N90-22936

EXOCENTRIC DIRECTION JUDGEMENTS IN COMPUTER-GENERATED DISPLAYS AND ACTUAL SCENES

Stephen R. Ellis, Stephen Smith, Michael W. McGreevy,
and Arthur Grunwald
NASA Ames Research Center
Moffett Field, California

INTRODUCTION

One of the most remarkable perceptual properties of common experience is that the perceived shapes of known objects are constant despite movements about them which transform their projections on our retina. This perceptual ability is one aspect of shape constancy (Thouless, 1931; Metzger, 1953; Borresen and Lichte, 1962). It requires that the viewer be able to sense and discount his or her relative position and orientation with respect to a viewed object. This discounting of relative position may be derived directly from the ranging information provided from stereopsis, from motion parallax, from vestibularly sensed rotation and translation, or from corollary information associated with voluntary movement.

The measurement of shape constancy usually involves requesting that the viewer make some estimate of the geometric properties of an object, such as the apex angle of a isosceles triangle. Significantly, shape constancy does not disappear during static, monocular viewing, but its basis under these conditions must be different, since sensed motion is not involved. In a static image, shape constancy amounts to the recognition that each of a variety of views of the objects in the scene are all views of the same objects. This perceived constancy may be based on consciously or unconsciously accessed information concerning alternative views of the objects. These "memories," however, need not be of complete objects, since perceived constancy may be based on recall of only some salient features, such as parallelism of significant planes of the object.

In the absence of information directly providing range and orientation, as when viewing realistic pictures, the viewer's relative position with respect to an object can be only indirectly inferred from the projection of the object itself and its surround. The information in the projected lines of sight in the optic array can be used to infer the relative position of the viewer only if the viewer has at least a partial internal 3D model of the viewed objects and their surround (Grunwald and Ellis, 1986; Wallach, 1985). Thus, "shape constancy" in static, monocular scenes is somewhat circular, since the necessary shape information required to infer relative viewing position is itself the shape of the object in question. Nevertheless, shape constancy can be obtained through an interactive process if the viewer has a variety of static views of the same scene or object from different viewing positions and is able to construct appropriate correct hypotheses regarding the shapes. Because of inherent regularities in the world, viewers are usually quite good at forming appropriate shape hypotheses in natural environments (Gregory, 1966). But they can be tricked (Ittelson, 1952; Hochberg, 1987).

Shape constancy may be generalized to constancy of interrelations among objects in a spatial layout. Just as the shape of an object ordinarily appears constant when a viewer moves with respect to it, so too do the spatial interrelations among objects generally appear constant during corresponding movement of a viewer (Pirenne, 1970; Wallach, 1985; also see Ellis, Smith, and McGreevy,1987; Goldstein 1987). Piaget's decentering task, which requires that one imagine how a scene would appear from an external viewpoint, is an experimental scenario that particularly exercises this type of constancy (Piaget, 1932).

The Piaget decentering judgement is formally similar to that required of someone using a map to establish viewer orientation with respect to some exocentric landmark. When based on a map in which there is a marker representing the viewer's position, this judgement constitutes an exocentric direction judgement (Howard; 1982). In recent experiments we have examined a specific instance of this judgement by presenting subjects with computer-generated, perspective views of three-dimensional maps that have two small marker cubes on them (fig. 1). One marker represented the subject's assumed position on the map, i.e. his or her reference position. The other represented a target position. The subject's task was to make an exocentric direction judgement and estimate the relative azimuth of the target direction with respect to a reference direction parallel to one axis of the ground reference. In the previous experiments this reference was typically a full grid.

Interpretations of recent systematic measurements of these exocentric judgements have suggested that the observed patterns of error can be analytically described in terms of an external world coordinate system rather than a viewing coordinate system centered and aligned with the view direction. (McGreevy and Ellis, 1986; McGreevy, Ratslaff, and Ellis, 1985). In these experiments in which scenes were viewed from the center of projection direction, errors were observed in which the subjects exhibited a kind of equidistance tendency in that they judged the target cubes to be closer to the axis crossing the reference axis than they actually were. The same bias appeared independent of viewing direction, and thus the patterns of direction judgement error exhibited a kind of position constancy; that is, the errors were functions of the physical positions of the targets and not the subject's view of them.

Since the subjects were not allowed freedom to move the display's eye point during the individual judgements, position constancy would have to be based on assumed properties of the objects and features of the scene. The most likely feature that could provide the basis for this constancy is the ground reference meshed grid. Since the subjects may reasonably make the correct assumption that the grid axes are orthogonal, the grid can provide information about the compressive and expansive perspective effects due to the viewing parameters and allow the viewer to discount them. The information is provided most directly in the projected angle between the reference axis and the crossing axis. (Attnaeve and Frost, 1969; Ellis, Smith, and McGreevy, 1987).

