth shaded ellipsoids with and without highlights, and flat shaded polyhedral surfaces with clearly defined edges to provide a more appropriate set of features for stereoscopic matching. Observers were required to adjust a stereoscopic probe dot at numerous positions on each display until it appeared to rest on the depicted surface. When the objects were presented with correct binocular disparity, the adjusted depths were almost perfectly accurate for surfaces with edges, and were underestimated by approximately 30% for those with smooth shading. However, when identical shaded images were presented to each eye (i.e., with zero disparity), all of the observers'
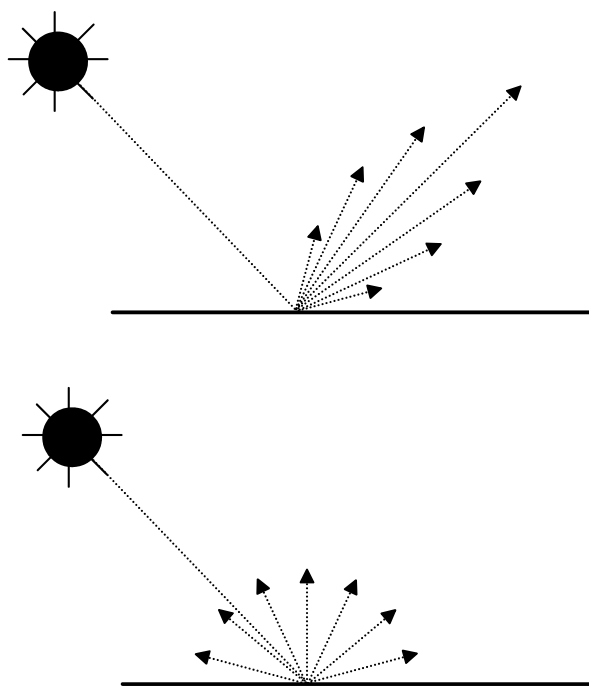

Figure 2 -- Two different components of surface reflectance. The upper diagram shows a shiny surface with specular highlights where the light reflected off each local region is concentrated around a single primary direction. The lower diagram shows a matte Lambertian surface where the reflected light is scattered equally in all directions.
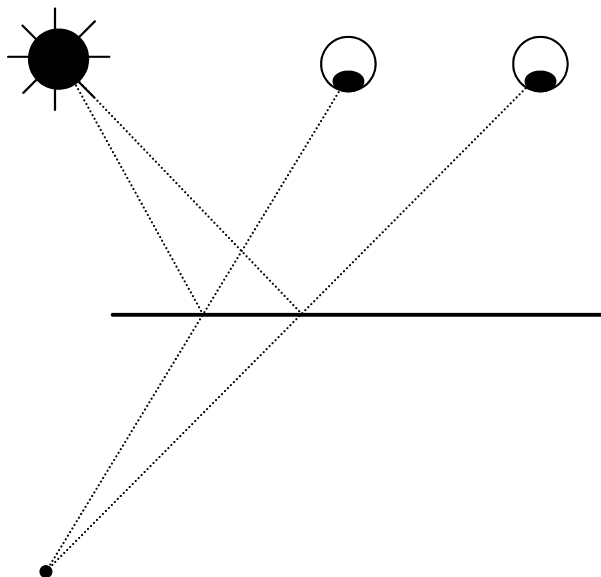
Figure 3 -- Because a specular highlight appears at different positions for each eye, its virtual binocular image can be spatially displaced from the observed surface.

probe settings were positioned at or near the plane of the display screen.

Although this study demonstrates that shading disparity can provide perceptually salient information for binocular stereopsis (see also Koenderink, Kappers, Todd, Norman & Phillips, 1996; Norman, Todd & Phillips, 1995), there are some methodological problems with the response task that make it difficult to compare performance across the different conditions. One likely strategy for performing these judgments would be to adjust the probe until its local disparity matches that of nearby surface features, without necessarily involving a judgment of apparent depth. This would produce perfect performance for surfaces with edges and appropriate disparity, and no depth at all when identical images are presented to each eye – exactly the same pattern that was obtained by Bülthoff and Mallot (1988). Their results have not been confirmed, however, when other techniques are employed to measure the perceived 3D structure of surfaces from edge based stereo (e.g., Johnston, 1991; Norman, Todd, Perotti & Tittle, 1996; Tittle, Todd, Perotti & Norman, 1995) or monocular shading (e.g., Erens, Kappers & Koenderink,1993; Mingolla & Todd, 1986; Reichel, Todd & Yilmaz, 1995; Todd & Mingolla, 1983; Todd & Reichel, 1989).

Another popular response task for measuring the perception of 3D shape involves judgments of local orientation at many different probe points on an object's surface (e.g., see Koenderink, van Doorn & Kappers, 1992, 1994, 1995, 1996; Koenderink & van Doorn, 1995; Koenderink, et. al., 1996; Mingolla & Todd, 1986; Norman, et. al., 1995; Stevens, 1983; Todd, Koenderink, van Doorn & Kappers, 1996). This can be accomplished in a variety of ways, but the most common technique is for observers to a adjust the 3D orientation of a circular disk, called a gauge figure, until it appears to rest in the tangent plane at some designated location. For stereoscopic displays, the gauge figure is presented monocularly so that the adjustment cannot be achieved by matching the disparities of nearby texture elements. Although the specific depth of the gauge figure is mathematically ambiguous in that case, most observers report that it appears firmly attached to the surface, and that they have a high degree of confidence in their adjustments.

For the research described in the present article, a gauge figure adjustment task was employed to investigate how stereoscopic form perception is influenced by several different aspects of image structure. Smoothly curved objects were presented both with and without a random surface texture to provide identifiable features for the process of stereo matching. They were also presented both with and without specular highlights, and with varying directions of illumination. Observers' judgments were obtained at numerous probe points on these objects so that it would be possible to detect any subtle variations in perceived structure across the different combinations of surface reflectance and illumination.

Methods

Apparatus. The stimuli were created and displayed on a Silicon Graphics Crimson VGXT workstation with hardware texture-mapping capabilities and stereoscopic viewing hardware. The displays were viewed through LCD (liquid crystal) shuttered glasses that were synchronized with the monitor's refresh rate. The different views of a stereo pair were displayed at the same position on the monitor screen, but they were temporally offset. The left and right lenses of the LCD glasses shuttered synchronously with the display at an alternation rate of 60 Hz, so that each view could

only be seen by the appropriate eye. The spatial resolution of the monitor was 1280 X 1024 pixels, which subtended 25.2 by 20.3 degrees of visual angle when viewed at a distance of 76 cm. Head movements were restricted by using a chin rest.

