Multiscale Evaluation of 3D Shape Perception in Computer Graphics

by

Peter H. Brown

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill

2002

Approved by:

Christina A. Burbeck, Advisor

Stephen M. Pizer, Reader

Mark Hollins, Reader

Russell M. Taylor, II, Reader

PETER H. BROWN. Multiscale Evaluation of 3D Shape Perception in Computer Graphics.
(Under the direction of Christina A. Burbeck.)

ABSTRACT

This dissertation describes a tool and associated analysis techniques (collectively called the Multiscale Test of Perceived Shape, or MTOPS) for measuring observers' perceptions of 3D surface orientation in computer-graphics visualizations at multiple spatial scales. Using this tool and techniques, I demonstrated that such multiscale measurements are both possible and fruitful. Specifically, I showed the following:

- Observers consistently made different surface-orientation settings at different scales, indicating both that perceived orientation changed across scale and that MTOPS was able to measure that change.
- By comparing observers' settings to the stimulus geometry, I could measure the across-scale effects of adding a given depth cue to a visualization. Specifically, I found the following:
  - Adding stereo to a non-stereo presentation caused the accuracy of observers' settings to improve more at large scale than at small.
  - Adding head-motion parallax to a stereo presentation caused a modest improvement in observers' accuracy, mostly at small scale.
  - Adding automatic object motion to a non-stereo presentation improved observers' accuracy substantially, at many scales. Adding user-controlled object motion to a non-stereo presentation gave a less consistent, and for most users smaller, accuracy benefit.
- The MTOPS tool and techniques were sufficiently sensitive to capture differences in those multiscale depth-cue effects due to surface complexity and observer knowledge.

Unless the LORD builds the house,
those who build it labor in vain.
Unless the LORD guards the city,
the guard keeps watch in vain.

Psalm 127:1, New Revised Standard Version

All things come of thee, O LORD,
and of thine own have we given thee.

Offertory sentence, 1928 Book of Common Prayer,
after 1 Chronicles 29:14b

Acknowledgments

Bless the Lord, O my soul,
    and all that is within me,
    bless his holy name.
Bless the Lord, O my soul,
    and do not forget all his benefits—
who forgives all your iniquity,
    who heals all your diseases,
who redeems your life from the Pit,
    who crowns you with steadfast love and mercy.

Psalm 103:1-4, New Revised Standard Version

Those to whom I owe a debt of gratitude in the completion of this dissertation are truly more than I can name. I will try anyway, and depend on the forbearance of those I omit not to take it personally.

Second on any list must come my wife Laura, for her long-tried support both professional and personal. Her readings and comments on both structure and argument have pulled me out of many a blind alley and contributed substantially to the readability of the resulting document; without the steadfastness of her emotional and spiritual support, I could not have brought this project to completion.

That my children, Hannah and Nathan, have not been equally long-suffering is due only to the fact that they were not yet born when I started. Like Laura, they have endured many hours of Daddy's distraction, as I wrestled with the problems both technical and emotional that caused this process to stretch over eleven years.

Also in the family category are my parents, who taught me to love truth and to persevere in its pursuit, and my parents-in-law, who managed to remain supportive even when they really had no idea what could possibly be taking me so long. My father-in-law did not live to see me graduate, but he died believing it would happen.

On the professional side, thanks almost equal to my wife's are due to my advisor, Christina Burbeck, both for her patience far beyond the call of professionalism and for her legions of perceptive criticisms. She first suggested this topic to me, she pulled me all but bodily through the writing of the paper which formed the core of Chapters 4 and 3, and her contributions to my understanding of human vision are too many to enumerate. My two prior advisors, David Beard and Frederick P. Brooks, Jr., are also due their fair share of credit for my growth as a scholar to the point where I can actually produce Ph.D.-level work. Special thanks are due to Stephen Pizer among my committee, both for for his important contributions to the geometric ideas in this dissertation and for enduring me longer than anyone else. Mark Hollins, Gary Bishop, and Russell Taylor each contributed helpful comments and insights.

I would be remiss as well not to thank the observers who volunteered their time and effort. Without their patience, this work could not have been done.

Many others contributed in ways both major and minor. I thank Joseph Grace for his counsel and encouragement over several years; Betty Coan for help with scanning figures; and the Converse College Department of English for the loan of a record player.

I also give thanks to and for the many whose prayers on my behalf over so many years surely availed much. I especially give thanks for the members of Vox Humana and the Thursday night Bible study at the Church of the Holy Family, Chapel Hill, NC; for the members of the choir and the Monday night prayer group at St. Christopher's Church, Spartanburg, SC; and for the many others both at Holy Family and at St. Christopher's whose prayers and actions have made the Body of Christ more than a metaphor.

Last, first, and always, I give thanks to the only Light of the world, who gave us eyes to see, and in whose light we see light.

Contents

List of Figures

Chapter 1

Motivation and Overview[1]

1.1 Major Claims

The main claim of this dissertation is that measuring perceived 3D shape in computer-graphics visualizations at multiple spatial scales is both possible and fruitful. The thesis demonstrates this using a novel tool and associated analysis techniques (collectively called the Multiscale Test of Perceived Shape, or MTOPS) for making multiscale measurements of perceived surface orientation.

I showed that multiscale measurements of perceived shape are possible. Observers consistently made different surface-orientation settings at different spatial scales, indicating that perceived surface orientation changed across scale.

I showed that multiscale measurements of perceived shape are fruitful. By comparing observers' surface-orientation settings to the stimulus geometry, I was able to measure the across-scale effects of adding a given depth cue to a visualization. Specifically, I found the following:

- Adding stereo to a non-stereo presentation caused the accuracy of observers' settings to improve more at large scale than at small.
- Adding head-motion parallax to a stereo presentation caused a modest improvement in observers' accuracy, mostly at small scale.
- Adding automatic object motion to a non-stereo presentation improved observers' accuracy substantially, at many scales. Adding user-controlled object motion to a non-stereo presentation gave a less consistent, and for most users smaller, accuracy benefit.

1.2 Motivation

For as long as 3D computer graphics have been used to visualize information about 3D shape (e.g., in presenting scientific data or in computer-aided design), computer-graphics researchers have evaluated those visualizations to find which gives the best user performance. Since 3D computer graphics allow 3D information to be communicated using the same depth cues humans use in everyday

[1]Some of the material in this chapter (specifically, much of Section 1.2.1) was previously published in [13], ©1999 Elsevier Science; it is used by permission.

vision, different computer-graphics presentations of the same data correspond to different depth-cue conditions, and comparing user performance on two presentations is equivalent to comparing user performance on the relevant task under the two depth-cue conditions. Such comparisons have enabled practitioners to measure the benefit of using a particular computer-graphics technique in a scientific visualization (e.g., [47]). This ability is particularly important if the technique under study is computationally expensive (e.g., providing object motion) or requires special hardware (e.g., the virtual-reality environments studied in [2, 88]). The ability to measure benefit further allows comparison of several competing depth-cue conditions to determine which is best for the task at hand (as in [75, 87, 132]). The tool and techniques presented in this dissertation, collectively known as the Multiscale Test of Perceived Shape (MTOPS), open up a new dimension in the evaluation of computer-graphics presentations of 3D shape information: spatial scale.

Spatial scale has long been known to be a pervasive and important factor in human visual performance [41], and its use in modeling and rendering has a long history. Such techniques as MIP-mapped textures [140], wavelet-based multiresolution analysis [29, 76], m-reps [91], and spatial-frequency-based approaches to artifact masking with textures[33], level-of-detail modeling [93], and antialiasing [82] are all implicitly or explicitly multiscale. But despite the use of multiscale approaches to modeling and rendering, spatial scale has not previously been used as an explicit variable in evaluation of 3D techniques. By using spatial scale in evaluation, MTOPS offers to bring answers to a range of evaluative questions which previous methods simply could not address.

1.2.1 Advantages of a Multiscale Measure

MTOPS is based on the premise that humans perceive 3D shape in a multiscale fashion, creating a small-scale representation for local judgments and a large-scale representation for judgments concerning the object as a whole[2]. Multiscale 3D shape perception is evident in the perceived similarity of shapes at one scale, even when the objects differ at other scales [61]. For example, people refer to an L-shaped piece of pipe as an “elbow” because of the obvious geometric similarity between the pipe elbow and our own. Clearly, however, this similarity does not extend either to larger or to smaller scales; the plumbing has little resemblance to an entire person, and the smooth fine structure of a pipe is not much like the skin, hair, and musculature of a human arm. Submarine sandwiches owe their name to the same kind of geometric similarity, although no one would confuse the fine structure of the sandwich with that of the submersible warship. The scale-specificity of these similarities clearly does not impede their perception; rather, people find the perception so compelling that it molds the language[3].

Because human observers perceive 3D objects at multiple spatial scales, measurements made at a single scale cannot capture everything the observer knows about the surface structure, even if those

[2]This idea is also certainly not new, although it is quite difficult to find explicit references to it in the spatial-vision literature. References to multiscale vision abound, and there are plenty of papers on shape; but the shape papers are either concerned with 2D shape (e.g., [17, 138]) or not clearly multiscale (e.g., [4, 78]). Multiscale 3D shape, it seems, gets mentioned in passing by papers which are primarily addressing something else [32, 61].

[3]Koenderink and van Doorn [61] confine their discussion to similarities in the surface mesostructure (which they define as that structure of an object which is smaller than its overall structure, or megastructure, by up to approximately two orders of magnitude). The perception of scale-specific similarities is not limited to the mesostructure, however, as the submarine example makes clear. Indeed, similarities are readily perceived between the mesostructure of one object and the megastructure of another; elbow noodles are separate objects, but this does not detract from their “elbowness.”

measurements are made at many locations on the surface. The experimenter can combine the local measurements into a global surface, as Koenderink *et al.* did [65], and analyze large-scale patterns in those small-scale data, but as Koenderink *et al.* noted, this surface is simply not the same as a "perceived surface" (even though it may resemble the original stimulus). The problem is one of inference; we simply do not know how the observer's large-scale and small-scale percepts relate. The representations underlying these percepts may have different characteristics; in particular, they may differ in their sensitivity to various depth cues. To know what the observer sees at different spatial scales, we need to measure the scale-specific percepts directly. Making such measurements means replacing a single-scale measure of perceived shape with measurements at multiple scales.

Measurements at multiple scales may provide important information for practitioners, because visual tasks—including those tasks that scientific visualizations are designed to support—exhibit scale-specific characteristics as well. It is usually possible, thinking about a task, to get at least an approximate idea of the scale of the important information. For example, in a molecular modeling application, anything smaller than an atom or larger than the molecules involved can usually be ignored, and the chemist will normally want to focus the most attention on structures somewhere in the middle of that range, at the size of some significant substructure of a molecule. Of course, it is the task which defines what constitutes a significant substructure. This is true across tasks; what is significant is defined by the task, and visual tasks generally have characteristic scale ranges within which the interesting information occurs. The precision with which interesting scale ranges can be identified can be expected to vary with how well the task is understood, but it is often possible to get a rough idea even for exploratory work; in the example above, it is possible at least to bound the interesting scales without specifying the task at all beyond "molecular modeling".

Measurements at multiple scales are exactly what MTOPS provides. Evaluating the effect of a computer-graphics technique on perceived 3D shape with MTOPS produces not a single measure showing the benefit of the new technique, as previous methods do, but an across-scale pattern of the benefit of adding the new technique. Thus, for example, when MTOPS was used to evaluate the benefit of adding stereo to a non-stereo presentation, the results (presented in chapter 4) show a much greater accuracy benefit from stereo at large scale than at small. For a visualization task that places a premium on understanding small-scale shape, these results suggest that stereo may not be as helpful as it would be for a large-scale task. For another example, when MTOPS was used to evaluate the benefits of adding object motion to a static presentation, the results (presented in chapter 5) show that the across-scale pattern of benefit from active (user-controlled) motion varied widely between individual observers. Previous methods might have detected a difference in the amount of benefit between observers (depending on the scale at which the measurement was taken); no previous method could have detected the variation in the across-scale patterns.

1.3 Overview

Following this introduction and a review of the pertinent literature, this dissertation describes the MTOPS tool and analysis techniques in detail. It then describes three experiments performed to validate the MTOPS tool and techniques; one measuring the effect of adding stereo to a non-stereo presentation, one measuring the effect of adding head-motion parallax to a stereo presentation, and one measuring the effect of adding object motion (both user-controlled and automatic) to a static presentation. It concludes with a general discussion and description of future work.

Chapter 2 of the dissertation reviews the pertinent literature and locates the MTOPS work within it. After a quick review of geometric terms, I review many of the depth cues used by human observers; then I review also the various models of how the information from different depth cues may or may not be integrated to form a unified percept of a surface. I describe the types of evidence for the importance of shape in human vision, discuss models of cue integration and the shape representation underlying human vision, and review prior studies that have found evidence of multiscale effects in particular depth cues.

Chapter 3 presents the MTOPS tool and techniques in detail. First, it describes the orientation probe used by observers to indicate perceived shape at a scale, and then it explains the stimulus surfaces used in experiments. Then it describes the MTOPS analysis techniques. The most fundamental of those techniques is the orientation path, a set of orientation measurements taken at different scales, linked into a path to show the evolution of the orientation across scale. I describe how, by studying the relationship of the orientation path formed from observers' settings to the orientation path calculated directly from the object, across-scale patterns in perception of the object's shape can be revealed. Further measures provide summaries of the ways in which accuracy and precision vary across scale, over all test locations on the surface of the object; the bias measure summarizes the accuracy of observers' settings, and the variability measure summarizes the repeatability of the observers' judgments.

Chapter 4 details an experiment in which the MTOPS probe and techniques were used to measure the effect of adding stereo to a non-stereo presentation. The chapter presents results showing that observers could consistently judge different surface orientations at different spatial scales and that the multiscale data generated by MTOPS could reveal multiscale differences between observers' percepts of a stimulus and the stimulus geometry. Those multiscale differences, in turn, indicate the multiscale effects of adding the stereo cue to the non-stereo presentation. The MTOPS probe and techniques are shown to be sensitive enough to capture differences in those multiscale effects due to the complexity of the object and to capture differences in those multiscale effects due to some subjects' prior knowledge of the stimulus geometry.

Chapter 5 describes two more experiments, in which the MTOPS tool and techniques are applied to depth cues of particular computer-graphics interest. In the first, MTOPS is used to measure the effect of adding head-motion parallax to a stereo presentation; in the second, MTOPS is used to measure the effect of adding two kinds of object motion to a static presentation. The second experiment also uses objects far more complex than the objects used in the earlier experiments. Taken together, the experiments demonstrate that multiscale measurement of perceived surface orientation with MTOPS remains possible and fruitful with cue combinations and surfaces quite different from those studied in Chapter 4. Specifically, the experiments find that the improvement in accuracy due to adding head-motion parallax to a stereo presentation, reported in previous work, is absent at medium and large scales; that the improvement in accuracy due to adding object motion to a static, non-stereo presentation happens across many scales; and that the accuracy benefit due to adding automatic object motion to a static presentation is far more consistent than the accuracy benefit due to adding user-controlled motion.

Chapter 6 lists the contributions of the dissertation. It concludes with discussion and some opportunities for future work.

Chapter 2

Background

This chapter locates the MTOPS work within the context of the relevant literature. It begins with a quick description of some geometric terms used throughout the dissertation. The following section describes the depth cues relevant to the MTOPS work. Next is a description of some of the models that attempt to describe how the information from the various depth cues is integrated into a single percept of a surface, some of the problems with those models, and some of the implications for humans' mental representation of 3D shape from vision. The final section deals with spatial scale; it covers the types of evidence for its importance, some multiscale models of shape perception, and prior evidence for the existence of multiscale effects in particular depth cues.

2.1 Geometric Measures

Many of the applications of computer-graphics visualizations attempt to convey a metric understanding of shape (see, for example, Forrest's feeling comments in [37]). Most studies of human vision also perform their measurements of perceived shape in Euclidean 3-space; some measures have been developed to meet the specific requirements of vision which are not as commonly used in other applications of geometry. This section is a brief review of those measures.

2.1.1 Slant and Tilt

Many studies of shape perception measure perceived surface orientation (e.g., [129, 50, 65, 81, 87, 108, 107, 110]). This orientation is measured relative to the observer's viewpoint. Because 3D vision is anisotropic—two dimensions in the world correspond directly to dimensions on the retina, but structure in the depth dimension must be inferred—it is common practice, following Gibson [38], to decompose the surface orientation into two angles, one measured in the picture plane and the other measuring the inclination of the surface with respect to the picture plane. The angle in the picture plane is the tilt; it is conventionally measured clockwise from a vector in the picture plane pointing to the observer's right. The inclination of the surface with respect to the picture plane is called slant.

The orientation of any visible surface can thus be expressed by a tilt value between -180 and 180 degrees and a slant value between 0 and 90 degrees. Tilt, as a circular quantity, is periodic; any convenient 360-degree range can be used. Slant is not periodic.

2.1.2 Shape Index and Curvedness

Studies of perceived curvature (e.g., [32, 89, 94, 27, 128]) often make use of a local shape index defined by Koenderink [55]. The shape index (see Figure 2.1) assigns a value in the interval $[-1, 1]$ to a surface at a point, based on the ratio of the principal curvatures at the point. (The principal curvatures at a point are the normal curvatures measured in the directions of maximum and minimum curvature.) The curvedness expresses the flatness of the local patch, so that together the shape index and the curvedness determine the osculating quadric—and thus the local curvatures—at any point. Put another way (following Koenderink's original description), the shape index captures the "qualitative" shape of the local patch—that quality of the shape that can be described non-numerically, as "elliptic" or "cylindrical" or "hyperbolic"—and the curvedness provides the appropriate scaling factor to quantify the shape and make the shape index fit the surface.

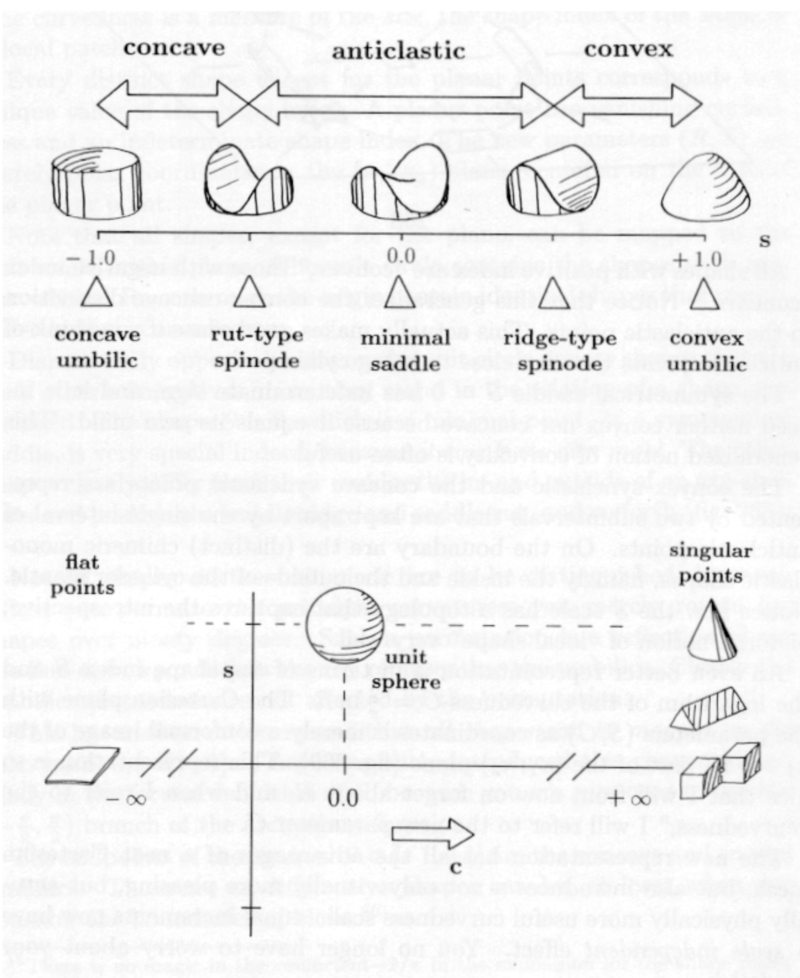

Figure 2.1: Koenderink's shape index, from [55].

2.2 Shape Cues

Humans do not have direct perceptual access to the 3D shape of objects around us; the 3D geometric information is, at best, ambiguously specified by the light energy available to us at any given moment. Instead, people infer the 3D structure of our surroundings from different shape cues, which exploit in more or less heuristic fashion certain regularities of visual experience. The most important of these cues form the foundation of 3D computer graphics, allowing the computer to communicate 3D shape information using (largely) flat displays. The cues most important to computer graphics are the following[1]:

- Accommodation
- Eye vergence
- Occlusion
- Object contour
- Linear perspective
- Relative size
- Texture
- Surface shading
- Cast shadows
- Stereopsis
- Motion parallax
- Head-motion parallax

Of these, the first two can be classed as physiological cues, having to do with the the muscles of the eye. The next seven are known as pictorial cues, since they are present in a single, static picture. Stereopsis, which exploits the differences between two static images, is in a class by itself. The last two are dynamic cues, since they rely on the changes in an image caused by the motion of the object viewed or of the viewer.

2.2.1 Physiological Cues

Accommodation is the cue resulting from the lens of the eye focusing at different depths. It is not simulated in computer graphics; it is not currently practical to track the depth at which a user's eye is focused. In scientific visualizations images are always presented with everything in focus; there is no reason to put an object in the image unless the viewer might want to look

[1]This is not an exhaustive list of the cues used in computer graphics. For example, it does not deal with depth-blurring, in which distant objects in the computer-graphics scene are blurred to simulate the blurring effect of light passing through air. The list does, however, contain the most important cues, and it contains all the cues dealt with in this dissertation.

at it. Fortunately for computer-graphics practitioners, accommodation is a weak depth cue, so the mismatch between accommodation and other cues is seldom a problem (although it can cause trouble in virtual environments, as will be discussed in Section 2.2.3).

Vergence is the cue resulting from the angle between the lines of sight of the two eyes looking at an object. Vergence is automatically simulated in stereo displays and is not simulated otherwise. Humans do perceive 3D structure despite conflicting vergence information (e.g., in a flat picture), so computer-graphics practitioners generally ignore vergence except as a source of eyestrain in stereo displays (this problem is described in Section 2.2.3).

2.2.2 Pictorial Cues

Projective Cues

The first two of the pictorial cues, occlusion and object contour, are direct consequences of projecting a 3D scene onto a 2D image plane (or a retina). They are generated by any type of projection.

Occlusion is the apparent overlap of one object (or part of an object) on another in an image; the occluding object (or part) appears nearer than the occluded object (or part). Occlusion provides only limited information, since it specifies only the order of the objects in depth. It is, however, one of the strongest depth cues; unlike other depth cues, which can be overridden, I know of no case where an occluded object can be made to appear to be in front of its occluder. Information from other depth cues can sometimes control occlusion, as in the stereo-occlusion experiments of Shimojo *et al.* [103] and Nakayama *et al.* [84], but even then the effect is to change the grouping of image regions into perceived objects, rather than to re-order the original objects in depth.

The object contour is formally defined by Koenderink [55] as the projection of the rim into the image; the rim is defined as the locus on the object surface dividing a visible region of surface from an invisible one. The outline of an object is thus all contour; contour occurs inside the outline of an object as well, however, wherever a part of the surface hides another part. Figure 2.2 shows one of the objects used in the MTOPS experiments, with its contour.

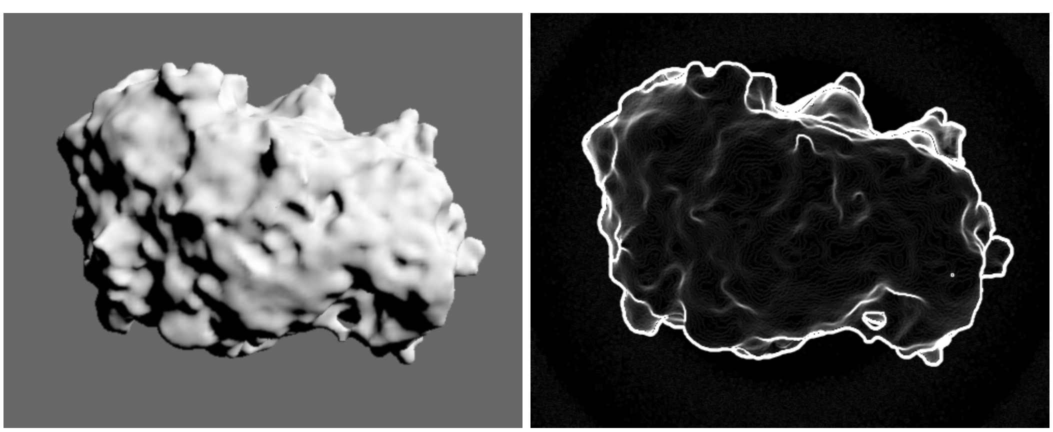

Figure 2.2: A random object (on the left), and its contour (on the right). The approach of Saito and Takahashi [100] was used to find the contour.

Each location on the contour corresponds to a location on the rim where the viewing ray is tangent to the surface (for a physical surface, since physical surfaces do not have infinitely thin corners); the shape of the contour conveys a good deal of information about the shape of the surface in the vicinity of that location on the rim. Koenderink shows that the curvature of the contour at a given location on the contour is related to the surface curvature at the corresponding point on the rim; specifically, convex sections of contour correspond to convex elliptic regions on the surface, concave sections of contour correspond to hyperbolic regions on the surface, and inflections on the contour correspond to inflections (parabolic curves, which separate hyperbolic from elliptic regions) on the surface. Occlusions, too, are indicated by features on the contour. Places where the contour just ends indicate a local occlusion (as of a bump on the front of an object hiding the bump's back side); the contour's end corresponds to a place on the surface where the viewing direction is along the rim. T-junctions in the contour indicate a bilocal occlusion (one object occluding part of another is always bilocal). In addition, Koenderink describes the characteristic events (in terms of catastrophe theory) by which the contour changes as the viewpoint changes (either due to the object moving or the observer moving).

The contour is quite a rich source of 3D shape information for human observers. In cartoons and line drawings, a few lines—largely indicating the contour—are often enough to make an object instantly recognizable. Halverston [42] notes that Paleolithic artists and today's children share a tendency to represent objects in terms of line drawings, generally representing contour information. Biederman and Ju [5] found that familiar objects were commonly recognized faster from outline drawings—again, heavily weighted towards contour information—than from photographs. Koenderink *et al.* [64] found that for an object as complex as a female torso, observers could infer a surprisingly detailed surface from a cartoon drawing; in fact, there was little difference between observers' perceived surface orientation settings on a cartoon rendering (showing mostly the contour) and the settings on a fully-lit photograph of the actual object.

Perspective Cues

Like the projective cues, linear perspective and relative size result from the mapping of a 3D scene onto a 2D image (or retina); unlike the projective cues, however, linear perspective and relative size are specific to a perspective projection. Linear perspective is the convergence in the 2D image of lines known to be parallel in the 3D scene; relative size is the scaling, in the image, of objects according to their depth, so that similar objects at different depths appear to have different sizes. See Figure 2.3 for an illustration of both.

While perspective is surely responsible for many fascinating 3D effects, perspective projection—and thus the perspective cues—is not always desirable in visualization applications. Wright [141] found that perspective projection was inferior to orthonormal projection for chemists doing molecular modeling. Wanger *et al.* [132] found that orthogonal projections were superior to perspective projections for an orientation-matching task and a size-matching task; they also found that perspective projections were quite valuable in an object-positioning task. They noted that the orientation-matching and size-matching tasks could be reduced to 2D problems by orthogonal projection but that the object-positioning task could not. This result underscores the need for proper task analysis in designing visualizations; if, as here, a task can be performed without requiring the user to understand the 3D geometry involved, 3D techniques may be actually counterproductive.

Figure 2.3: Linear perspective and relative size. As the wall recedes to the left, the parallel lines of mortar appear to converge. All the bricks are assumed to be the same size, but the closer ones appear larger than those farther away.

Texture

When vision researchers describe the use of surface markings–optical texture–as a depth cue, they are generally concerned with the gradients in the density and size of texture elements in an image as a uniformly textured surface recedes in depth or changes its angle with respect to the viewing direction. Following Gibson [38], a good deal of research has been done on the perception of surface shape from texture gradients; Interrante [47] gives a good review. Texture gradients provide surface orientation information; textures that make strong use of perspective cues seem to be very effective [47], although a strong impression of shape can be conveyed by other types of optical texture, such as contour-line textures [110, 117, 125] or principal-direction textures [48].

Optical texture plays a significant role in visualizations of 3D surface shape. The use of a regular mesh pattern to show the shape of a 3D function, as in Figure 2.4, is widespread in mathematics texts; the use of contour or hachure lines in topographic maps is universal. In computer-based visualizations, grids (or sometimes checkerboards) are used both to show mathematical functions (as in Figure 2.5) and as a scale indicator on a ground plane underneath untextured data (as in Figure 2.6). Optical textures are also important in the display of transparent surfaces [48, 72, 95]. Optical textures are not much used to display opaque surfaces, however, despite the arguments of Forrest [37], Schweitzer [101], and Ware [134] over the years. This may be due to studies such as those of Todd and Reichel [125], Norman *et al.* [87], and Christou and Koenderink [22], in which a textured surface was not perceived more veridically than the same surface presented with smooth shading[2].

Shading Cues

The last two pictorial cues, surface shading and cast shadows, arise from the way light falls on a surface and is reflected to a viewer, irrespective of the color or texture of the surface itself. Surface

[2] It should, however, be noted that none of the three studies cited added a regular texture to a shaded, complex surface, as proposed by (for example) Forrest. It should also be noted that Todd and Mingolla [120] and Curran and Johnston [25] did find that texture improved 3D shape perception over that available from shading alone

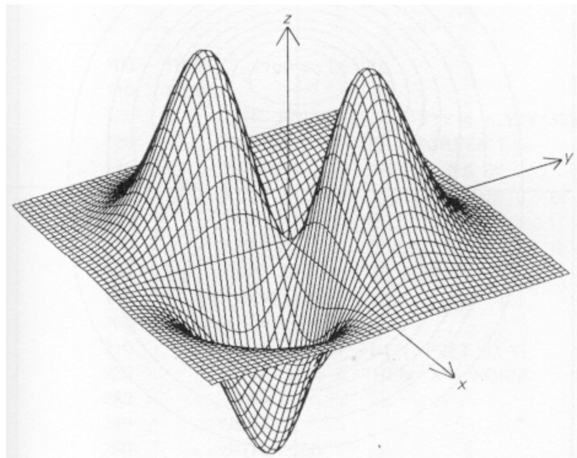

Figure 2.4: The function $z = xy\exp(-\frac{1}{2}(x^2+y^2)), |x| \leq 3.5, |y| \leq 3.5$, shown by a mesh pattern. Taken from Edwards and Penney [30].

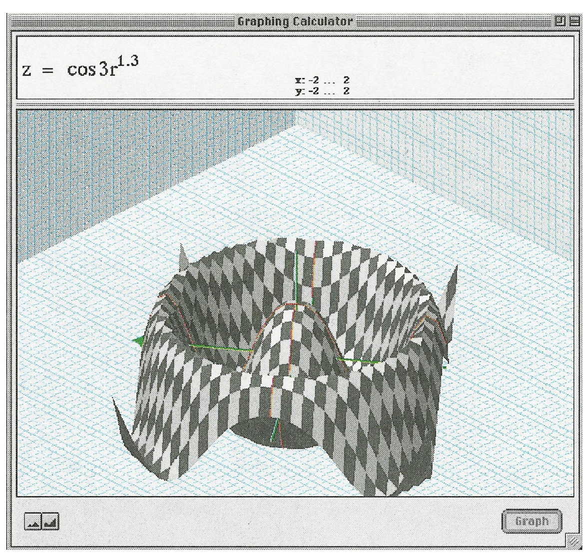

Figure 2.5: A checkerboard overlaid on a cosine function to show its shape. Taken from Card *et al.* [20].

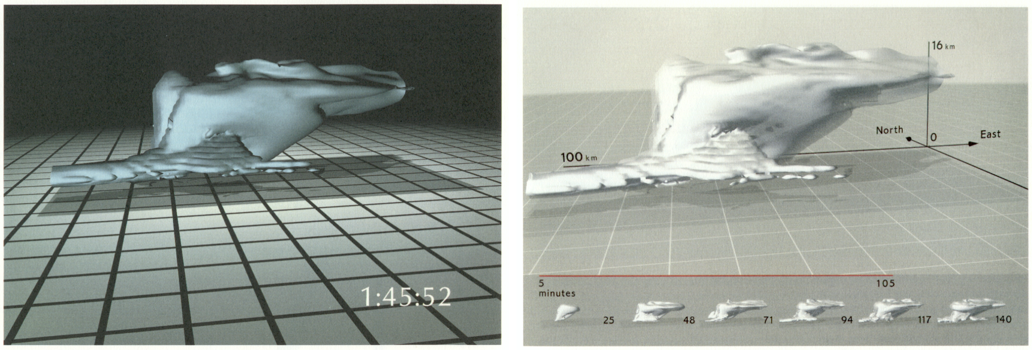

Figure 2.6: The image on the left is originally from the NCSA video "Study of a Numerically Modeled Severe Storm", described by Wilhelmson *et al.* in [139]; the grid on the ground plane shows the scale (one grid square represents 10×10 km). Tufte's re-visualization of the same data in [127], on the right, changes much but retains the grid. Both images shown here are taken from Tufte [127].

shading is the amount of light reflected from a given location on a surface towards a given viewpoint; cast shadows are the regions where light is prevented from falling on the surface by another object or another part of the surface.

Computer graphics has efficient algorithms for approximating the surface shading of an object, given a light source, a surface, and a viewpoint [36]. The basic approach is to handle specular highlights separately, do Lambertian (diffuse) shading, and add a constant (ambient) term to compensate for the non-physical assumptions being made (e.g., interreflections, cast shadows, and true specular reflections are neglected, and light sources are usually points rather than areas). The result is fairly inaccurate in a physical sense; critiques can be found in [61] and [57], among others. There are standard algorithms to address some of the shortcomings—for example, radiosity [39, 36] handles interreflections, and ray tracing [1, 36] does cast shadows and true specular reflections—but these involve a much higher computational cost, and so are commonly not used. Standard computer-graphics surface shading has survived because it does give results that look good enough to be useful; as will be seen, its inaccuracies have less experimental effect than might be supposed.

Since the amount of light reflected from a given surface location towards a given viewpoint depends on the surface's orientation with respect to the light source and the viewpoint, surface shading encodes orientation information directly. In practice, however, partly because humans are poor judges of absolute brightness and partly because the light reflected depends on several other factors[3] as well as surface orientation, patterns of surface shading are more informative than the surface shading at a single point; surface shading is thus a better cue to relative surface orientation and to curvature than to absolute surface orientation. Surface orientation—whether relative or absolute—and curvature information is direct shape information, however, so in theory surface shading could be a rich source of shape information. In practice, however, surface shading turns out to be a useful but fairly weak cue.

Psychophysical researchers have found that human observers are quite poor at recovering shape information from patterns of surface shading alone. Erens *et al.* [32] showed their observers shaded

[3]Some of the salient factors in addition to surface orientation are the reflectance of the surface, the brightness, position, and extent of the light source and the reflectance and orientation of other surfaces that take part in interreflections with the given surface location.

quadratic surfaces, systematically excluding all other sources of shape information, and asked the observers to classify the shape of each surface into one of eight intervals along Koenderink's shape index [55]. The observers' performance was barely better than chance. Adding a cast shadow to the stimulus to indicate the direction of the illuminant did help, but even then observers were able to make the correct classification less than 30% of the time (chance would be 12.5%).