Accordingly, removal of the crossing axis should remove the most direct information that allows the viewer to discount the geometric consequences of his or her particular viewing direction. Thus, a display used for the same kind of exocentric direction judgements, but lacking the crossing axis, should not exhibit position constancy. Direction judgement errors should now depend upon the viewing direction, since the source of information that allowed subject to directly determine the direction of the viewing vector has been removed. Experiment 1 examines this possibility.

EXPERIMENT 1

Methods

The eight paid subjects who participated in the experiment viewed a spatial layout made from a ground-plane reference and two slowly and irregularly tumbling wire-frame cubes marking positions on the reference and target positions on the plane. The techniques of data collection and viewing and display of the geometric projection were made identical to those used in previously described analytical and experimental studies (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986)

A ground reference of irregularly spaced, parallel lines aligned with the reference direction was constructed with randomized spacing (fig. 1). To assure presentation of the correct lines of sight, the subject's eye was located at the center of projection. Two symmetrically placed viewpoint locations which were rotated clockwise and counterclockwise 22° with respect to a reference direction were used (left stations: -22°; right station: 22°). Both had a depression of 22° below the horizon. The target cubes were randomly placed at 72 equally spaced target azimuths. The subject showed his or her estimates of the target cube azimuth angle with respect to the reference direction by controlling a dial drawn electronically on the CRT with the method of adjustment.

Results

Analysis of variance of the errors in target azimuth showed a statistically significant interaction between viewing station and true azimuth, ($F = 2.413$, $df = 71,497$, $p < .001$); hence, the azimuth error curves of left and right station appear to depend upon viewpoint.

Figure 2 shows the overall average error in the azimuth angle estimate for the left and for the right station plotted on circular graphs in which the direction of the error is shown as a directed arc. The across-subject means are good summaries of the data since the standard errors were only 1-4°. For both stations a systematic relationship between the azimuth error and the true azimuth angle is clearly recognized. Local minima in the errors, which are indicated by reversals in the directions of the error arcs, are not exactly where an actual grid-crossing axis would be, but are somewhat shifted toward a position orthogonal to the viewing axis. The largest direction errors are near ±45° and ±135° azimuth, and the error patterns for the symmetrically placed view stations are themselves approximate mirror images.

Discussion

The symmetrical pattern of mean error clearly shows a dependency on view direction and demonstrates a breakdown of position constancy in the error pattern, thus confirming the initial hypothesis that removal of the crossing axis should break down the position constancy. This breakdown is particularly evident near ±90 target azimuth since these are generally not minimums as they were for previous experiments with gridded ground references (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986). Thus, it is likely that the subjects are at least partially responding to the actual projected geometric properties of the scene which are seen from the separate viewpoints.

The breakdown of position constancy would be consistent with an alternative hypothesis which arises from previous analyses of errors in estimation of depicted directions in pictures (Ellis, Smith, and McGreevy, 1987; Gogel and Da Silva, 1987), and raises the classical question of the extent to which perception of an object's true geometric properties can be made to depend upon its projected retinal image (Thouless, 1931; Beck and Gibson, 1955; Gilensky, 1955; Gogel and Da Silva, 1987). According to this hypothesis, errors in judged direction in pictures are modeled as functions of the interrelations of actual lines of sight to viewed objects. For viewing situations in which pictures are viewed from the geometric center of projection, this analysis may be restricted to hypothesizing that the error in estimated target azimuth e is proportional to the difference between the depicted and projected azimuth angles y and y', respectively, i.e., $e = k(y' - y)$. Here the depicted angle y is measured with respect to the reference direction, clockwise positive, and the projected angle on the retina y' is measured with respect to the corresponding projection of the reference direction, clockwise positive. Positive errors correspond to clockwise errors. This formulation makes clear that not only should viewing direction affect the pattern of direction estimation, but also that symmetrically placed viewpoints should produce symmetrical patterns of direction errors.

The actual error data departs in significant ways from that expected based on this hypothesis. For example, the hypothesis implies that all direction errors for a view from the left station should be clockwise (fig. 3). The actual error data corresponding to this condition are both clockwise and counterclockwise, as shown by the circular plots of the error data. These error data could be modeled, as previously suggested, by introducing a 22° shift which produces an appropriate vertical shift in the theoretical function (McGreevy and Ellis, 1986; McGreevy, Ratzlaff, and Ellis, 1985). But this shift would be equivalent to asserting that the subject is responding to a potential projection rather than the one he or she actually sees. Since the data show evidence of viewpoint dependence and symmetry, the use of a theoretical function that suggests position constancy in the error data seems inappropriate. Accordingly, alternative theoretical explanations may be sought.

Binocular Conflict

One possible influence on the direction judgements that the subjects were requested to make is the binocular stimulus which they viewed. This stimulus was essentially the picture surface which provided fixed accommodative and vergence demands as well as disparity and motion parallax cues to its physical distance. These cues tell the viewer that all objects are at an approximately equal egocentric distance, i.e., on the picture surface. Thus, if exocentric direction were to be based solely on egocentric ranges estimated from the binocular information, all targets would be at the same distance. In the reference system used, all targets would appear at azimuth positions perpendicular to the view direction; e.g., for a left view station they would appear either at 112° or 68°.