Stimuli. All of the objects employed in this study were defined initially as a polygonal mesh composed of 5120 individual triangles that were evenly spaced over the surface of a sphere. In order to create more complicated structures, a series of sinusoidal perturbations was added to this sphere at random orientations. This procedure provides an analytically defined surface normal and position at each vertex on the triangular mesh, which are used to compute visual images of the objects with appropriate patterns of shading and texture. Three such objects were selected for the experiment. Two of these are shown in Figure 4 as textured stereograms, and the third is depicted in Figure 5 with texture, smooth Lambertian shading, and smooth shading with specular highlights.

The texture patterns (see Figure 4) were generated from an image of red granite obtained from the Pixar Corporation. Each polygon in the triangular mesh was rotated to a frontoparallel orientation and positioned at random in the image of red granite to define its individual texture pattern. This ensured that the pattern of texture on the depicted surface was statistically homogeneous and isotropic -- i.e., that equal areas of the surface contained equal amounts of texture (see Todd & Mingolla, 1984). To achieve antialiasing for polygons depicted at different orientations in depth, the texture was blended by the SGI hardware using a MIP-mapped trilinear filter.
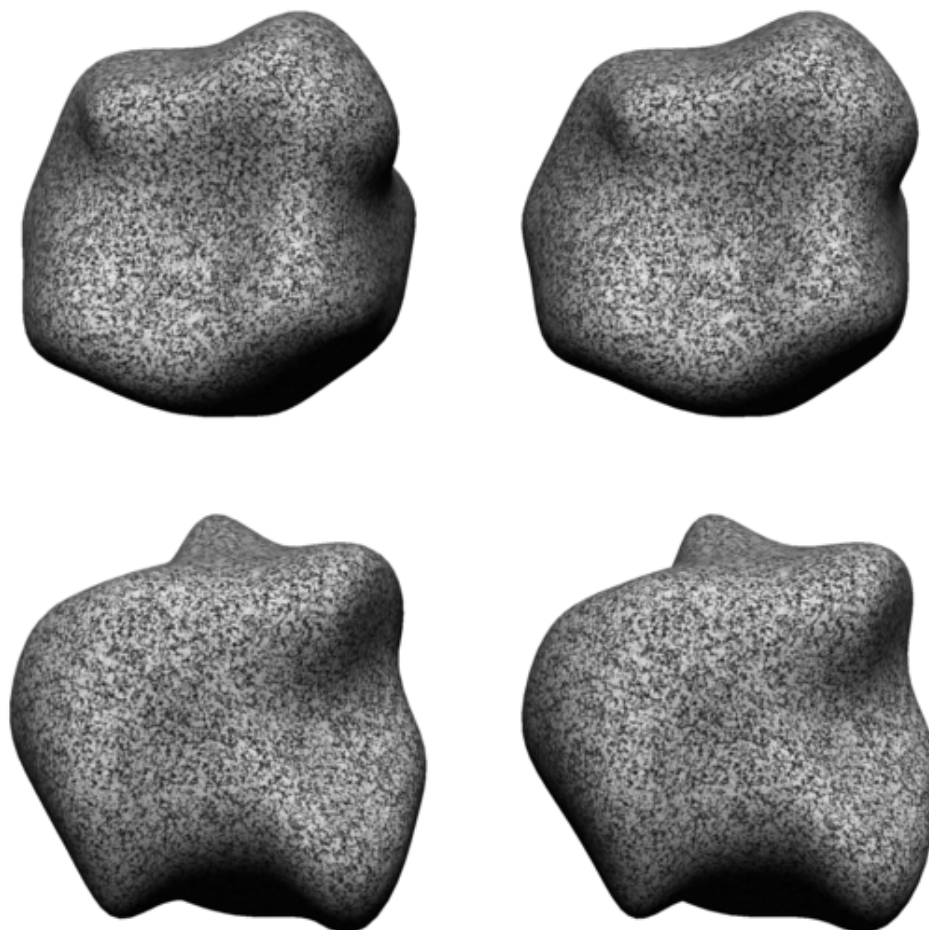