Other studies have found that observers can make consistent depth and surface orientation judgments when the shading information is supplemented with contour information. In many cases, inconsistencies between judgments at adjacent locations are found to be accounted for by the variability in judgments at a single location, so the judgments are consistent with being samples from a single "perceived surface" (e.g., [26, 65, 68, 119])[4]. Observers' settings indicated that the objects appeared flattened in the depth direction, generally by an amount specific to each observer; such flattening has been observed both with depth-probing tasks [15, 68] and with surface-orientation-probing tasks [22, 26, 81, 65, 68, 119]; it has been observed on computer-generated surfaces [15, 26, 81, 120], in photographs [22, 65, 119], and even on real objects viewed directly [67].

In comparison with other depth cues, surface shading appears to be fairly weak. Todd and Mingolla [120] and Curran and Johnston [25] presented observers with stimuli where the shading information was inconsistent with the texture information; Todd and Mingolla found that the texture information dominated the shading information, and Curran and Johnston found the opposite. Koenderink *et al.* [64] added shading to a contour rendering and got only a modest increase in the perceived relief of the stimulus; Bülthoff and Mallot [15] and Norman *et al.* [87] added stereo information to a shaded stimulus and got substantial increases in veridicality. Norman *et al.* report a similar veridicality advantage for a stimulus with object motion over a stimulus with shading and contour alone.

Even when other cues are present, however, changes in surface shading are enough to perturb the perceived shape systematically, if subtly. Several studies [64, 63] perturbed the illuminant direction on surfaces defined by contour, cast shadows, and surface shading, and showed that the change in the illuminant direction caused small but consistent changes in the observers' perception of the shape. Christou and Koenderink [22] found the same effect for a surface defined by contour, shading, cast shadows, and texture; Todd *et al.* [119] found a similar effect for surfaces defined by contour, shading, cast shadows, and stereo.

Cast shadows convey information about the relationship of the shadowed area to the shadowing surface. Ware [134] notes that cast shadows are particularly effective for showing height above a ground plane.

Cast shadows are a far stronger cue than surface shading, but only when the geometry is complicated enough that non-local relationships (such as those between shadowing and shadowed surfaces) are important. Wanger *et al.* [132] found cast shadows to be the most helpful cue of all on object-positioning and size-matching tasks (the other cues they tested were object texture, ground-plane texture, object motion, perspective projection, and elevation above the ground plane); the shadows helped immensely in determining depth and size of an object relative to the depths and sizes of other objects. Kersten *et al.* [53] found that observers' perceptions of the track of a moving ball could

[4] In fact, as Koenderink *et al.* point out [65], there is no basis for believing that such a "perceived surface" has any existence in the observer's actual visual system, but the consistency of the judgments makes it possible to construct such a surface as a very neat summary of an observer's judgments across the entire stimulus surface.

be varied systematically by varying the path of the ball's shadow, even if the shadow information contradicted the information provided by the size of the ball's image; again, the motion of the ball was a non-local relationship. Mingolla and Todd [81] found that cast shadows were quite helpful in judgments of illuminant direction, and not helpful at all in judgments of local surface orientation; on the other hand, their stimulus surfaces were ellipsoids, lit from the front, so cast shadows could provide little information about locations on the ellipsoid's front surface.

2.2.3 Stereo

Stereo vision relies on the fact that objects visible to both eyes are projected onto the two retinas differently; the disparities between the projections are a rich source of information about the 3D geometry of the objects. The information is so rich that many people speak of stereo vision as "3D vision," as if there were no other depth cues.

Stereo in computer graphics is a powerful cue, but an expensive one. Displaying in stereo involves calculating a separate image for each eye, which generally doubles the rendering cost, and special display hardware is required to show each image only to its proper eye.

There are other pitfalls with stereo display, as well. When displaying on a standard monitor, the edge of the screen can clip an object which otherwise appears to be in front of the monitor; the occlusion destroys the depth effect [134]. Another problem is the mismatch between vergence and focus; stereo displays provide correct vergence information but do not simulate accommodation (the display remains at a fixed depth, with everything displayed in focus). Since the vergence and focus movements of the eyes are coupled by the visual system, the mismatch between the vergence information and the focus information can cause significant eyestrain if the objects presented are not confined in apparent depth [133, 134].

Stereo properly displayed, however, is a much more powerful cue than the pictorial cues. Bülthoff and Mallot [15] found that adding stereo information to their shaded stimuli drastically increased the depth perceived by their observers. Norman *et al.* [87] found a similar result for judgments of surface orientation; the addition of stereo to a shaded presentation cut the errors in their observers' orientation settings quite substantially. De Vries [27] found that observers looking at random-dot stereograms of quadric surfaces could classify those surfaces into eight intervals along Koenderink's shape index with ease—the same task at which Erens *et al.*'s observers, having only shading, were so poor. De Vries found, in fact, that observers' thresholds for discriminations on the shape index were generally between 2.5% and 5% of the length of the shape-index scale; thresholds were lower for approximately-cylindrical shapes, and higher for hyperbolic and elliptic shapes.

Human observers are able to make extraordinarily fine distinctions in stereo disparity, perceiving disparity differences down to approximately ten seconds of arc, albeit only at the fixation plane; however, stereo thresholds rise rapidly as one or both targets move away from the fixation plane, the usual case in everyday scenes [79]. Perhaps because of this dropoff, the precision of surface orientation judgments based on stereo information is considerably less than might be supposed from the fineness of the disparity judgments; Norman *et al.* [87], for example, reported mean errors of around 15 degrees in their observers' orientation settings using stereo (but not object motion).

2.2.4 Motion Cues

Motion cues to depth result from the paths in the retinal image of objects in the environment as the observer moves or as the objects move. Gibson [38] used the name "optic flow" for the collection of these paths; Koenderink [54] described the mathematics by which 3D shape information could in principle be extracted from that flow. There are two basic motion cues: motion parallax, caused by the motion of objects in the environment, and head-motion parallax, caused by the motion of the observer. They are closely related; Koenderink's mathematics makes no distinction between egomotion and object motion[5].

Motion cues are strong cues, comparable to stereo in strength. Rogers and Graham [97] show that both motion cues give a depth percept comparable to that obtained from stereo. Norman *et al.* [87] found that the accuracy of observers' surface orientation judgments on a rotating object viewed without stereo was somewhat greater than on the same object viewed with static stereo; the combination of object rotation and stereo improved the observers' accuracy further still. Van Damme [128] measured observers' thresholds for discriminations on the shape index using head-motion parallax; the results were remarkably similar to those already cited for stereo [27], right down to the lower thresholds for approximately-cylindrical shapes.

Some data exist to suggest that restricting the amplitude of motion restricts the effectiveness of the motion cues. Durgin [28], for example, found that neither motion parallax nor head-motion parallax allowed observers to make depth judgments on real objects as veridical as those made with stereo. But the motions in Durgin's experiments were limited to a maximum of 7.5 degrees of rotation; substantially larger rotational amplitudes were used in the experiments, cited above, of Rogers and Graham (30 degrees) and Norman *et al.* (24 degrees). Van Damme's rotational amplitude was 9.6 degrees, which is smaller than those of Rogers and Graham and of Norman *et al.*, but is still larger than Durgin's.

Head-motion parallax provides the observer with strictly more information than object motion, since in head-motion parallax the observer has proprioceptive information about the movement in addition to the information in the optic flow. In practice, however, the differences between the two cues, while significant, are not huge. Rogers and Graham [97] found that the difference between the depth percept due to egomotion and that due to object motion was less than 20% on average. Van Damme [128] found no difference between head-motion parallax and object rotation in judgments of orientation; he later found, however, that observers' judgments of curvedness were substantially more accurate with head-motion parallax than with object rotation.

2.3 Cue Integration

Humans get 3D shape information from all the cues just listed (and more) yet our percepts of objects are single. Researchers have attempted to construct models explaining how the 3D information from all the various cues gets integrated into a single percept. Maloney and Landy [77] proposed a model, later revised [71], in which the 3D information from various depth cues is calculated basically independently and then combined by a weighted sum. Nakayama and Shimojo [83] proposed a model in which the cues are combined by Bayesian inference to give the most probable 3D configuration.

[5] Koenderink's mathematics requires only two images, so it actually applies to stereo as well.

Both these models are couched in terms of recovering a depth field across the image, although they could be reconfigured to recover orientation or curvature instead.

One problem for the additive model of Maloney and Landy is that depth cues are known to be context-sensitive. Koenderink *et al.* found that for many (not all) subjects, optically superimposing the viewpoints of an observer's two eyes increased the perceived relief of a picture compared to monocular viewing [66] but decreased the perceived depth of a real scene compared to monocular viewing [67]. Tittle *et al.* [115] found that 3D shape inferred from object motion is not constant under changes in the object's orientation; they also found interactions between stereo and motion. A later study by Tittle, Perotti, and Norman [114] found that the integration of stereo and motion data on a shape-discrimination task was not the same as the integration on a spatial-frequency-discrimination task. Such results make it less likely that the results of depth cues are indeed calculated mainly independently, as Maloney and Landy proposed.

A more serious problem for both the additive and Bayesian models is that they assume a common measure (usually depth) in which the information from all the cues can be integrated. Johnston and Passmore [50] found that the effect of illuminant direction on a slant-discrimination task was opposite to the effect of the same manipulation on a curvedness-discrimination task, suggesting that neither the slant perception nor the curvature perception could be derived from the other.

In addition, the precision with which humans can judge depth, orientation, and curvature is quite low for any of the three to serve as a fundamental measure. Koenderink *et al.* [68] measured perceived orientation and perceived depth independently, with only pictorial cues available; they found that the precision of the depth judgments was an order of magnitude less than the precision of the orientation judgments, so the orientation judgments could not have been derived from the perceived depths. Independent measurement of the orientation and depth when object motion and stereo were available [58] showed roughly equal precision in orientation and depth.

Orientation discrimination under full cue conditions was studied by Todd and Norman [122]. They found Weber fractions (the denominator was slant) for orientation of around 8%, an order of magnitude higher than the Weber fractions common in other sensory discriminations; making the stimuli more complex raised the Weber fractions substantially further (to around 25%). They observe that such high Weber fractions make it unlikely that local orientation could be a primitive measure. Norman and Todd [86] extended the earlier study to higher slants; they found that a constant threshold of 9 or 10 degrees fit the data from both studies approximately as well as a Weber fraction. Johnston and Passmore [50] found slightly lower orientation-discrimination thresholds—most between 5 and 10 degrees—but their stimulus surface was a sphere, which was considerably more regular than the potato-like surfaces used in both the Todd and Norman study and the Norman and Todd study. Since the orientation difference between two visible surfaces is always less than 180 degrees, a 10-degree discrimination threshold is not very precise.

Van Damme measured curvedness discrimination with head-motion only (without stereo) and found a Weber fraction of around 20% (with object motion only, the fraction rose to 35%). Johnston and Passmore [50] found a somewhat lower Weber fraction, around 10%; as in their orientation-discrimination experiment, their stimulus surfaces were spheres, which were simpler than the quadratic surfaces used by van Damme. Phillips and Todd [89] replicated the shape-index thresholds of de Vries [27] and van Damme [128], but on smooth surfaces with all cues available; the fact that the thresholds were still between 2.5% and 5% of the shape-index scale indicates that humans can discriminate only 20–40 distinct quadric shapes.

The imprecision evidenced by such results suggests a need for a human spatial organization that is not dependent on metric judgments. Koenderink [56] argued that the fundamental human spatial coding is not Euclidean but what has been called an "affine-in-depth" representation [60], an argument made more fully by Koenderink, van Doorn, and Kappers in [69]. An affine representation preserves parallelism and distance ratios in a given direction, but may not preserve angles or absolute distances. Specifically, Koenderink *et al.* argue that human spatial coding largely preserves relationships in the image plane but that the depth dimension is subject to scalings of a type not found in the two picture-plane dimensions. Koenderink *et al.* also cite evidence of shears—another affine transformation—involving all three dimensions.

Some support for an affine-in-depth representation comes from the anisotropic errors observers made in the surface-orientation studies performed by Koenderink and his colleagues [22, 26, 65, 66, 67, 68, 64, 58, 63, 87, 119]. In every study, the tilt portions of the settings were both far more accurate and far more precise than the slant portions of the settings. Typically, the tilt part of the settings was correct to within a few degrees while the slant part had Weber fractions of 20–30%; in addition, the slant accounted for more than 90% of the variance in the settings. The observers were thus far more accurate and far more precise in that part of the settings involving judgments in the picture plane than in the part involving judgments in depth. Mingolla and Todd [81] also reported more accuracy in tilt (in the image plane) than in slant (in depth), although the effect in that study was not so strong as in the studies performed by Koenderink and his colleagues.

Further support for affine-in-depth representation comes from results in the inference of shape from motion. Todd and Bressan [118] and Koenderink and van Doorn [60] demonstrated that it is mathematically possible to extract affine shape from the optic flow given a motion sequence of only two images (so the result could in fact be applied to stereo as well). Adding a third image allows the reconstruction of a metric result, but only (as van Veen [130] points out) if the observers can extract acceleration information from the image sequence; humans are known to be inaccurate at extracting acceleration information [40, 104, 116]. Todd and Bressan [118] found that observers were far more accurate at tasks for which an affine representation is sufficient (judging planarity and rigidity of moving dot configurations) than they were at tasks for which an affine representation is not sufficient (comparing lengths of moving non-parallel lines, and judging angles in 3D with the angles specified by moving lines). Both Todd and Bressan and Todd and Norman [121] (who experimented with smoothly curved surfaces specified by moving random-dot displays) found that observers' performance did not improve substantially when the motion sequence was extended to more than two images. Tittle *et al.* [115] found that observers were both more accurate and more subjectively confident on tasks that are supported by an affine representation of visual space (two different surface-orientation matching tasks) than on tasks that an affine representation would not support (adjusting a dihedral angle in depth and adjusting the depth of a cylinder to make its cross-section circular).

Additional support for an affine-in-depth representation of visual space can be found in studies that have made direct measurements of the metric of visual space. Wagner [131] had observers make judgments on distances and angles defined by stakes in an open field; he found that both types of judgments evidenced substantial compression of the depth dimension relative to the horizontal dimension in the picture plane. Todd *et al.* [124] had observers perform a set of linked bisection tasks in depth and found that the bisections of differently-oriented lines were consistent with each other—a result consistent with an affine space, but accidental in any non-affine space.

While the affine-in-depth representation is sufficient to account for many visual phenomena, it is not a complete account of visual behavior. Todd *et al.* [124] found that their observers would see straight lines as curved; thus, although the bisections were indeed internally consistent, implying that the perceived space was intrinsically affine, the mapping from physical space to perceived space was in fact not an affinity. Other studies have also shown curvature in perceptual space. Koenderink *et al.* [70], using an exocentric pointing task, found that the space indicated by the observers' settings was elliptic at small distances and hyperbolic at large distances; similar results, complete with the reversal of the effect as distance increases, have been known at least since Helmholtz. Hecht *et al.* [43] reviewed much of that work. Hecht *et al.* also found by experiment that the depth scaling for dihedral-angle judgments was not the same as the depth scaling for distance judgments for the same stimuli. They posited that the structure of visual space is in fact task-dependent.

Koenderink [56] and Stevens [109] argued that depth-cue integration, when it is actually needed, is done in a task-specific fashion. Some aspects of 3D visual performance are certainly task-dependent; for example, in Wanger *et al.*'s study [132], the relative usefulness of the cues depended strongly on the task under study. Ware [134], following up on this idea, made a start towards identifying a set of elementary tasks that are as common as possible.

Task-dependence raises the possibility that perceived depth or surface orientation might be specific to a particular operationalization of surface shape. Happily, there is significant evidence that many of the possible operationalizations of 3D shape yield results that are consistent with each other up to an affine transformation. For example, Koenderink *et al.*, comparing orientation judgments to local depth comparisons [68, 58], found that the representations generated by each experimental task were related by an affine transformation. Van Doorn and Koenderink [129] compared observers' settings using several different orientation probes and found those settings also consistent with each other. If different operationalizations of shape yield consistent results, it is no longer necessary to separate the perceived 3D structure measured with an orientation probe from the same object measured by a punctate probe in depth. Thus, it is reasonable to study perceived 3D shape using an orientation probe like MTOPS and expect the results to have some applicability to other tasks.

Nor is human visual performance strictly limited to the tasks directly supported by an affine or affine-in-depth visual space. Many researchers (e.g., [15, 28, 87, 96, 115, 123]) have done studies in which they could assess the veridicality of perceived 3D shape in a metric space. The results have showed repeatedly that, even if humans' basic spatial organization is affine or affine-in-depth, adding to the depth information available can improve the veridicality of observers' performance on metric tasks. For a significant class of applications, including many scientific visualizations, giving the user the best possible understanding (best in a metric sense) of the 3D geometry is the point of the exercise. Since it is possible to improve veridicality by adding visual information, it is appropriate to measure veridicality—as MTOPS does—with the intent of improving performance on metric tasks.

2.4 Scale and Shape

2.4.1 Evidence for the Importance of Spatial Scale

Over the last three decades, the study of scale-specific effects in human vision has been phenomenally active. Graham's book-length treatment of the subject [41], published in 1989, contains references to over a thousand prior works—and Graham limited her scope to effects at or near the threshold of visibility. The present treatment will therefore be limited to describing the basic experimental approaches and touching on some of the more salient results.

Scale-specificity in vision runs very deep. Graham argues that the cortical receptive fields described by Hubel and Wiesel (e.g., [45, 46]) and others (for example, [8, 31, 92]) exhibit a certain degree of scale-selectivity simply by virtue of having excitatory and inhibitory regions; the cell will respond more strongly to a stimulus the size of the excitatory region than to a stimulus much larger or smaller. While the neurons in a single region of cortex do not exhibit a wide variety of receptive-field sizes (for example, Kelly and Burbeck [51] note that the neurons mapped by Hubel and Wiesel vary in receptive-field size by only a factor of 2.5), Koenderink and van Doorn [62] have shown how a complete set of basis functions, supporting blurring (and thus a Gaussian-based scale space) among other image operations, could be built up out of the receptive fields that have indeed been discovered.

Psychophysical results have shown that scale-specificity does indeed permeate human vision. A wide variety of visual phenomena, including contrast thresholds, spatial-displacement thresholds, distance judgments, and scanning order, all exhibit scale-specific properties.

The scale-specificity of low-level phenomena such as contrast thresholds has been studied exhaustively. Adaptation studies, following the lead of Blakemore and Campbell [9], have shown that the detection threshold for a stimulus of given spatial frequency is affected more by an adapting stimulus of a nearby spatial frequency than by an adapting stimulus of a sharply-different spatial frequency. Summation studies, such as Campbell and Robson's [19], have found that the detection threshold of a complex waveform can be predicted from the detection thresholds of its Fourier components; these results suggest that stimuli of different spatial frequencies can be processed independently.

The Fourier formulation of the analysis turns out not to be crucial. Koenderink and van Doorn [59] have explained many of these same phenomena using an aperture-based scale space instead of the Fourier formulation. Bijl [7] extended that analysis to produce an aperture-based model that not only accounts for the original spatial-frequency data but predicts the detection thresholds of Gaussian-modulated blobs as well.

Higher-level visual phenomena also exhibit scale-specific properties. Toet and Koenderink [126] found that the threshold distance for detecting a displacement was proportional to the size of the stimulus entities being displaced. Burbeck and Hadden [16] found that the presence of a flanking line in a distance-judgment task changed observers' judgments, and that the area over which the flanking line had its effect scaled with the distance being judged. Burbeck *et al.* [18] found similar effects for observers asked to bisect a rectangle with wiggly edges; again, the accuracy of the judgments was consistent with the positions of the edges being integrated over an area approximately proportional to the width of the figure.

Scanning order, another high-level phenomenon, also displays scale-specific qualities. Watt [137] found that scanning of a stimulus image proceeded from coarse to fine spatial scales. Schyns and Oliva [102], using composite images made of the fine-spatial-frequency information from one image and the coarse-spatial-frequency information from another, found that observers could scan coarse-to-fine or fine-to-coarse, depending on how the task was specified, although the default tendency seemed to be scanning coarse-to-fine.

2.4.2 Evidence of Scale-Specific Effects in Depth Cues

Although the research has not gone far towards detailing the characteristics of scale-specific phenomena in 3D vision, the existence of such phenomena has been reported, specifically in the important depth cues of stereo and motion. Interactions between stereoacuity, fusion, and spatial frequency have been reported since the 1970's (Smallman [105] gives a review). Rohaly and Wilson [99] showed spatial-frequency-dependent effects in disparity averaging. Disparity averaging occurs when a stereogram containing two disparities separated by a small visual angle causes an observer to perceive a single surface at a depth between the depths indicated by the two disparities. Using two sine-wave gratings to provide the two disparities, Rohaly and Wilson found that the occurrence of disparity averaging depends on spatial-frequency similarity between the two gratings; they also found that, when the gratings were two octaves apart in frequency, the perceived depth interpolated smoothly between the depths of the two gratings, depending on the relative contrast of the gratings. These data suggest that disparity components at different spatial frequencies can be processed to some extent independently, even if they are later averaged.

Evidence of multiscale processing has been found in motion, as well. Spatial-frequency selectivity in motion detection has been found in quite a number of studies; Nishida *et al.* [85] give a review. Rogers and Graham [98] found that motion thresholds for perceiving depth depend on the spatial frequency of the moving stimulus. Hogervorst *et al.* [44] performed a masking experiment in 3D and found that observers' ability to detect depth from head-motion parallax is sensitive to masking noise in a spatial-frequency-specific manner; depth-detection thresholds are affected more if the spatial frequencies of the moving stimulus and the masking noise are similar than if the frequencies are different.

The evidence of pervasive multiscale processing in human vision underscores the need for multiscale measurement of perceived shape. In particular, the evidence of scale-specific behaviors in important depth cues suggests that multiscale evaluation of the effect of a particular depth cue—such as that performed with MTOPS—may well be important in determining the usefulness of that cue for communicating 3D shape.

Chapter 3

Tool and Techniques[1]

3.1 Rationale

Scale-specific perception has been an important field of spatial-vision research for over thirty years, since the pioneering work of Campbell and Robson [19] suggested that stimuli of different spatial frequencies can be processed independently. Subsequent work recognized the need for specificity in location and spatial frequency, and the Gabor patch became a popular stimulus ([41] has a good review). In parallel with these developments was research on the role of aperture size on perception (e.g., [6, 126, 137]). While a Gabor patch acts as a bandpass filter on the stimulus structure, aperture acts as a low-pass filter; consequently, a small aperture includes not only small-scale information but large-scale information as well. Both spatial frequency and aperture size have their advantages in studying low-level visual processes [6, 41, 59]; higher-level spatial vision shows signs of being less dependent on spatial frequency and more dependent on aperture size, however [18, 16, 126]. Furthermore, aperture size corresponds to a natural perceptual construct—a region on an object—and spatial frequency does not. Consequently, I use aperture size as my scale parameter.

I chose to measure perceived surface orientation rather than depth or curvature because surface orientation offers readier support for precise, relatively rich measurements of multiscale percepts. Perceived depth [15, 58] is a more impoverished measure than perceived surface orientation and is not as amenable to scaling. Although perceived curvature is theoretically a richer measure than surface orientation and could support scaling, practical methods of measuring perceived curvature (as in, e.g., [32, 50, 125]) have not been able to exploit this richness.

Previous measurements of 3D surface orientation have not dealt explicitly with the scale of the observers' judgments; judgments were made at a single scale or without clear control of the scale (e.g., [65, 81, 108, 110]). Human observers perceive 3D objects in a multiscale fashion, however; as noted in Section 1.2.1, a single-scale measure cannot give an adequate understanding of what the observer perceives at multiple scales. To learn about observers' multiscale perceptions requires measurements at multiple scales.

I call such a multiscale set of 3D orientation measurements at a given location the orientation path for that location. This orientation path shows how the perceived surface orientation at a location

[1] Much of the material in this chapter (except for Sections 3.3.2 and 3.4.2) was previously published in [13], ©1999 Elsevier Science; it is used by permission.

changes as the observer considers a larger and larger portion of the surface around the location. I compare the orientation path to an objective one computed from the surface geometry to assess the veridicality of the observers' judgments. Comparison of orientation paths obtained under different display conditions shows the effectiveness of various depth-cue conditions across spatial scale.

3.2 The Probe

The probe used was a top-like figure, as shown in Figure 3.1. It consisted of a stick normal to the surface [108], ending in a circle. The 2D projection of the circle indicated the orientation up to two reflections; the angle of the normal stick removed this remaining ambiguity [65, 110]. In prior studies using similar probes, the probe's circle was on the surface; the observer's task was to orient the circle so that it appeared to lie on the surface, a natural task given the relatively smooth surfaces and relatively small probes used in those studies. My interest in multiscale measurements on rougher stimuli required a change in both probe and task.

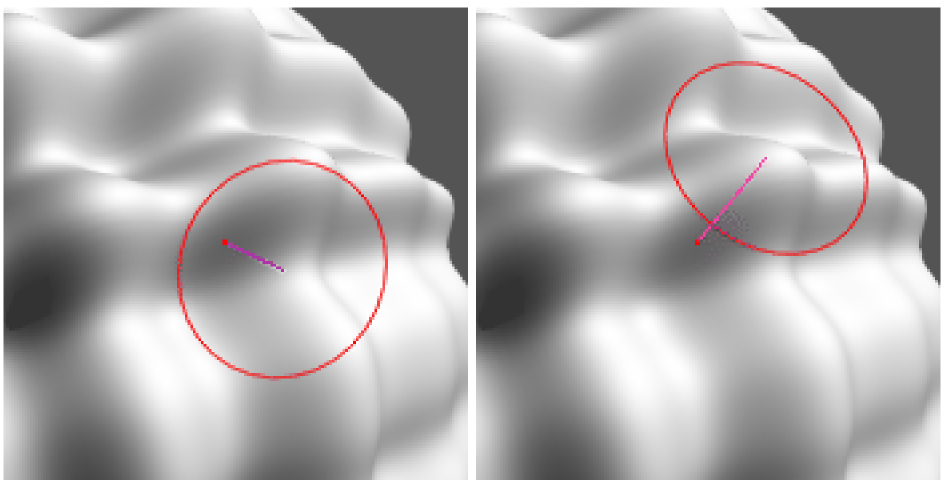

Figure 3.1: Observers indicate the orientation of the surface underneath the probe by pivoting the probe in two dimensions about the tip of the normal stick where it touches the surface.

The probe circles were raised off the surface; observers manipulated the probe's orientation by moving a computer mouse. The observers' task was to make the probe's top circle parallel to the region of surface that was the same size as the circle and centered on the probe's tip.

The size of the circle thus indicated the scale at which the observer was to make a judgment; that scale could be varied by changing the size of the circle. The length of the probe's normal line and the radius of its top circle were scaled together, giving the impression of a single probe being scaled as a whole. To maintain the probe's wireframe appearance across scale, the widths of the lines were not scaled.

The probe was displayed as part of a computer-graphics scene that contained the stimulus surface. The probe's colors were chosen to make the probe visible against my grayscale stimuli. The large circle was red; the normal line was magenta. The probe was rendered so that it cast no shadows and was never occluded by the surface. Cast shadows and occlusions are strong depth cues; I did not want the probe to add information to the stimulus.

Scaling the probe as a whole resulted in the larger circles being raised farther from the surface. This relationship between probe size and probe height had two advantages. First, it reduced the likelihood of intersections between a large-scale probe and small-scale surface irregularities; such intersections would result in two visible objects occupying the same location, an inconsistent stimulus. Second, previous studies show that in 2D judgments, spatial relationships at larger distances are made at coarser resolutions than those at smaller distances [18, 16]; therefore, increasing the distance from the circle to the surface may coarsen the resolution of the comparison, consistent with the task.

Raising the circles off the surface did introduce some uncertainty into the task, because there was no indication on the surface of the region whose orientation was to be judged. Identifying this region was thus itself a judgment, with no guarantee that exactly the same region of surface would be used on successive trials. The uncertainty in identifying the region could be expected to be exacerbated in regions of high slant or in areas where the surface is rough or highly curved. It is not clear, however, how an indication of the extent of a region could be added to the surface without the indicator itself conveying information about the surface geometry. It should also be noted that, as will be shown in Chapter 6.1, the overall accuracy of MTOPS settings is similar to the overall accuracy of settings made with probes (such as Koenderink *et al.*'s) that do lie on the surface.

3.3 Stimulus Surfaces

The MTOPS tool and techniques are applicable to a wide variety of stimulus surfaces, not just those used in the experiments presented in Chapters 4 and 5. But the surfaces used in those experiments were important to the aims of the experiments, so they will be described here.

My stimulus objects were constructed to have different structures across scale, to demonstrate the ability of the MTOPS probe to capture scale-specific perceptions. To this end, the stereo experiment used a smooth sphere and a rough sphere; the head-motion parallax experiment used the same rough sphere again. The object-motion experiment used two objects of greater complexity. These were defined as isosurfaces of multiscale solid-noise functions; the functions were coerced to make the objects ellipsoidal at large scale.

All my stimulus objects were rendered by ray tracing. The surfaces were all gray and purely Lambertian; all appeared to be lit by a single white light source above and to the right of the observer.

The dimensions of all surfaces are given in pixels, where a pixel is a unit of distance defined to be equal to the size of one pixel on the screen of the monitor used to display the surface. Because the size of the virtual space in all experiments was dependent on the pixel size of the monitor used, the pixel is a convenient unit for expressing distances in that space. For the two spheres, the pixel is 1/96 inch (0.26 mm). For the solid-noise surfaces, the pixel is 0.29 mm.

3.3.1 Spheres

Two spheres were used in experimentation, a smooth sphere and a rough sphere. The smooth sphere, shown in Figure 3.2, had a radius of 200 pixels (5.29 cm).

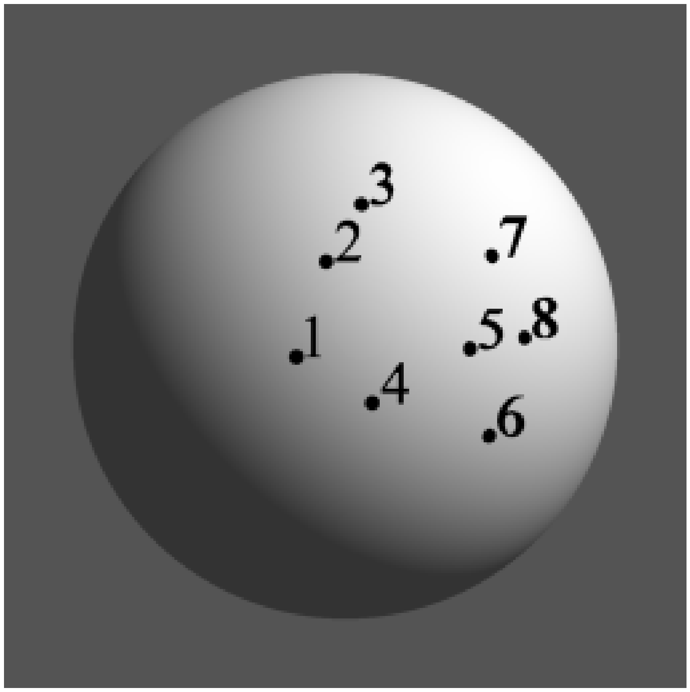

Figure 3.2: The smooth sphere, with indications added to show the eight test locations where the probe tip was placed during the experiments. The test locations are numbered from left to right.

The rough sphere, shown in Figure 3.3, had the same global shape and size as the smooth sphere. To the smooth sphere were added sinusoidal fluctuations in radius, defined as a function of central angle with appropriate constraints to avoid tearing the surface at the poles. The amplitudes of the sinusoids were five or six pixels; interactions among the several sinusoids produced peaks and valleys with amplitudes up to 29 pixels. The object was rotated so that it was seen from a reasonably generic viewpoint.

3.3.2 Solid-Noise Objects

The objects used in the object-motion experiment were isosurfaces of a 3D density function. The function used was a multiscale generalization of the solid-noise function described in [73]. Lewis constructs his solid-noise function by blurring together randomly-placed impulse functions of random amplitude; the average number of impulses per unit volume is related to the size of the blurring kernel. Different functions can be obtained by giving a different seed to the pseudo-random number generator used. I varied the scale of the Lewis solid-noise function by varying the size of the blurring kernel; I used different seeds to produce different surfaces.

I constructed the multiscale solid-noise function from a weighted sum of five single-scale solid-noise functions; each single-scale function was weighted by its scale. The first four single-scale functions were Lewis solid-noise functions, with kernel sizes of 16, 32, 64, and 128 pixels. The fifth had a kernel size of 512 pixels; it was like a Lewis function, but with non-random noise impulses. It served to make the overall isosurface roughly ellipsoidal; in the absence of the smaller-scale components, the isosurface became an ellipsoid with a major axis of 373 pixels and equal minor axes

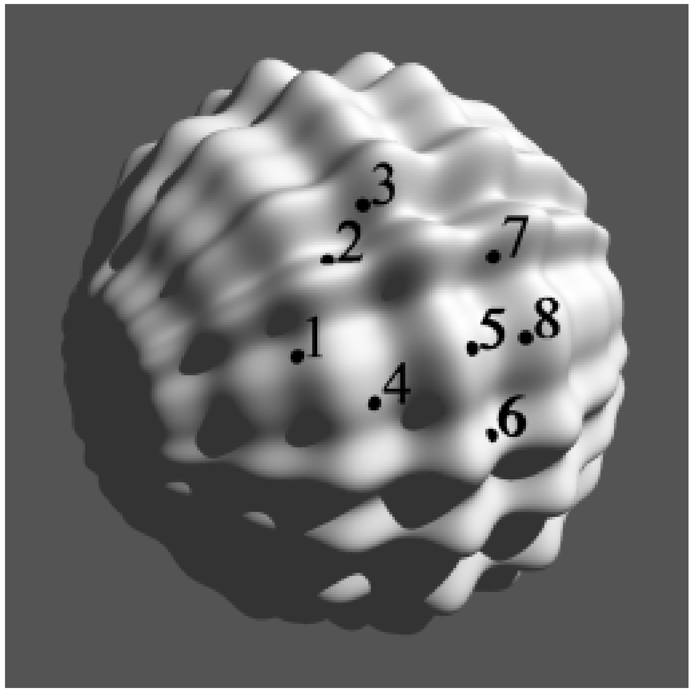

Figure 3.3: The rough sphere, with indications added as in Figure 3.2 to show where the probe tip was placed during the experiments.

of 264 pixels.

Two objects were used in the object-motion experiment, one for training observers and the other for the actual experiment. The training object is shown in Figure 3.4. The experimental object is shown in Figure 3.5.

3.4 Analysis Techniques

3.4.1 Orientation Paths

The most fundamental of the MTOPS analysis techniques is the orientation path, which shows the evolution of an observer's perceived orientation across scale for a given location and depth-cue condition. For a given location and condition, the observer made five settings at each probe size. I represented each setting as a unit vector in the direction of the probe's normal line; the mean of the five settings (i.e., the normalized mean of the corresponding unit vectors) was the reported datum for that probe size. Linking the mean vectors obtained with the various probe sizes gave the orientation path for that location, condition, and observer.

The orientation path characterizes change in perceived surface orientation across scale. A compact orientation path indicates that there was little change in perceived orientation across scale. A drawn-out path indicates a substantial difference between the perceived orientations at different scales, and shows the pattern of those differences.