This binocular information is at odds with the monocular information that is drawn on the display, e.g., the size changes of the cubes as its depicted distance changes. The viewer is in a sense being presented with two simultaneous but conflicting stimuli, one binocular and the other monocular. One may suppose that the resulting perception is a combination of the two. Conflicts of this type have been studied in classical experiments (Beck and Gibson, 1955; Gogel, 1977) in which monocular and binocular stimuli are superimposed and viewed. Significantly, the finding has been that for some simple stimuli, the binocular depth sensation spreads to determine the apparent position of the monocularly viewed component of the visual field. Accordingly, it is reasonable to suspect a similar process acting in the present experiment in which the binocular information in the picture surface causes the apparent positions of all targets to be attracted to a plane normal to the view direction. This process provides a hypothetical mechanism of the equidistance tendency observed in previous experiments. Its effects could be expected to be dominating were it not for the opposing influence of the remaining monocular depth cues provided by familiar shapes in the image.

Familiar Shape

Assumptions regarding the physical properties of objects in pictures are necessary for picture perception because of the inherent ambiguity of the pictorial information. Though the images used for Experiment 1 are relatively impoverished in this respect, the viewer may introduce useful assumptions such as that the reference lines dropped from the cubes markers are parallel and equal and are themselves perpendicular to the ground reference. Other important assumptions would be that the marker cubes remain equal in depicted size and that the lines in the ground reference are all parallel and coplanar.

These assumptions allow the clarification of the ambiguities inherent in the picture and can account for residual viewpoint-independent aspects of the errors. For example, despite the absence of a crossing axis, the pattern of mean direction error reported reverses direction in a manner similar to that found in earlier experiments with gridded ground references. This judgement bias has been described as an "equidistance" since the errors indicated the perceived space is collapsed toward the crossing axis, compressing the space in a picture. The clear observation of this bias without a crossing axis shows that the crossing axis itself cannot be its cause.

Inspection of the circular plots of the direction error in figure 3 shows that zero crossings of the direction error are not as closely associated with the ±90° target positions in the present experiment as they were in similar experiments using a complete grid. In fact, there is substantial error at these positions. For the most part the actual zero crossings are along axes rotated towards positions orthogonal to the direction of view and hence parallel to the surface of the picture. That they are not completely rotated orthogonal to the view vector is probably due to distance cues based on the changing sizes of the cubes and reference lines which both provide relative distance information.

In fact, it is probably correct to argue that shape assumptions are the principal basis for the construction of a perceived space from the line-of-sight information provided by a picture. The properties of this inferred virtual space are opposed, however, by the properties of the physical space of the picture surface which, as mentioned earlier, provide a mechanism to produce the pattern of direction errors that have been recorded. A simple test of this hypothetical mechanism would be to repeat the previous experiment in a real scene, a situation where there is no binocular conflict. Experiment 2 investigates this possibility.

EXPERIMENT 2

Methods

Eight paid subjects viewed physical objects with the viewing geometry used in Experiment 1. The marker cubes were physically reproduced with PVC pipe and positioned in a parking lot adjacent to the Life Science Building at the Ames Research Center. The details of data collection and stimulus presentation are contained in a San Jose State University thesis (Smith, 1986). Conditions in Experiment 1 were generally duplicated, although electronically produced apertures and dials were replaced by actual objects with similar functions. A microcomputer randomized the sequence of conditions for each subject and timed and collected the responses.

The subjects viewed the stimulus scenes binocularly from about 61 cm behind and centered in the viewing windows. At the 28-m viewing distance the reference cube subtended an average 5.2°. The cubes markers provided a significant stereoscopic stimulus since the binocular disparity of the target varied between 6.6 to 9.8 ft around the reference cue. This maximum disparity difference of 3.2 ft is about 50 times the stereo threshold, but within typical values for fusion area for the retinal excentricities used. Subjects were required to make azimuth judgments of 24 equally spaced, randomly presented target positions. Two viewing directions (±22° left and right viewing stations, respectively) and two square window apertures (30° and 60° FOV) were used. The dependent variable again was the error in judging target azimuth direction.

The distance between the two observation stations was 21 m. Rather than have subjects walk this distance as often as a completely random schedule would dictate, each subject stayed at one direction of viewing for at least 16 trials (one block). For each direction of viewing, the factorial combination of 24 target cube directions, two window sizes, and two repetitions were randomly assigned to six blocks of 16 trials. Each subject was presented with 12 blocks of trials (six at each direction of viewing). The total of 192 trials required about 3 hr to complete.

Results

The azimuth error data were analyzed by variance with repeated measures on target azimuth, window aperture, and viewing direction. Variation in the amount of background information by changing window size did not significantly affect judgments of azimuth error and did not interact with any other factor. As in Experiment 1, the two-way interaction between azimuth of the target cube and view direction was statistically significant ($F(23,138) = 3.861$, $p < .001$).