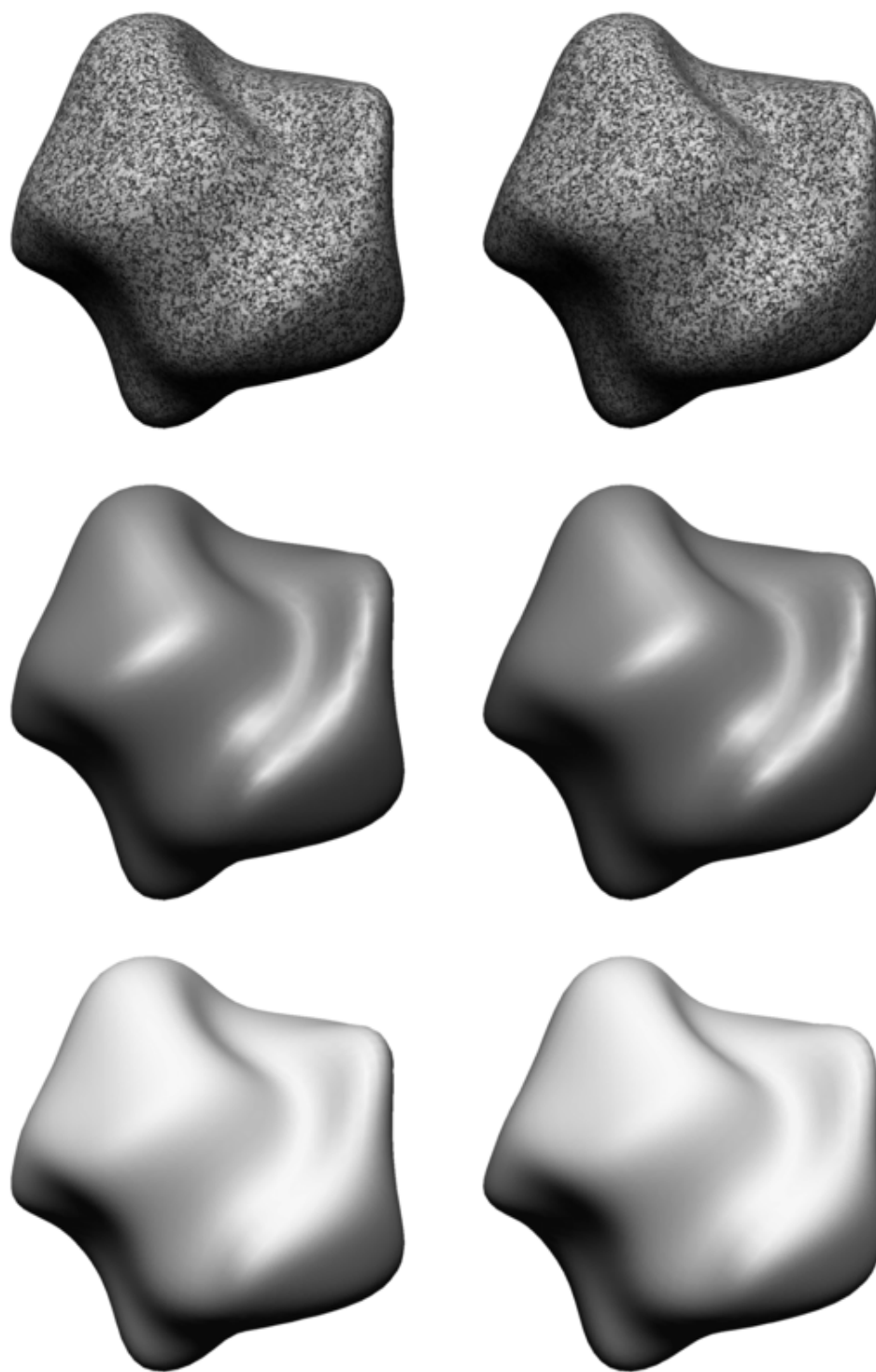**Figure 4** -- Textured stereograms of two stimulus objects employed in the present experiment.

Figure 5 -- Three stereograms of a single stimulus object with different types of surface reflectance. From bottom to top, the objects are depicted with matte Lambertian shading, specular highlights and random texture.

The patterns of surface shading were created using a standard computer graphics reflectance model, in which the image intensity I(r,g,b) at any given surface point is determined by the following equation:

$$I(r,g,b) = I_i(r,g,b) [k_a(r,g,b) + k_d(r,g,b) (\mathbf{L} \cdot \mathbf{N}) + k_s(r,g,b) (\mathbf{H} \cdot \mathbf{N})^{20}]$$

where $\mathbf{N}$ is a unit surface normal at that point, $\mathbf{L}$ is a unit vector toward a point light source, $\mathbf{H}$ is a unit vector that bisects the angle between $\mathbf{L}$ and the direction of view, $I_i(r,g,b)$ is the intensity of incident illumination in the three primary colors, and $k_a(r,g,b)$, $k_d(r,g,b)$, and $k_s(r,g,b)$ are their ambient, diffuse and specular reflectance coefficients. It is assumed in this equation that for each primary color, I_i must be less that 255, and the sum of the coefficients k_a, k_d and k_s must be less than one, so that the resulting range of image intensities does not exceed the eight bit resolution of the monitor. In the SGI implementation of this reflectance model, the values of I(r,g,b) are computed at the vertices of the triangular mesh and then linearly interpolated through the interiors of each projected polygon.

Over different experimental sessions, each object was presented with three different patterns of surface reflectance: In the Matte and Shiny conditions, the surfaces were presented without any texture, and the three components of I_i were all set to 255. In the matte condition, all of the reflectance coefficients were zero, except for the blue ambient and diffuse components, which were 0.3 and 0.7, respectively. In the shiny condition, the blue component of ambient reflectance was 0.3, the blue component of diffuse reflectance was 0.4, and all components of specular reflectance were 0.3. The resulting images in that case appeared as shiny blue surfaces with white specular highlights. For the Textured condition, all of the ambient reflectance components were 0.3, all of the diffuse components were 0.7, all of the specular components were zero, and the different components of I_i were determined by the color of the texture map at each point .There were two possible directions of illumination employed in these displays, both of which were slanted 30 degrees relative to the observers' line of sight. In the Side Illumination condition, the light source direction was tilted 90 degrees with respect to vertical, and in the Overhead Illumination condition the tilt of the light source was zero.

Procedure -- The task on each trial was to adjust the slant and tilt of a circular gauge figure centered at a given probe point so that it appeared to rest within the tangent plane of the surface at that point (see Koenderink, van Doorn & Kappers, 1992, 1995). Slant is defined in this context as the angle between the surface normal and the line of sight, while tilt is the direction of the surface depth gradient within the frontoparallel plane. The gauge figure simulated a small circle in 3-dimensional space with a perpendicular line at its center whose length equaled the radius of the circle. These appeared in the image as a red ellipse with a small line along the minor axis, whose lengths and orientations could be manipulated using a hand held mouse. When adjusted appropriately, all of the observers were able to perceive this configuration as a circle oriented in depth in the tangent plane with a line perpendicular to it in the direction of the surface normal. Although the objects were presented stereoscopically, the gauge figures were presented to one eye only, so that observers could not perform the adjustment by matching the disparities at the edges of the gauge figure with those of nearby texture elements on the depicted surface.

There were eighteen different experimental conditions involving all possible combinations of the three basic stimulus manipulations -- 3 objects X 3 types of surface reflectance X 2 directions of illumination. During a single experimental session 110 surface points for a given display were probed three times each in a random sequence, which took approximately 45 minutes. The vector average of these three repeated measurements was then used to estimate the judged surface orientation at each probe point. To further assess the test-retest reliability of this task, each observer participated in six additional sessions to obtain a second set of judgments for one of the three objects in all of the different illumination and reflectance conditions. Thus, each observer made a total of 7920 adjustments over 24 experimental sessions.

Observers -- Three of the authors (J.T., J.K. and A.K.) participated as observers and performed the sessions in a different random order. All had normal or corrected to normal vision, and extensive experience with the gauge figure adjustment task.