To visualize these data, I projected them onto 2D. Since they all lie on a hemisphere, I use the

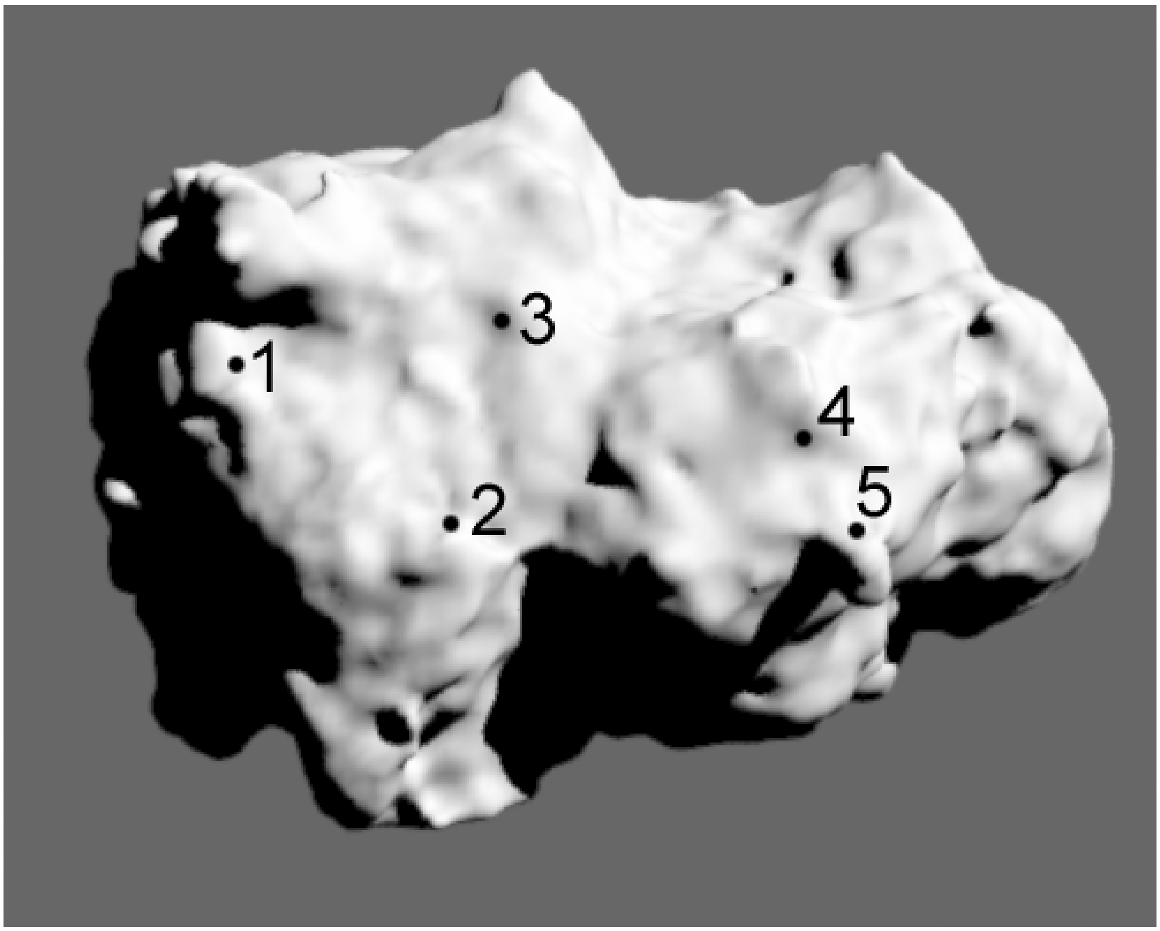

Figure 3.4: Training surface for the object-motion experiment. As in previous figures, indications have been added to show where the probe tip was placed during experiments.

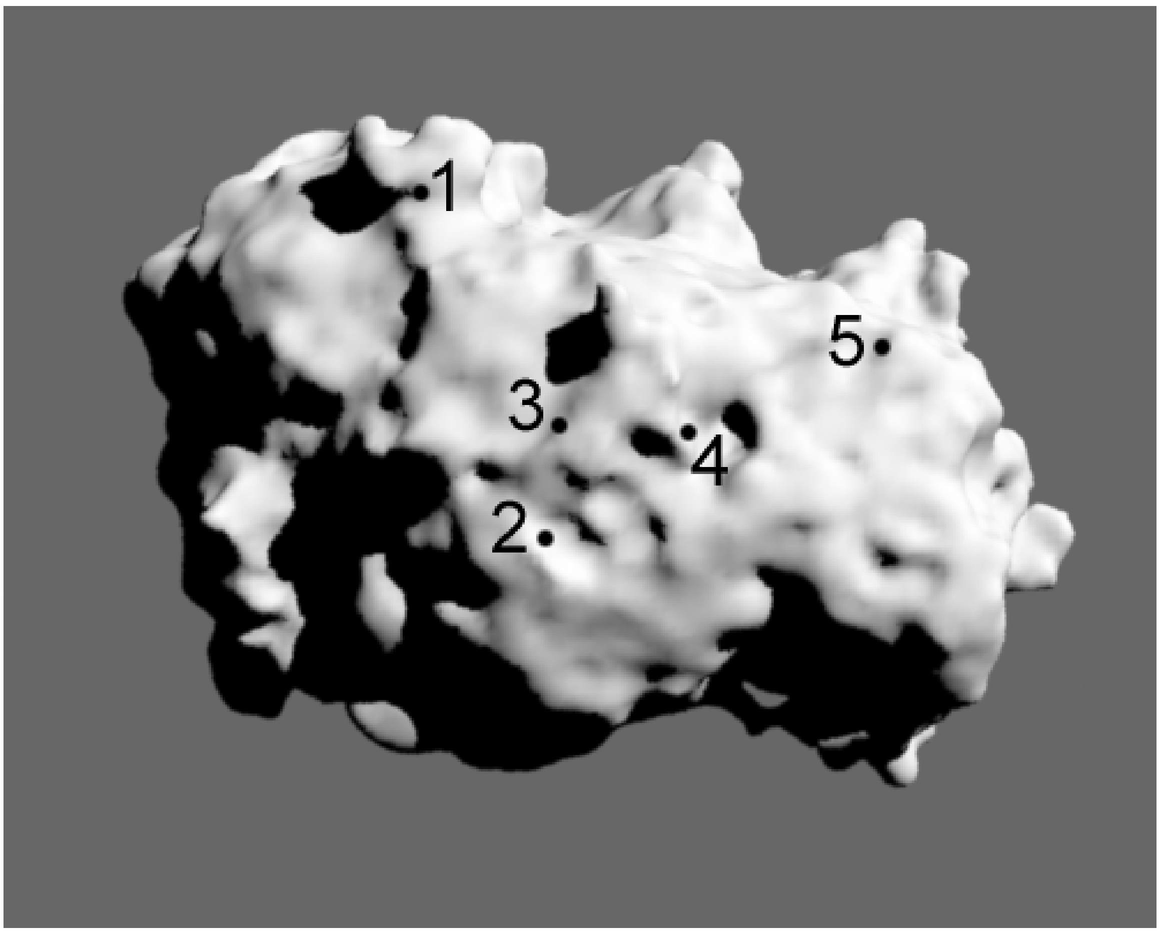

Figure 3.5: Experimental surface for the object-motion experiment, with indications added to show where the probe tip was placed during experiments.

Wulff projection, an equal-angle spherical projection which maps a hemisphere onto a unit disk[2]. I treated the center of the disk as the origin; because the center corresponds to the normal to the display screen, the coordinates of a given vector's projection indicate how much that vector deviates from the display-screen normal. The x-coordinate represents the vector's deviation to the observer's left or right; the y-coordinate, its deviation up or down. Although the projection is non-linear, a displacement of 0.1 in this projection corresponds roughly to an angular deviation of 10°.

3.4.2 Objective Surface Normals

The MTOPS probe and the orientation path gave me the ability to measure an observer's percept across spatial scales and to compare percepts obtained from different visualizations of the same object. To assess the veridicality of those percepts, however, required a set of objective surface normals, representing the surface geometry at multiple scales, from which I could construct objective orientation paths.

A defining requirement for the objective surface normals was that they must represent the surface geometry. Thus, they should depend only on the object itself, not on extrinsic variables like the viewpoint of the observer. Therefore, the scale parameter should be defined in the same 3D space as the object, without reference to image space, since image space is necessarily viewpoint-dependent.

Blurring with a spherical Gaussian fulfills both these requirements and offers the unique advantage of causality. I therefore based my scale parameter on the spherical Gaussian blob.

Defining an objective normal measurement based on a spherical Gaussian blob can be done in many ways, however. To ensure that the results did not depend on my choice of implementation, I used two quite different approaches and compared the results. As will be seen, the results proved to be very similar.

σ-Normals

The first approach to calculating objective surface normals was based on the observation that surfaces in nature do not occur by themselves; they are properties of objects. Many treatments of surface shape are rooted in object properties; a few of the better-known examples are Blum's medial axes [10], Leyton's process grammars [74], and Pizer *et al.*'s m-reps [91]. The treatment with the most influence on my construction is Koenderink's in [55].

Koenderink mentions several possible definitions of an object (with consequent definitions of surface), but the one he spends the most time discussing is the definition of an object as a region of space within which the density of some physical property exceeds some threshold value. The surface is then the level set of space where the density is within some finite tolerance of the threshold value; this has non-zero thickness (although it can be as thin as desired down to the resolution limit of one's measurements). Koenderink also suggests that the natural way to arrive at a multiscale treatment of such a physical object is to blur its density.

A mathematically-defined object (such as my stimulus objects) imposes a partition of space into

[2]The Wulff projection maps a point (x, y, z) on the surface of a unit sphere centered at the origin to a point (xW, yW) such that $xW = x/(1 + z)$ and $yW = y/(1 + z)$.

the space inside the object and the space outside the object.[3] This partition can be described as a membership function, with a value of 1 inside the object and 0 everywhere else. The surface of such a mathematical object is the locus where the membership function changes from one value to the other. One can then measure the density of this membership function within a given volume and blur the membership function, rather than the density, to get a multiscale treatment of the object.

I define the surface normal at a given location and scale as the vector through that location in the direction of greatest change in the density of the membership function at the specified scale of measurement. As the scale approaches zero, this membership-gradient surface normal will approach the conventional surface normal. As the scale increases, protrusions or indentations on the object's surface will be included in the volume of interest, and these lumps and holes of object density will push and pull the membership-gradient surface normal in different directions; thus, in general, the membership-gradient surface normal at a given location will change as the scale of measurement changes.

Specifically, I approximated membership-gradient surface normals as follows. For each scale at which a surface normal was desired, I started with a spherical Gaussian blob centered at the test location; the scale of measurement was defined by the standard deviation σ of the Gaussian blob. I truncated the Gaussian blob at 2σ. I multiplied the truncated Gaussian by the membership function and found the center of gravity of the product. The direction from this center of gravity to the test location was the direction of the normal vector for scale σ at that test location. Figure 4 shows a 2D example of this calculation; the extension to 3D is straightforward. I call the vectors calculated in this way the σ-normal vectors. I calculated σ-normal vectors for a range of σ's spanning the sizes of the probe radii, and from these I interpolated the objective orientation path of the surface at a given location.

s-Normals

The second approach to objective surface normals is an attempt to treat surface normals purely as a surface property, by blurring the conventionally-measured surface normals. In this approach, I approximated the surface by the same polygon mesh used in rendering it. I took the average normal of the outward-facing polygon normals from the polygon mesh, weighted by a spherical Gaussian cut off at 2σ (in fact, the same spherical Gaussian blob used for the σ-normal of the same scale and location). I call the vectors calculated in this way s-normal vectors, and I developed objective orientation paths from them in the same manner as from the σ-normals.

The s-normals are, in fact, quite close to the σ-normals for a given choice of test location and scale. Figure 3.7 compares the objective orientation paths derived using the two methods on the rough sphere; Figure 3.8 does the same for the experimental solid-noise object. As can be readily seen in those figures, the paths are almost identical. While the individual normals are not coincident, the results presented in Chapters 4 and 5 will show that the differences are too small to affect the conclusions of my experiments.

[3]A physical object, however defined, also imposes such a partition, but only at a scale large enough that the thickness of the surface can be ignored.

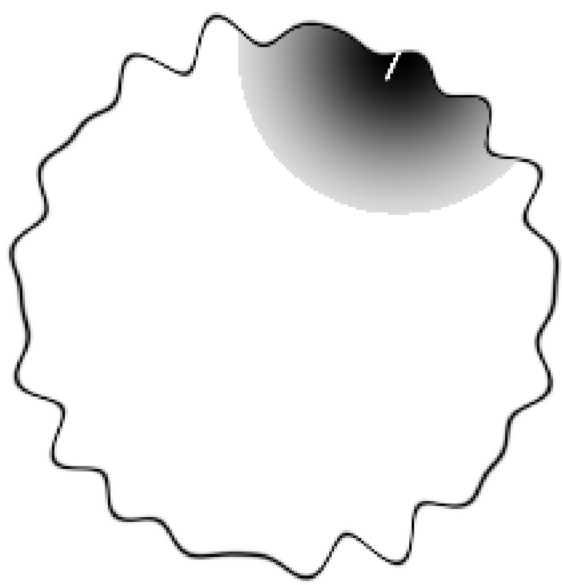

Figure 3.6: The calculation of a σ-normal vector in 2D. In 2D, the starting point is a circular Gaussian blob of standard deviation σ, centered at the test location. I multiply the Gaussian by a membership function of the object; the direction of the σ-normal vector is the direction from the center of gravity of the product to the original test location. In this case, the object is a cross-section of the rough sphere, taken along the plane $\phi = 0$; the Gaussian has $\sigma = 60$ pixels. The short white line is the vector linking the center of gravity to the test location. This figure illustrates the general case, where the σ-normal vector is not parallel to the conventional surface normal at the test location; the conventional surface normal is the limit of the σ-normal vectors as σ approaches 0.

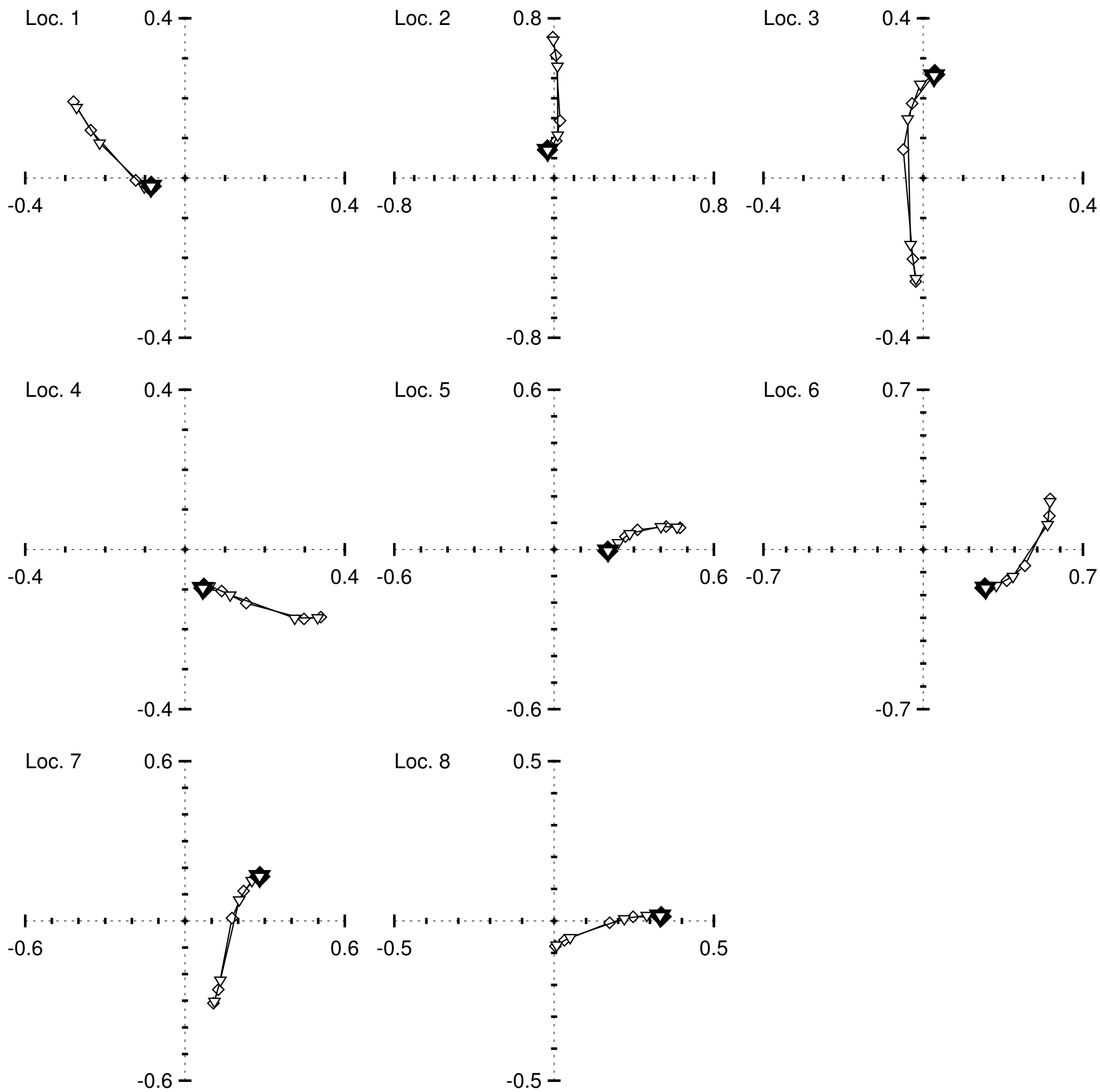

Figure 3.7: Objective orientation paths for the rough sphere, calculated using both σ-normals and s-normals. The diamond symbols are σ-normals; the upside-down triangles are s-normals. The large-scale end of each orientation path is indicated by a larger, darker symbol.

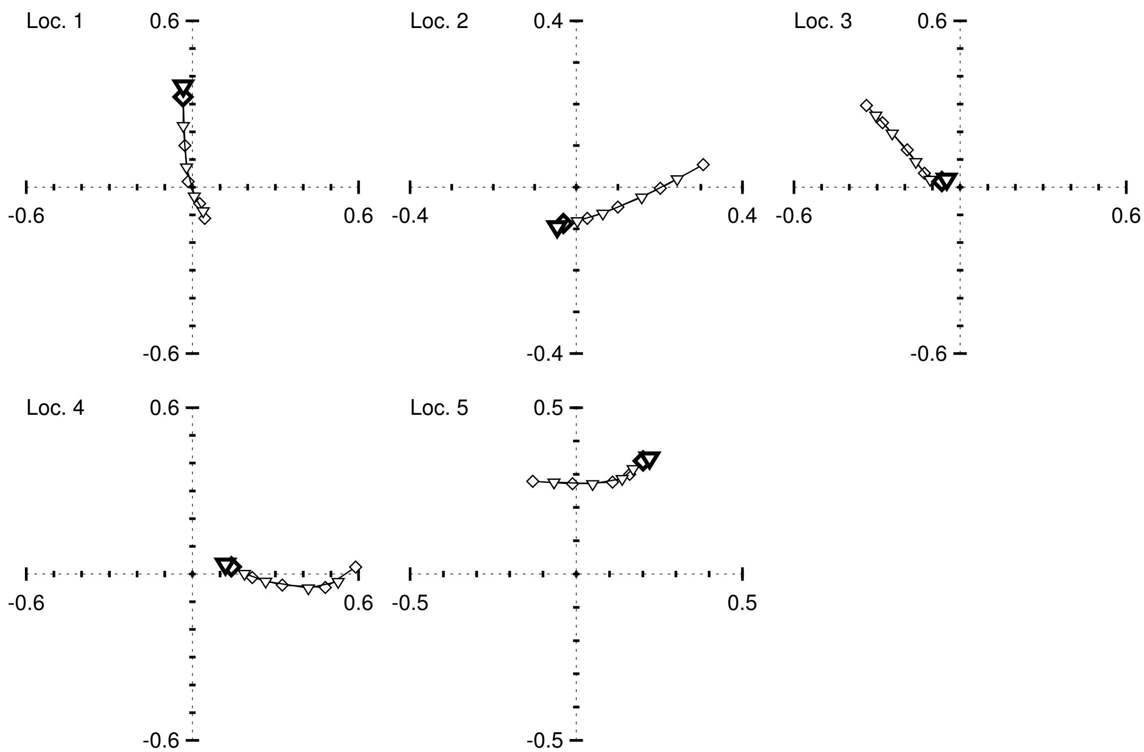

Figure 3.8: Objective orientation paths for the solid-noise experimental object, calculated using both σ-normals and s-normals. The diamond symbols are σ-normals; the upside-down triangles are s-normals. The large-scale end of each orientation path is indicated by a larger, darker symbol.

3.4.3 Bias

The overall veridicality of an observer's perceived surface orientations can be inferred by calculating the average difference between the observer's orientation paths and the corresponding objective paths. The difference between an observer's orientation path and the corresponding objective one is the bias of the observer's percept at that location. I use, as the measure of bias, the angular distance between the measured and objective orientation paths at each probe radius and the equivalent values of σ on the objective paths.[4] The average of these biases across locations gives a measure of the overall veridicality, or accuracy, of the observer's percept.

In interpreting bias results, patterns of change in the bias—both change across scale and change across depth-cue conditions—are far more useful than absolute bias numbers. The consistency of these patterns across observers indicates their reliability. Because this dissertation places very little weight on absolute bias numbers—all the important results are couched in terms of how bias changes across scale or across depth-cue conditions—I give no explicit tolerance measure for bias.

3.4.4 Variability

I do give an explicit measure of variability in the settings, however, for a different purpose. Because my multiscale probe is a novel measurement tool, it is important to determine whether the task required to use the probe is reasonable perceptually; specifically, I want to know whether increasing the size of the probe impaired the repeatability of observers' settings. To answer this question requires a measure of variability[5].

For spherical data like my observers' probe settings[6], the literature contains many measures of variability, none of which enjoys universal acceptance. Some, like the length of the resultant vector $\vec{R}$, give a measure of scatter but little intuition as to how the vectors are really distributed over the sphere [34]. Others, such as κ from the von Mises-Fisher or Kent distributions [35, 52] or the confidence cones about the mean, are specific to particular distributions. With no central limit theorem for spherical data, no spherical distribution offers the level of universality found in the normal distribution for data on the line.

The variability measure I report for an observer's settings at a given probe size is equivalent to the central angle of a cone centered on the mean setting and containing 68% of the individual settings. I calculated the variability of each observer's settings for each probe size as follows. I found the angular difference between each setting and the mean setting for that observer at that location and probe size. I took the magnitudes of those differences for all test locations; I report the value greater than 68% of the individual differences from the mean.

[4] The correspondence of σ with a probe radius of the same value is the simplest plausible correspondence. I found that as long as probe radii were in constant proportion to the corresponding values of σ, it made little difference to the results of my experiments if the values of σ were anywhere from half to twice their corresponding probe radii.

[5] It should be noted that the measure of variability I report does not offer a measure of the variability of the bias. The variability I report measures the repeatability of the individual settings; unlike the bias, the variability is independent of the choice of objective normals.

[6] My data is more precisely hemispherical—all the vectors point towards the observer—but spherical and hemispherical data are analyzed the same way.

Chapter 4

Experiment: Stereo[1]

4.1 Introduction

I conducted an experiment to validate the MTOPS probe and analysis techniques and to explore the across-scale effect of adding stereo to a non-stereo presentation; I found that multiscale measurement of perceived orientation is both possible and fruitful. Observers' settings indicated that perceived surface orientation is consistently different at different scales and that MTOPS is able to measure perceived surface orientation across scale. Observers' settings were no less consistent at large scale than at small, indicating that the MTOPS task is perceptually reasonable. Comparing observers' settings to the stimulus geometry showed the following:

- Observers were much more accurate on a simpler object.
- On the more complex object, adding stereo caused observers' accuracy to improve much more at large scale than at small.

Finally, for observers with prior knowledge of the object, the across-scale effect of adding stereo was quite different from the effect for naïve observers; the MTOPS probe and techniques were sensitive enough to capture this difference.

4.2 Methods

The probe and task used are described in Section 3.2. In this experiment, I used four probe sizes, with radii of 6, 12, 24, and 60 pixels. The stimulus objects used were the two spheres described in Section 3.3.1, shown in Figures 3.2 and 3.3.

I displayed the images on a Silicon Graphics Onyx RealityEngine2 workstation, using a 1280 $\times$ 1024-pixel monitor with a resolution of 96 $\times$ 96 dots per inch (37.8 $\times$ 37.8 dots/cm). Each of my two stimulus objects was rendered in perspective from a viewpoint 2000 pixels (52.9 cm) away. 2000 pixels approximated the distance from the screen to the observer, and the field of view of the projection transformation matched the angle subtended by the screen from 2000 pixels away; the

[1]The bulk of the material in this chapter (excepting only Section 4.1) was previously published in [13], ©1999 Elsevier Science; it is used by permission.

rendering of the stimulus object thus placed the object's center approximately in the plane of the screen.

I used two depth-cue conditions, stereo and non-stereo renderings of each stimulus object. In the stereo condition, the observers used a pair of CrystalEyes LCD-shutter glasses to see a stereo image from two alternately-displayed images. In the non-stereo condition, the observers viewed a single image on the screen with both eyes, without the glasses[2].

4.2.1 Procedures

I curtained off the experimental area to block off visual distractions, but I did not otherwise darken it. The observers were seated approximately 50 cm from the display; their heads were unrestrained, and viewing was with both eyes. These conditions kept the essential features of ordinary views of computer-graphics displays while remaining controlled enough not to add unnecessary noise to the data.

I used eight test locations on each stimulus surface (the locations marked in Figures 3.2 and 3.3). I recorded five settings for each location, probe size, and stimulus surface in the stereo and non-stereo conditions. Trials with different probe sizes and locations were interleaved. Trials with different surfaces and different depth-cue conditions were not interleaved; observers made their settings on the rough sphere in the non-stereo condition first, then on the rough sphere using stereo viewing, then on the smooth sphere in the non-stereo condition, and finally on the smooth sphere using stereo.

A typical trial ran as follows. The observer was shown the stimulus object with the probe superimposed on it. The probe appeared with its tip in one of the eight points on the surface; its initial attitude was randomly generated for each trial, with the constraint that the angle between its normal stick and the line of sight could not exceed 80 degrees. Using the mouse, the observer adjusted the angle of the probe until the probe's top disk appeared parallel to a similarly-sized surface patch centered on the probe's tip. When the observer was satisfied with the probe's position, he or she would hit the space bar, recording the final probe setting and starting the next trial. The observers were instructed to emphasize accuracy over speed. They were encouraged to take breaks every 40 trials; they were allowed to take breaks more often if they wished.

The four observers were adult volunteers. Two observers (LFB and JHA) were naïve observers, with no prior knowledge of the stimulus objects. They had limited (LFB) or no (JHA) prior experience with the task; they were provided with a list of tips (found in Appendix A) intended to make the task easier. The other two observers (the informed observers) knew before the experiment that both objects were spheres at large scale, and were thoroughly familiar with the task. All four observers have normal or corrected-to-normal vision; observer PHB (one of the informed observers) is mildly stereo-deficient.

[2] Koenderink *et al.* [66] have shown that binocular viewing of a single image tends to reduce the apparent depth of objects in the image. However, this is the way most people actually look at computer-graphics images.

4.3 Results

Variability of the Settings

Because my probe is a novel measurement tool, it is important to determine whether the task required to use the probe is reasonable perceptually. I approached this question by asking whether increasing the size of the probe seriously increased the variability of the observers' settings.

Using the variability measure described in Section 3.4.4, I plot variability of the observers' settings in Figures 4.1 and 4.2 for the two stimuli. Figure 4.1 shows the naïve observers' data, and Figure 4.2 shows the informed observers' data. Overall, the variability was constant or tended to decline with increasing probe size. Increasing the probe size clearly did not seriously increase the variability of the observers' settings; orientation judgments made with the large probe were at least as robust as those made with the small probe.

The variability of the observers' settings did depend on the stimulus and the presentation condition. The variability was generally larger for the rough sphere than for the smooth sphere, as expected from the differences in complexity of the stimuli. The variability also tended to be smaller for stereo than for non-stereo, suggesting that the stereo percept may have been more stable.

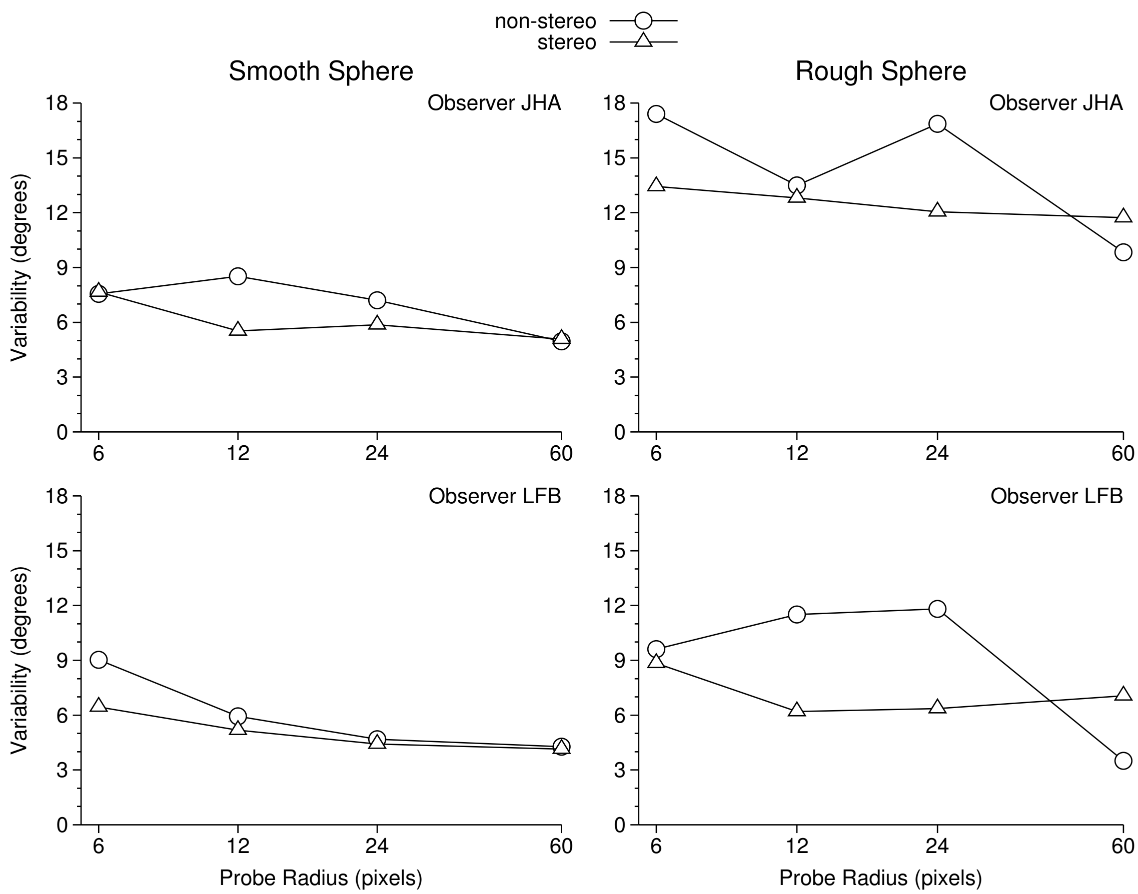

Figure 4.1: Variability in orientation settings for the naïve observers, for the two stimulus objects.

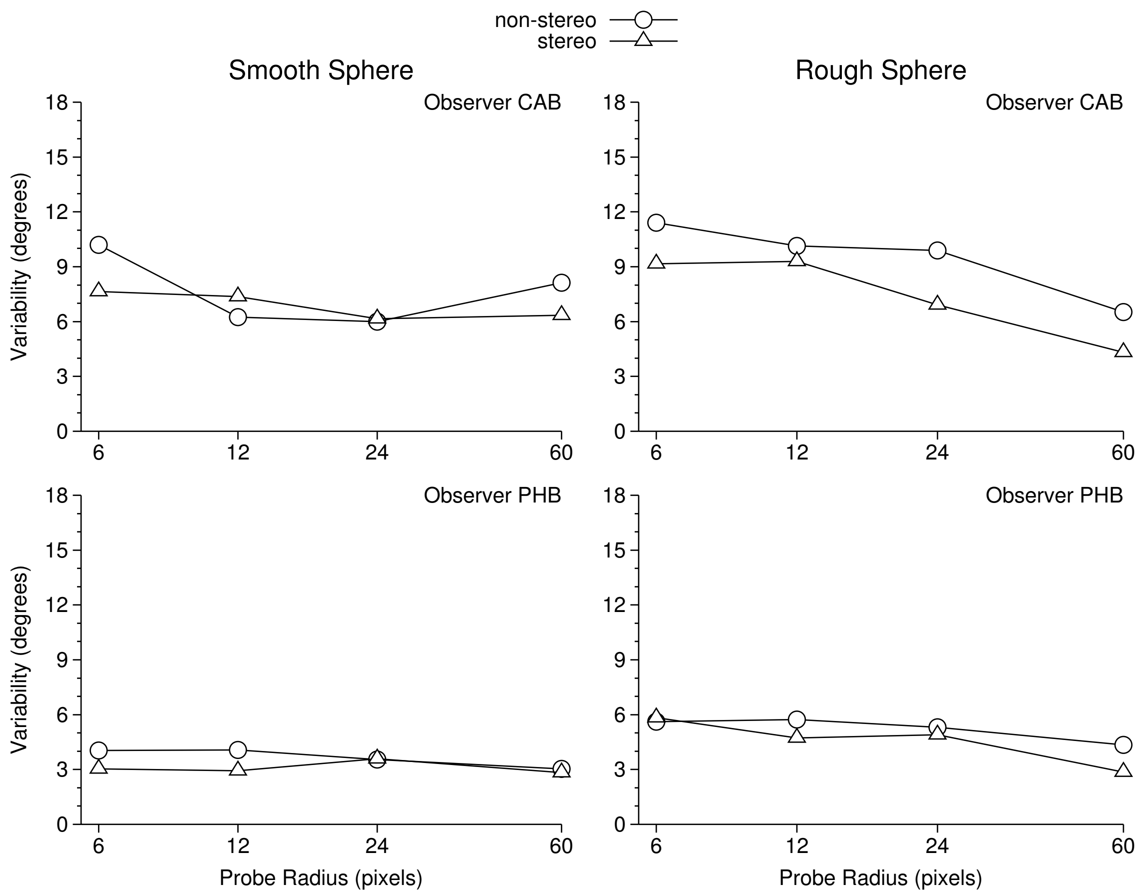

Figure 4.2: Variability in orientation settings for the informed observers, for the two stimulus objects.

4.3.1 Orientation Paths

Figures 4.3 and 4.4 show the two naïve observers' orientation paths for each of the eight test locations. The left column shows one observer's data for the smooth sphere; the other naïve observer's data for this stimulus were similar. The right two columns show the rough-sphere data for both naïve observers. Figures 4.5 and 4.6 are similar but show the informed observers' data. In both figures, each orientation path shown links the settings for the four scales in order; the largest scale is indicated by a filled data symbol.

Shown with each measured orientation path is the corresponding objective orientation path, constructed from σ-normals as described in Section 3.4.2. The objective orientation paths, developed by linking σ-normals, are represented in Figures 4.3–4.6 by bold lines terminating in $\times$'s at the large-scale end. For the smooth sphere these paths have zero length because the sphere has the same surface orientation at all measured scales. Because of the rough sphere's greater complexity, its objective orientation paths extend over a substantial range, reflecting the change in the surface orientation across scale.