The nature of the statistical interaction that was observed between viewpoint and target azimuth is clarified by circular plots in figure 4. This figure illustrates the underlying symmetry in the error data, which is similar to that in Experiment 1. It also shows the absence of the "equidistance tendency" and generally smaller size of the errors.

Discussion

The azimuth error observed in Experiment 2 does not exhibit the "equidistance tendency." Thus the results confirm the supposition that the binocular conflict or other cues to the picture surface such as motion parallax could be the cause of the bias. In Experiment 1 the azimuth errors for displays viewed from the correct geometric eye point were generally away from the reference axes and towards the crossing axis. This equidistance tendency has been called a "telephoto bias" since it resembles the pattern of error that would be induced if the view of the spatial configuration were distorted by a telephoto lens. In fact, it was not a true "telephoto bias" and equidistance tendency is a better description because the reported spatial compression was not aligned with the actual view direction, but with the axes, or implicit axes, in the scene itself. In contrast to the relatively large bias in Experiment 1, the errors in Experiment 2 are smaller and away from the crossing axes rather than towards them. The residual error pattern, however, does continue to exhibit a symmetrical dependence on view positions, supporting the conclusion from Experiment 1 that the error pattern does not exhibit position constancy. The new error pattern in Experiment 2 needs an explanation.

The bias pattern is not similar to what would be expected if it were due to the difference between the size of the projected and depicted azimuth angles. If the difference between depicted and projected angle were the cause of the observed error, the errors would be expected to resemble the traces in figure 3. As in Experiment 1, the results do not closely resemble these traces, so new alternatives need to be considered to explain both the smaller average size of the error and the particular pattern itself.

Since correct three-dimensional interpretation of the array of lines of sight to the objects in view depends upon both a correct internal model and a correct estimate of viewing direction, errors in either of these assumptions can be a source of systematic bias . Systematic errors in the internal model would result in apparent loss of perceptual rigidity when the object was rotated or translated. These kinds of distortions are not expected and were not reported as the cubes tumbled in the wind during Experiment 2. Accordingly, the biases found in this experiment might be attributed to incorrect estimation of the viewing direction. A classical error of this kind is called "slant overestimation" (Sedgwick, 1986) and corresponds to overestimation of the amount of depression of the viewing vector.

Figure 5 shows a family of theoretical azimuth error curves for different overestimates of the viewing vector depression together with the data from Experiment 2. These curves are constructed on the assumption that the viewer makes an error in the interpretation of the projected target angle, in a sense, by looking up its 3D characteristics in the wrong table. For example, the trace labeled "elevation = -40" shows the expected azimuth errors from a subject who, when looking a scene from a left viewing station (azimuth = 22.5°) with a -22.5° elevation angle, assumes that the actual elevation is -40°, and looks up the 3D interpretation of the projected angles that he or she does see in the wrong table, i.e., the one for a -40° elevation. Interestingly, the hypothesis that azimuth error could be influenced by the difference between depicted target angle and its projection, which was described in the discussion of Experiment 1, is really a special case of this kind of slant overestimation. The hypothesis discussed in Experiment 1 is equivalent to asserting that the overestimation is equal to the complement of the actual depression angle.

Figure 5 also shows the azimuth error data from Experiment 2 combined for both view stations by reflecting the data from the right view station so as to allow averaging with that of the left station. The combined data are then replotted in cartesian form for comparison with the theoretical curves. The experimental data exhibit several features inconsistent with a slant overestimation. In particular, the errors are smaller, not markedly sinusoidal, and not biased in the correct directions. The elevation overestimation hypothesis predicts, for example, that from the left viewing station, errors for depicted angle between 0° and 180° should be clockwise whereas the data show a predominant counterclockwise bias for these conditions. In fact, the data may suggest an elevation underestimation. Clearly, further experiments in which errors in exocentrically judged azimuth and estimates of viewing direction elevation and azimuth are both collected are needed to evaluate the role of viewing direction misjudgement as an explanation for the pattern of azimuth error.

Summary

1. Errors in exocentric judgements of the azimuth of a target generated on an electronic perspective display are not viewpoint-independent, but are influenced by the specific geometry of their perspective projection.

2. Elimination of binocular conflict by replacing electronic displays with actual scenes eliminates a previously reported "equidistance tendency" in azimuth error, but the viewpoint dependence remains.

3. The pattern of exocentrically judged azimuth error in real scenes viewed with a viewing direction depressed 22° and rotated ±22° with respect to a reference direction could not be explained by overestimation of the depression angle, i.e., a slant overestimation.

REFERENCES

Attneave, F., and Frost, R. (1969) The determination of perceived tridimensional orientation by minimum criteria. *Percept. and Psychophys., 6,* 391-396.

Beck, Jacob and Gibson, James J. (1955) The relation of apparent shape to apparent slant in the perception of objects. *J. Exp.Psychol.*, 50, 125-133.