Results

We began our analysis of the data by measuring the consistency of observers' judgments over multiple trials for individual probe points. The striped bars of Figure 6 show the average deviation in degrees between each individual judgment at a given probe point and the average for that point over several repeated trials. When collapsed over all conditions, the average spread of the adjusted normals was approximately 5 degrees, though there were significant differences among the three types of surface reflectance, $F(2,2) = 11.1$, $p<.05$. A post hoc analysis revealed that the observers' judgments of textured surfaces were significantly less variable than those obtained for purely shaded displays ($p<.05$), but that there were no significant differences between the shiny and matte conditions.

We also measured the relative magnitudes of these variations in different directions. This was accomplished by mapping each judgment onto a unit sphere, where slant and tilt are represented as latitude and longitude respectively. The total deviation of a given judgment from the average of several sessions is given by the arclength between them on the unit sphere. To break this up into directional components, it is useful construct a
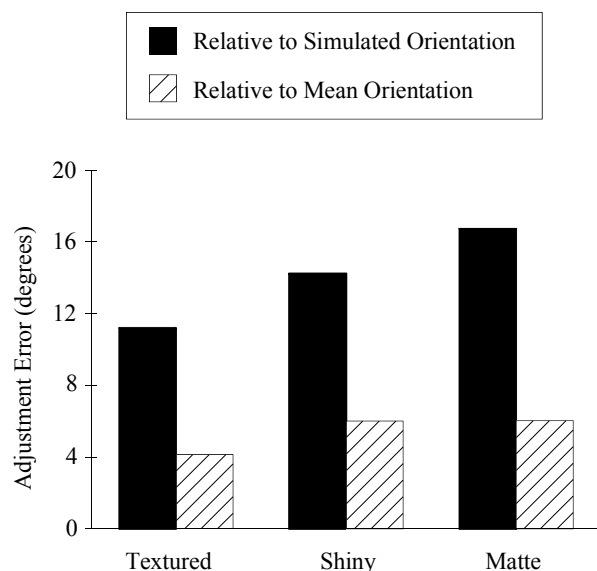

Figure 6 -- The average deviation in degrees between each individual judgment at a given probe point and 1) the simulated orientation at that point (solid bars), or 2) the average for that point over several repeated trials (striped bars).
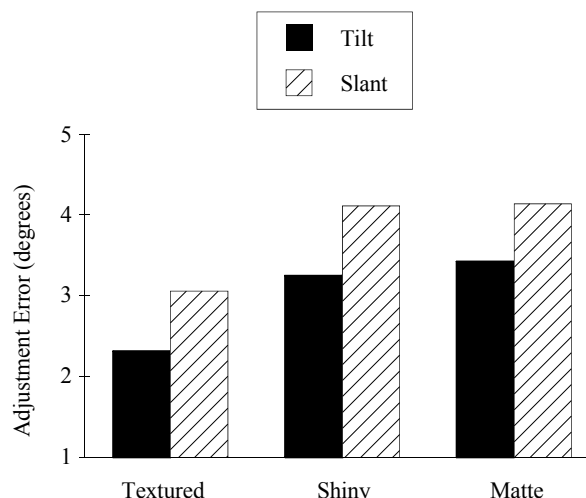

Figure 7 -- The slant and tilt components of arclength on a unit sphere for the average deviation in degrees between each individual judgment at a given probe point and the average for that point over several repeated trials.

spherical right triangle, where the total deviation is the hypotenuse, one leg lies on a line of longitude (i.e., the deviation in slant), and the other leg is perpendicular to that. For the sake of convenience, we will refer to this last component as the deviation in tilt, though the reader should be advised that this is a nonstandard usage of that term. Figure 7 shows the relative magnitudes of these different components for each type of surface reflectance. An analysis of variance revealed that there was a significant anisotropy in the observers' judgments $F(1,2)=182.07$, $p<.01$, such that the average deviation in slant was about 25% larger than the deviation in tilt.

Another possible method for measuring the variance in different directions is to compare the average depth gradient at each point with the gradients obtained for each individual judgment in the direction of average slant and in a perpendicular direction (see Koenderink, van Doorn & Kappers, 1992, 1996). The anisotropy of slant and tilt in the present data as revealed by the gradient measure is about 4 times larger than the one shown in Figure 7 for our analysis of the surface normals. This result is predictable from the mathematical relationship between the two measures, which is described by the following equation:

$$\frac{\Delta G_{\tau}}{\Delta G_{\sigma}} = \frac{\tan(\Delta N_{\tau})}{\tan(\Delta N_{\sigma})}[\cos(\sigma) - \sin(\sigma)\tan(\Delta N_{\sigma})]$$

where σ is the average slant setting at a given probe point, ΔN_s and ΔN_t are the slant and tilt components of variation in the adjusted normals, and ΔG_s and ΔG_t are components of variation in the adjusted gradients. If ΔN_s is relatively small as in the present experiment, then the gradient and normal measures will produce similar anisotropies for small surface slants, but the two measures will deviate as slant is increased.

To measure the accuracy of individual judgments, we calculated the angle between each orientation setting at a given probe point and the simulated surface normal at that point. The solid bars of Figure 6 show the average errors obtained for each type of surface reflectance. One important aspect of these data that deserves to be highlighted is that the average difference between the judged and simulated orientation at each point was significantly larger than the average deviation among multiple judgments, $F(1,2)=435.6$, $p<.01$, and that there was a significant interaction of this effect with variations in surface reflectance, $F(1,2)=22.2$, $p<.01$. These findings indicate that the errors cannot be due solely to random noise, and that there was a systematic pattern of distortion in the observers judgments.

To better understand the nature of these distortions, we computed the pattern of pictorial relief that is most consistent with the total set of attitude measurements for each observer and each condition using the method of surface reconstruction developed by Koenderink et. al. (1992). The basic procedure is as follows: For each vertex V_i in the lattice of probe points, a gradient vector $\mathbf{G}_i$ is computed from the observer's slant and tilt settings (s_i and t_i) using the following equation:

$$\mathbf{G}_i = (\cos t_i, \sin t_i) \tan s_i$$

Similarly, for each pair of adjacent vertices ($i = n$, m), a 2D connection vector is computed such that:

$$\mathbf{C}_{nm} = (X_n - X_m, Y_n - Y_m)$$

The depth difference D_{nm} between the two probe points can then be determined from the scalar product:

$$D_{nm} = \mathbf{C}_{nm} \cdot (\mathbf{G}_n + \mathbf{G}_m)/2.$$

The individual depths Z_n and Z_m are estimated -- up to a translation in depth -- from the least squares error solution of the over determined set of simultaneous equations ($Z_n - Z_m = D_{nm}$) defined by the depth differences for all pairs of adjacent vertices.