The observers' orientation paths resembled the corresponding objective paths in many respects, showing that the observers were able to judge surface orientation across scale. For the smooth sphere, the objective paths had zero length; the observers' paths were very short, tending to zigzag back and forth over small distances in apparently random fashion. For the rough sphere, most of the observers' orientation paths stretched over a much larger range, reflecting the richness of the rough sphere's surface structure evident in the long objective orientation paths. The drawn-out observers' paths showed both that the perceived surface orientation changed across scale and that my scaled set of probes measured that change.

The relationship between the observers' orientation paths and the corresponding objective orientation paths varied across test locations on the rough sphere. At some test locations, the observers' orientation paths roughly superimposed on the corresponding objective ones (e.g., at location 2 for all observers); this indicated a high degree of veridicality across scale in the observers' percepts. Other locations showed deviations from such veridicality, commonly manifested by data points clustering near a single value on the observer's orientation path with no corresponding clustering on the objective orientation path. At some test locations, perceived surface orientation was fairly constant for the three smallest probe sizes (their data points clustered) and then changed markedly at the largest scale; the relatively smooth change in the objective orientation across scale at these locations was thus not reflected in the percept. At other test locations, all four data points clustered; at these locations, the perceived orientation was constant across scale although the objective surface orientation changed. Such clustering of data points could not stem from a failure on the part of the probe, because some orientation paths were superimposed on their objective paths or were otherwise extended. Instead, the clustering suggested that, at some locations, the observer's perception of the surface supported orientation judgments only at certain scales. The variety of relationships between observers' and objective orientation paths meant that there were differences in the veridicality of the observers' percepts across test locations, which the probe captured.

Figure 4.7 shows the naïve observers' orientation paths for locations 6 and 8, with the individual settings marked to show the scatter. The scatter for these observers at these two locations is typical.

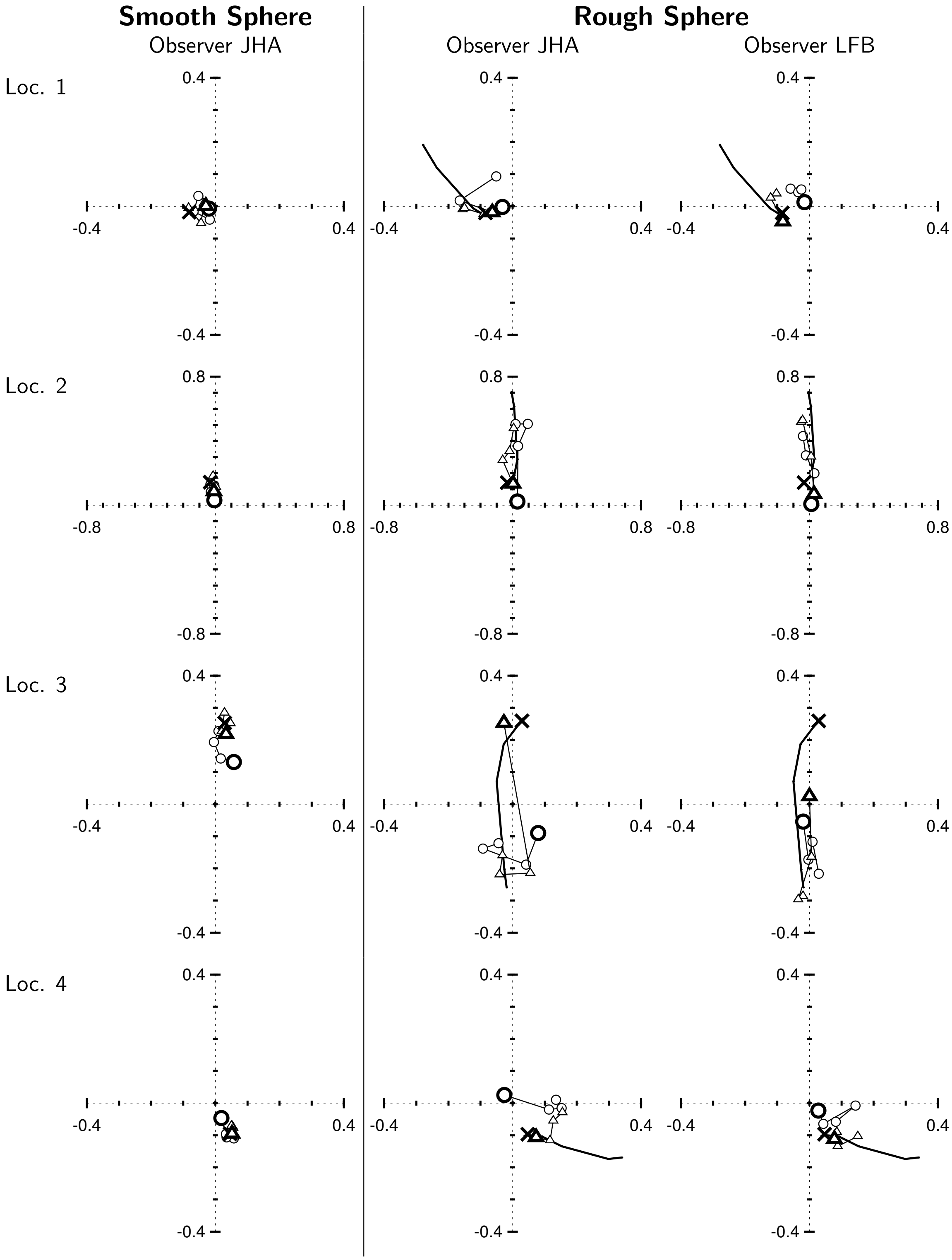

Figure 4.3: The naïve observers' orientation paths for test locations 1–4. Non-stereo data are shown by circles and stereo data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

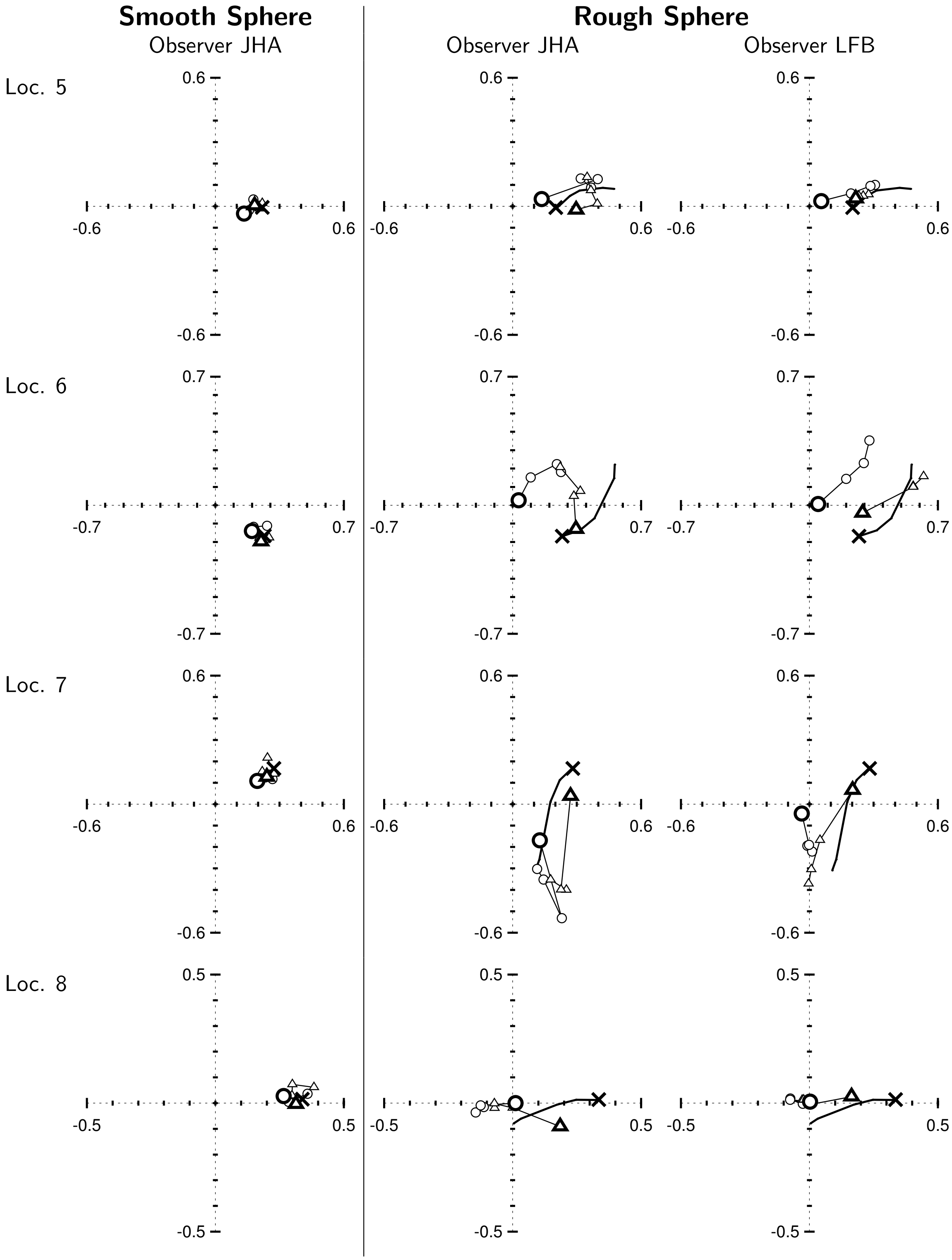

Figure 4.4: The naïve observers' orientation paths for test locations 5–8. Non-stereo data are shown by circles and stereo data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

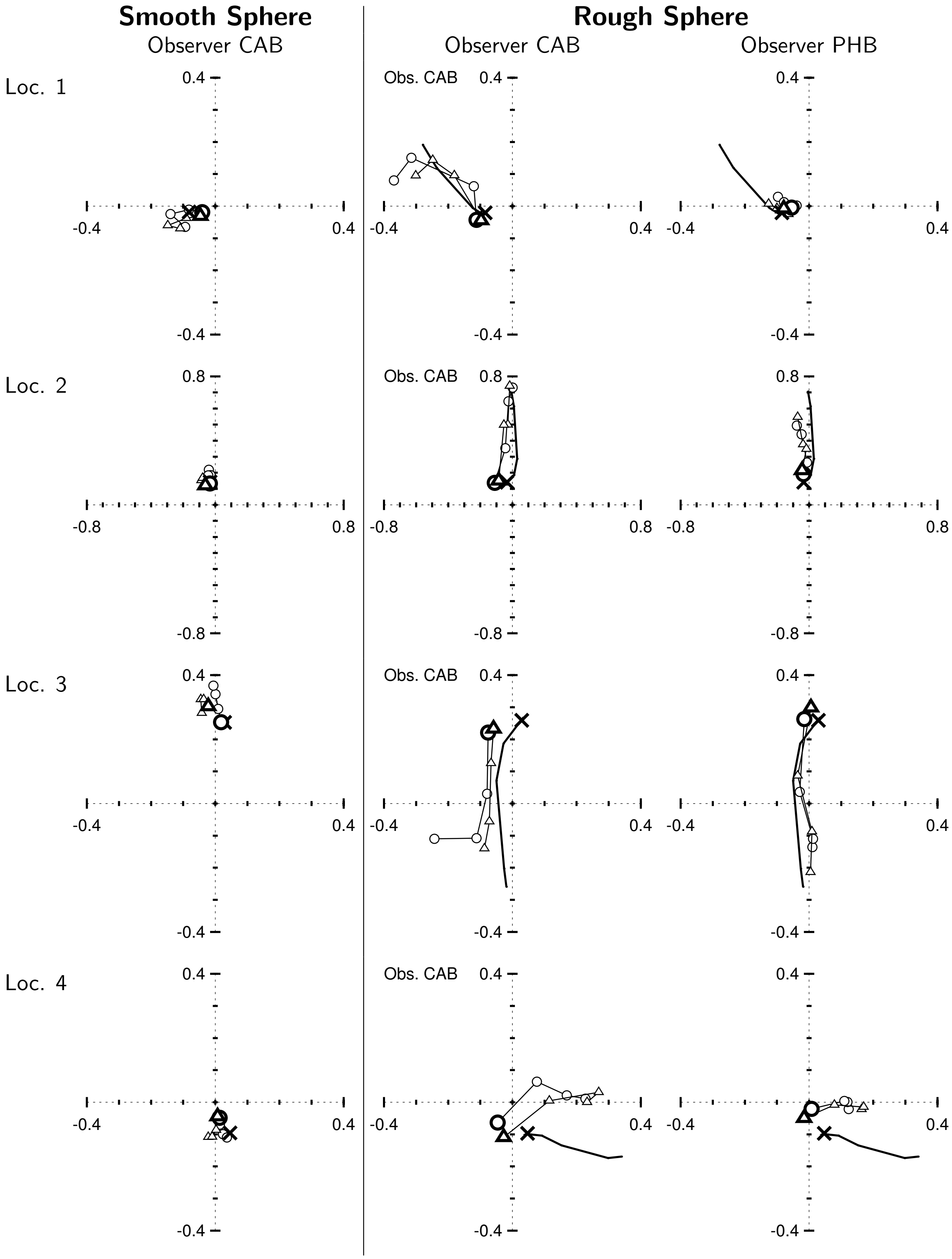

Figure 4.5: The informed observers' orientation paths for test locations 1–4. Non-stereo data are shown by circles and stereo data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

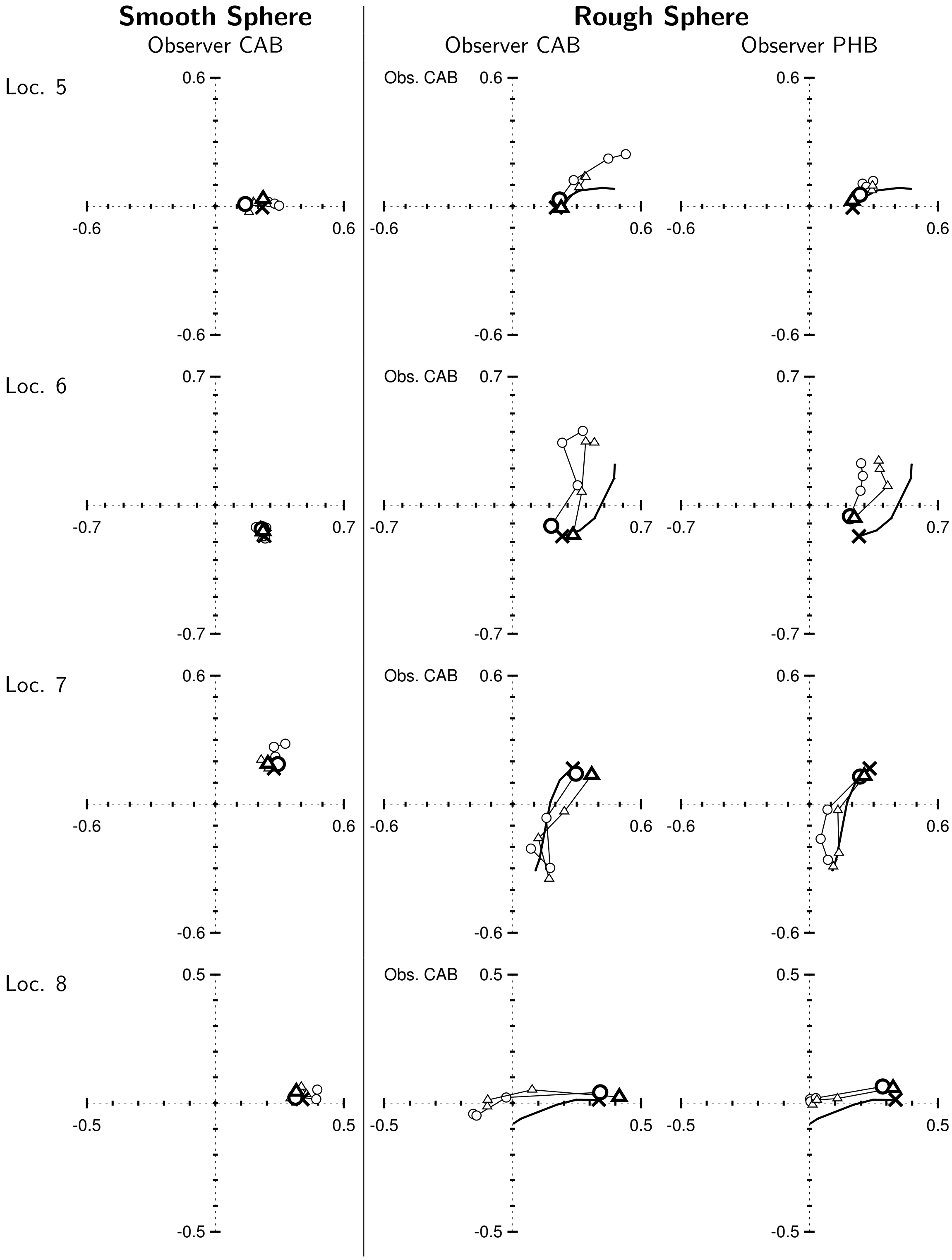

Figure 4.6: The informed observers' orientation paths for test locations 5–8. Non-stereo data are shown by circles and stereo data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

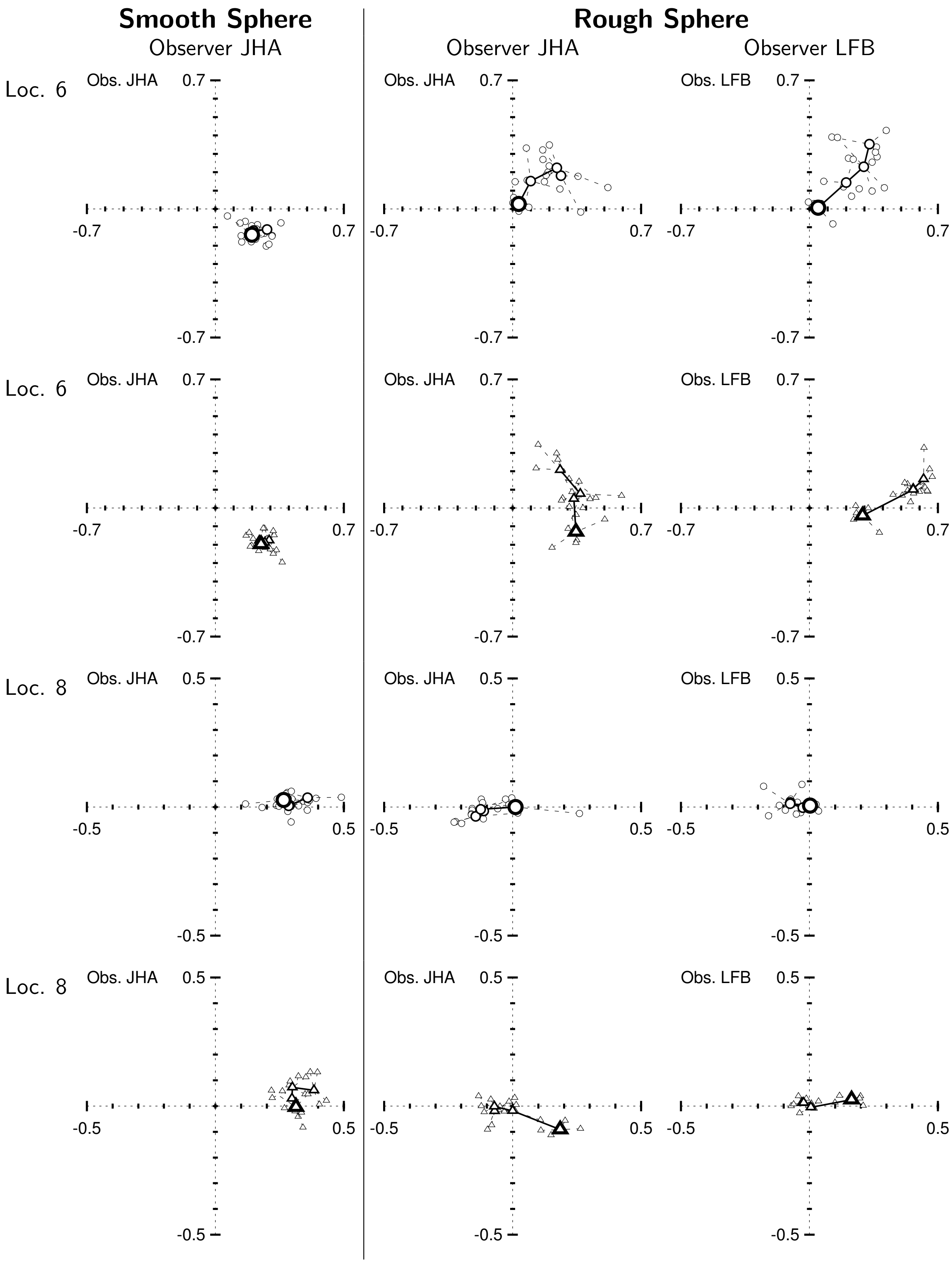

Figure 4.7: The naïve observers' orientation paths for test locations 6 and 8, showing the scatter of the individual settings. The orientation paths are shown as in Figures 4.3–4.6, with circles for non-stereo data, triangles for stereo data, and the large-scale end of each path marked with a larger, darker symbol. The individual settings are shown by smaller, lighter symbols; each individual setting is linked to its mean setting by a dashed line.

4.3.2 Veridicality Across Locations

The veridicality of an observer's perceived surface orientation can be inferred from the distance between the observer's orientation paths and the corresponding objective paths, as described in Section 3.4.3.

Figure 4.8 shows the average biases in perceived surface orientation for the two stimuli under the two presentation conditions. The biases for the rough sphere were calculated using both σ-normals (shown in the graph on the left side of the figure) and s-normals (shown in the graph on the right side of the figure). Because accuracy and bias are inversely related, a negative slope shows increasing accuracy across scale. Figure 4.8 shows the data for the two naïve observers. (The informed observers' data are discussed separately.)

The naïve observers' settings were far more accurate (had smaller biases) for the smooth sphere than for the rough sphere. For the smooth sphere, accuracy was roughly constant across scale. For the rough sphere, the effect of scale depended on the presentation condition. In the non-stereo condition, accuracy decreased (bias increased) with increasing scale; in the stereo condition, accuracy remained constant or increased (bias remained constant or decreased) as scale increased.

To capture the effect of stereo on the overall veridicality of the perceived surface orientation, I took the difference between the non-stereo and stereo data graphed in Figure 4.8. Figure 4.9 shows the difference between the average bias in the non-stereo condition and the corresponding bias in the stereo condition for the two naïve observers and two stimulus objects. For the smooth sphere, the addition of stereo improved the accuracy only slightly, although it changed the subjective percept dramatically. The stereo effect may have been so small because the observers' accuracy was near its maximum, with biases in some cases under 5 degrees; other researchers have reported discrimination thresholds of 5–10 degrees for surface orientation [50, 86]. There was no consistent effect of scale for the smooth sphere. By contrast, for the rough sphere, the addition of stereo improved the accuracy substantially. Furthermore, this effect depended strongly on scale. The primary effect of stereo for the naïve observers was to enhance perception of the large-scale structure of the stimulus.

Learning Effects

Because the observers did the stereo trials after the non-stereo trials, I used another bias-based measure to distinguish between improvements in accuracy due to the depth-cue conditions and improvements in accuracy due to the observers' increasing familiarity with the stimulus object. Because the improvement in accuracy occurred at large scale, I considered only the settings made with the largest probe size. For a given observer, I took the first trial with that probe at each location and found the angular difference between the observer's setting and the objective surface normal. Averaging these across locations gave me the average bias for all the first trials. I then repeated this for the second trial with the largest probe at each location, for the third trial with the largest probe at each location, and so on. Because all the first trials were done before any of the second trials, all the second trials were done before any of the third trials, and so on, the resulting graph shows how the observer's bias at the largest scale changed as the experiment wore on. In particular, if an improvement in accuracy is due to the change in depth cue, it should be concentrated at the point in the experiment where the depth cues changed; on the other hand, if the improvement in accuracy is due to increasing familiarity with the object, it can be expected to be spread out more through

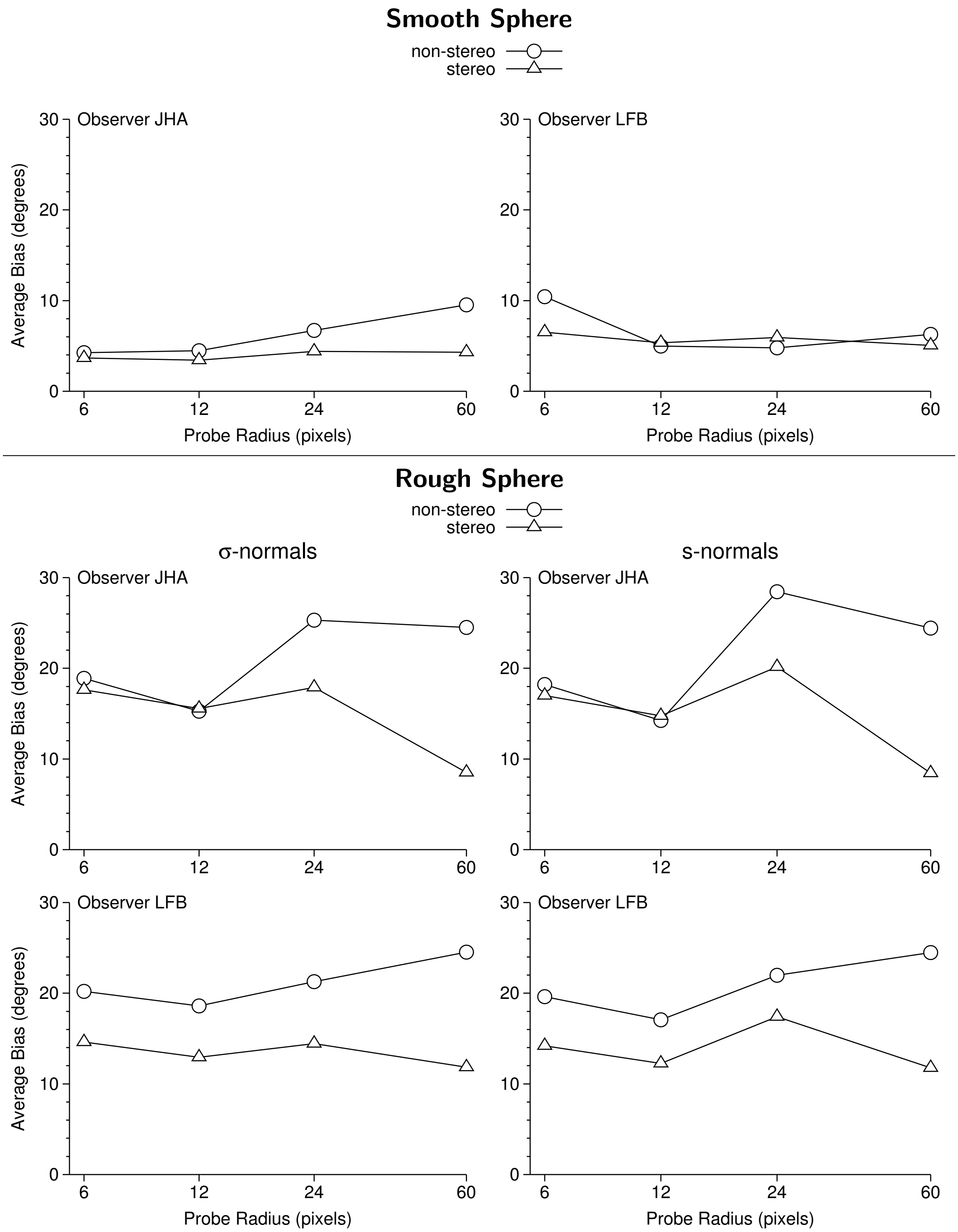

Figure 4.8: Average bias in the naïve observers' perceived surface orientation, for both stimulus objects. The rough-sphere graphs in the left column were calculated using σ-normals; the rough-sphere graphs in the right column were calculated using s-normals.

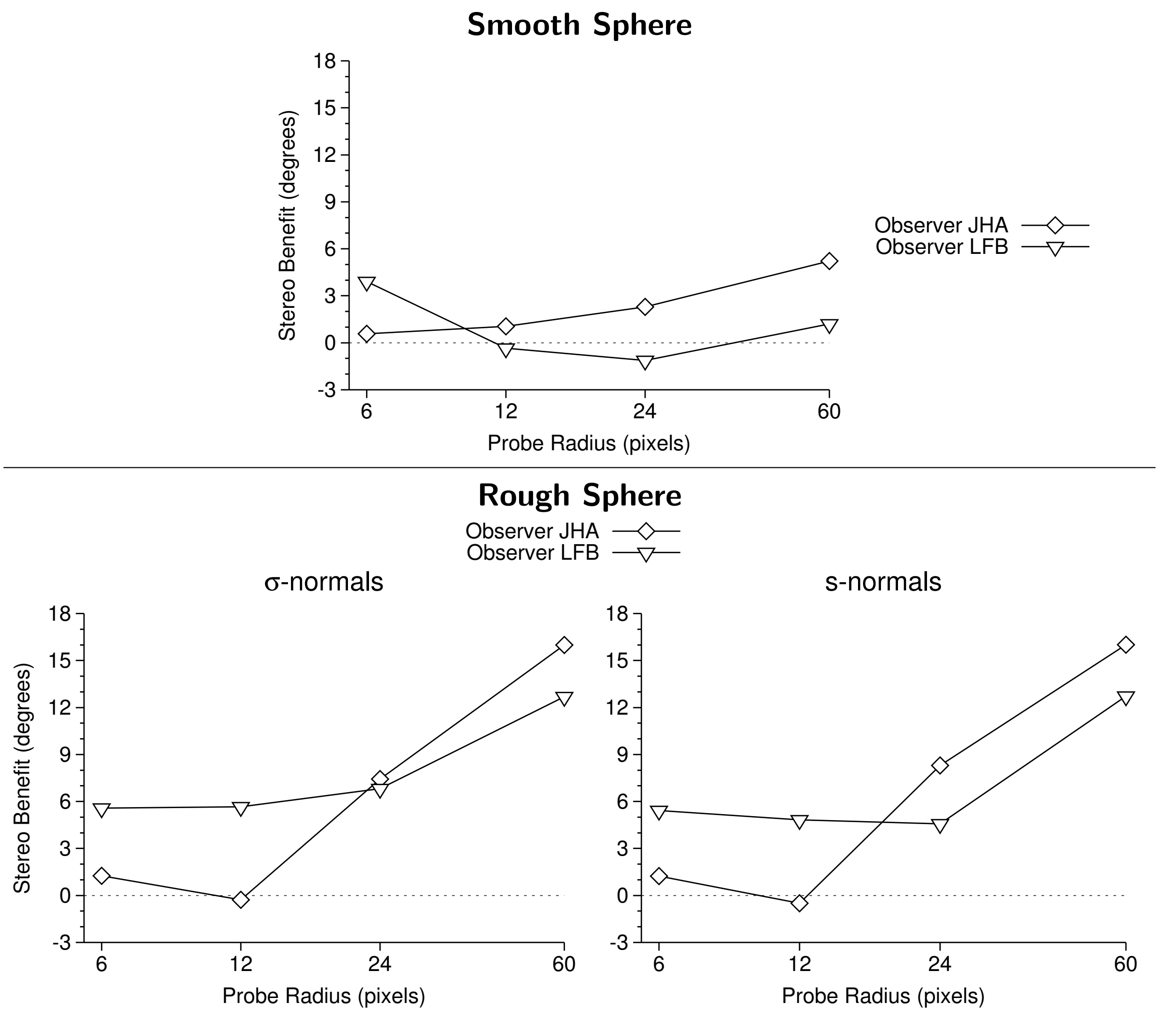

Figure 4.9: Improvement in accuracy of perceived surface orientation due to the addition of stereo to the display, for the two naïve observers.

the experiment.

Figure 4.10 shows that the improvement in the naïve observers' accuracy on the rough sphere was indeed due to the addition of stereo, not to the observers' increasing familiarity with the object. The improvement in accuracy is highly concentrated between the fifth and sixth trials, at the change between non-stereo and stereo.

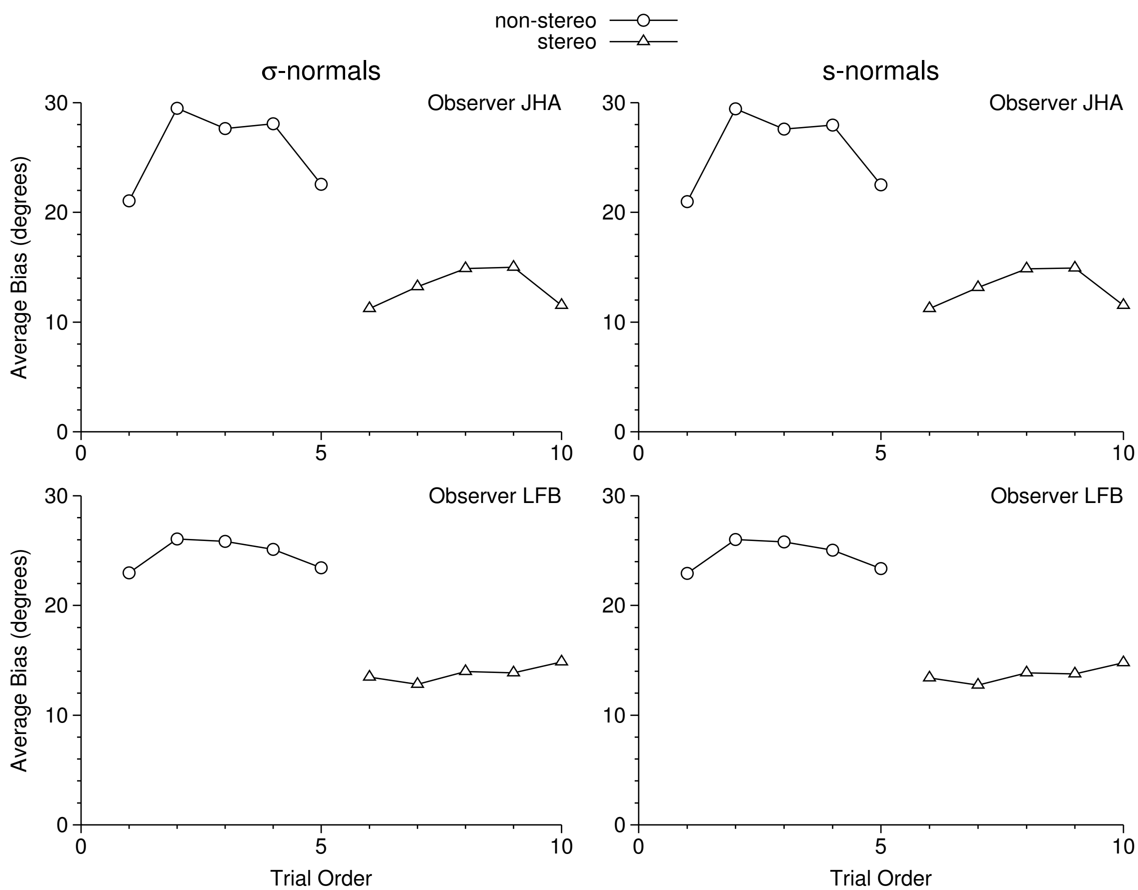

Figure 4.10: The naïve observers' large-scale bias as the experiment progressed. The improvement in accuracy is highly concentrated at the change between non-stereo and stereo.

The informed observers' data, shown in Figure 4.11, exhibit a different pattern than that for the naïve observers, shown in Figure 4.8. At large scale, the informed observers' accuracy was substantially higher than that of the naïve observers and much less dependent on stereo. The informed observers knew that the rough sphere was, in fact, spherical at large scale; their data reflect this knowledge. The difference between the informed observers' data and that of the naïve observers also shows that the probe is sensitive enough to capture such an effect.