Borresen, C. Robert and Lichte, William H. (1962) Shape constancy: dependence upon stimulus familiarity. *J. Exp.Psychol., 63,*92-97.

Ellis, Stephen R., Smith, Stephen, and McGreevy, Michael W. (1987) Distortions of perceived visual directions out of pictures. *Percept. Psychophys.*, 42, 535-544.

Gilinsky, A.S. (1955) The effect of attitude upon perception of size. *Amer. J. Psychol., 68,* 173-192.

Gogel, Walter C. (1977) The metric of visual space. in *Stability and constancy in visual perception*, W. Epstein ed. New York, Wiley.

Gogel, Walter C. and Da Silva, Jose Aparecido (1987) Familiar size and the theory of off-sized perceptions. *Percept. Psychophys., 41,* 318-428.

Goldstein, E. Bruce (1987) Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *J. Exp.Psychol., 13,* 256-266.

Gregory, Richard L. (1966) *Eye and brain: the psychology of seeing*. New York: World University Library

Grunwald, Arthur, and Ellis, Stephen R. (1986) Spatial orientation by familiarity cues. in *Training, human decision making and control.* J. Patrick and K.D. Duncan, eds. North-Holland, Elsevier, 257-279.

Ittelson, W.H. (1951) Size as a cue to distance. *Amer. J. Psychol., 64,* 54-67.

Ittelson, W.H. (1952) *The Ames demonstrations in perception*, Princeton: Princeton University Press.

Hochberg, Julian. (1987) Machines should not see as people do, but must know how people see. *Computer Vision, Graphics, and Image Processing* (in press).

Howard, Ian (1982) *Human Visual Orientation,* New York: Wiley

Flavell, J.H. (1963) *The developmental psychology of Jean Piaget.* Princeton: Van Norstand.

Metzger, Wolfgang (1953) *Gesetze des Sehens,* Frankfort am Main: Waldeemar Kramer

McGreevy, Michael W., Ratzlaff, Cordell, and Ellis, Stephen R. (1985) Virtual space and two-dimensional effects in perspective displays. *Proc. 21st Ann. Conf. on Manual Contr.*, June, 1985.

McGreevy, Michael W. and Ellis, Stephen R. (1986) The effect of perspective geometry on judged direction in spatial information instruments. *Human Factors*, 28, 439-456.

Sedgwick, H. A. (1986) Space perception in *Handbook of Perception and Human Performance vol 1*. K. R. Boff, L. Kaufman, J.P. Thomas eds, New York: Wiley, 21:19 - 21:23

Piaget, J. (1932) *The origins of intelligence in children*, London: Keagan Paul.

Pirenne, M. (1970) *Optics, painting, and photography*, Cambridge: Cambridge Univ. Press

Smith, Stephen (1986) Viewing in three-dimensional visual space: the two-dimensional effect and direction judgement errors. Unpublished MA thesis, San Jose State Univ., San Jose, CA.

Thouless, Robert H. (1931) Phenomenal regression to the real object. *British J. Psychol., 21*, 338-59.

Wallach, Hans. (1985) Perceiving a stable environment. *Sci. Amer., 252*, 118-124.