Each reconstructed surface from observers' judgments in different conditions was compared to the actual structure of the simulated object using an analysis of linear regression. The values of r^2 obtained from this procedure indicate the accuracy of the reconstructions up to a scaling in depth, and the slopes of the best fitting regression lines indicate the accuracy of the depth scaling. Table 1 shows the average values of these parameters for each observer and each type of surface reflectance. One of the first things to note in this table is that, on

Table 1 -- The r^2 values and slopes of regression for the correlations between simulated and reconstructed shapes. The value in each cell is an average of the six possible combinations of objects and directions of illumination.

	r^2			Slope		
	Textured	Shiny	Matte	Textured	Shiny	Matte
A.K.	0.872	0.795	0.762	0.751	0.698	0.584
J.K.	0.929	0.856	0.813	0.650	0.571	0.514
J.T.	0.915	0.900	0.864	0.706	0.699	0.573
Mean	0.905	0.850	0.813	0.702	0.656	0.557

average, the structure of the depicted objects accounted for approximately 85% of the total variance in depth on the reconstructed surfaces. An analysis of variance revealed, however, that there were small but significant differences among the different types of surface reflectance $F(2,2)= 14.7$, $p<.05$, such that the correlations were somewhat higher than average in the textured conditions and somewhat lower than average in the matte shading conditions. Another interesting aspect of the data is that the overall magnitude of pictorial relief in the reconstructed surfaces varied systematically across the different types of surface reflectance $F(2,2)=32.6$, $p<.01$. In general, the textured surfaces produced slightly more relief than those that were smoothly shaded, and shiny surfaces produced slightly more relief than matte surfaces.

In examining the detailed structure of the reconstructed surfaces, we noticed in many instances that the deviations from the simulated objects involved partwise distortions, in which some local regions had exaggerated relief relative to others (cf. Todd, et. al., 1996). These effects can be observed most clearly by examining the patterns of residuals for the linear correlations between the simulated and reconstructed surfaces. Figure 8 shows a gray scale map of the residuals for Astrid Kappers for all three objects with matte Lambertian shading and overhead illumination. These are superimposed on stereograms of the actual depicted objects in order to reveal their relation to the underlying surface. Note in each case that the patterns are highly structured, thus suggesting the residuals were due to systematic distortions of the observer's judgments rather than random noise. Moreover, it is also interesting to note in these patterns that the iso-residual contours seem to be aligned with perceptually distinct surface parts such as hills and ridges.

In order to analyze the reliability of these distortions, we performed test-retest correlations of the reconstructed surfaces for the subset of six conditions that each observer repeated over two experimental sessions. The results of this analysis are shown in Table 2. It is clear from these data that the observers' judgments were remarkably reliable. The average r^2 over all of the different conditions was 0.975, so that the variance between the reconstructed surfaces from different experimental sessions was less than 3%.

Table 2 -- The r^2 values and slopes of regression for the test-retest correlations between repeated sessions of the same condition. The value in each cell is an average of the two possible directions of illumination.

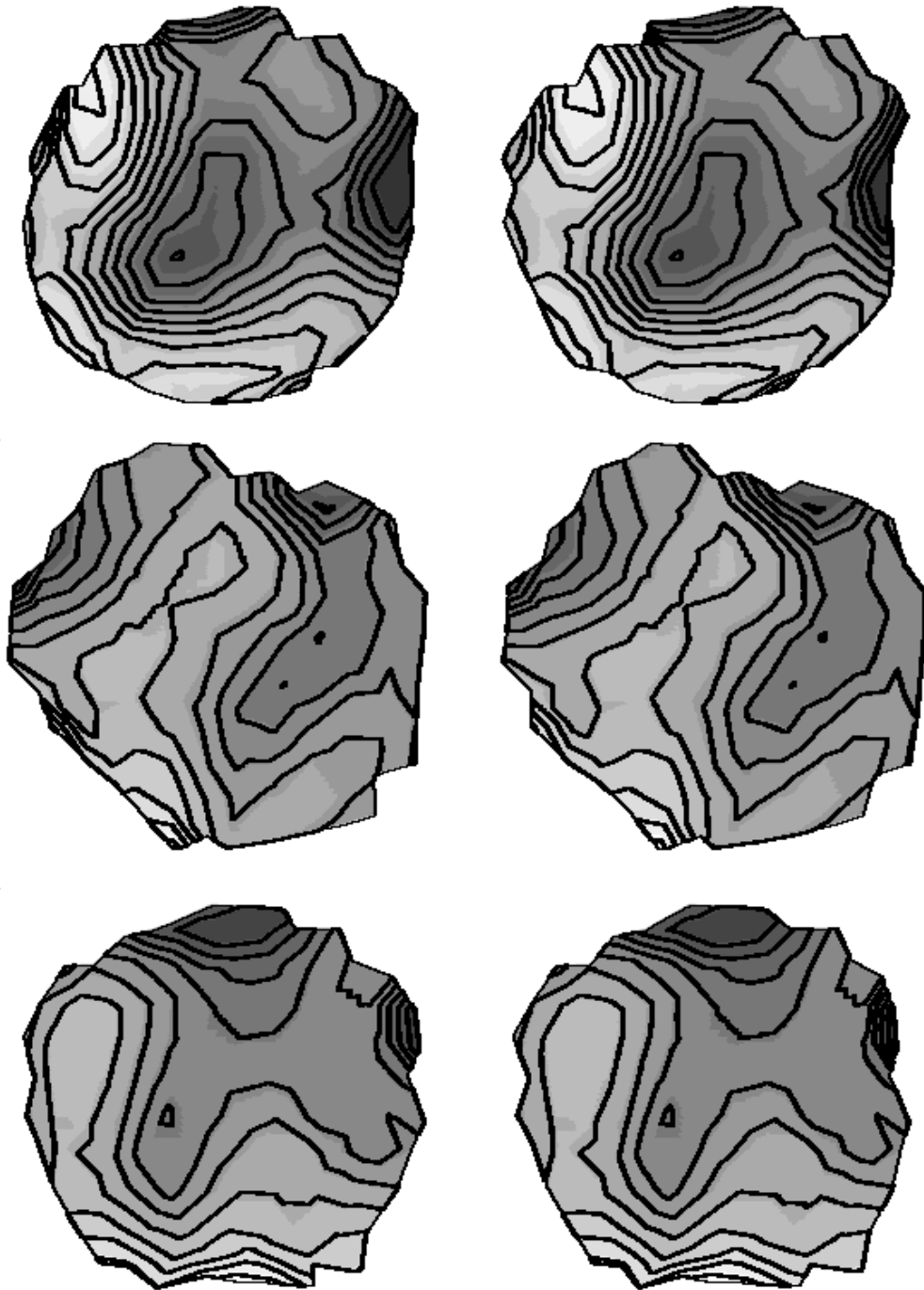|	r^2			Slope		
	Textured	Shiny	Matte	Textured	Shiny	Matte
A.K., Shape 1	0.992	0.978	0.952	0.961	0.930	0.880
J.K., Shape 2	0.952	0.961	0.979	0.947	1.022	0.992
J.T., Shape 3	0.972	0.969	0.966	0.972	0.946	1.019
Mean	0.979	0.972	0.973	0.960	0.966	0.963