4.4 Discussion

This experiment shows that a multiscale approach to measuring perceived 3D shape is both possible and fruitful. The patterns of relationship between the observers' orientation paths and the corre-

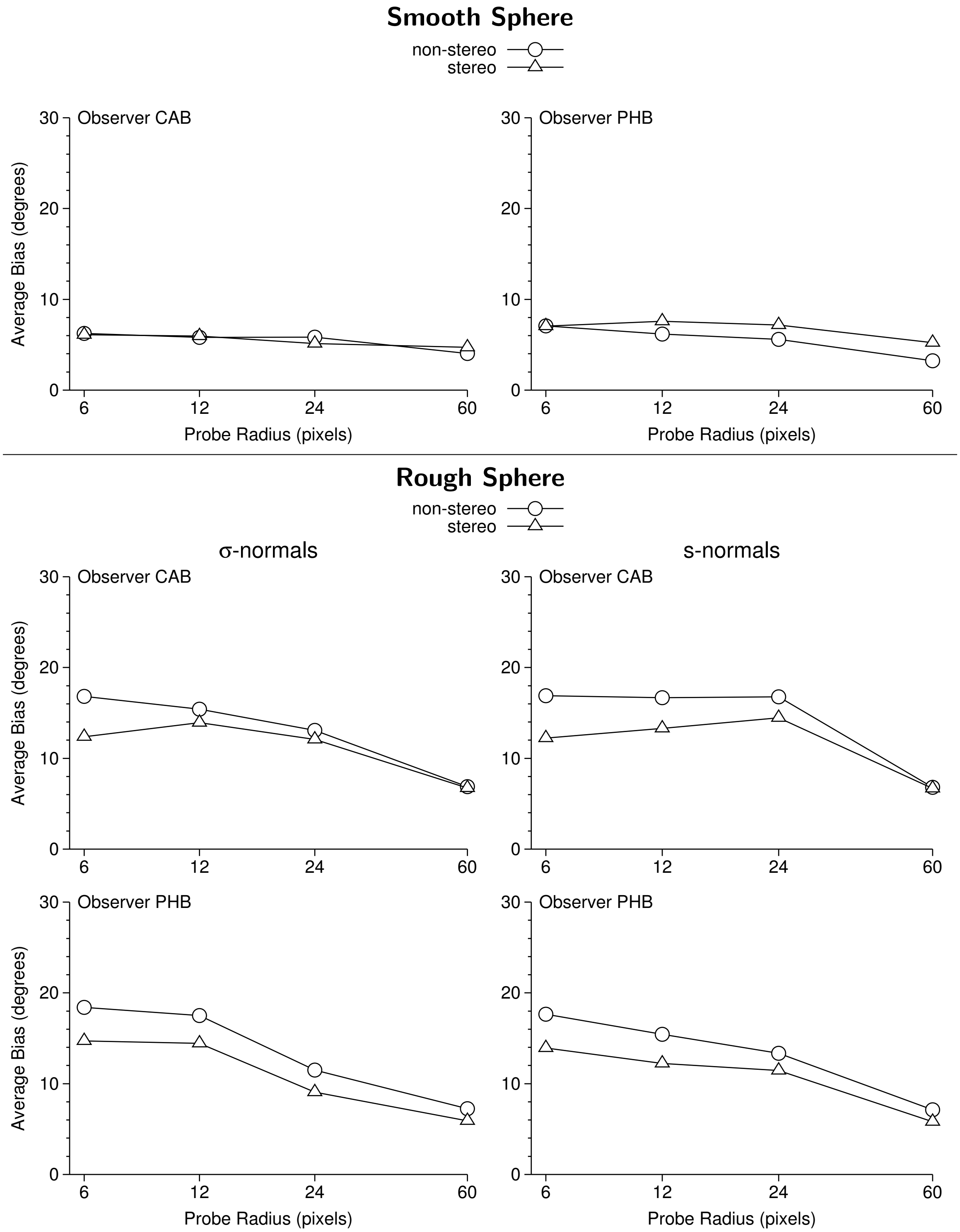

Figure 4.11: Average bias in the perceived surface orientation for the informed observers, for both stimulus objects. The rough-sphere graphs in the left column were calculated using σ-normals; the rough-sphere graphs in the right column were calculated using s-normals.

sponding objective paths, especially, tell a great deal about the observers' percepts. For example, if an observer's orientation path superimposed on the corresponding objective path, the observer was able to judge the surface shape in that vicinity with a high degree of veridicality. If the observer's orientation path roughly paralleled the objective path—a pattern seen more in the informed observers' data than in the naïve observers'—then the observer's perceived orientation differed from the actual by an approximately constant angular offset. When the observer's path differed more markedly from the objective path, the most common difference was a clustering of the observer's data points with no corresponding cluster on the objective path, suggesting that the observer's perception of the surface at those test locations may have supported orientation judgments only at certain scales.

I have also shown that changes in the orientation paths capture changes in the percept due to the addition of a depth cue. Comparing the orientation paths obtained with and without stereo to the corresponding objective paths showed that, for the rough sphere, the primary effect of stereo was at large scale.

Clearly, this varies from stimulus to stimulus. I found almost no effect of stereo on the smooth sphere, and a substantial effect on the rough sphere. More study is needed to characterize the multiscale effects of stimulus surfaces on stereo; differences in task may play a role, as well.

Chapter 5

Motion

5.1 Introduction

Following the stereo experiment described in Chapter 4, I performed two further experiments, demonstrating the ability of MTOPS to measure the across-scale effect of adding cues other than stereo and showing that MTOPS can produce measurements on objects other than spheres (whether smooth or rough). I chose to study motion cues since motion is well-known to be a compelling source of 3D information for human observers.

The rise of virtual reality in scientific visualization (e.g., [12, 14, 23, 24, 113]) made head-motion parallax an important cue from a computer-graphics standpoint. Previous work with path-tracing tasks [3, 136] found a substantial benefit to head-motion parallax with stereo over stereo alone. Norman *et al.* [87], using a surface-orientation task very like the MTOPS task, found similar advantages to object motion with stereo over stereo alone. The first experiment presented in this chapter measured the across-scale effect of adding head-motion parallax to a stereo display.

Another important visualization technique is object motion, which has been used in visualizations for almost four decades [49, 141, 75, 47]. Unlike stereo or head-motion parallax, object motion does not require special hardware for head tracking or for display; thus, object motion is more suitable than stereo or head-motion parallax for visualizations delivered in situations where such hardware cannot be assumed (for example, over the World Wide Web). Rogers and Graham [97, 98] found that non-stereo object motion is a source of shape information comparable to stereo; Norman *et al.* [87], in the study mentioned above, found a similar effect. The second experiment presented in this chapter studied the across-scale benefits of adding two different kinds of object motion to a static, non-stereo display.

5.2 Head-motion Parallax

5.2.1 Methods

Virtual-reality displays can be partitioned into three major types. The first and earliest type is the classic head-mounted display, originally implemented by Ivan Sutherland [111] as part of his pioneering virtual-reality system; the head-mounted display is still the most common type of VR display. Because of the size and weight constraints on a head-mounted device, head-mounted displays have always pushed the limits of display technology and have always suffered from poor resolution relative to standard monitors. The second major display type addresses this problem by moving the image away from the viewer's eyes, using images projected on large screens to produce an immersive experience. The pioneer example, NCSA's CAVE [24], uses screens surrounding the viewer; UNC's Bunker View [106] used a single large screen in front of the viewer. The third main type is what Ware *et al.* [135] call "fish tank virtual reality" (although the technique had been used for at least a decade previously—McKenna [80] traces the history). Fishtank virtual reality uses a standard monitor for display, tracking the viewer's movements and updating the image accordingly to create a virtual space inside, or just in front of, the monitor.

Of the three types, fishtank virtual reality offers the most straightforward means to study the effects of adding head-motion parallax to a stereo presentation in a realistic setting. The vast majority of static stereo presentations in real-world visualizations are displayed on standard monitors. By using that same standard monitor for the head-motion parallax condition, fishtank VR avoids confounds from the differences between the monitor and a head-mounted display or between the monitor and a projection-based display[1] Accordingly, I used fishtank VR for the head-motion condition in this experiment.

The two depth-cue conditions in the head-motion parallax experiment were thus a static stereo condition and a fishtank-VR head-motion condition, also with stereo. In both conditions, stereo was provided by means of a Tektronix LCD stereo plate affixed to the front of the monitor; observers wore passive circular-polarized glasses to see the image in stereo. In the head-motion condition, an ADL-1 mechanical tracker arm was used to track the position of the observer's head, and the stereo image on the monitor was updated using the output of the tracker. In the static-stereo condition, the tracker was not used.

Images were displayed on a 1280×1024-pixel monitor. The monitor's resolution was 96 dots per inch (37.8 dots/cm) both horizontally and vertically.

The approach to image generation was similar to that used in UNC's Bunker View [106]. A set of images were precomputed, corresponding to viewpoints along a horizontal line parallel to and 2000 pixels (52.9 cm) away from the monitor's display surface. Head-tracking allowed the display system to choose the pair of viewpoints best approximating the actual positions of the viewer's eyes at any given moment; the precomputed images corresponding to those viewpoints could then be displayed. Observers viewed the images through a large horizontal slot in a cardboard box. As in the Bunker View system, the slot served to keep observers from moving vertically, in order to keep observers' actual eye positions close to the line of viewpoints; the left and right ends of the slot kept observers from moving horizontally beyond the ends of the line of viewpoints.

Only one stimulus surface was used in this experiment, the rough sphere described in Section 3.3.1 and pictured in Figure 3.3. 512 images were precomputed, corresponding to viewpoints spaced every four pixels (1.06 mm) along the line of viewpoints. The rightmost viewpoint was 1024 pixels (27.1 cm) to the right of the monitor's center; the leftmost was 1024 pixels to the left of the monitor's

[1] Pausch *et al.* [88] approached this problem from a different direction, by using a head-mounted display bolted to a post for a static stereo display, and a normal (unbolted) head-mounted display to add head motion to the presentation. Using a head-mounted display forced them to use a task for which display resolution was not important; fishtank VR avoids that constraint.

center. Measured from the center of the screen, the viewpoints spanned an angle of 54.2 degrees horizontally, with an angle between viewpoints no greater than 0.115 degrees.

For each image, the distance between the viewpoint on the line of viewpoints and the center of the screen matched the distance between the viewpoint of the perspective transformation and the center of the object, and the field of view of the transformation matched the size of the screen. Subjectively, the object appeared to be fixed in the center of the screen; it seemed to rotate smoothly as an observer moved back and forth along the line of viewpoints. Lag was not subjectively noticeable.

The probe and task used for this experiment were those described in Section 3.2. The probe sizes were the same as in the stereo experiment described in Chapter 4: four sizes, with radii of 6, 12, 24, and 60 pixels.

The three observers were adult volunteers. All were naïve observers. They were provided with an overview of the project and a list of tips for observers, found in Appendix B. All three observers had normal or corrected-to-normal vision.

Procedures

The procedures in this experiment were largely similar to those in the stereo experiment described in Section 4.2.1. The experimental area was curtained off to block visual distractions, but it was not otherwise darkened. The observers were seated approximately 60 cm from the display. Viewing was binocular. Observers' heads were not restrained, although (as noted above) viewing the screen through the slot in the box did constrain their motion somewhat. As in the stereo experiment, I intended to retain the essential features of the normal viewing of computer-graphics displays while providing enough of a controlled experimental environment to avoid adding unnecessary noise to the data.

I used the same eight test locations as in the stereo experiment (those indicated in Figure 3.3); each observer recorded five settings for each location, probe size, and stimulus surface in the static stereo and head-motion conditions. Trials with different probe sizes and locations were interleaved, and trials with different depth-cue conditions were not interleaved. A typical trial was the same as that described in Section 4.2.1 for the stereo experiment.

The major procedural difference between the head-motion experiment and the stereo experiment was that the head-motion experiment was designed to test explicitly for learning effects; that is, it was designed to distinguish explicitly between the effect of adding head-motion parallax to a stereo presentation and the effect of observers' increasing familiarity with the object and the task. To this end, the eight test locations were split into two groups of four. The first group, the locations marked 1, 3, 6, and 8 in Figure 3.3, were the static-first test locations; observers made settings at those locations first in the static stereo condition and then in the head-motion condition. The second group, the locations marked 2, 4, 5, and 7 in Figure 3.3, were the motion-first locations; observers made settings at those locations first in the head-motion condition and then in the static stereo condition. The trials were arranged in the following order:

1. Observers performed 16 practice trials in the static stereo condition, one trial with each of the four probe sizes at each of the four static-first locations. These were purely for familiarization; no feedback was provided.

2. Observers performed 80 experimental trials in the static stereo condition at the static-first test locations.

3. Observers performed 16 practice trials in the head-motion condition, one trial with each of the four probe sizes at each of the four motion-first locations. Again, no feedback was provided.

4. Observers performed 80 experimental trials in the head-motion condition at the motion-first test locations.

5. Observers performed 80 experimental trials in the head-motion condition at the static-first test locations.

6. Observers performed 80 experimental trials in the static stereo condition at the motion-first test locations.

5.2.2 Results

Variability of the Settings

As in the stereo experiment described in the previous chapter, I begin by assessing whether the task is perceptually reasonable. Figure 5.1 shows the variability (measured as described in Section 3.4.4). All observers exhibit an overall decrease in variability with increasing scale; as in the stereo experiment, orientation judgments made with a large probe were at least as robust as those made with a small probe.

There was no consistent difference in variability between the depth-cue conditions. Further, the variabilities in this experiment are consistent with the variabilities found in the stereo experiment with the rough sphere.

Orientation Paths

Figure 5.2 shows the orientation paths for the test locations 1–4; Figure 5.3 shows the orientation paths for test locations 5–8. The objective orientation paths shown here are developed from the σ-normals.

The majority of the observers' orientation paths in this experiment stretch over a substantial range, indicating a substantial change in perceived surface orientation across scale. This is consistent with the results from the stereo experiment on the rough sphere, in which most of the paths were also extended.

The patterns of difference between the observers' orientation paths and the corresponding objective paths varied across test locations, and in some instances they varied from observer to observer; many of the patterns were similar to those found in the rough-sphere part of the stereo experiment. A common pattern in this experiment (and also in the stereo experiment) was data points clustering near a single value—typically the largest-scale value—on the observer's orientation path (e.g., at location 5 for all observers); this indicated that the perceived surface orientation at this location did not change much across scale, even though the objective surface orientation did. At some other test locations (e.g., location 3 for all observers), perceived surface orientation was close to constant for the three smallest scales (the three smallest-scale data points clustered together), and then changed

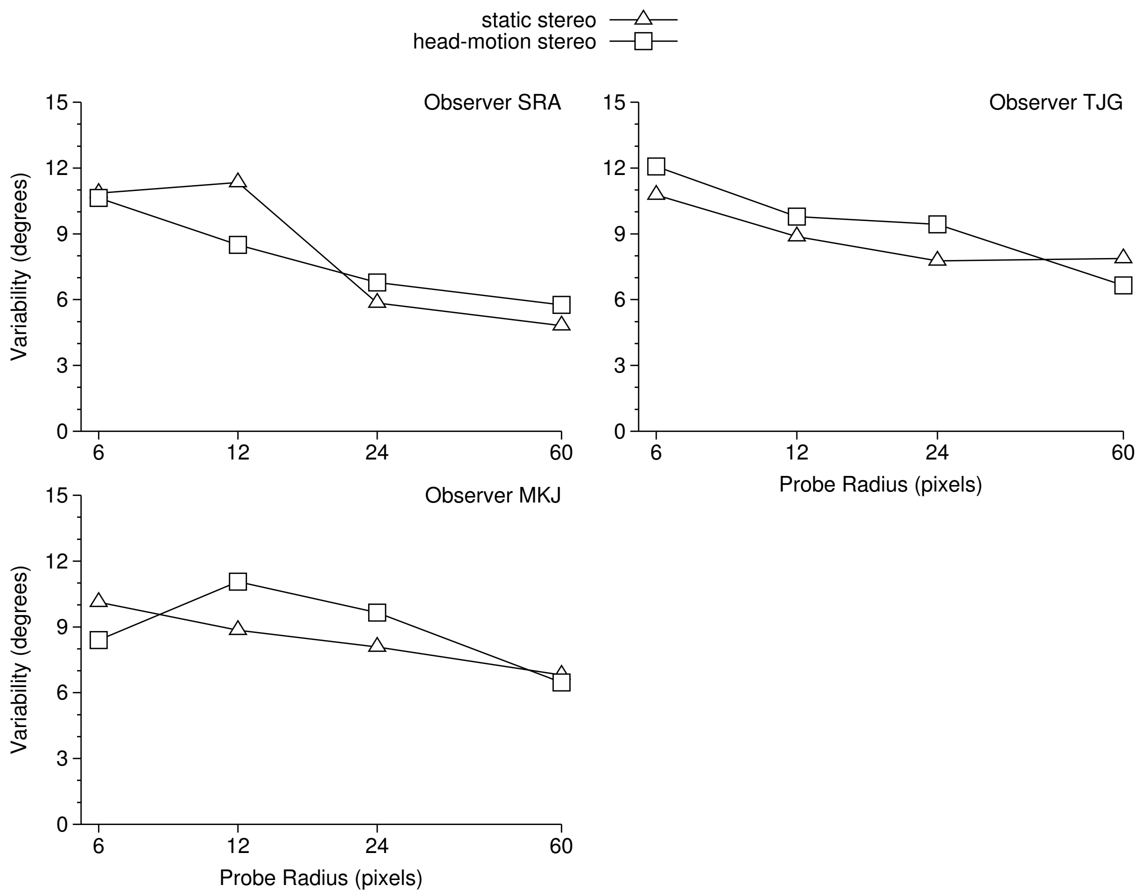

Figure 5.1: Variability in orientation settings, across all test locations.

sharply at the largest scale; as in the stereo experiment, this pattern indicated that the relatively smooth change in the objective orientation across scale was not present in the percept. Another pattern common to both experiments (e.g., observer MKJ's location 7 in this experiment; observer PHB's location 6 in the stereo experiment) was an observer's orientation path roughly paralleling the corresponding objective path, indicating that the observer's perceived surface orientation differed from the objective orientation by an approximately constant angular offset. While not as common as in the stereo experiment, approximate superimposition of the observer's path on the corresponding objective path (indicating a high degree of veridicality in the perceived surface orientation across scale) did happen in this experiment (e.g., observer SRA's locations 2 and 8).

Figure 5.4 shows the orientation paths for all observers for locations 6 and 8, with the individual settings marked to show the scatter. The scatter at these two locations is typical.

Veridicality Across Locations

Figure 5.5 shows the biases for all three observers, averaged across locations as described in Section 3.4.3. In both depth-cue conditions, accuracy increased (biases decreased) with increasing scale. This is the same accuracy pattern exhibited by the informed observers on this same surface in the stereo experiment; it is consistent with the accuracy pattern of the naïve observers with stereo in that experiment. The biases in this experiment were notably higher at the smallest probe sizes than the corresponding biases in the stereo experiment, although the biases are comparable between the two experiments at larger scales.

Figure 5.6 shows the effect on accuracy of adding head-motion parallax to the static stereo presentation, by graphing the difference between the static-stereo biases and the head-motion biases. The addition of head motion improved accuracy most at small scales; all three observers got their greatest benefit at one of the two smallest probe sizes, and all three had negligible or even negative benefit (adding motion made the settings *less* accurate) at the largest probe size.

Separate analysis of the static-first and motion-first test locations revealed no evidence of learning effects in this experiment. The differences in results between the two groups of test locations in this experiment were entirely consistent with the inter-location differences found in the stereo experiment.

5.2.3 Discussion

This experiment shows that a multiscale approach to measuring perceived 3D shape is not limited to adding stereo to a non-stereo presentation. With head-motion parallax, the same patterns were present in the orientation paths as were present in the stereo experiment; many orientation paths were extended in the head-motion parallax experiment, indicating across-scale changes in perceived surface orientation. The same patterns of difference between the observers' and objective orientation paths were in evidence in both experiments, offering the same types of information on the observers' percepts.

In addition, this experiment shows that multiscale measurements in general, and the MTOPS tool in particular, can detect different across-scale patterns of benefit. The increased accuracy due to adding head-motion parallax to a stereo presentation in this experiment was almost all at small scales; this increase in accuracy was completely absent at the largest scale. The small-scale benefit is consistent with the results, mentioned earlier, of Arthur *et al.* [3], Norman *et al.* [87], and Ware

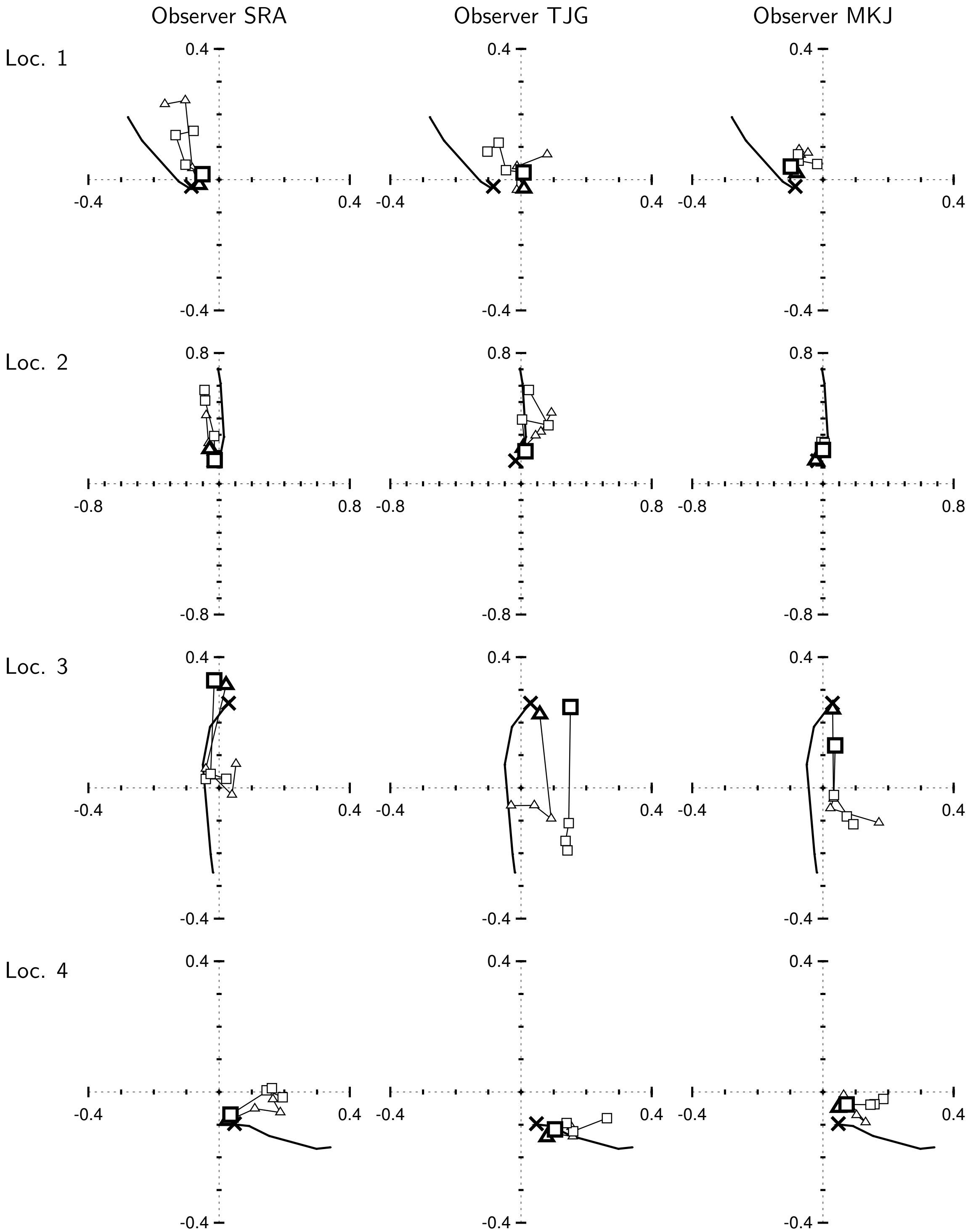

Figure 5.2: The observers' orientation paths for test locations 1–4. Static stereo data are shown by triangles and head-motion stereo data by squares; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

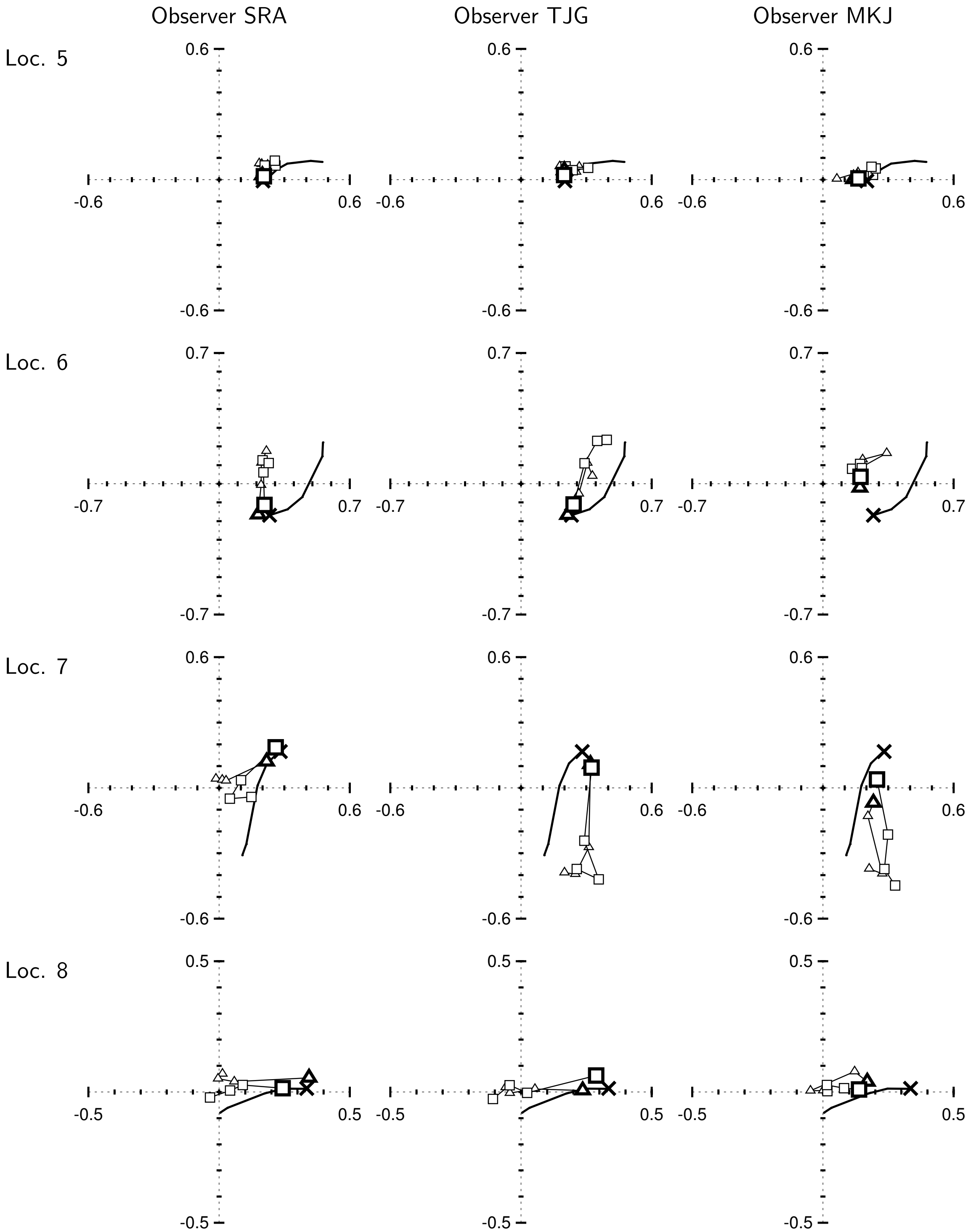

Figure 5.3: The observers' orientation paths for test locations 5–8. Static stereo data are shown by triangles and head-motion stereo data by squares; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

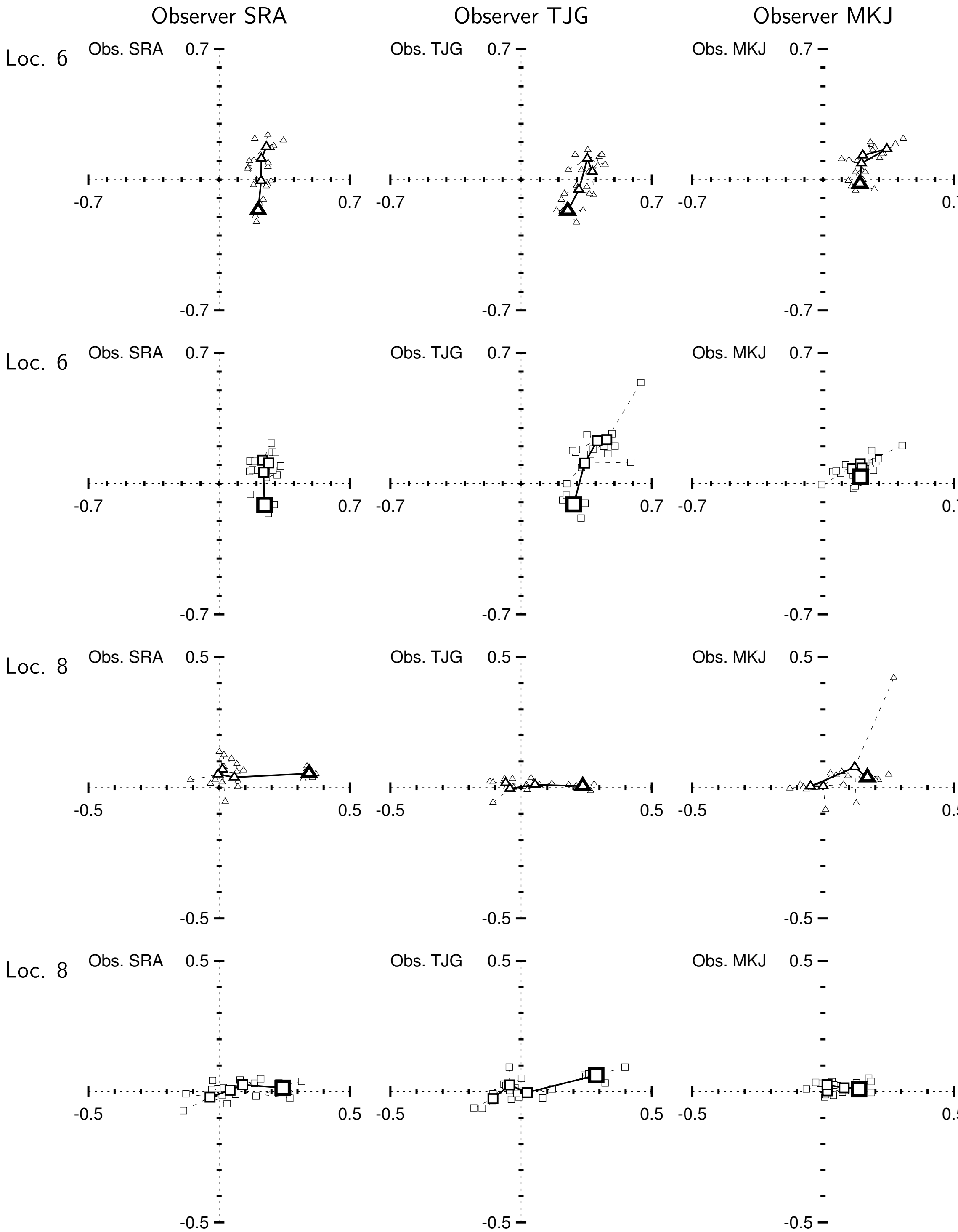

Figure 5.4: The orientation paths for locations 6 and 8, with the individual settings plotted to show the scatter. The orientation paths are shown as in Figures 5.2 and 5.3, with circles for non-stereo data, triangles for stereo data, and the large-scale end of each path marked with a larger, darker symbol. The individual settings are shown by smaller, lighter symbols; each individual setting is linked to its mean setting by a dashed line.

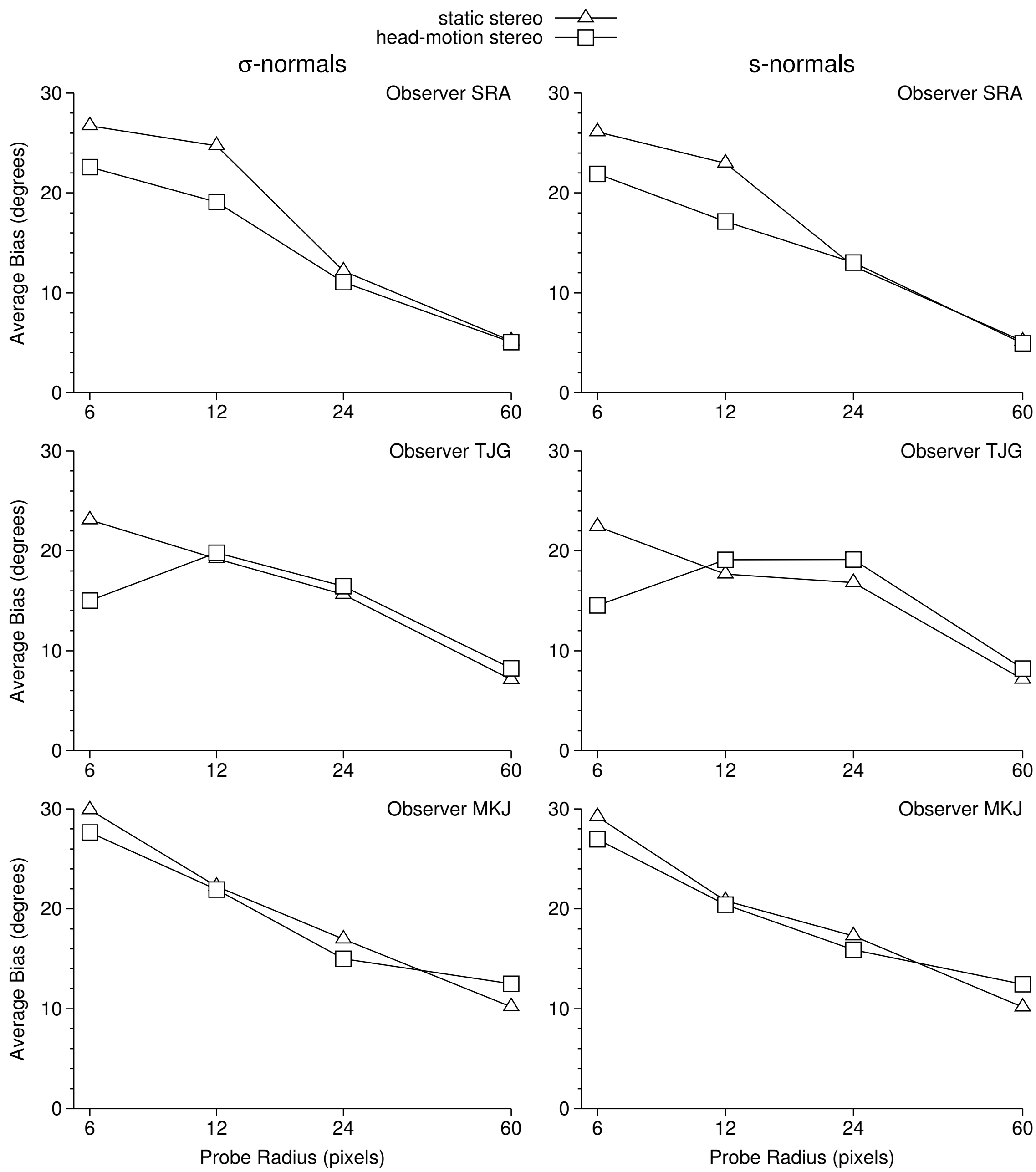

Figure 5.5: Bias in the perceived surface orientation, averaged across locations, for all observers. The graphs on the left were calculated using σ-normals; those on the right were calculated using s-normals.