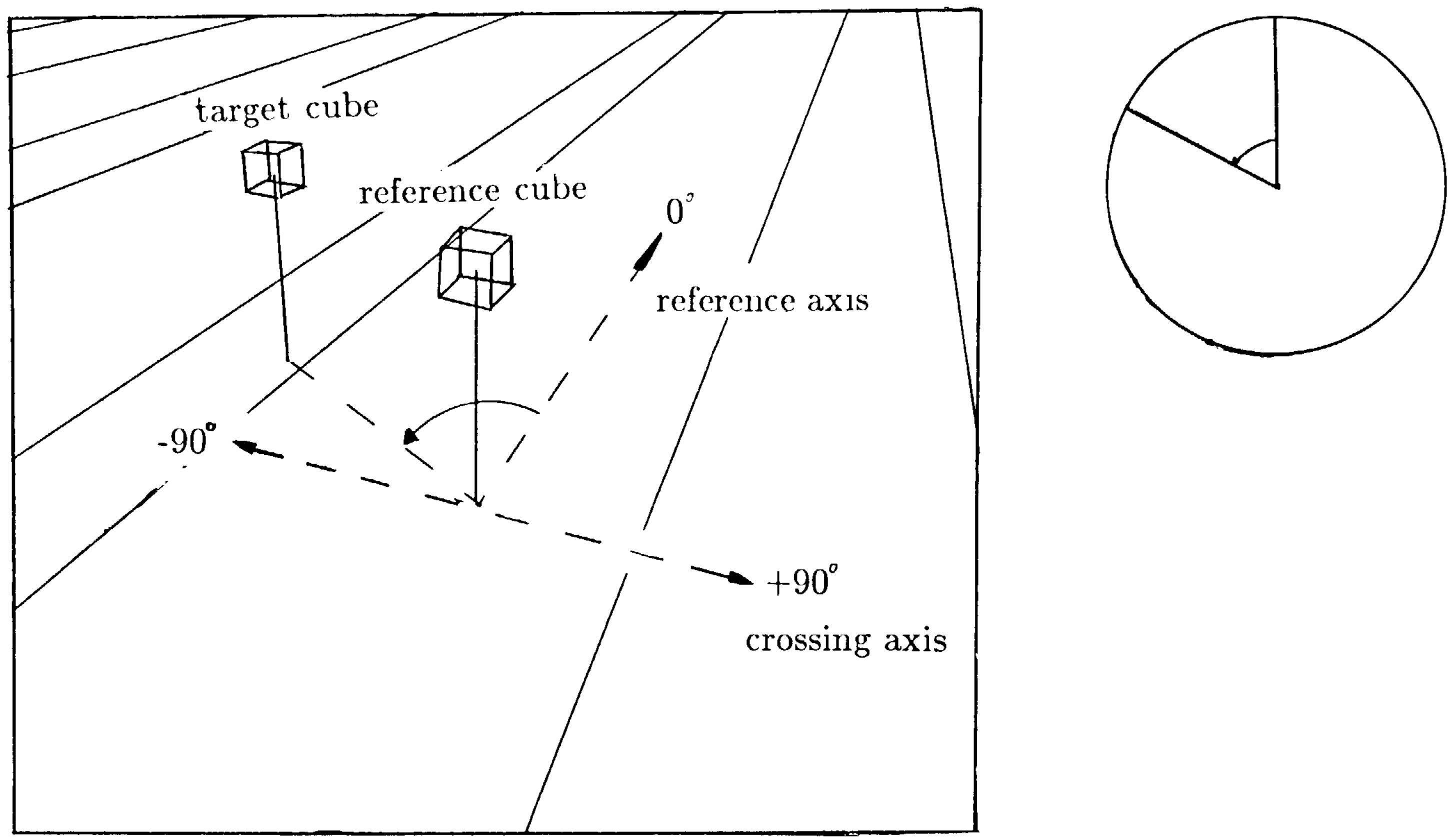

Figure 1. Schematic of the direction judgement task. The subject adjusted the angle shown on the dial at the right until it appeared equal to the azimuth angle of the target cube. Dashed line, labels, and arrows did not appear on the display. The ground reference in previous experiments was a full grid.