Figure 8 – A gray scale map of the residuals produced by Astrid Kappers from the linear correlations between judged and simulated depth for all three objects with matte Lambertian shading. White contours against dark regions indicate positive values, and black contours against light regions indicate negative values. These patterns are superimposed on stereograms of the depicted objects in order to reveal their relation to the underlying surfaces.

Table 3 -- The r^2 values and slopes of regression for the correlations between different re-flectance conditions. The value in each cell is an average of the six possible combinations of objects and directions of illumination.

	r^2			Slope		
	Textured vs. Matte	Textured vs. Shiny	Shiny vs. Matte	Textured vs. Matte	Textured vs. Shiny	Shiny vs. Matte
A.K.	0.869	0.916	0.928	0.777	0.931	0.893
J.K.	0.893	0.927	0.899	0.780	0.864	0.962
J.T.	0.931	0.954	0.953	0.806	0.976	0.857
Mean	0.898	0.932	0.927	0.787	0.924	0.904

Additional analyses were performed to examine the constancy of observers' judgments across the different experimental conditions. For each combination of observer, object, and direction of illumination, an analysis of linear regression was performed to compare the reconstructed surfaces for each pairwise combination of reflectance conditions. Table 3 shows the r^2 values obtained from this analysis and the slopes of the best fitting regression lines. The results reveal a high degree of constancy for the relative depths of different surface parts (i.e., the average r^2 was .919), but that there were systematic differences among the overall magnitudes of depth scaling for the textured, shiny and matte conditions. In general, the reconstructed surfaces for the textured condition had about 10% more relief than those obtained for shaded surfaces with specular highlights, and the reconstructed shiny surfaces had about 10% more relief than those with matte Lambertian shading.

Table 4 shows the average r^2 values and the slopes of regression for the linear correlations between the two different directions of illumination when all other factors were held constant. An analysis of variance revealed that there were significant differences among the r^2 values obtained for the different types of surface reflectance, $F(2,2)=81.5$, $p<.01$. Note that there was almost perfect illumination constancy for the textured surfaces, but that the correlations were reduced somewhat for the purely shaded displays, especially those with matte Lambertian reflectance functions. Although it is sometimes argued that human observers have a bias to assume overhead illumination in the perceptual analysis of image shading, we found no evidence to support that hypothesis in the present data for stereoscopic displays (see Reichel & Todd, 1990; Todd et al, 1996).

Table 4 -- The r^2 values and slopes of regression for the correlations between different directions of illumination. The value in each cell is an average of the three objects.

	r^2			Slope		
	Textured	Shiny	Matte	Textured	Shiny	Matte
A.K.	0.986	0.958	0.911	0.962	0.956	0.972
J.K.	0.960	0.942	0.867	0.970	0.924	1.097
J.T.	0.965	0.944	0.897	0.997	0.972	0.993
Mean	0.970	0.948	0.892	0.976	0.951	1.021

Table 5 -- The r^2 values and slopes of regression for the correlations between observers. The value in each cell is an average of the six possible combinations of objects and directions of illumination.

	r^2			Slope		
	Textured	Shiny	Matte	Textured	Shiny	Matte
A.K. vs J.K.	0.919	0.867	0.799	0.809	0.729	0.749
J.K. vs J.T.	0.866	0.816	0.784	1.026	1.109	0.995
J.T. vs A.K.	0.753	0.742	0.778	0.941	0.900	0.961
Mean	0.846	0.808	0.787	0.925	0.913	0.902

One final series of regressions was performed to examine the similarity of the reconstructed surfaces for different observers. Table 5 shows the average r^2 values and slopes of regression for the different types of surface reflectance. It is especially interesting to compare these findings with the correlations between the simulated and reconstructed shapes shown in Table 1. The fact that the correlation between observers is consistently lower than the correlation with the actual depicted object indicates that that the partwise distortions in the surface reconstructions of different observers are largely independent of one another.

To summarize the results from the various regression analyses, Figure 9 shows a histogram of the average r^2 values for the linear correlations between of the reconstructed surfaces and the actual simulated objects, and those obtained for the reconstructed surfaces over different observers, different directions of illumination or different experimental sessions. Note in the figure that about 15% of the variance in the pattern of judged relief is due to partwise distortions in the 3D structure of the depicted object. These distortions are highly reliable over different experimental sessions or directions of illumination, but they vary considerably among different observers. In general, the judged pattern of relief is most reliable for textured surfaces, and least reliable for smoothly shaded Lambertian surfaces.
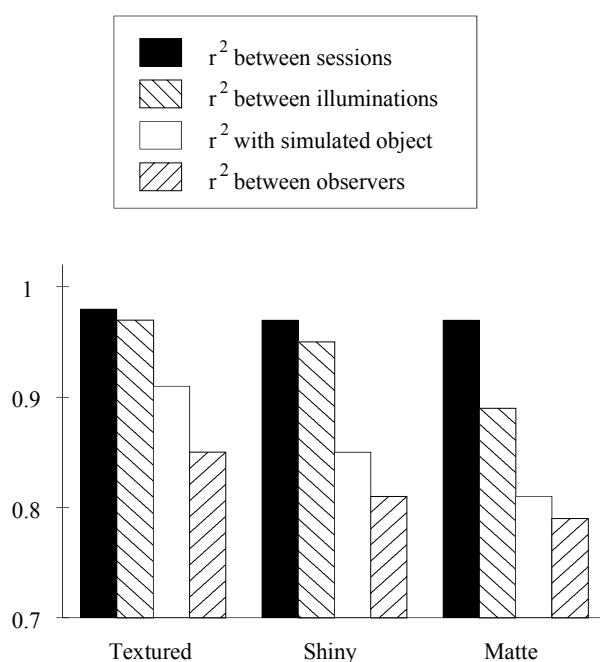

Figure 9 -- The average r^2 values for various types of correlations performed for the textured, shiny and matte reflectance conditions.