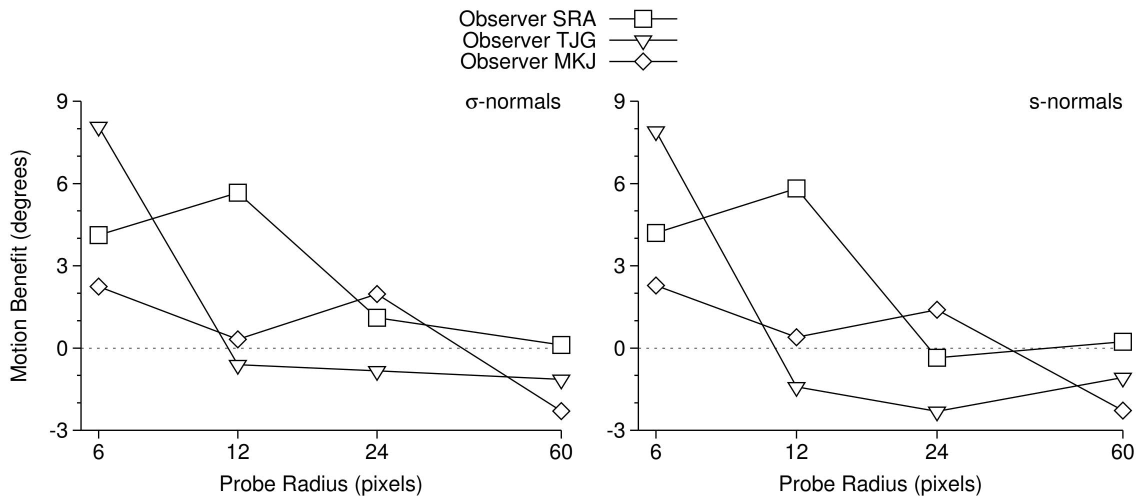

Figure 5.6: Improvement in accuracy due to adding head motion to the display, for all three observers. The graph on the left is based on biases calculated using σ-normals; the graph on the right is based on biases calculated using s-normals.

and Franck [136], all of whom used small-scale tasks[2]. The absence of large-scale benefit in this experiment is new, however, and underscores the advantage of multiscale measurement.

5.3 Object Motion

The simplest form of object motion is object rotation. Visualizations based on rotation, in turn, come in two major types: automatic motion (called auto-motion in this chapter), and user-controlled rotation (referred to in this chapter as manual motion). With auto-motion, the computer rotates the object, commonly without control inputs from the user. The user may have control of the acceleration or the velocity of the rotation but does not have positional control of the object. With manual motion, the user has direct positional control over the rotation of the object, commonly by means of a slider (physical or virtual) or a joystick controlled by hand [11, 21, 47, 75].

Lipscomb [75] finds that users prefer manual motion to automatic motion for molecular modeling. Auto-motion, however, is easier to deliver on a remote computer; not only is the user interface much easier to supply (no slider or joystick is needed), but the presentation of the data can readily be packaged as a cine loop or a video. The second experiment measured the effect of adding manual motion to a static presentation and the effect of adding auto-motion to the same presentation; the experiment also compared those effects.

[2]Norman *et al.* used the gauge figure developed by Koenderink *et al.* [65], and Norman *et al.* include in their paper an illustration showing that the gauge figure was small compared to their stimulus surfaces. In the two path-tracing studies (Arthur *et al.* and Ware and Franck), the paths are thin structures crossing each other; Ware [134] points out that the major cause of errors in path-tracing tasks is path crossings. Following one thin structure where it crosses behind or in front of another thin structure is a small-scale task, so the most error-prone part of the path tracing took place at small scale. Both path-tracing studies concentrate on error rates.

5.3.1 Methods

The probe and task used for this experiment were those described in Section 3.2. Five probe sizes were used, with radii of 8, 11, 16, 22, and 32 pixels. In this experiment, the probe was controlled by means of a Logitech WingMan Extreme Digital 3D joystick with the centering spring removed. Each joystick angle corresponded directly with an orientation of the probe.

The object-motion experiment used three depth-cue conditions. The first was a static, non-stereo condition. The second was a manual motion condition, in which the observer used the joystick's throttle lever to rotate the object about a vertical axis through the object's center; each throttle position corresponded directly with a particular rotational position of the object. The third depth-cue condition was a automatic-motion condition, in which the object was rotated automatically with no observer control of the rotation.

The images were displayed on a ViewSonic P775 monitor; the picture size was 1024×768 pixels. At that picture size, the monitor's resolution was 34.1 dots/cm (86.7 dots per inch) both horizontally and vertically.

Object rotation was simulated by displaying precomputed images, portraying the object at evenly spaced rotations about the vertical axis through its center. Two stimulus surfaces were used, the two solid-noise surfaces described in Section 3.3.2 and shown in Figures 3.4 and 3.5. Sixty-one images of each stimulus surface were precomputed, covering rotations from +15 degrees to -15 degrees at half-degree intervals.

The same images were used for both manual and automatic rotation, so that the visual information in the two depth-cue conditions differed only in the type of the motion. Because the images were evenly spaced every half-degree of rotation, the automatic motion appeared to jerk at the ends of the rotation; on the other hand, the even spacing made the manual control more straightforward. Aside from the ends, the rotation appeared smooth in the automatic-motion condition; all the observers found it easy to produce a smooth-looking rotation in the manual-motion condition, as well.

The frame rate of the system was approximately 40 Hz (25 ms/frame) when the object was stationary. The frame rate was around 25 Hz (40 ms/frame) when the object was rotating (whether with user motion or auto-motion).

I measured system latency using a variant of the camera method described in Swindells *et al.* [112]; specifically, the joystick was driven in a circle by a phonograph turntable running at 78 rpm, and I used a Sony DSC-P50 digital camera to take pictures showing the simultaneous positions of the turntable and the normal stick on a somewhat-modified probe. Because the turntable was driving the joystick at a constant speed, the lag time could be inferred from the difference between the angular position of the joystick and the angular position of the probe in a given picture. The results indicate that system lag is 1.5 to 2 frames, only slightly above the 1.5-frame inherent lag suggested by Arthur *et al.* [3].

Each image was computed from a viewpoint 2000 pixels (58.6 cm) distant; as in previous experiments, this approximated the actual distance from the observer to the screen. The field of view of the perspective transformation matched the image size on the monitor, so that the center of the object appeared to be approximately in the plane of the screen.

Procedures

The procedures in this experiment were largely similar to those in the previous two experiments. Again, the experimental area was curtained off to block visual distractions, but it was not completely darkened. The observers were seated approximately 60 cm from the display, with unrestrained heads; viewing was binocular, without stereo. Trials with different probe sizes and test locations were interleaved, and trials with different depth-cue conditions were not interleaved.

The observers were seven adult volunteers. Two observers (PHB, the author, and LFB) had extensive experience with the task and were thoroughly familiar with the surfaces from pilot work. The other five were naïve observers. All observers were provided with the project overview and observers' tips found in Appendix C. All observers had normal or corrected-to-normal vision.

Five test locations were used on each stimulus surface, the same locations indicated in Figures 3.4 and 3.5. In each depth-cue condition, observers made fifty practice settings, with feedback, on the surface shown in Figure 3.4 (two settings at each test location and probe size), and then 125 experimental settings, without feedback, on the surface shown in Figure 3.5. Five observers (LFB, PHB, JRLE, LNG, and DOH) made their settings first in the static condition, then in the manual-motion condition, and finally in the auto-motion condition. To control for order effects between the motion conditions, the other two observers (LAA and CBF) made their settings first in the static condition, then in the auto-motion condition, and finally in the manual-motion condition. LFB, PHB, JRLE, LNG, and DOH were thus manual-first observers, and LAA and CBF were auto-first observers.

A typical trial in this experiment differed only slightly from those in the previous two experiments. As noted above, probe adjustment in this experiment was by means of a joystick, not a mouse. At the end of each trial, the observer squeezed the joystick trigger to record the setting and pressed one of the buttons on top of the joystick to begin the next trial. A physical model of a probe, made of wire, was available to the observers at all times.

5.3.2 Results

Variability of the Settings

As in the previous two experiments, I begin by assessing whether the task is perceptually reasonable. Figure 5.7 shows the variability (measured as described in Section 3.4.4). In general, variability remained approximately constant across scale, although the across-scale patterns do differ from observer to observer and from condition to condition for the same observer. There is no evidence here to suggest that large-scale settings were systematically less robust than small-scale settings, however.

There is also no consistent difference in variability between depth-cue conditions. The two observers with the highest variability overall (JRLE and LAA) were more consistent (lower variability) with object motion than without it. Two other observers (CBF and LFB) were generally more consistent (lower variability) in the static condition than in either motion condition, however, and the balance of the observers made equally robust settings in all conditions.

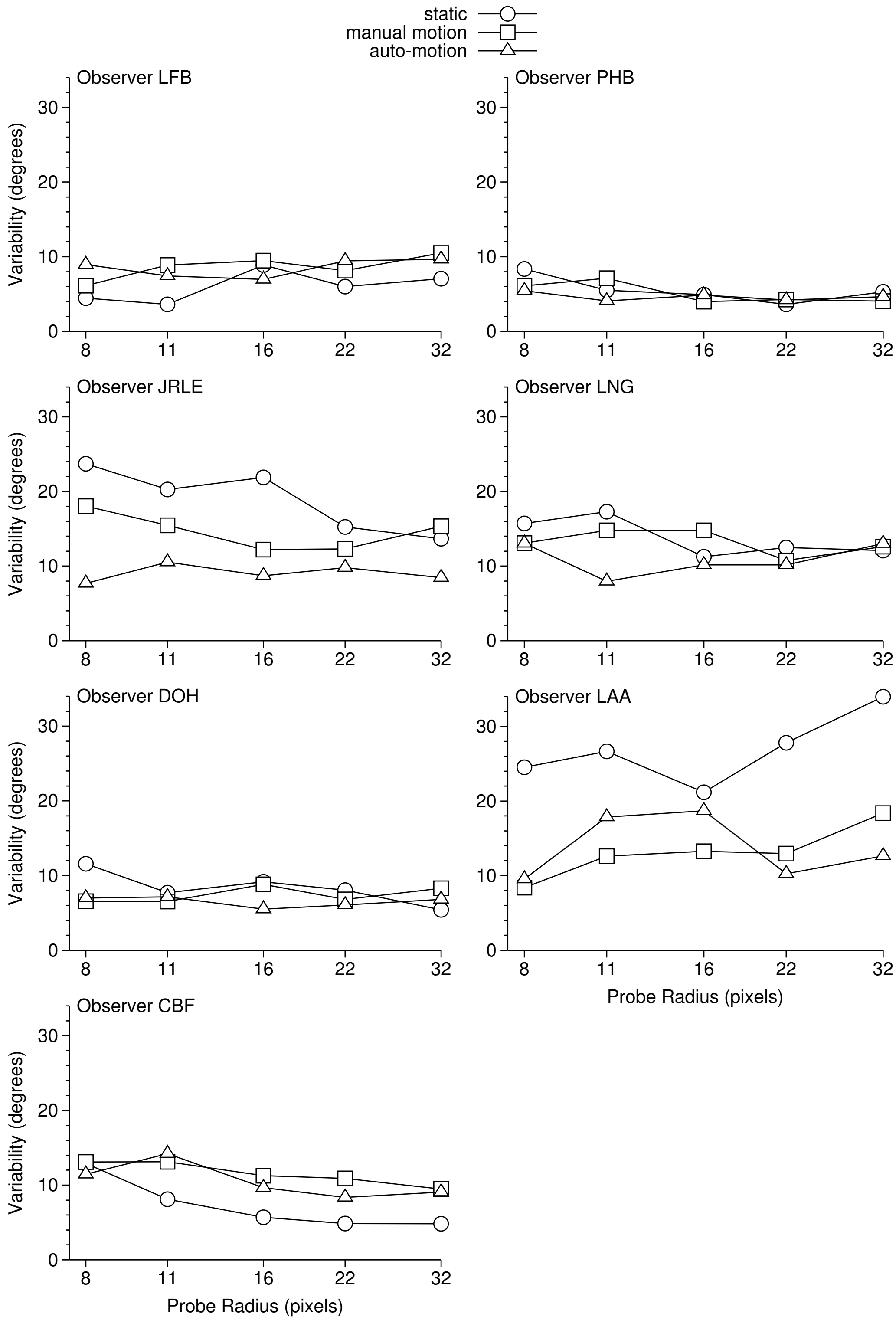

Figure 5.7: Variability in orientation settings for all observers.

Orientation Paths

Figures 5.8–5.12 present the orientation paths for the object-motion experiment. The vast majority of the observers' orientation paths in this experiment followed the patterns described in the stereo experiment: observers' orientation paths that were approximately superimposed on the corresponding objective orientation paths, observers' data points clustering with no corresponding cluster in the objective orientation path, and observers' paths parallel to the corresponding objective paths.

One observer (LAA), at two locations (1 and 5) in the static condition only, did have orientation paths that extended over a significant range but wandered randomly with no apparent relation to the objective orientation path. This suggests that observer LAA was aware that the surface orientation changed across scale at those locations but was not able to perceive the actual direction of that change. Since this did not occur for LAA at other test locations or in other depth-cue conditions at locations 1 and 5 and since it did not occur for other observers, this does not indicate a probe failure; rather, it indicates only that LAA found the surface at locations 1 and 5 confusing in the static condition.

Figures 5.13 and 5.14 show the orientation paths for four observers for locations 2 and 3, with the individual settings marked to show the scatter. The scatter at these two locations for these observers is typical.

Veridicality Across Locations

The bias for all observers is presented in Figures 5.15 (calculated using σ-normals) and 5.16 (calculated using s-normals). Observers' accuracy was, in general, approximately constant across scale; there was no clear, consistent pattern of biases increasing or decreasing as scale increased.

There is no obvious difference between the bias results for the auto-first observers (LAA and CBF) and the bias results for the motion-first observers. There is also no consistent difference between the experienced observers and the naïve observers; while one of the experienced observers (PHB) does have somewhat lower biases than the naïve observers, the other experienced observer (LFB) does not.

The major accuracy effects were effects of the depth-cue conditions. I show the effects of the two types of object motion on the veridicality of observers' settings by graphing the difference in bias between the manual motion condition and the static condition and between the auto-motion condition and the static condition. The effect of the manual motion condition is shown in Figure 5.17 for all seven observers; the effect of automatic motion is shown in Figure 5.18. The same differences are presented observer-by-observer, for both types of motion, in Figure 5.19 (based on the biases calculated with σ-normals) and Figure 5.20 (based on the biases calculated with s-normals).

The most obvious, consistent effect is that auto-motion improved accuracy substantially for every observer but DOH. The across-scale pattern of that improvement varied from observer to observer, but almost every observer did benefit at almost every probe size.

The effect of manual motion is far less consistent. Three of the seven observers (LFB, DOH, and CBF) got improvement from manual motion only at some probe sizes. There was no consistent pattern to the probe sizes at which those three observers did benefit; manual motion helped LFB most at the smaller scales, CBF most at larger scales, and DOH most in the mid-to-large scales.

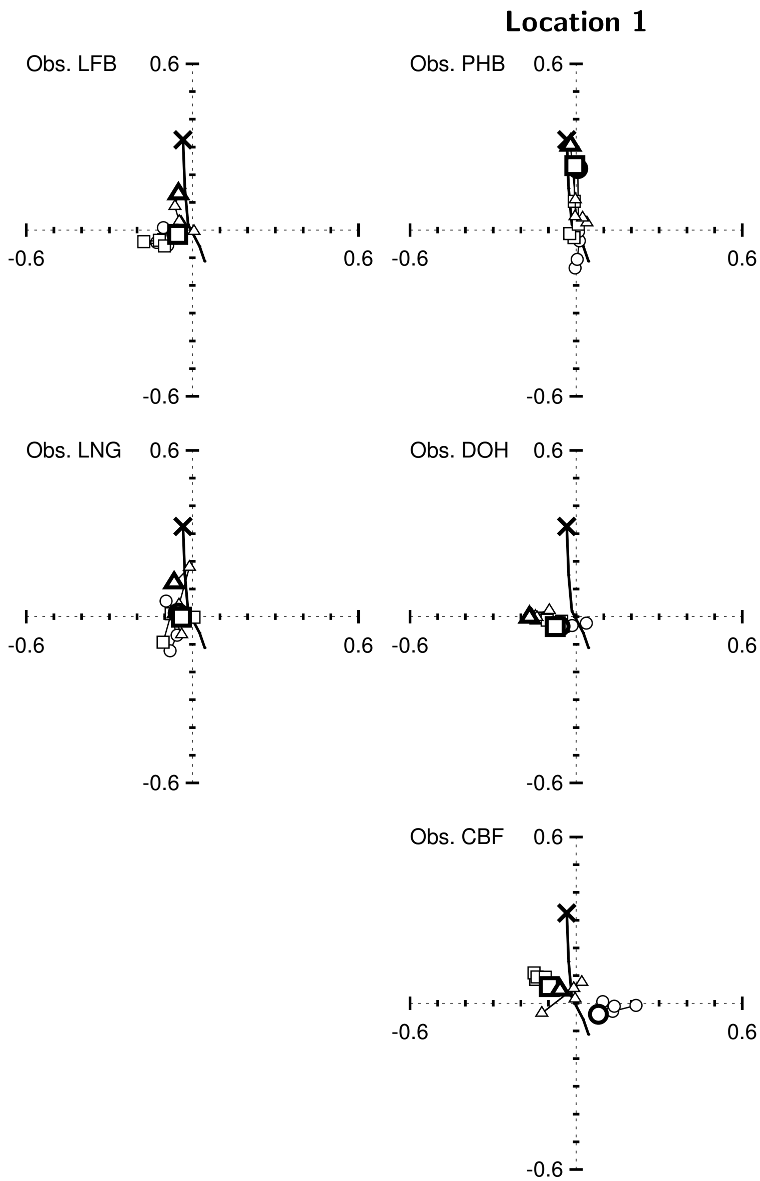

Figure 5.8: The orientation paths for location 1, for all observers. Static data are shown by circles, manual-motion data by squares, and auto-motion data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an $\times$.

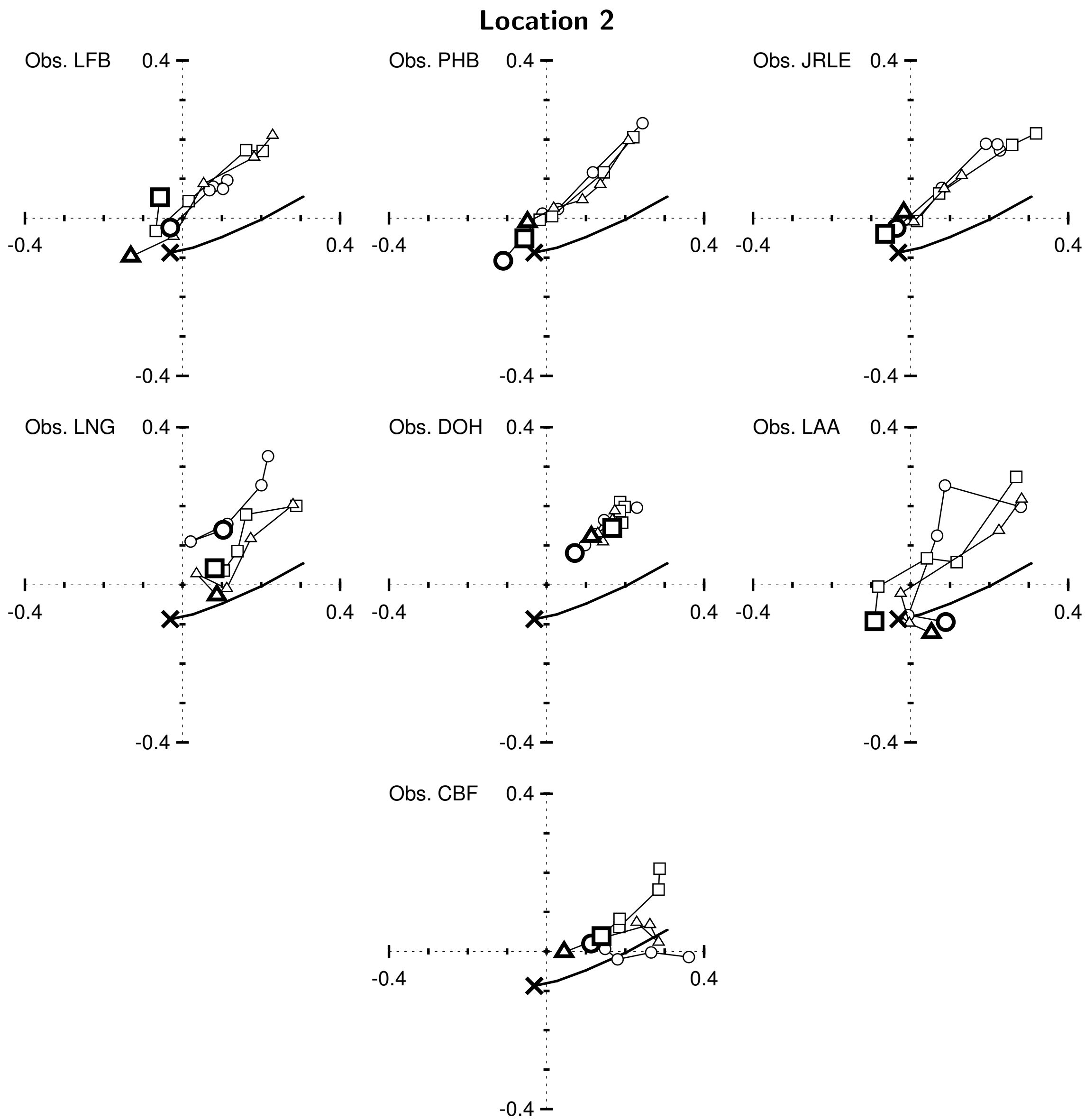

Figure 5.9: The orientation paths for location 2, for all observers. Static data are shown by circles, manual-motion data by squares, and auto-motion data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

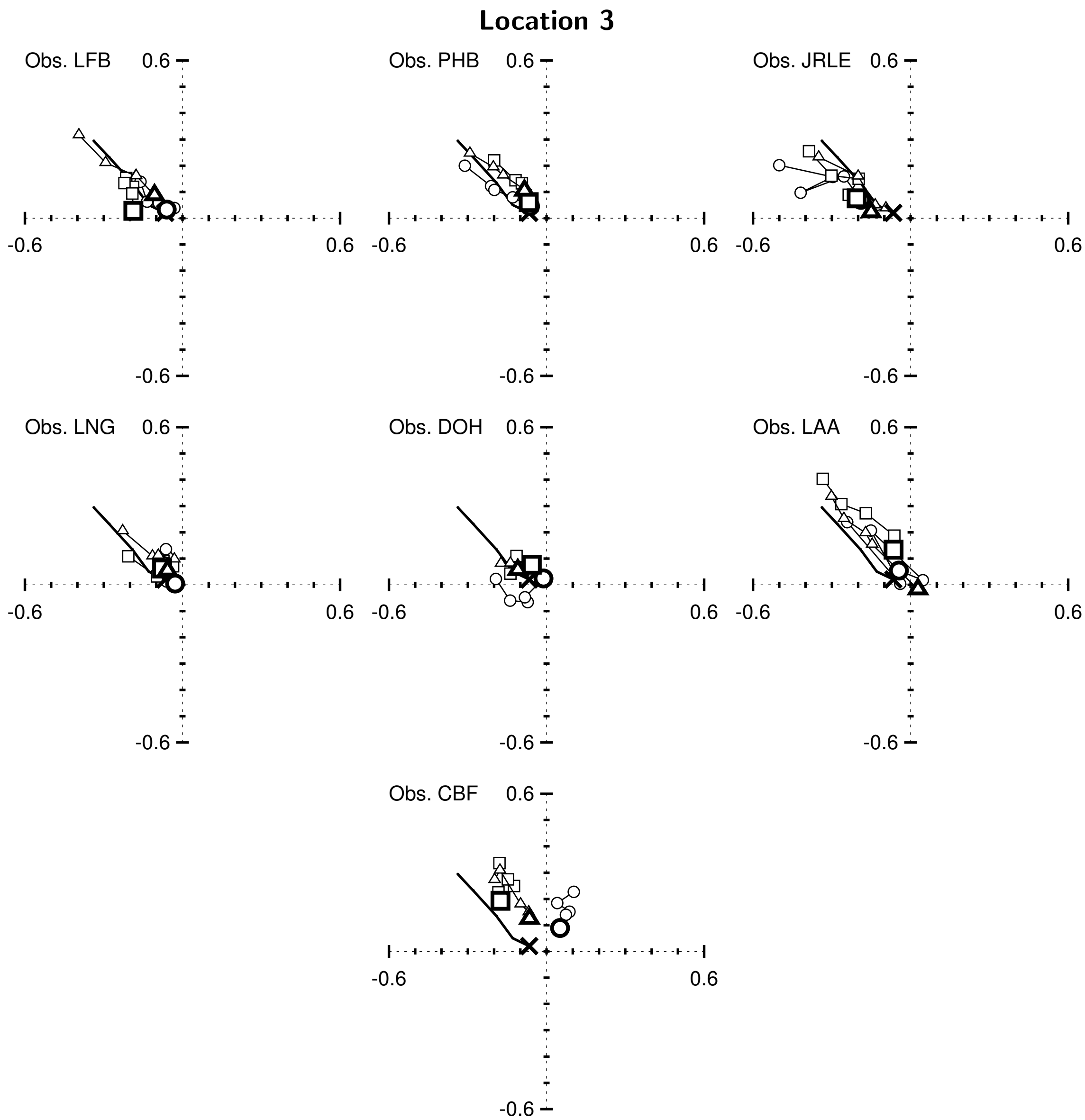

Figure 5.10: The orientation paths for location 3, for all observers. Static data are shown by circles, manual-motion data by squares, and auto-motion data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

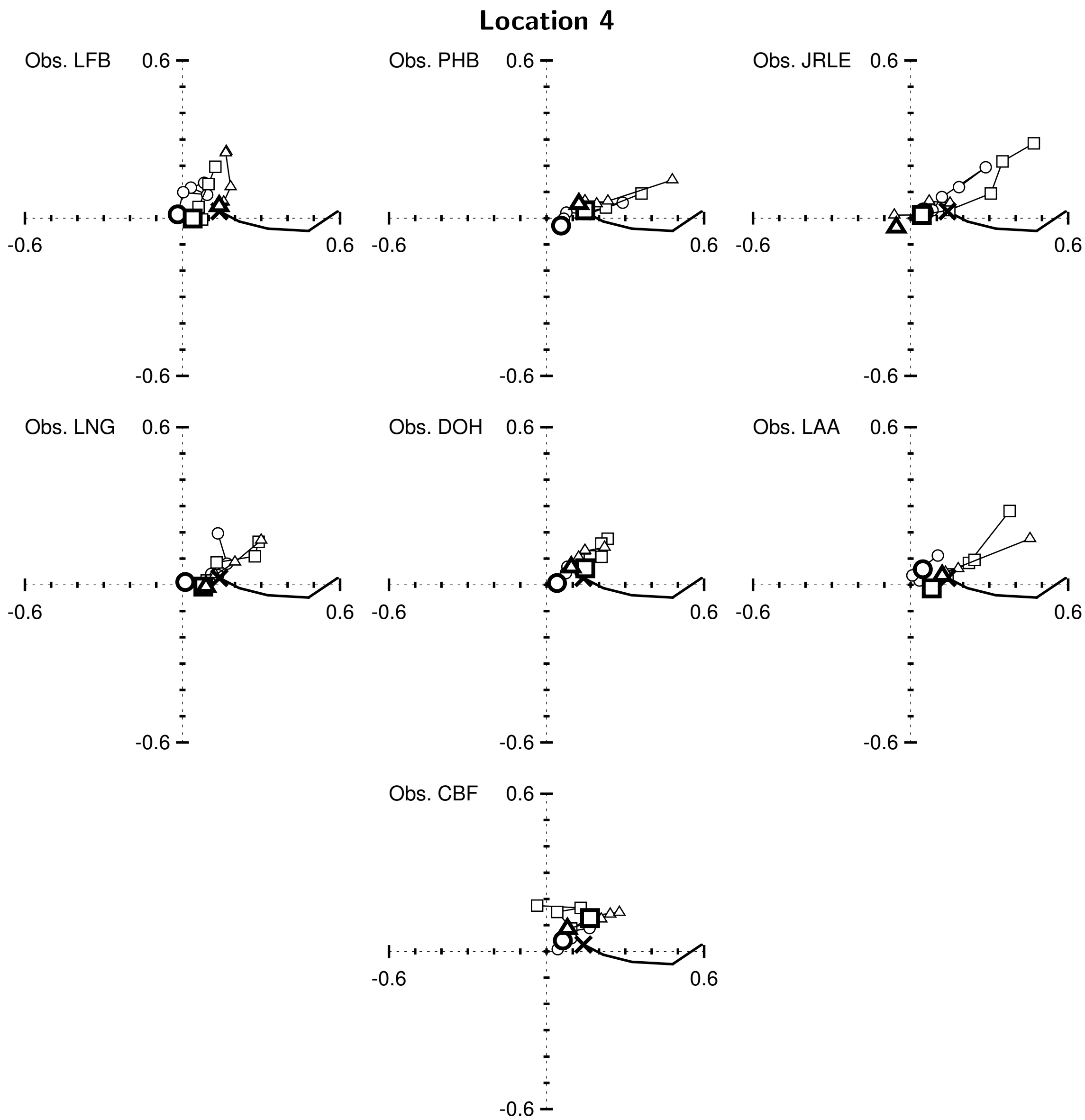

Figure 5.11: The orientation paths for location 4, for all observers. Static data are shown by circles, manual-motion data by squares, and auto-motion data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an ×.

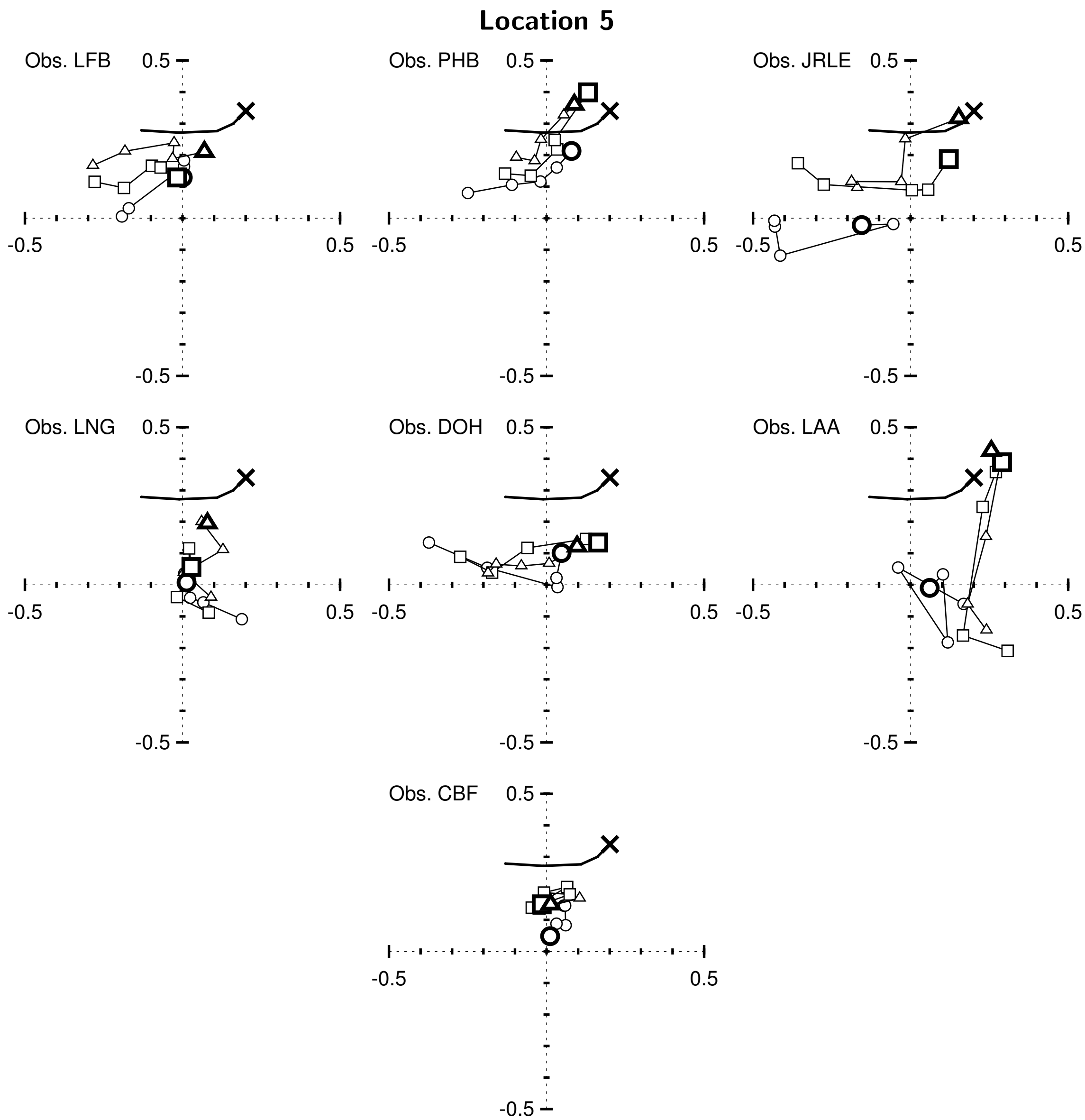

Figure 5.12: The orientation paths for location 5, for all observers. Static data are shown by circles, manual-motion data by squares, and auto-motion data by triangles; the thicker lines with no symbols indicate the objective orientation paths. The large-scale end of each observed orientation path is marked by a larger, darker symbol; the large-scale end of the objective orientation path is shown with an $\times$.

