34

Direction judgement error in computer generated displays and actual scenes[1]

Stephen R. Ellis, Stephen Smith, Arthur Grunwald and Michael W. McGreevy

NASA Ames Research Center
Moffett Field, CA 94035

Introduction

Shape constancy

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 correct for his relative position and orientation with respect to a viewed object. This discounting of relative position may be derived directly from the ranging information provided by stereopsis, motion parallax, vestibularly sensed rotation and translation, or corollary information associated with voluntary movement. Some correction may even be possible directly based on purely gibsonian higher order psychophysical variables.

Significantly, shape constancy, which usually involves requesting that the viewer make some estimate of the geometric properties of an object, does not disappear during static, monocular viewing. 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

---

[1] Preliminary versions of the results included in this paper have been reported at the *NASA Ames—U.C. Berkeley Conference on Spatial Displays and Spatial Instruments*, Asilomar, California, September, 1987.

recall of only some salient features, such as parallelism of significant planes of the object.

In situations in which information directly providing range and orientation is absent, as during viewing realistic pictures, the viewer's relative position with respect to an object can only be indirectly inferred from the projection of the object itself and its surround. But the information in the projected lines-of-sight in the optic array can be used to infer the relative position of the viewer with respect to the pictured objects only if the viewer has at least a partial internal 3D model of the viewed objects and their surround (Grunwald and Ellis, 1986; Grunwald *et al.*, 1988; 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 correct hypotheses regarding the shapes. Due to 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).

Position constancy

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; but also see Ellis *et al.*, 1987; Goldstein 1987). Piaget's decentering task which requires that one imagine how a scene would appear from an external view point is an experimental scenario that particularly exercises this type of constancy (Piaget, 1932; Flavell, 1963).

The Piaget decentering judgement is formally similar to that required of someone using a map to establish his orientation with respect to some exocentric landmark. When based on a map in which there is a marker representing the viewer's position, i.e., the "you-are-here" marker (Levine, 1984) 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 (see Figure 1). One marker represented the subject's assumed position on the map, i.e., his 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 of two sets of orthogonal parallel lines.

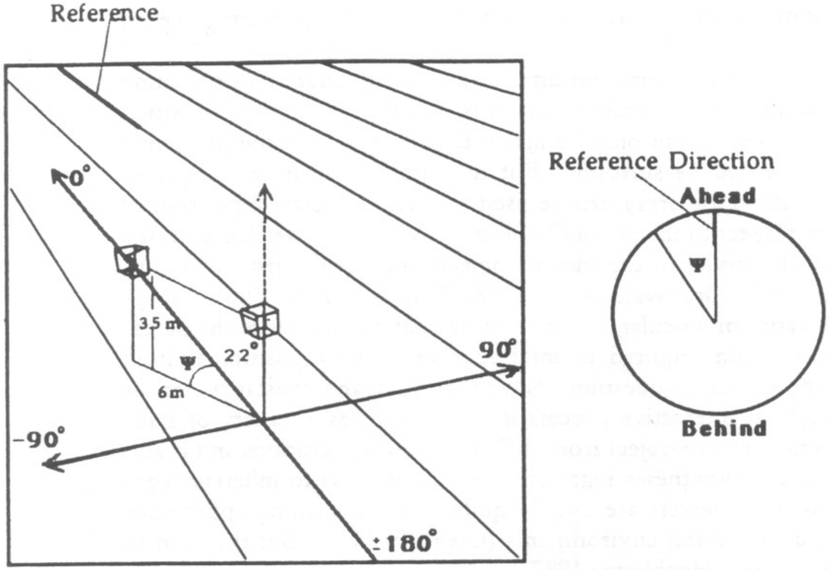


Figure 1. Schematic illustration of the direction judgement task. The subject adjusted the angle Ψ shown on the dial at the right until it appeared equal to the azimuth angle Ψ of the target cube. Dotted lines, labels and arrows did not appear on the display.

Position constancy during judgement of exocentric direction

Interpretations of recent systematic measurements of these exocentric judgements have suggested that the observed patterns of error could 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 *et al.*, 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.

Geometric mechanisms

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 grid. Since the subjects may correctly assume that the grid axes are orthogonal and in the same plane, the grid can provide information about the compressive and expansive perspective effects of the viewing parameters and allow the viewer to determine them.

Alternative sources of the same information could be sought in the convergence angle between parallel lines at the horizon. Since the horizon was generally not visible due to placement of the clipping planes, this angle would be hard to determine. Thus, the information sufficient to determine the view direction pitch, Θ, and yaw, Ψ, is provided most directly in the projected angle between the reference axis and the crossing axis. For objects viewed along the principal direction of the projection this angle, A′O′B′ in Figure 8, may be expressed as:

$$\angle A'O'B' = \sin(2\Psi)\tfrac{1}{2}(\sin^2\Theta - 1)((\cos^2\Psi + \sin^2\Theta\sin^2\Psi)(\sin^2\Psi + \sin^2\Theta\cos^2\Psi))^{-1/2} \qquad (1)$$

This projected angle is invariant with distance and magnification distortion and also approximately describes the projected angles for right angles near the principal direction of the projection and in the ground reference plane (see Appendix for development of this equation; for related work see Attnaeve and Frost, 1969; Ellis *et al.*, 1987). When the pitch down is greater than about −70 deg, the right most term is weakly modulated by changes in yaw and approximately is equal to 1 so that the resulting equation is simplified to: $\frac{1}{2}(\sin\Theta - 1)\sin(2\Psi)$. In fact, this simplified expression is the dominant factor for most of the range of Θ and Ψ and allows the geometry to provide a relatively fixed association between the grid crossing angle, assumed to be 90 deg, and its projection. The interpretation of the projected angle only breaks down for very shallow depression angles when the projection can become indeterminate due to small amounts of measurement noise. Thus, measurement of the projected angle can provide a means to infer and possibly discount view direction.