Discussion

In considering the results of the present study it is important to keep in mind the limitations that are inherent in all matching paradigms. For example, if an observer adjusts a gauge figure so that it is perfectly aligned with the tangent plane of a depicted surface, that would not necessarily indicate that the observer has an accurate perception of local orientation. In order to draw such a conclusion, it would be necessary to assume that the orientation of the gauge figure is perceived veridically, which need not be the case. Thus, a reconstructed surface computed from these adjustments should not be misinterpreted as a precise model of an ob-

server's perceptual representation of 3D form. We do think it reasonable to assume, however, that the endpoints of adjusted slant (i.e., 0º or 90º) are perceived veridically, and that intermediate settings are monotonically related to perceived orientation.

In order to compare performance across different experimental conditions, it is also necessary to assume that the relationship between adjusted and perceived orientation remains stable over time for each individual observer. One way of testing this assumption we have employed in previous investigations is to measure the spread of observers' adjustments to individual probe points over repeated trials (e.g., see Figures 6 and 7; Koenderink et. al., 1992; Todd et. al., 1996). In the present experiment, we also added an additional manipulation to measure the consistency of the reconstructed surfaces over multiple experimental sessions. A regression analysis of these data revealed that the judged patterns of pictorial relief had almost perfect test-retest reliability. That is to say, when a given object was judged in different experimental sessions separated by several days, the linear correlation between the two reconstructed surfaces had an average r^2 of .97.

When we examined the specific patterns of distortion in observers' judgments, they seemed to involve two distinct types. First, there were errors in the overall magnitude of judged relief, which are revealed most clearly by the slopes of the best fitting regression lines between the reconstructed surface depths and the actual simulated objects (see Table 1). These ranged from about .56 for the smoothly shaded Lambertian surfaces to approximately .7 for those with texture. Second, there were also errors in the relative magnitudes of judged relief for different parts of the same object. These are revealed by the values of r^2 for the linear correlations between the simulated and reconstructed surfaces, and by the structure of their residuals.

It is important to point out in this context that while piece-wise distortions of judged surface structure have been reported previously by Koenderink, van Doorn & Kappers (1996) and Todd et. al. (1996), they have not been noticed by most other researchers within the vast literature on 3D form perception. One possible reason for this is that the typical methodologies employed to measure observers' perceptions do not have suffi-

cient power to detect complex patterns of distortion. In many experiments, observers' judgments of an entire object are limited to only one or two parameters, such as its height, width or depth. In order to provide a reasonably precise measure of judged 3D shape that is applicable to arbitrary surfaces, it is obviously necessary to provide observers with a much larger number of degrees of freedom in making their judgments. Another relevant factor that is likely to influence the occurrence of piece-wise distortions in judged surface structure is the overall complexity of the depicted objects. Perhaps it is the case, for example, that such distortions can only occur for objects that contain perceptually distinct parts (e.g., see, Biederman, 1987; Hoffman & Richards, 1984; Richards, Koenderink, & Hoffman, 1987). If this hypothesis is correct, then the patterns of distortion for simple surfaces such as cylinders, ellipsoids or pyramids, which have no distinct parts, may be quite different from those obtained with more complex objects like the ones shown in Figures 4 and 5.

One of our goals in designing this experiment, was to identify the specific aspects of local surface structure that define the boundaries between perceptually distinct parts. In our previous investigations relating to this issue (Koenderink, van Doorn, & Kappers,1996; Todd, et. al., 1996), the stimuli were created from computer scanned photographs, which do not allow a direct comparison of observers' judgments with the actual 3D structure of a depicted object. To correct this limitation in the present study, the stimulus displays were all generated from mathematically defined surfaces, so that the local differential structure at each probe point could easily be determined with a high level of precision. Unfortunately, the data were less revealing than we had hoped about the perceived parts of these objects. Although the reconstructed surfaces for individual observers were highly reliable over multiple experimental sessions, there were large variations among different observers in their part-based patterns of distortion.

Another important goal of the present experiment was to evaluate the effects of various types of surface reflectance on stereoscopic form perception. In a previous investigation to address this issue, Bülthoff & Mallot (1988) found that observers' judgments were perfectly veridical for surfaces with polygonal texture, but that the judged pattern

of relief was systematically underestimated by approximately 30% for smoothly shaded curved surfaces. A similar difference between smoothly shaded and textured surfaces was also obtained in the present experiment, although the magnitude of this effect was somewhat smaller. There were, however, several other aspects of the data for which our results differed from those reported by Bülthoff & Mallot (1988): First, we did not obtain veridical performance with textured surfaces; second, the judged surfaces were systematically distorted in a piece-wise manner; and third, there were significant differences in the magnitude of perceived relief for surfaces depicted with specular and Lambertian reflectance functions.

There are a number of methodological differences between these studies that could potentially account for their discrepant results. One such difference, that we suspect could influence the occurrence of piece-wise distortions, is that the objects used in the present experiment contained relatively complex arrangements of hills, valleys, and ridges, whereas the ones used by Bülthoff & Mallot (1988) were all ellipsoids of revolution about the line of sight. Another potentially important difference in the design of these studies involves the specific response tasks the observers were required to perform -- i.e., observers adjusted a monocular orientation probe in the present experiment, whereas a stereoscopic depth probe was used by Bülthoff & Mallot (1988). Koenderink, Kappers, Todd, Norman and Phillips (1996) have recently demonstrated that these response tasks produce similar results for smoothly shaded surfaces, but there are good reasons to suspect that this may not be true for surfaces that are textured. It is likely in that case that observers would adjust the stereoscopic depth probe so as to match its local disparity with that of nearby texture elements, without necessarily knowing its precise metric structure relative to other elements in the scene. Previous research has shown that observers can exhibit hyperacuity at detecting relative disparity of targets presented near the fixation plane (e.g., Ogle, 1953; Westheimer & McKee, 1978), but most investigations of perceived metric structure from edge based stereo have found it to be systematically distorted (e.g., Johnston, 1991; Norman, et. al., 1996; Tittle, et. al., 1995; Todd, et. al., 1996).