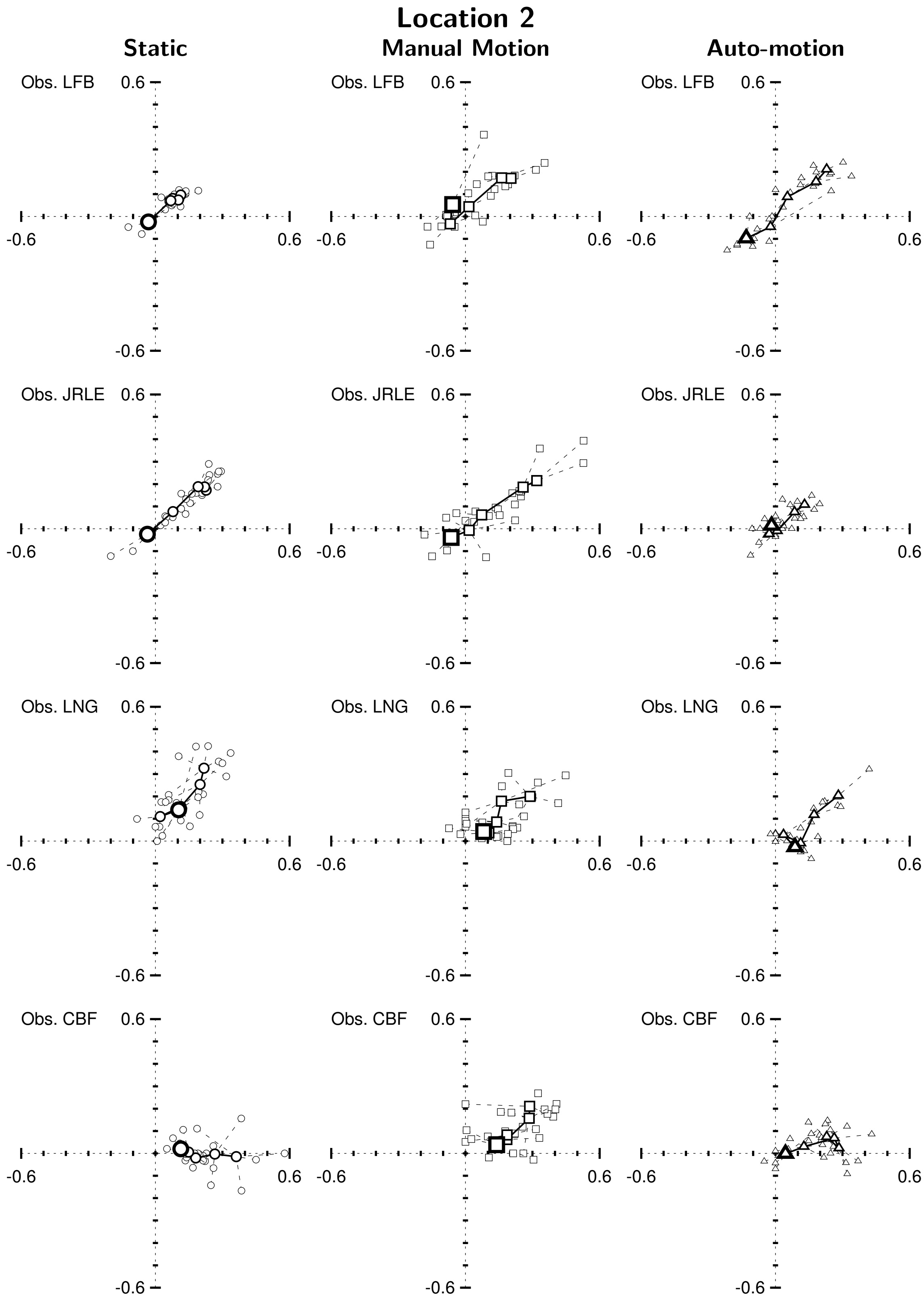

Figure 5.13: Four observers' orientation paths for location 2, with the individual settings plotted to show the scatter. The orientation paths are shown as in Figures 5.8–5.12, with circles for static data, squares for manual-motion data, triangles for auto-motion data, and the large-scale end of each path marked with a larger, darker symbol. The individual settings are shown by smaller, lighter symbols; each individual setting is linked to its mean setting by a dashed line.

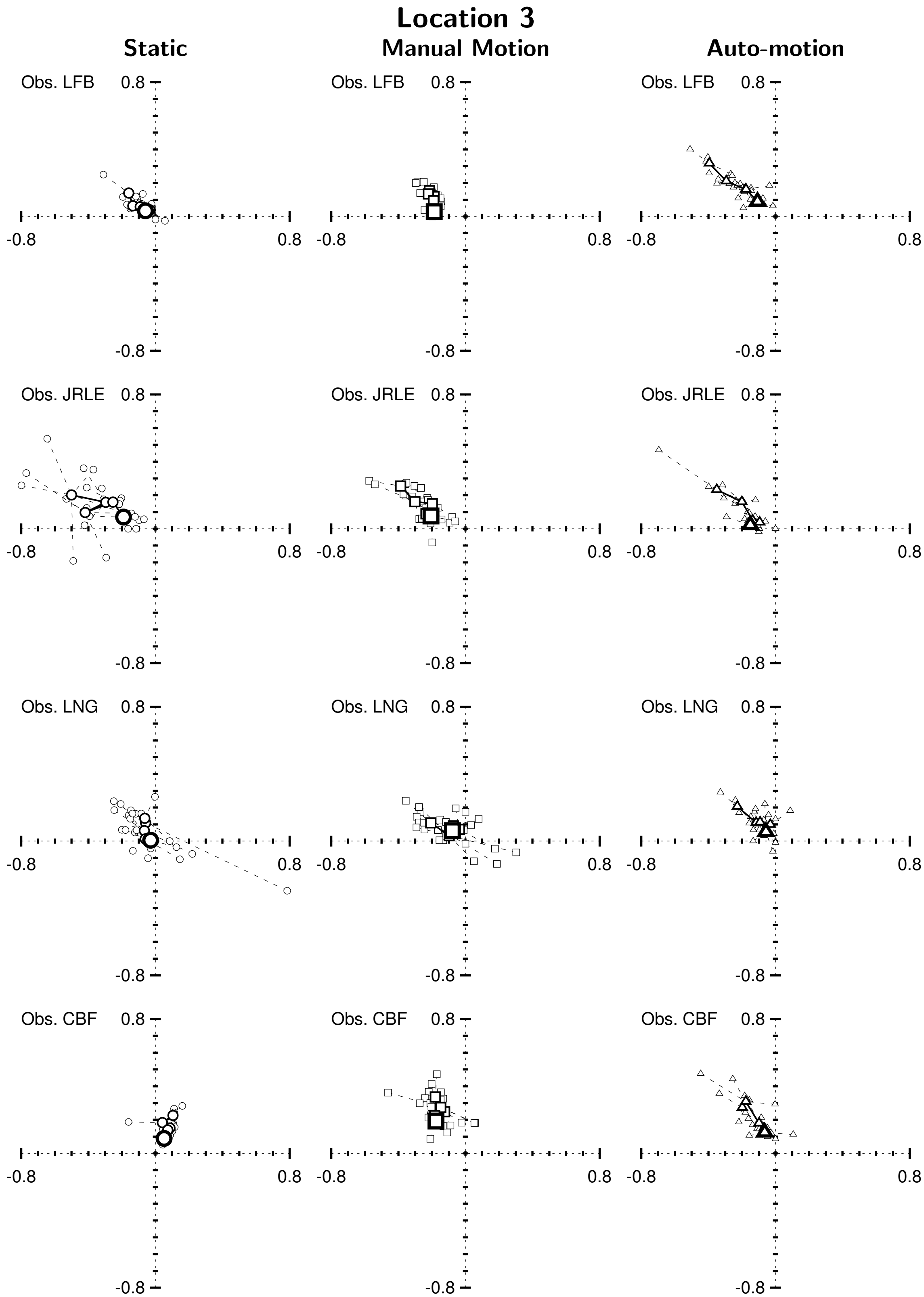

Figure 5.14: Four observers' orientation paths for location 3, with the individual settings plotted to show the scatter. The orientation paths are shown as in Figures 5.8–5.12, with circles for static data, squares for manual-motion data, triangles for auto-motion data, and the large-scale end of each path marked with a larger, darker symbol. The individual settings are shown by smaller, lighter symbols; each individual setting is linked to its mean setting by a dashed line.

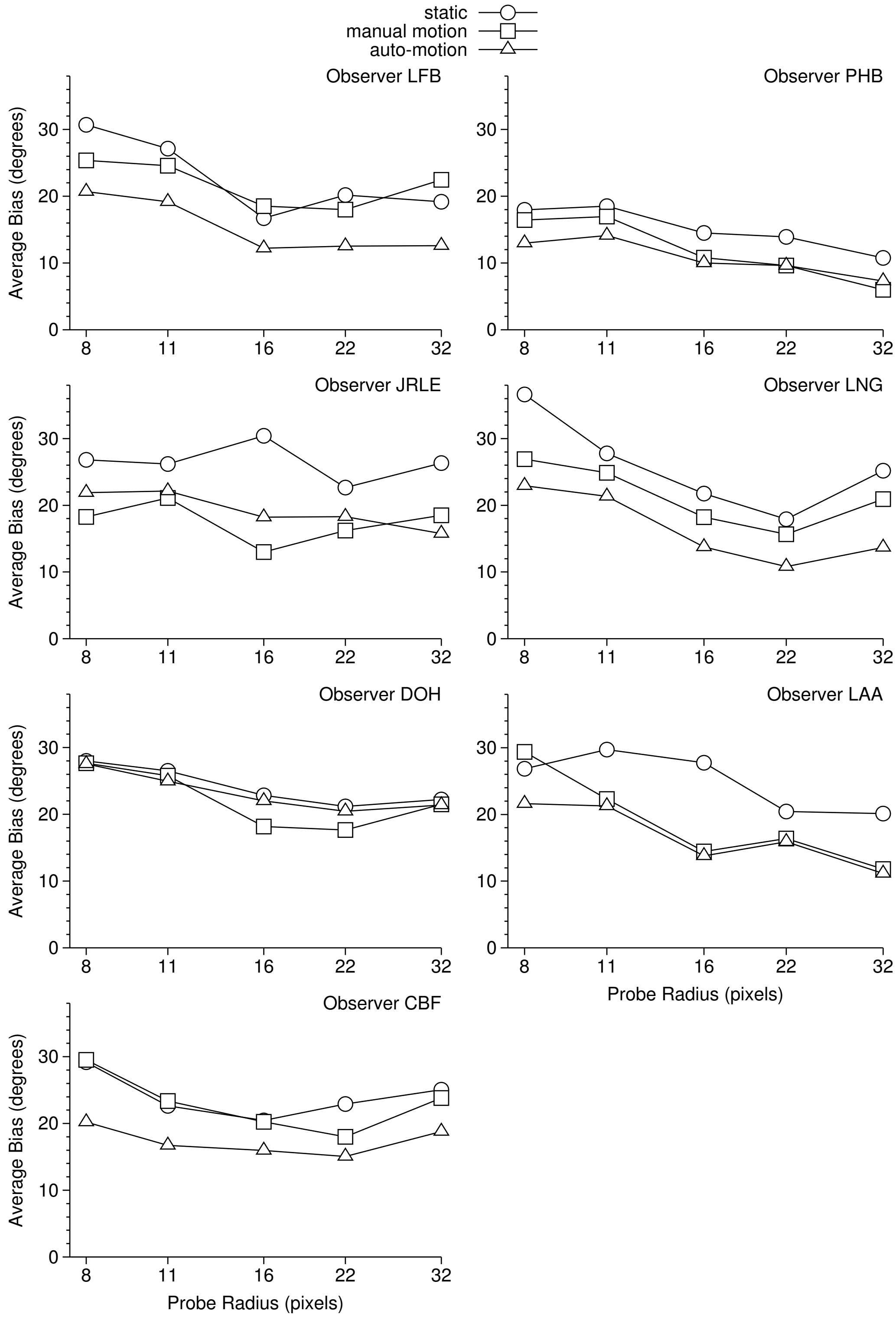

Figure 5.15: Bias in the perceived surface orientation, averaged across locations, for all observers. These biases were calculated using σ-normals.

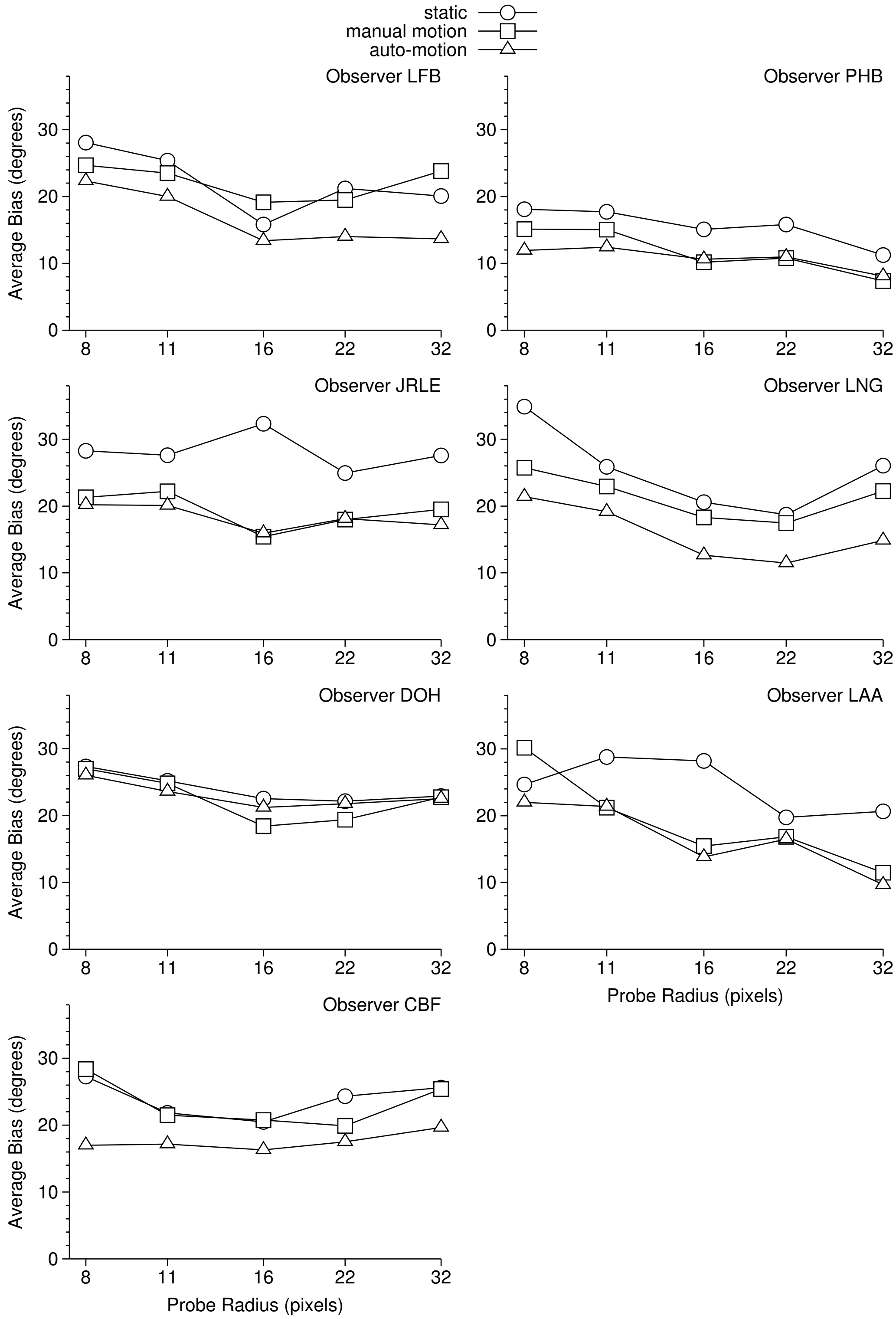

Figure 5.16: Bias in the perceived surface orientation, averaged across locations, for all observers. These biases were calculated using s-normals.

The other four observers (PHB, JRLE, LNG, and LAA) did get a substantial accuracy benefit from manual motion. There was no consistent across-scale pattern to the benefit for those four observers; PHB and LNG got nearly constant benefit across scales, and JRLE and LAA got much more benefit at 16 pixels than at 11 or 22. For LNG, the improvement in accuracy with manual motion was less than with auto-motion; for LAA, the improvement with manual motion was no greater than with auto-motion. For PHB, the accuracy benefit of manual motion was no greater than the benefit of auto-motion; if the biases are calculated with σ-normals, the benefit of manual motion was less than that of auto-motion. Finally, of the four observers who got consistent benefit from manual motion, only JRLE showed greater benefit from manual motion than from auto-motion, and then only if the biases are calculated using σ-normals.

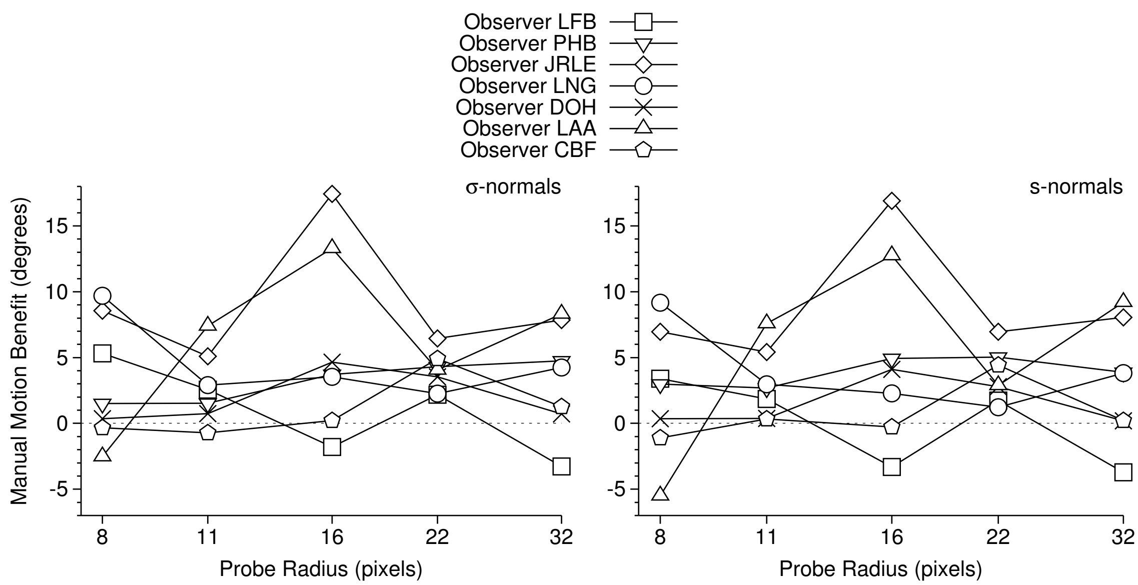

Figure 5.17: Improvement in accuracy of perceived surface orientation due to the addition of manual motion to the display, for all observers.

5.3.3 Discussion

The object-motion experiment underscores both the possibility and fruitfulness of multiscale measurement in another task and on another surface than those measured previously. The patterns of the observers' orientation paths show that on this very different stimulus surface, with different depth cues available, observers' perceived surface orientation still changed consistently across scale, and the relationships between the observers' orientation paths and the corresponding objective paths continue to provide information about the veridicality of the observers' percepts. Observers' extended orientation paths, parallel to or superimposed on the objective paths, indicate that observers were able to appreciate the change of orientation across scale. Clusters of two, three, or more data points indicate that at some test locations, observers' percepts supported judgments only at some scales. Random movements of the orientation path across scale are, with a few exceptions noted previously, limited to areas where observers' data points clustered; in the vast majority of cases, extended scale paths show some comprehension of the 3D structure of the stimulus surface.

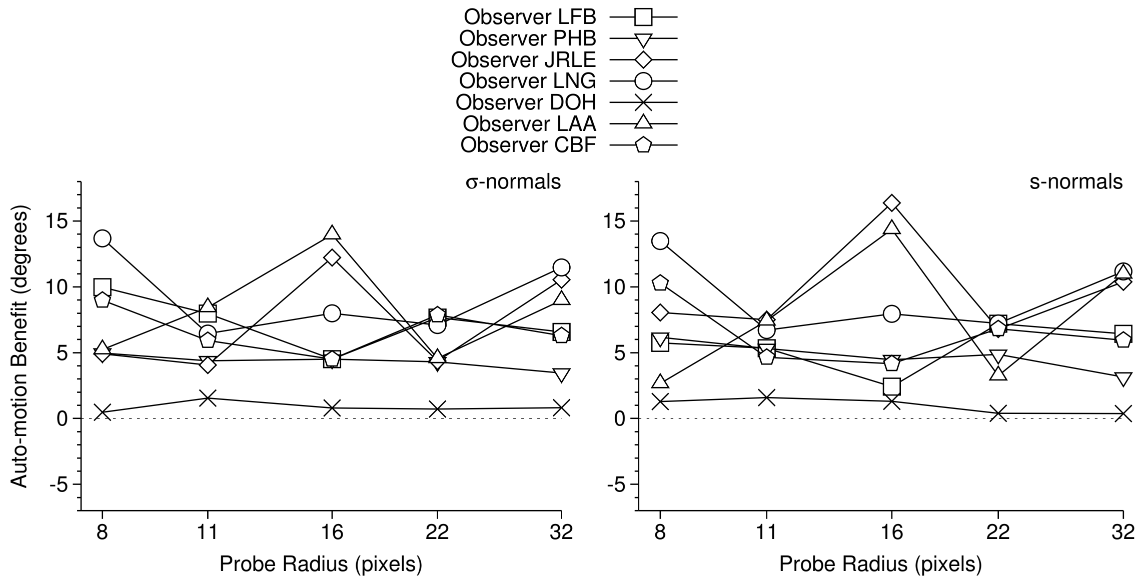

Figure 5.18: Improvement in accuracy of perceived surface orientation due to the addition of automatic motion to the display, for all observers.

The object-motion experiment also demonstrates the use of MTOPS to evaluate more than one competing depth-cue condition against a common baseline. While manual motion did improve the accuracy of many of the observers compared to a static presentation, auto-motion gave a more consistent accuracy benefit, and for most observers auto-motion gave a larger benefit than did manual motion.

The observers may have gotten a more consistent accuracy benefit from auto-motion, but the preferences they expressed favored manual motion. Three of the seven observers volunteered (unprompted) that they liked manual motion better than auto-motion, although two of those three were in fact more accurate using auto-motion. No observer expressed a preference for auto-motion. The disagreement between observers' preferences and their actual performance underscores the need for direct testing in visualizations where veridicality is important.

Lipscomb's chemists [75] also overwhelmingly preferred manual motion to automatic motion; they may have had better performance reasons for their preference, however. Lipscomb's users were doing tasks that required picking and manipulation, which are difficult to do on a rotating object. MTOPS requires only that the observer rotate the probe, which was made manageable (if not necessarily easy) by nesting the rotation of the probe inside the coordinate system of the rotating object. The task difference between MTOPS and Lipscomb's chemists underscores the necessity of applying psychophysical results to actual visualizations in the context of other human-factors considerations.

The accuracy benefit found in the object-motion experiment is not inconsistent with the benefits found in the stereo and head-motion experiments. The stereo experiment added stereo to a non-stereo presentation and found an accuracy benefit chiefly at large scale; the head-motion experiment further added head motion to a stereo presentation and found that accuracy improved at small scale. Taken together, these patterns suggest that adding head-motion stereo to a static, non-stereo

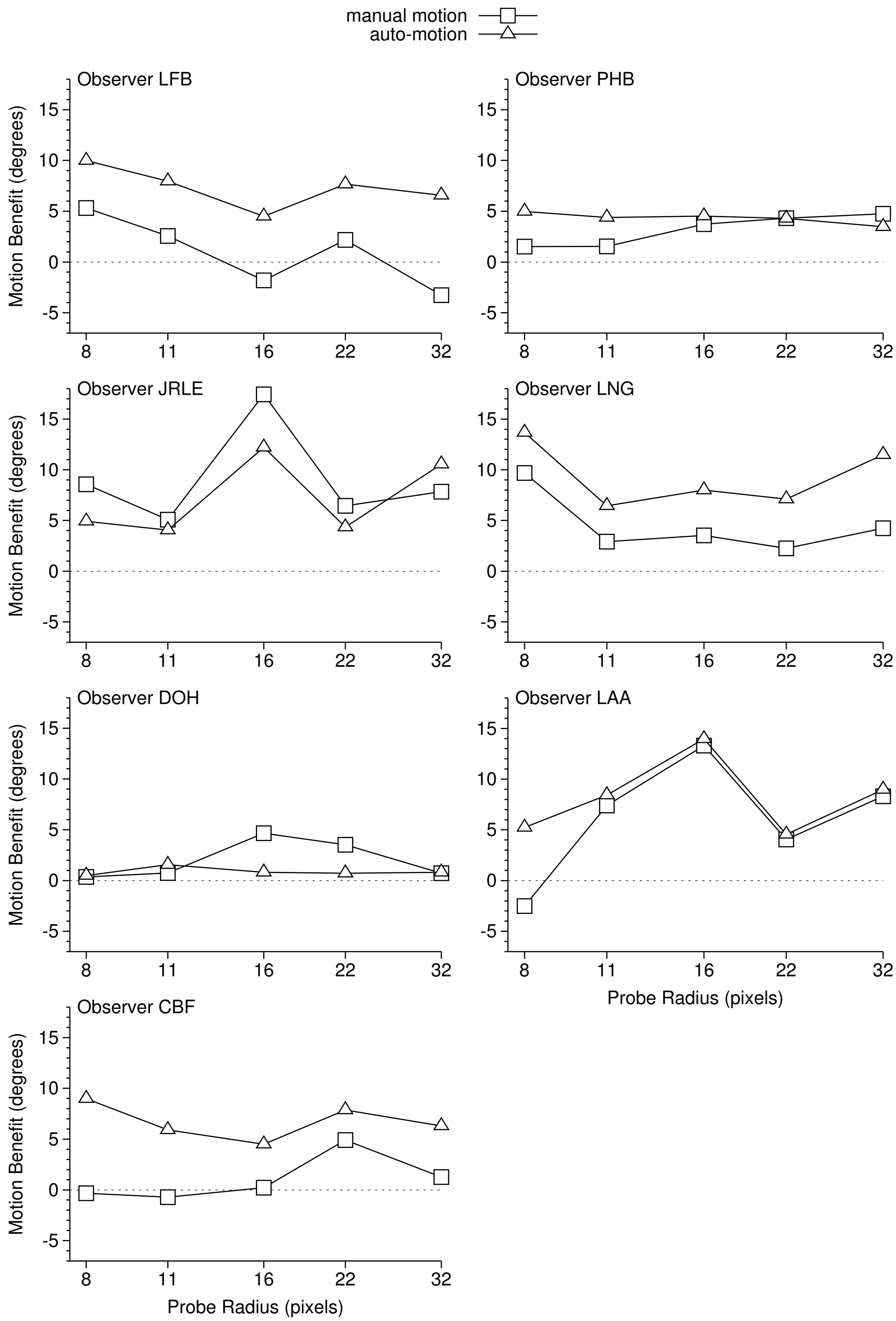

Figure 5.19: Improvement in accuracy of perceived surface orientation due to the addition of motion to the display, for all observers and both types of motion. This figure is based on the biases calculated using σ-normals.

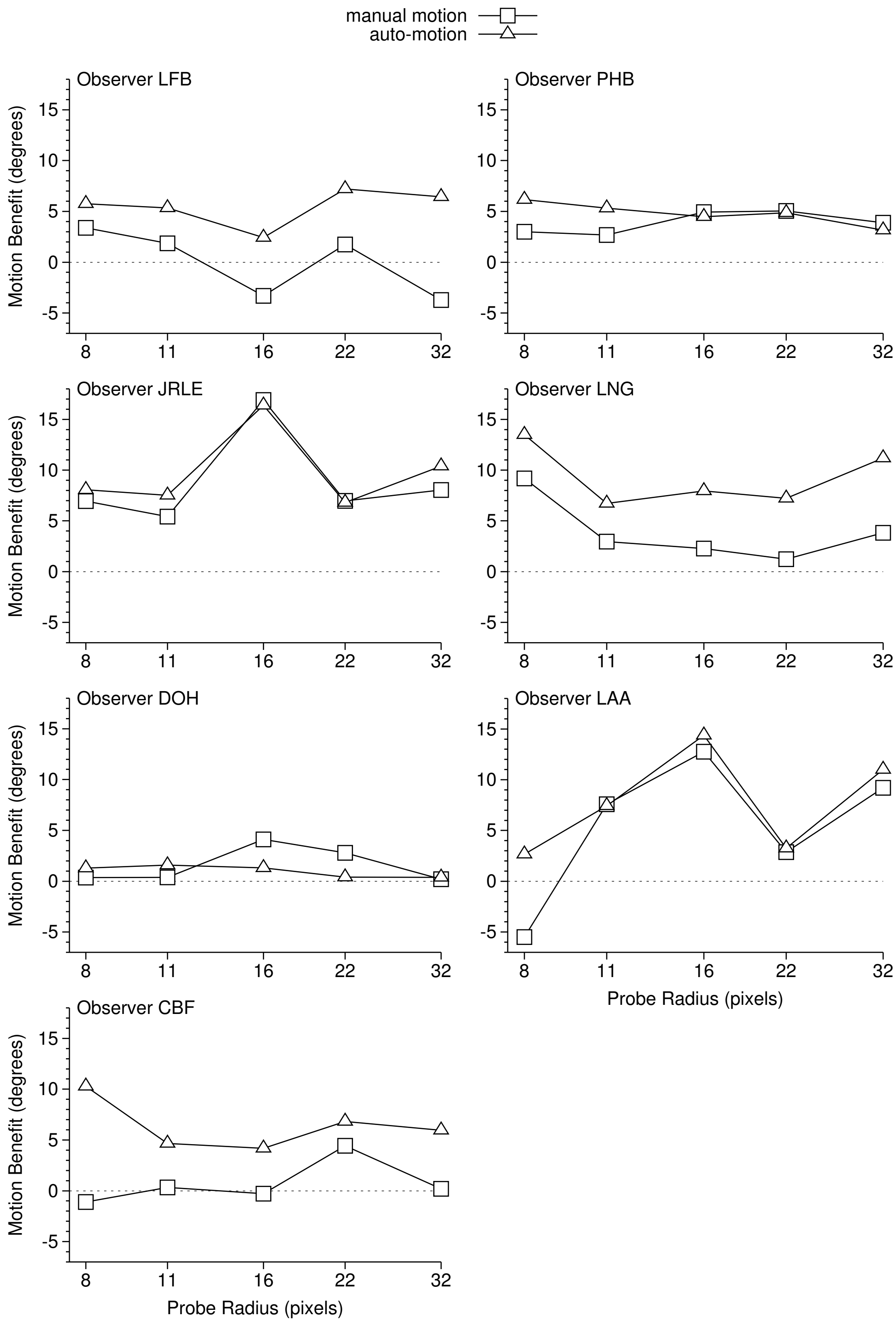

Figure 5.20: Improvement in accuracy of perceived surface orientation due to the addition of motion to the display, for all observers and both types of motion. This figure is based on the biases calculated using s-normals.

presentation would give an accuracy benefit at a wide variety of scales; the object-motion experiment found that adding object motion to a static, non-stereo presentation did indeed give an accuracy benefit at a wide variety of scales. Any conclusions must be tentative, however, because of the substantial differences between the stimulus surfaces.

The orientation paths from the two experiments presented in this chapter call attention to another opportunity. I have earlier described how the relationship between observers' orientation paths and the corresponding objective orientation paths described the veridicality of a given observer's perceived orientations. For many of the test locations, however, the relationship between the observers' orientation paths and the corresponding objective paths is similar across observers. For example, at location 4 on the rough sphere, all the observers clustered too high at small scale and then moved to a more or less veridical position at the largest scale. At location 2 on the solid-noise object, nearly all the orientation paths start out well above the objective path, well up in the upper-right quadrant, and follow a track approximately parallel to the objective path. Such results suggest the effect reported by Koenderink *et al.* [64, 58, 63] and Todd *et al.* [119], in which observers share systematic departures from veridicality on distinct regions of a surface due to the precise interactions of the available cues (in these cases, mainly shading) with the surface. In the MTOPS case, such systematic departures from veridicality offer information about precisely how the available depth cues interact with a given surface across scale. None of these experiments—even taken together—has enough test locations, or enough variations on the available depth cues, to fully exploit this information. But the opportunity is there for those who would pursue it.

Chapter 6

Contributions and Discussion

6.1 Contributions

This dissertation has demonstrated, using the MTOPS tool and techniques, that multiscale measurement of perceived 3D shape is both possible and fruitful. Such information promises to be important in improving the computer-graphics display of 3D shape information.

The major contributions of this dissertation are the following:

- A demonstration that multiscale measurement of perceived 3D shape is possible. In three different experiments, presented in Chapters 4 and 5, observers consistently made different surface-orientation settings at different scales, indicating both that perceived orientation changed across scale and that MTOPS was able to measure that change.
- A demonstration that, by comparing observers' surface-orientation settings to the stimulus geometry, one can measure the across-scale effects of adding a given depth cue to a visualization. I did this for three cues of interest to the computer-graphics community, as follows:
  - Adding stereo to a non-stereo presentation caused the accuracy of observers' settings to improve more at large scale than at small. This result is described in detail in Chapter 4.
  - Adding head-motion parallax to a stereo presentation caused a modest improvement in observers' accuracy, mostly at small scale. The experiment measuring this is presented in Section 5.2.
  - Adding automatic object motion to a non-stereo presentation improved observers' accuracy substantially, at many scales. Adding user-controlled object motion to a non-stereo presentation gave a less consistent, and for most users smaller, accuracy benefit. The experiment which produced this result is presented in Section 5.3.

Subsidiary contributions are listed by chapter.

After giving a motivation and summary in Chapter 1, I reviewed the pertinent literature in Chapter 2. I touched on the relevant geometric measures, the various cues used by humans to infer 3D shape, the models of how humans integrate information from those cues into a unified 3D percept, and the reasons why measurement at multiple spatial scales could be predicted to be important for the study of perceived 3D shape.

In Chapter 3, I described the MTOPS tool and techniques. The contributions of that chapter include the following:

- A probe capable of measuring perceived orientation at multiple spatial scales.
- A method for constructing orientation paths, which show the evolution of orientations (whether perceived or objective) across scale.
- σ-normals and s-normals, two methods for calculating objective surface normals at multiple scales.
- Bias and variability measures to summarize the accuracy of observers' judgments and the repeatability of observers' settings.

In Chapter 4, I used MTOPS to study a depth cue of interest to the computer-graphics community; I measured the across-scale effect of adding stereo to a non-stereo presentation, on two different objects. Contributions in that chapter include the following:

- Using the variability measure, I found that the observers' large-scale settings on the novel MTOPS task were no less repeatable than their small-scale settings.
- I showed that comparing the observers' orientation paths with the objective orientation paths told much about the observers' percepts at a given location, and I described three characteristic patterns in the observers' orientation paths. These patterns were
    1. Approximate superimposition of the observer's orientation path on the corresponding objective path, indicating a high degree of veridicality in the settings across scale.
    2. An observer's path roughly paralleling the corresponding objective path, indicating that the observer's settings differed from the veridical by an approximately constant angular offset.
    3. A cluster of two or more points on the observer's path, with no corresponding cluster on the objective path, suggesting that the observer's perception of the surface at that test location may have supported orientation judgments only at certain scales.
- I found that the probe was sensitive enough to capture the effect of observer knowledge; the biases across scale of the informed observers were quite different from the patterns across scale of the naïve observers.
- Using the bias measure, I showed that the effect of stereo depended on the stimulus surface; stereo on the smooth sphere did not improve the accuracy of observers' settings, but stereo on the rough sphere improved accuracy substantially.

In Chapter 5, I used MTOPS to study depth cues that are important in virtual-reality applications, measuring the effect of adding head-motion parallax to a stereo presentation and two different kinds of object motion to a static, non-stereo presentation.

- I showed that observers' large-scale settings were no less repeatable than small-scale settings, confirming the pattern described in Chapter 4.

- I found that observers' orientation paths in these two experiments exhibited the same three characteristic patterns described in the stereo experiment. Different observers tended to have similar patterns at the same test location, in some cases even across experiments, suggesting that the patterns are due more to the interaction between the available depth cues and the surface than to inter-observer variation.

6.2 Discussion and Future Work

6.2.1 Comparisons with Previous Measures

By comparing settings at comparable scales, it is possible to assess the gross accuracy of observers' settings with the MTOPS probe relative to those made with related, prior probes. The accuracy of MTOPS settings is similar to the accuracy of judgments made in the most relevant previous work, indicating that the multiscale capabilities of MTOPS do not exact a price in the accuracy of the settings.