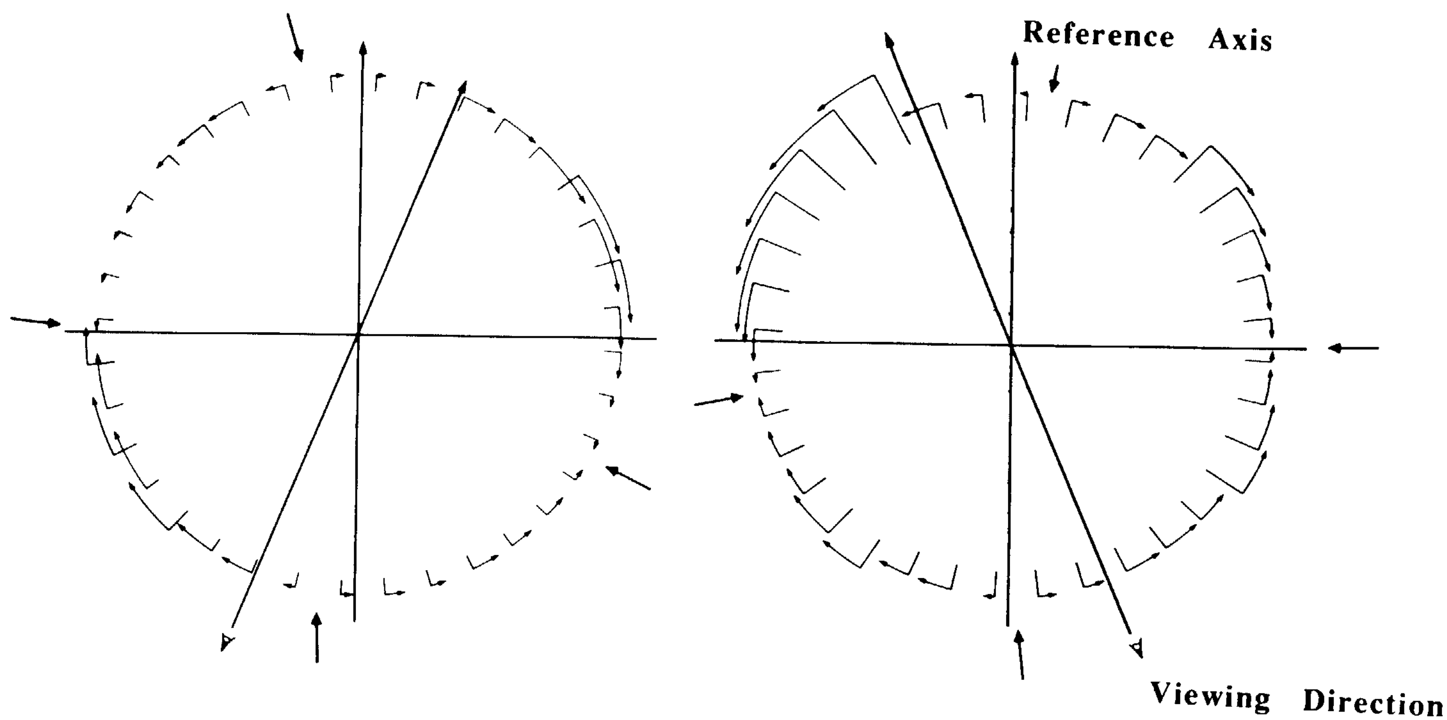

Figure 2. Circular plot of mean azimuth error for direction judgement experiment using a display with a ground reference of parallel lines. The length of the arc corresponds to the mean error in estimating target azimuth. Arrows show target azimuths, where azimuth errors were at local minimum.

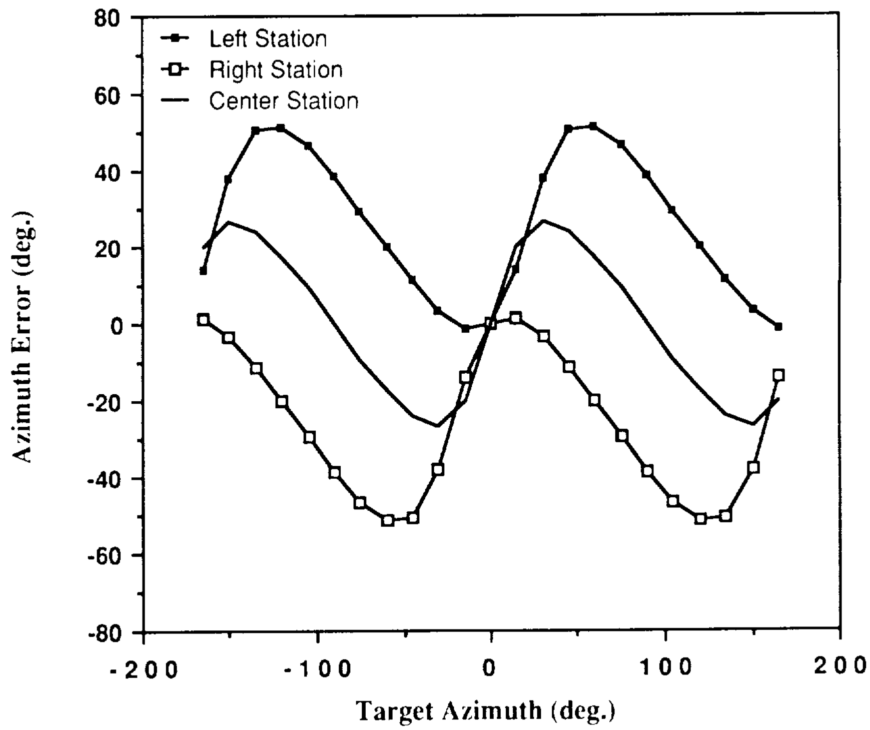

Figure 3. Predicted azimuth errors. If the subjects' direction errors were entirely determined by the difference between the true depicted value of a target's azimuth angle and its projection, errors like those shown in this figure would be expected. The three traces show the expected error pattern if the depicted targets were viewed from a left (22.5°), right (-22.5°), or center (0°) viewing azimuth.