In comparing the influences of different type of surface reflectance on stereoscopic form perception, we were particularly interested in how performance would be affected by the presence of specular highlights. It is generally assumed in the theoretical analysis of binocular stereopsis that a matched pair of image features between two stereoscopic views must be the optical projections of the same physical point in 3-dimensional space. Although this assumption is satisfied for the binocular correspondence of changes in image intensity that arise from Lambertian shading or texture, that is not the case for specular highlights. Because highlights change as a function of the viewing direction, they can appear at different 3D positions for each eye, so that their virtual binocular images are displaced relative to the observed surface (see Figure 3). The results of the present experiment provide clear evidence, however, that specular highlights can be a useful source of information for the process of stereoscopic form perception rather than a source of unwanted noise. Over all of the various performance measures we examined, observers judgments for smoothly shaded shiny surfaces were significantly more accurate and reliable than those obtained for Lambertian surfaces.

One possible method for exploiting this information in the computational analysis of binocular stereopsis has been proposed by Blake and Bülthoff (1990, 1991), who showed that relative disparity between a specular highlight and a nearby surface marking varies as an inverse proportion of the magnitude of surface curvature, except in certain degenerate cases. We believe it is unlikely, however, that observers in the present experiment would have been able to perform such an analysis, since all of our specular displays were smoothly shaded with no identifiable surface markings. Another relevant issue to consider in this context is the extent to which performance depends on the ability of observers to identify a highlight as such in order to correctly interpret its binocular disparity. It is important to keep in mind that the specular highlights in our displays had a different chromatic structure than the Lambertian components of shading (i.e., the highlights were white, whereas the Lambertian components were blue). An interesting question for future research is whether there is any deterioration of performance for monochro-

matic images in which the specular and Lambertian components might be more difficult to perceptually segregate (e.g., see Figure 5).

One final aspect of the observers' judgments that deserves to be considered is that they remained largely invariant over the different possible directions of illumination. Indeed, the linear correlations between the judged patterns of relief for surfaces depicted with overhead and side illumination produced average r^2 values of .97, .95 and .89, respectively, for the textured, shiny and matte conditions. It is interesting to note while evaluating this issue that there have been a number of conflicting reports in the literature. For example, in the monocular perception of shape from shading Koenderink, van Doorn and Christou (1995) found that the judged pattern of relief for spherical objects was sheared slightly toward the direction of illumination, but no such effects were obtained by Todd and Reichel (1989) for observers' judgments of shaded ellipsoids. Similarly, for stereoscopic displays, we found no effects of illumination tilt in the present experiment, though Todd, et. al. (1996) obtained large variations in the judged pattern of relief over changes in illumination slant. It will remain for future research to sort out these findings in order to obtain a more complete understanding of how different patterns of illumination may influence observers' perceptions of 3D form.

References

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

Blake, A., & Bülthoff, H. H. (1990) Does the brain know the physics of specular reflection. *Nature, London*, *343*, 165-168.

Blake, A., & Bülthoff, H. H. (1991) Shape from specularities: Computation and psychophysics. *Philosophical Transactions of the Royal Society of London B*, *331*, 237-252.

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

Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993) Perception of local shape from shading. *Perception & Psychophysics*, *54*, 145-156.

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

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*, 1351-1360.

Julesz, B. (1971) *Foundations of cyclopean perception*. London: The University of Chicago Press Ltd.

Koenderink, J. J., and van Doorn, A. J. (1995). Relief: pictorial and otherwise. *Image and Vision Computing*. *13*, 321-334.

Koenderink, J. J., van Doorn, A. J., and Christou, C. (1995). Shape constancy under variation in illumination. *Perception*, *24* (Supplement), 34.

Koenderink, J. J., van Doorn, A. J., and Kappers, A. M. L. (1992). Surface perception in pictures. *Perception and Psychophysics*, *52*, 487-496.

Koenderink, J. J., van Doorn, A. J., and Kappers, A. M. L. (1994). On so-called "paradoxical monocular stereoscopy". *Perception*, *23*, 583-594.

Koenderink, J. J., van Doorn, A. J., and Kappers, A. M. L. (1995). Depth relief. *Perception*, *24*, 115-126.

Koenderink, J. J., van Doorn, A. J., and Kappers, A. M. L. (1996). Pictorial surface attitude and local depth comparisons. *Perception & Psychophysics*, *58*, 163-173.

Koenderink, J. J., Kappers, A. M. L., Todd, J. T., Norman, J. F. & Phillips, F. (1996) Surface and attitude probing in stereoscopically presented dynamic scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 869-878.

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

Norman, J. F., Todd, J. T., Perotti, V. J. & Tittle, J. S. (1996) The visual perception of 3D length. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 173-186.

Norman, J. F., Todd, J. T., & Phillips, F. (1995) The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, *57*, 629-636.

Ogle, K. N. (1953) Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America*, *43*, 906-913.

Reichel, F. D. & Todd, J. T. (1990) Perceived depth inversion of smoothly curved surfaces due to image orientation. *Journal of Experi-*

mental Psychology: Human Perception and Performance, 16, 653-664.

Reichel, F. D., Todd, J.T., Yilmaz, E., (1995) Visual discrimination of local surface depth and orientation. *Perception & Psychophysics, 57*, 1233-1240.

Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1987) Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America A, 4*, 1168-1175.

Stevens, K. (1983) Slant and Tilt: The visual encoding of surface orientation. *Biological Cybernetics, 46*, 183-195.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995) The systematic distortion of perceived 3D structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 663-678.

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 and Performance, 22*, 695-706.

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

Todd, J. T., & Mingolla, E. (1984). The simulation of curved surfaces from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 734-739.

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

Westheimer, G. & McKee, S. P. (1978) Stereoscopic acuity of moving retinal images. *Journal of the Optical Society of America, 68*, 450-455.

Acknowledgments