Of the several measures of perceived 3D orientation in the literature [65, 81, 108, 110], the one most closely related to the MTOPS probe is the Koenderink top described in [65]. Of the series of studies performed with that probe, five (de Haan *et al.* [26], two by Koenderink *et al.* [67, 58], Norman *et al.* [87], and Todd *et al.* [123]) assessed the veridicality of their observers' settings; of those, only those by Norman *et al.* and Todd *et al.* reported a measure of angular error comparable to the bias measure of MTOPS[1]. Norman *et al.* measured perceived orientation on a potato-like surface, with five different kinds of surface shading and/or texture, with and without stereo, and with or without automatic object motion (in all twenty possible combinations). Todd *et al.* used the same kind of potato-like surface, with textured, shiny, or matte surfaces presented in stereo. The mean errors reported in those studies (comparable to MTOPS's bias measure) are listed in Table 6.1 with the corresponding MTOPS biases. (I have omitted Norman *et al.*'s condition with both stereo and object motion, since none of the MTOPS experiments tested that. From Todd *et al.*, I have listed only the matte results, since the MTOPS surfaces were all matte.) Because both Norman *et al.* and Todd *et al.* used a small probe, the MTOPS biases given are those for the smallest probe in each experiment. In general, the errors from the two prior studies are similar to the MTOPS errors, particularly when one takes surface complexity into account; the MTOPS biases steadily increase with increasing surface complexity. The potato objects used in the two earlier studies are clearly more complex than the smooth sphere, somewhat less complex (at least at small scale) than the rough sphere, and clearly simpler than the solid-noise object.

6.2.2 Surface Complexity

Interactions between surface complexity and shape cues are evident both in MTOPS data and in prior work. In the stereo experiment, for example, it is clear that surface complexity did interact with the stereo cue; the addition of stereo improved observers' accuracy substantially on the rough sphere but improved accuracy very little on the smooth sphere. In addition, observers were not, in

[1] The other three assessed veridicality with an affine correlation between the gradient vectors calculated from observers' settings and the gradient vectors calculated from the surface. Norman *et al.* and Todd *et al.* also reported correlations, which are generally similar to the correlations reported in the other studies.

Condition	Norman *et al.* (and Todd *et al.*) Potato	MTOPS Smooth sphere	Rough sphere	Solid-noise
No stereo, no motion	20–30	4–11	16–20	18–37
Stereo, no motion	12–18 (16)	4–7	12–30	
Motion, no stereo	10–18			12–27

Table 6.1: Errors reported by Norman *et al.* and by Todd *et al.*, compared to MTOPS biases. Todd *et al.*'s reported error is in parentheses; it is a single number rather than a range because Todd *et al.* averaged their reported error across observers.

general, either as accurate or as precise on the more complex objects as they were on the smooth sphere. As the surface complexity increased from the smooth sphere to the rough sphere and the solid-noise object, the median bias in the static condition (the only condition common to all three surfaces) increased from 5.82 degrees on the smooth sphere to 17.4 degrees on the rough sphere and to 22.9 degrees on the solid-noise object; the median variability in the static condition increased from 5.96 degrees on the smooth sphere to 9.85 degrees on the rough sphere and to 9.18 degrees on the solid-noise object. Effects of surface complexity have been reported previously as well; examples include Todd and Norman [122], who found that orientation discrimination was over twice as bad for patches on curved surfaces as for planar patches, and Phillips and Todd [90], who found that shape-index discrimination thresholds were about 50% higher for judgments on turbulent patches than for judgments on non-turbulent patches.

In each of these cases, while it is clear which surface is the more complex, it is not clear what aspects of "surface complexity" contribute to the worsened performance. Interactions between surface complexity and depth cues can be scale-specific (as in the stereo experiment), and surface complexity itself can be at least partly characterized by the number and distribution of scales at which a surface's structure occurs. MTOPS's ability to make measurements at multiple scales makes it a promising tool for exploring this question.

I noted a similar but more local opportunity for future research in Section 5.3.3. The orientation paths at many test locations reveal systematic departures from veridicality, which offer information about the precise interaction between the available depth cues and the surface near that test location across scale. Studies with more test locations and more variations on the available cues offer the opportunity to explore those interactions.

It would be particularly interesting to apply the MTOPS probe and task to stimuli that obviously parse into major subparts (as a hand with its fingers—see Biederman [4] for further examples and discussion). It is not clear how observers would place a probe which overlaps the boundary between two parts of an object; observers' treatment of such problems could provide insight into the nature of the underlying shape representations.

6.2.3 More Opportunities

The MTOPS probe and techniques also open the way to learn more about the scale-specific properties of depth cues, independent of surface-complexity issues. I have shown that changes in observers' ability to judge surface orientation due to the addition of a depth cue are captured by changes in the orientation paths for stereo, for head-motion parallax, and for object motion; I have also shown that the changes are not constant across spatial scale. What about other cues, or parametric variations of these? Are Forrest, Schweitzer, and Ware correct to suggest that optical texture offers unexploited opportunities for communicating the 3D shape of opaque surfaces—and if they are right, at what scales are they right? What effect does the rotation angle in object motion have on the accuracy with which observers perceive the metric 3D shape of the object—and is there a consistent scale pattern to that effect? Is metric shape communicated more accurately by head-motion parallax not restricted to a single horizontal line of viewpoints—and if so, at what scales is this unrestricted head motion effective?

While single-scale measures of perceived shape can begin to answer these questions, single-scale measures cannot reveal across-scale patterns and may not capture effects at other than the scale of measurement. Even aperture-based measures, which act as low-pass filters on surface structure, do not always capture large-scale effects; for example, the small-scale settings in the stereo experiment did not reflect the large gain in large-scale accuracy on the rough sphere. Multiscale measurement offers a far more complete understanding of the effects of depth cues.

On a more applied level, MTOPS could be used to explore the characteristic scales of a particular task. By running an MTOPS experiment in parallel with a task study and comparing the improvement (or lack thereof) in task performance against the scales at which users' orientation judgments became more accurate, investigators can learn about the scales at which metric comprehension of shape is really important to performance on that particular task.

Perception of 3D shape is an important part of many applications of computer graphics, from medical imaging to molecular modeling. This dissertation has demonstrated that multiscale measurement of perceived shape has an important role to play in learning about 3D shape perception. Computer-graphics practitioners can exploit the resulting knowledge to build better presentations of 3D shape.

Bibliography

[1] Arthur Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the AFIPS Spring Joint Computer Conference*, volume 32, pages 37–45, 1968.

[2] Kevin W. Arthur. 3D task performace using head-coupled stereo displays. Master's thesis, University of British Columbia, 1993.

[3] Kevin W. Arthur, Kellogg S. Booth, and Colin Ware. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*, 11(3):239–265, July 1993.

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

[5] Irving Biederman and Ginny Ju. Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20:38–64, 1988.

[6] P. Bijl, Jan J. Koenderink, and A. Toet. Visibility of blobs with a Gaussian luminance profile. *Vision Research*, 29(4):447–456, 1989.

[7] Piet Bijl. *Aspects of Visual Contrast Detection*. PhD thesis, Universiteit Utrecht, 1991.

[8] P. O. Bishop, J. S. Coombs, and G. H. Henry. Receptive fields of simple cells in the cat striate cortex. *Journal of Physiology*, 231:31–60, 1973.

[9] C. Blakemore and F. W. Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, 203:237–260, 1969.

[10] H. Blum. A transformation for extracting new descriptors of shape. In W. Wathen-Dunn, editor, *Models for the Perception of Speech and Visual Form*, pages 362–380, Cambridge, MA, 1967. MIT Press.

[11] E. G. Britton, J. S. Lipscomb, and M. E. Pique. Making nested rotations convenient for the user. In *Computer Graphics (SIGGRAPH '78 Proceedings)*, volume 12, pages 222–227, August 1978.

[12] Frederick P. Brooks, Jr., Ming Ouh-Young, James J. Batter, and P. Jerome Kilpatrick. Project GROPE — Haptic displays for scientific visualization. In Forest Baskett, editor, *Computer Graphics (SIGGRAPH '90 Proceedings)*, volume 24, pages 177–185, August 1990.

[13] Peter H. Brown and Christina A. Burbeck. Measuring perceived 3D shape at multiple spatial scales. *Vision Research*, 39(10):1833–1845, 1999.

[14] Steve Bryson and Creon Levit. The virtual windtunnel: An environment for the exploration of three-dimensional unsteady flows. In *Proceedings of Visualization '91*, pages 17–24, 1991.

[15] Heinrich H. Bülthoff and Hanspeter A. Mallot. Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A*, 5(10):1749–1758, October 1988.

[16] Christina A. Burbeck and Scott Hadden. Scaled position integration areas: Accounting for Weber's law for separation. *Journal of the Optical Society of America A*, 10(1):5–15, January 1993.

[17] Christina A. Burbeck and Stephen M. Pizer. Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, 35(13):1917–1930, 1995.

[18] Christina A. Burbeck, Stephen M. Pizer, Bryan S. Morse, Dan Ariely, Gal S. Zauberman, and Jannick Rolland. Linking object boundaries at scale: A common mechanism for size and shape judgments. *Vision Research*, 36(3):361–372, 1996.

[19] F. W. Campbell and J. G. Robson. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197:551–566, 1968.

[20] Stuart Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufman Publishers, 1999.

[21] Michael Chen, S. Joy Mountford, and Abigail Sellen. A study in interactive 3-D rotation using 2-D control devices. In John Dill, editor, *Computer Graphics (SIGGRAPH '88 Proceedings)*, volume 22, pages 121–129, August 1988.

[22] Chris G. Christou and Jan J. Koenderink. Light source dependence in shape from shading. *Vision Research*, 37(11):1441–1449, 1997.

[23] James C. Chung. A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams. In David Zeltzer, editor, *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*, volume 25, pages 193–196, March 1992.

[24] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In James T. Kajiya, editor, *Computer Graphics (SIGGRAPH '93 Proceedings)*, volume 27, pages 135–142, August 1993.

[25] William Curran and Alan Johnston. Integration of shading and texture cues: Testing the linear model. *Vision Research*, 34(14):1863–1874, 1994.

[26] Erik de Haan, Roderik G. F. Erens, and André J. Noest. Shape from shaded random surfaces. *Vision Research*, 35(21):2985–3001, November 1995.

[27] Sjoerd C. de Vries. *Binocular Perception of Spatial Configuration*. PhD thesis, Universiteit Utrecht, 1993.

[28] Frank H. Durgin, Dennis R. Proffitt, Thomas J. Olson, and Karen S. Reinke. Comparing depth from binocular disparity with depth from motion. *Journal of Experimental Psychology: Human Perception and Performance*, 21:679–699, 1995.

[29] Matthias Eck, Tony DeRose, Tom Duchamp, Hugues Hoppe, Michael Lounsbery, and Werner Stuetzle. Multiresolution analysis of arbitrary meshes. In Robert Cook, editor, *SIGGRAPH 95 Conference Proceedings*, Annual Conference Series, pages 173–182. ACM SIGGRAPH, Addison Wesley, August 1995.

[30] C. H. Edwards, Jr. and David E. Penney. *Calculus and Analytic Geometry*. Prentice-Hall, second edition, 1986.

[31] Christina Enroth-Cugell and J. G. Robson. The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology*, 187:517–552, 1966.

[32] Roderik G. F. Erens, Astrid M. L. Kappers, and Jan J. Koenderink. Perception of local shape from shading. *Perception & Psychophysics*, 54(2):145–156, 1993.

[33] James A. Ferwerda, Sumanta N. Pattanaik, Peter Shirley, and Donald P. Greenberg. A model of visual masking for computer graphics. In Turner Whitted, editor, *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, pages 143–152. ACM SIGGRAPH, Addison Wesley, August 1997. ISBN 0-89791-896-7.

[34] N. I. Fisher, T. Lewis, and B. J. J. Embleton. *Statistical Analysis of Spherical Data*. Cambridge University Press, 1987.

[35] R. A. Fisher. Dispersion on a sphere. *Proceedings of the Royal Society A*, 217:295–305, 1953.

[36] James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley, Reading, Massachusetts, second edition, 1990.

[37] A. R. Forrest. On the rendering of surfaces. In *Computer Graphics (SIGGRAPH '79 Proceedings)*, volume 13, pages 253–259, August 1979.

[38] James J. Gibson. *The Perception of the Visual World*. Houghton Mifflin, 1950.

[39] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modelling the interaction of light between diffuse surfaces. In *Computer Graphics (SIGGRAPH '84 Proceedings)*, volume 18, pages 212–222, July 1984.

[40] Robert Gottsdanker, James W. Frick, and Robert B. Lockard. Identifying the acceleration of visual targets. *British Journal of Psychology*, 52(1):31–42, 1961.

[41] Norma Van Surdam Graham. *Visual Pattern Analyzers*. Oxford University Press, 1989.

[42] John Halverston. The first pictures: perceptual foundations of Paleolithic art. *Perception*, 21:389–404, 1992.

[43] Heiko Hecht, Andrea van Doorn, and Jan J. Koenderink. Compression of visual space in natural scenes and in their photographic counterparts. *Perception & Psychophysics*, 61(7):1269–1286, 1999.

[44] M. A. Hogervorst, M. F. Bradshaw, and R. A. Eagle. Spatial frequency tuning for 3-D corrugations from motion parallax. *Vision Research*, 40:2149–2158, 2000.

[45] David H. Hubel and Torsten N. Wiesel. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962.

[46] David H. Hubel and Torsten N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195:215–243, 1968.

[47] Victoria Interrante. *Illustrating Transparency: communicating the 3D shape of layered transparent surfaces via texture*. PhD thesis, University of North Carolina at Chapel Hill, 1996.

[48] Victoria Interrante, Henry Fuchs, and Stephen M. Pizer. Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics*, 3(2), April–June 1997. ISSN 1077-2626.

[49] Timothy E. Johnson. Sketchpad III: a computer program for drawing in three dimensions. In *Proceedings of the AFIPS Spring Joint Computer Conference*, volume 23, pages 347–353, 1963.

[50] Alan Johnston and Peter J. Passmore. Shape from shading. I: Surface curvature and orientation. *Perception*, 23:169–189, 1994.

[51] D. H. Kelly and Christina A. Burbeck. Critical problems in spatial vision. *CRC Critical Reviews in Biomedical Engineering*, 10(2):125–177, 1984.

[52] J. T. Kent. The Fisher-Bingham distribution on the sphere. *Journal of the Royal Statistical Society B*, 44:71–80, 1982.

[53] Daniel Kersten, Pascal Mamassian, and David C. Knill. Moving cast shadows induce apparent motion in depth. *Perception*, 26:171–192, 1997.

[54] Jan J. Koenderink. Optic flow. *Vision Research*, 26(1):161–180, 1986.

[55] Jan J. Koenderink. *Solid Shape*. The MIT Press, 1990.

[56] Jan J. Koenderink. Pictorial relief. *Philosophical Transactions of the Royal Society of London A*, 356:1071–1086, 1998.

[57] Jan J. Koenderink. Virtual psychophysics. *Perception*, 28:669–674, 1999.

[58] Jan J. Koenderink, Astrid M. L. Kappers, James T. Todd, J. Farley Norman, and Flip Phillips. Surface range and attitude probing in stereoscopically presented dynamic scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):869–878, 1996.

[59] Jan J. Koenderink and Andrea van Doorn. Invariant features of contrast detection: An explanation in terms of self-similar detector arrays. *Journal of the Optical Society of America*, 72(1):83–87, January 1982.

[60] Jan J. Koenderink and Andrea van Doorn. Affine structure from motion. *Journal of the Optical Society of America A*, 8(2):377–385, 1991.

[61] Jan J. Koenderink and Andrea van Doorn. Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A*, 13(3):452–463, March 1996.

[62] Jan J. Koenderink and Andrea J. van Doorn. Receptive field families. *Biological Cybernetics*, 63(4):291–297, 1990.

[63] Jan J. Koenderink, Andrea J. van Doorn, Chris Christou, and Joseph S. Lappin. Perturbation study of shading in pictures. *Perception*, 25:1009–1026, 1996.

[64] Jan J. Koenderink, Andrea J. van Doorn, Chris Christou, and Joseph S. Lappin. Shape constancy in pictorial relief. *Perception*, 25:155–164, 1996.

[65] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Surface perception in pictures. *Perception & Psychophysics*, 52(5):487–496, 1992.

[66] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. On so called “paradoxical monocular stereoscopy”. *Perception*, 23:583–594, 1994.

[67] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Depth relief. *Perception*, 24:115–126, 1995.

[68] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Pictorial surface attitude and local depth comparisons. *Perception & Psychophysics*, 58(2):158–173, 1996.

[69] Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Surfaces in the mind’s eye. In Roberto Cipolla and Ralph Martin, editors, *The Mathematics of Surfaces IX*, pages 180–193. Springer, 2000.

[70] Jan J. Koenderink, Andrea J. van Doorn, and Joseph S. Lappin. Direct measurement of the curvature of visual space. *Perception*, 29(1):69–79, 2000.

[71] Michael S. Landy, Laurence T. Maloney, Elizabeth B. Johnston, and Mark Young. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35(3):389–412, February 1995.

[72] Marc Levoy. Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37, May 1988.

[73] John-Peter Lewis. Algorithms for solid noise synthesis. In Jeffrey Lane, editor, *Computer Graphics (SIGGRAPH ’89 Proceedings)*, volume 23, pages 263–270, July 1989.

[74] Michael Leyton. *Symmetry, Causality, Mind.* MIT Press, Cambridge, MA, 1992.

[75] James S. Lipscomb. *Three-Dimensional Cues for a Molecular Computer Graphics System.* PhD thesis, University of North Carolina at Chapel Hill, 1981.

[76] Michael Lounsbery, Tony D. DeRose, and Joe Warren. Multiresolution analysis for surfaces of arbitrary topological type. *ACM Transactions on Graphics*, 16(1):34–73, January 1997. ISSN 0730-0301.

[77] Laurence T. Maloney and Michael S. Landy. A statistical framework for robust fusion of depth information. In *SPIE vol. 1199 Visual Communications and Image Processing IV*, pages 1154–1163, 1989.

[78] D. Marr and H. K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B*, 200:269–294, 1978.

[79] Suzanne P. McKee, Dennis M. Levi, and Samuel F. Bowne. The imprecision of stereopsis. *Vision Research*, 30(11):1763–1779, 1990.

[80] Michael McKenna. Interactive viewpoint control and three-dimensional operations. In David Zeltzer, editor, *Computer Graphics (1992 Symposium on Interactive 3D Graphics)*, volume 25, pages 53–56, March 1992.

[81] Ennio Mingolla and James T. Todd. Perception of solid shape from shading. *Biological Cybernetics*, 53:137–151, 1986.

[82] Karol Myszkowski, Przemyslaw Rokita, and Takehiro Tawara. Perception-based fast rendering and antialiasing of walkthrough sequences. *IEEE Transactions on Visualization and Computer Graphics*, 6(4):360–379, October-December 2000.

[83] Ken Nakayama and Shinsuke Shimojo. Experiencing and perceiving visual surfaces. *Science*, 257:1357–1363, September 1992.

[84] Ken Nakayama, Shinsuke Shimojo, and Gerald H. Silverman. Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18:55–68, 1989.

[85] Shin'ya Nishida, Timothy Ledgeway, and Mark Edwards. Dual multiple-scale processing for motion in the human visual system. *Vision Research*, 37(19):2685–2698, 1997.

[86] J. Farley Norman and James T. Todd. The discriminability of local surface structure. *Perception*, 25:381–398, 1996.

[87] J. Farley Norman, James T. Todd, and Flip Phillips. The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, 57(5):629–636, 1995.

[88] Randy Pausch, Dennis Proffitt, and George Williams. Quantifying immersion in virtual reality. *Proceedings of SIGGRAPH 97*, pages 13–18, August 1997. ISBN 0-89791-896-7.

[89] Flip Phillips and James T. Todd. Perception of local three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):930–944, 1996.

[90] Flip Phillips, James T. Todd, Jan J. Koenderink, and A. M. L. Kappers. Perceptual localization of surface position. *Journal of Experimental Psychology: Human Perception and Performance*, 23(5):1481–1492, 1997.

[91] Stephen M. Pizer, Andrew L. Thall, and David T. Chen. M-reps: A new object representation for graphics. Technical Report 99-030, University of North Carolina at Chapel Hill, 2000.

[92] F. Poggio, R. W. Doty, Jr., and W. H. Talbot. Foveal striate cortex of behaving monkey: Single-neuron responses to square-wave gratings during fixation of gaze. *Journal of Neurophysiology*, 40(6):1369–1391, November 1977.

[93] Martin Reddy. *Perceptually Modulated Level of Detail in Virtual Environments*. PhD thesis, University of Edinburgh, 1997.

[94] Penny Rheingans. *Dynamic Explorations of Multiple Variables in a 2D Space*. PhD thesis, University of North Carolina at Chapel Hill, 1993.

[95] Penny Rheingans. Opacity-modulating triangular textures for irregular surfaces. In *IEEE Visualization '96*. IEEE, October 1996. ISBN 0-89791-864-9.

[96] James C. Rodger and Roger A. Browse. Choosing rendering parameters for effective communication of 3d shape. *IEEE Computer Graphics & Applications*, 20(2):20–28, March/April 2000. ISSN 0272-1716.

[97] Brian Rogers and Maureen Graham. Motion parallax as an independent cue for depth perception. *Perception*, 8:125–134, 1979.

[98] Brian Rogers and Maureen Graham. Similarities between motion parallax and stereopsis in motion perception. *Vision Research*, 22:261–270, 1982.

[99] Ann Marie Rohaly and Hugh R. Wilson. Disparity averaging across spatial scales. *Vision Research*, 34(10):1315–1325, 1994.

[100] Takafumi Saito and Tokiichiro Takahashi. Comprehensible rendering of 3-D shapes. In Forest Baskett, editor, *Computer Graphics (SIGGRAPH '90 Proceedings)*, volume 24, pages 197–206, August 1990.

[101] D. Schweitzer. Artificial texturing: An aid to surface visualization. In *Computer Graphics (SIGGRAPH '83 Proceedings)*, volume 17, pages 23–29, July 1983.

[102] Philippe G. Schyns and Aude Oliva. Flexible, diagnosticity-driven, rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception*, 26:1027–1038, 1997.

[103] Shinsuke Shimojo, Gerald H. Silverman, and Ken Nakayama. An occlusion-related mechanism of depth perception based on motion and interocular sequence. *Nature*, 333(6170):265–268, May 1988.

[104] William A. Simpson. Temporal summation of visual motion. *Vision Research*, 34(19):2547–2559, October 1994.

[105] Harvey S. Smallman. Fine-to-coarse scale disambiguation in stereopsis. *Vision Research*, 35(8):1047–1060, 1995.

[106] Andrei State, Suresh Balu, and Henry Fuchs. Bunker View: Limited-range head-motion-parallax visualization for complex biomedical data sets, 1994. Extended abstract submitted to VBC '94.

[107] Kent A. Stevens. Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, 46:183–195, 1983.

[108] Kent A. Stevens. Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics*, 33(3):241–250, 1983.

[109] Kent A. Stevens. Constructing the perception of surfaces from multiple cues. *Mind & Language*, 5(4):253–266, 1990.

[110] Kent A. Stevens and Allen Brookes. Probing depth in monocular images. *Biological Cybernetics*, 56:355–366, 1987.

[111] I. E. Sutherland. A head-mounted three-dimensional display. In *AFIPS Conference Proceedings*, volume 33, pages 757–764, 1968.

[112] Colin Swindells, John C. Dill, and Kellogg S. Booth. System lag tests for augmented and virtual environments. In *Proceedings of the 13th annual ACM symposium on user interface software and technology*, pages 161–170. ACM Press, 2000. ISBN 1-58114-212-3.

[113] Russell M. Taylor, II, Warren Robinett, Vernon L. Chi, Frederick P. Brooks, Jr., William V. Wright, R. Stanley Williams, and Eric J. Snyder. The Nanomanipulator: A virtual reality interface for a scanning tunnelling microscope. In James T. Kajiya, editor, *Computer Graphics (SIGGRAPH '93 Proceedings)*, volume 27, pages 127–134, August 1993.

[114] James S. Tittle, Victor J. Perotti, and J. Farley Norman. Integration of binocular stereopsis and structure from motion in the discrimination of noisy surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, 23(4):1035–1049, 1997.

[115] James S. Tittle, James T. Todd, Victor J. Perotti, and J. Farley Norman. Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3):663–678, 1995.

[116] James T. Todd. Visual information about moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, 7(4):795–810, 1981.

[117] James T. Todd and Robin A. Akerstrom. Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, 13(2):242–255, 1987.

[118] James T. Todd and Paola Bressan. The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, 48(5):419–430, 1990.

[119] James T. Todd, Jan J. Koenderink, Andrea J. van Doorn, and Astrid M. L. Kappers. Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, 22(3):695–706, 1996.

[120] James T. Todd and Ennio Mingolla. Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance*, 9(4):583–595, 1983.

[121] James T. Todd and J. Farley Norman. The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, 50(6):509–523, 1991.

[122] James T. Todd and J. Farley Norman. The visual discrimination of relative surface orientation. *Perception*, 24:855–866, 1995.

[123] James T. Todd, J. Farley Norman, Jan J. Koenderink, and Astrid M. L. Kappers. Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception*, 26:807–822, 1997.

[124] James T. Todd, Augustinus H. J. Oomes, Jan J. Koenderink, and Astrid M. L. Kappers. On the affine structure of perceptual space. *Psychological Science*, 12(3):191–196, May 2001.

[125] James T. Todd and Francene D. Reichel. Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review*, 96(4):643–657, 1989.

[126] A. Toet and Jan J. Koenderink. Differential spatial displacement thresholds for Gabor patches. *Vision Research*, 28(1):133–143, 1988.

[127] Edward R. Tufte. *Visual Explanations*. Graphics Press, Post Office Box 430, Cheshire, Connecticut 06410, 1997.

[128] W. J. M. van Damme. *Active Vision: exploration of three-dimensional structure*. PhD thesis, Universiteit Utrecht, September 1994.

[129] Andrea J. van Doorn and Jan J. Koenderink. How to probe different aspects of surface relief. In A. M. L. Kappers, C. J. Overbeeke, G. J. F. Smets, and P. J. Stappers, editors, *Studies in Ecological Psychology*, pages 115–130, 1996.

[130] Hendrik-Jan van Veen. *Visually Acquired Rotation About Rotating Objects*. PhD thesis, Universiteit Utrecht, 1996.

[131] Mark Wagner. The metric of visual space. *Perception & Psychophysics*, 38(6):483–495, 1985.

[132] Leonard R. Wanger, James A. Ferwerda, and Donald P. Greenberg. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics & Applications*, 12(3):44–58, May 1992.

[133] John P. Wann, Simon Rushton, and Mark Mon-Williams. Natural problems for stereoscopic depth in virtual environments. *Vision Research*, 35(19):2731–2736, 1995.

[134] Colin Ware. *Information Visualization: Design for Perception*. Morgan Kaufmann Publishers, 2000.

[135] Colin Ware, Kevin Arthur, and Kellogg S. Booth. Fish tank virtual reality. In *Proceedings of CHI '93*, pages 37–42, apr 1993.

[136] Colin Ware and Glenn Franck. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics*, 15(2):121–140, April 1996.

[137] R. J. Watt. Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America A*, 4:2006–2021, October 1987.

[138] Roger Watt. A computational examination of image segmentation and the initial stages of human vision. *Perception*, 23:383–390, 1994.

[139] Robert B. Wilhemson, Brian F. Jewett, Louis J. Wicker, Matthew Arrott, Colleen B. Bushell, Mark Bjuk, Jeffrey Thingvold, and Jeffery B. Yost. A study of the evolution of a numerically modeled severe storm. *International Journal of Supercomputer Applications*, 4(2):20–36, Summer 1990.

[140] Lance Williams. Pyramidal parametrics. In *Computer Graphics (SIGGRAPH '83 Proceedings)*, volume 17, pages 1–11, July 1983.

[141] William V. Wright. *An Interactive Computer Graphic System for Molecular Graphics*. PhD thesis, University of North Carolina at Chapel Hill, 1972.

Appendix A

Observers' Tips, Stereo Experiment

Task: Make the probe's top circle parallel to a patch of surface of the same size as the circle, centered on the tip of the probe. The key is to make sure that you're using a patch of surface as big as the probe's top circle.

Commands you can use:

Spacebar	Go to the next trial
Backspace	Return to the previous trial
q	Quit the program

Do's:

Focus on the circle, not the stick. The stick is there basically to tie the circle to the surface, so you can see where the center of your patch of surface should be.

Average out surface irregularities that are smaller than the probe size. You're trying to get an "average" orientation. Little lumps (little relative to the probe) won't contribute to that.

If you get tired, especially if your eyes get tired, take a break.

Don'ts:

Don't focus too much on the bottom of the stick, or you'll find yourself making purely local settings that don't reflect a patch that's as big as the probe.

Don't take too long on a single trial. We (and others) have found that spending a long time trying to wring that last little bit of accuracy from a setting actually makes the results worse. If you're spending more than twenty seconds on a trial, you're almost certainly not improving anything.

Things to note:

If the mouse freezes, quit and restart. (You may need help to restart; I'll be close by.) It will automatically restart where you left off.

The text window on the left will show you what trial you're on at any given time; it also gives abbreviated instructions. If something odd happens, glance over there to see if it says anything helpful.

You can take a break at any time. The program will remind you every 40 trials.

If you have suggestions for improving the experiment (including this sheet), please tell me: I'm always looking for ways to make the task easier. Your reactions are important!

You can quit the experiment at any time. If you feel your rights have been violated in any way, please let me know. If you can't get a satisfactory response from me, you can take up the matter with my advisor, Dr. Christina Burbeck.

Appendix B

Observers' Overview and Tips, Head-Motion Parallax Experiment

The purpose of this project is to investigate how humans perceive 3-D shapes presented in a computer-graphics environment. By comparing how people perceive the same shape presented with and without some source of information (e.g., stereo or motion), we can assess how the additional information actually affected the viewers' ability to perceive the shape accurately; we can, in turn, use that information to help design computer-graphics displays that do a better job of communicating shape information.

In this particular study, we will compare observers' perceptions of a shape with and without head-motion parallax. You will be asked to complete a number of trials; in each trial, a surface will be displayed in stereo on a monitor, and you will be asked to use the computer's mouse to position a probe indicating what you believe the orientation of the surface to be in a particular region. You will be asked to wear stereo glasses so that the display will be properly stereoptic; in half the trials, you will also be asked to wear a mechanical head-tracker so that the image on the screen can properly reflect your vantage point. In both conditions, there will be rest periods scheduled, and you may also take breaks at any time you like.

Your participation should last approximately two to three hours (depending on how fast you go). During that time, you will be asked to perform 320 trials in four sets; a set of 80 trials using stereo alone, then a set of 80 trials using stereo and the head-tracker, then a set of 80 more trials with stereo and the head-tracker, and finally a set of 80 more trials using stereo alone. Before the first set, you will be asked to complete a practice set of 16 trials using stereo alone; before the second set, you will be asked to complete another practice set of 16 trials using stereo and the head-tracker.

Your data will not be identified with your name either in the raw form or in any publication based on this data. Your initials may, however, be used to identify it; this is common practice in psychophysics. If you would prefer not to have your data identified with your initials, let the experimenter know and we will identify it otherwise.

Your participation in this study is entirely voluntary, and you can stop participating at any time for any reason; just tell the experimenter.

Multiscale Test of Perceived Shape: Observers' Tips

Task: Make the probe's top circle parallel to a patch of surface of the same size as the circle, centered on the tip of the probe. The key is to make sure that you're using a patch of surface as big as the probe's top circle.

Commands you can use:

Spacebar	Go to the next trial
Backspace	Return to the previous trial
q	Quit the program

Do's:

Pay attention to the circle, not the stick. The stick is there basically to tie the circle to the surface, so you can see where the center of your patch of surface should be.

Average out surface irregularities that are smaller than the probe size. You're trying to get an "average" orientation. Little lumps (little relative to the probe) won't contribute to that.

If you get tired, especially if your eyes get tired, take a break.

Don'ts:

Don't concentrate too much on the bottom of the stick, or you'll find yourself making purely local settings that don't reflect a patch that's as big as the probe.

Don't take too long on a single trial. We (and others) have found that spending a long time trying to wring that last little bit of accuracy from a setting actually makes the results worse. If you're spending more than thirty seconds on a trial, you're almost certainly not improving anything (unless you're waiting for the equipment to behave).

Things to note:

If the program freezes, wait for it to catch up. If it keeps on happening, enough to make the task a lot harder for you, quit and restart. (You may need help to restart; I'll be close by.) The program will automatically restart where you left off.

You can take a break at any time. The program will remind you every 40 trials.

If you're unhappy with the experiment, don't suffer. Let me know. If you can't get a satisfactory response from me, you can take up the matter with my advisor, Dr. Christina Burbeck (phone: 460-0539; address: 109 Dundee Ct., Cary, NC 27511). You can also simply quit at any time, for any reason.

If you have suggestions for improving the experiment, please let me know!

Appendix C

Observers' Overview and Tips, Object-Motion Experiment

The purpose of this project is to investigate how humans perceive 3-D shapes presented in a computer-graphics environment. By comparing how people perceive the same shape presented with and without some source of information (e.g., stereo or motion), we can assess how the additional information actually affected the viewers' ability to perceive the shape accurately; we can, in turn, use that information to help design computer-graphics displays that do a better job of communicating shape information.

In this particular study, we will compare observers' perceptions of a shape with and without the ability to move the object. You will be asked to complete a number of trials; in each trial, a surface will be displayed on a monitor, and you will be asked to use a joystick to position a probe indicating what you believe the orientation of the surface to be in a particular region. You may take breaks at any time you like.

Your participation should last approximately four to six hours (different people go at different speeds). During that time, you will be asked to perform 525 trials in six sets:

- a set of 50 practice trials without object motion
- a set of 125 experimental trials without object motion
- a set of 50 practice trials with user-controlled object motion
- a set of 125 experimental trials with user-controlled object motion
- a set of 50 practice trials with automatic object motion
- a set of 125 experimental trials with automatic object motion

Your data will not be identified with your name either in the raw form or in any publication based on this data. Your initials may, however, be used to identify it; this is common practice in psychophysics. If you would prefer not to have your data identified with your initials, let the experimenter know and we will identify it otherwise.

Your participation in this study is entirely voluntary, and you can stop participating at any time for any reason; just tell the experimenter.

Multiscale Test of Perceived Shape: Observers' Tips

Task: Make the probe's top circle parallel to a patch of surface of the same size as the circle, centered on the tip of the probe. The key is to make sure that you're using a patch of surface as big as the probe's top circle.

Commands you can use:

Spacebar	Go to the next trial
Backspace	Return to the previous trial
q	Quit the program

Do's:

Pay attention to the circle, not the stick. The stick is there basically to tie the circle to the surface, so you can see where the center of your patch of surface should be.

Average out surface irregularities that are smaller than the probe size. You're trying to get an "average" orientation. Little lumps (little relative to the probe) won't contribute to that.

If you get tired, especially if your eyes get tired, take a break.

Don'ts:

Don't concentrate too much on the bottom of the stick, or you'll find yourself making purely local settings that don't reflect a patch that's as big as the probe.

Don't take too long on a single trial. We (and others) have found that spending a long time trying to wring that last little bit of accuracy from a setting actually makes the results worse. If you're spending more than about thirty seconds on a trial, you're almost certainly not improving anything.

Things to note:

You can take a break at any time.

If you have a queston, please ask! I'll be close by. Similarly, **if you have any suggestions** for improving the experiment, please let me know!

If you're unhappy with the experiment, don't suffer. Let me know. If you can't get a satisfactory response from me, you can take up the matter with my advisor, Dr. Christina Burbeck (phone: (919) 460-0539; address: 109 Dundee Ct., Cary, NC 27511). You can also simply quit at any time, for any reason.