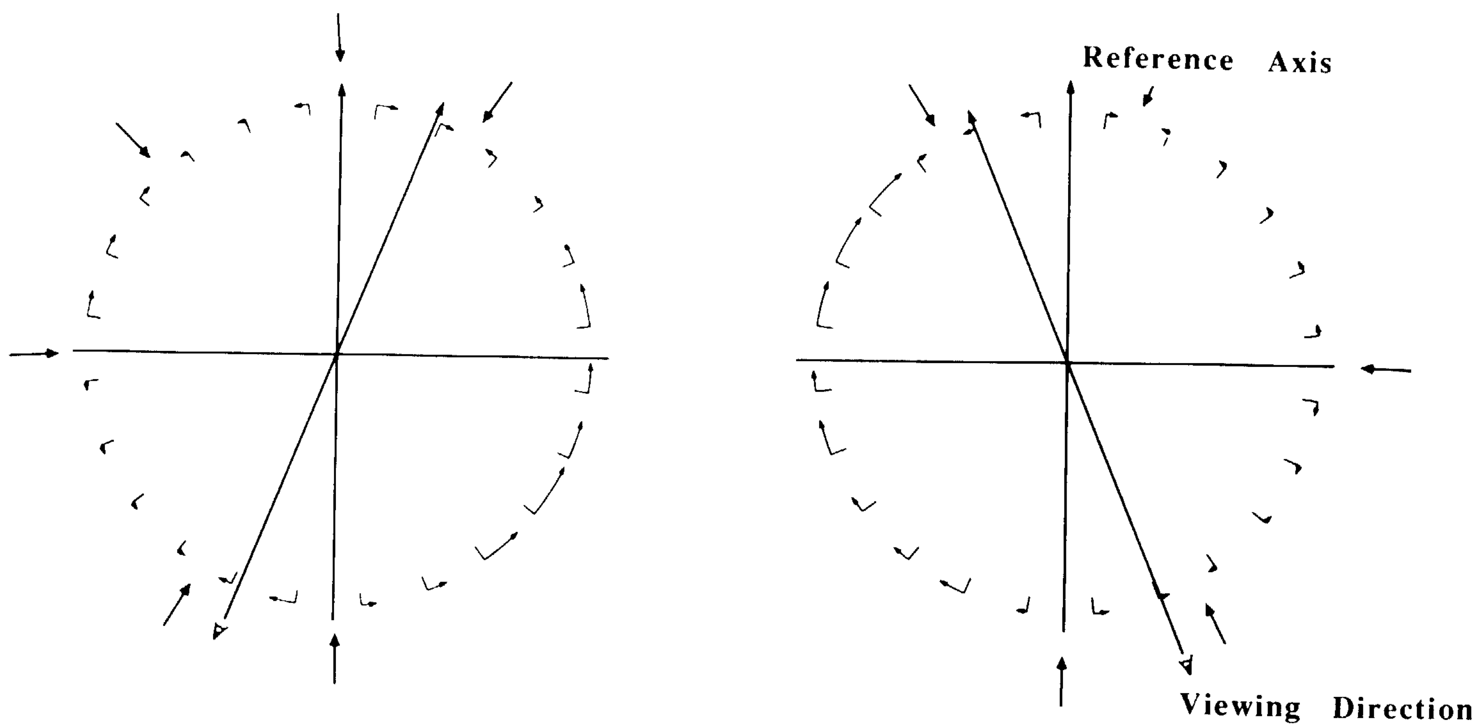

Figure 4. Circular plot of mean azimuth error for direction judgment experiment using an actual scene. Arrows show target azimuths where azimuth errors were at local minimum.

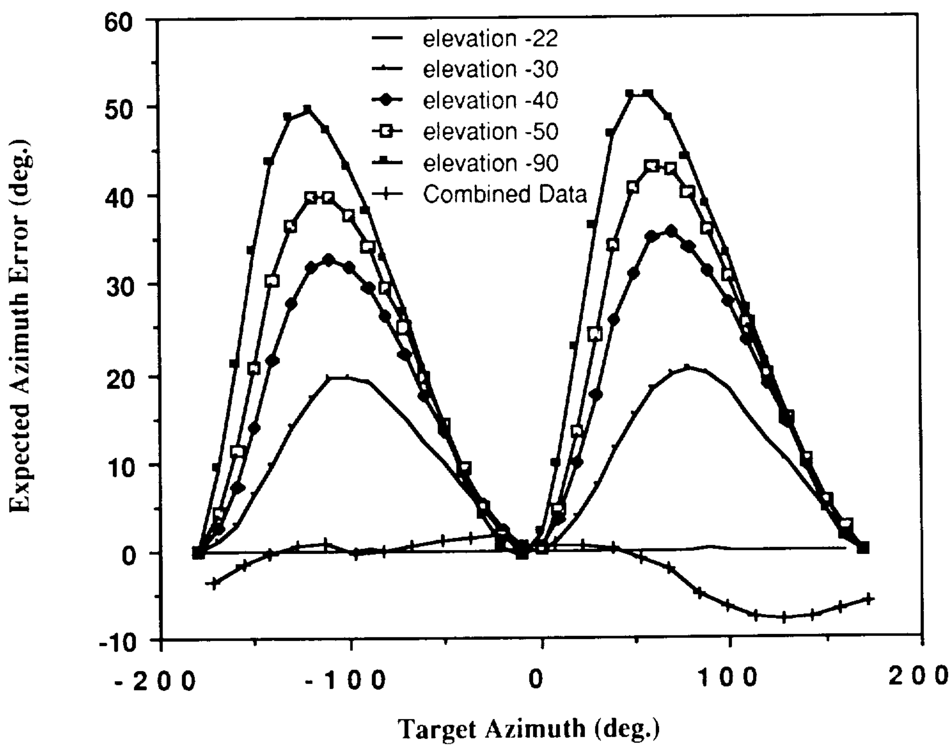

Figure 5. Plot of expected azimuth error if a subject misjudged the depression angle of the viewing direction. Errors are calculated for a left viewing station (azimuth = 22.5°) with a depression angle of -22.5°, assuming that the subject misjudged the depression to be the parameter of each of the curves. Error data from Experiment 2 are also plotted for comparison.