Experimental manipulation

Deletion of the crossing axis should remove this information that directly allows the viewer to correct for the geometric consequences of his particular viewing direction. Thus, the direction errors from a display used for the same kind of exocentric direction judgements but lacking the crossing axis should exhibit weakened position constancy. With such a display direction judgement errors should depend upon the viewing direction since the principal source of information that allowed the subject directly to determine the direction of the viewing vector has been removed. Experiment 1 examines this conjecture.

Experiment 1

Methods

Subjects

Eight paid subjects participated in the experiment, six of whom were aircraft pilots. All were selected from the Ames Research Center subject pool of aircraft pilots and non-pilots and had normal or corrected to normal visual acuity.

Apparatus and stimuli

The images presented to the subjects showed a spatial layout made from a ground plane reference and two slowly and irregularly tumbling wire-frame cubes (< 1 rpm) used to mark positions on the reference plane. One marked the reference position at the center and the other was the target. The viewing and display parameters of the geometric projection were made identical to those used in previous analytical and experimental studies (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986; Grunwald *et al.*, 1988). Most notably the reference cube was centered in the frustum of vision and subtended an average of about 5 deg of visual angle. The entire image was 28 × 28 cm and viewed from a 48 cm viewing distance. It was generated with a Silicon Graphics IRIS 2400 color raster graphics workstation controlled by mouse and keyboard input.

Figure 1 provides a schematic image of the stimuli illustrating the ground reference made only of randomized parallel line segments which was used to remove cues provided by compressive and expansive projection effects evident in the angles between orthogonal grid axes. The constrained randomization of the placement of the line segments and the slow tumbling of the cube markers were features intended to defeat specific object-based judgement strategies which might favour particular target positions. The subjects were thus encouraged to make a subjective estimate of the spatial layout.

Viewing stimulus geometry and procedure

The ground reference of parallel lines aligned with the reference direction was constructed with randomized modeled spacing at an average of 5 m and a modeled viewing distance of 28 m to the reference cube. To assure presentation of the correct lines of sight, the subject's eye was located at the center of projection of the image 48 cm from the screen. Two symmetrically placed viewpoint locations rotated clockwise and counterclockwise 22 deg with respect to a reference direction were used. Hereafter these viewing directions are referred to as "left" station and "right" station, respectively. Both had a viewing pitch of −22 deg, i.e., pitch down. The target cubes were randomly placed at each of 72 target azimuths with respect to the reference direction ranging between −177 deg (ccw) and +178 deg (cw) in 5 deg increments.

Each experimental series contained one set of 72 azimuth angles for the left and one set for the right station. The viewing station (left or right) and target azimuth were picked randomly without replacement from the series. Each subject performed two series of 144 trials each, or two repetitions per azimuth angle, per station in a 2 × 72 factorial design with repeated measures.

The subject was instructed to show his estimates of the target cube azimuth angle with respect to the reference direction by adjusting a dial drawn on the CRT to the right of the display. The 10 cm diameter dial, which was drawn electronically adjacent to the perspective viewport, was provided with a vertical red line corresponding to a red line in the perspective display that indicated the reference direction. Mouse buttons were used to rotate a pointer on the dial clockwise and counter clockwise for a method of adjustment. The dial adjustment required that the subject judge the target azimuth through a subjective compensation involving an inverse perspective transformation. The subjects were instructed to produce an angle that would be needed if they had to correlate the display information with a 2D map of the layout. To insure that all subjects understood the task, the frame of reference for the judgement and the meaning of the dial adjustment subjects were shown a demonstration program in which dial position and the azimuth of a target cube were both slewed to the *y* axis of the mouse.

Although no time limit was set for the response to each trial, the subjects were told not to take more than about 30 seconds per judgement. Azimuth direction error was calculated as estimated azimuth target minus true azimuth so that clockwise errors were positive.

Results

The subjects' estimates of depicted target azimuth were subjected to an analysis of variances, with repeated measures on subjects. The analysis showed a statistically significant interaction between viewing station and true azimuth, ($F = 2.413$, $df = 71,497$, $p < 0.001$), hence the azimuth error curves of left and right station appear to depend upon viewpoint.

Figure 2 shows the error in the azimuth angle, averaged over all eight subjects, for the left and for the right stations plotted on circular graphs in which each arc length corresponds to both the magnitude and direction of an error. The across subject means are good summaries of the data since the associated standard errors were only 1–4 deg. For both stations a systematic relationship between the azimuth error and the true azimuth angle, is clearly recognized. The errors are virtually zero on the reference axis. Secondary zeros are not exactly where an actual grid crossing axis would be but are shifted. Those zero crossings on the side in the direction of the view vector

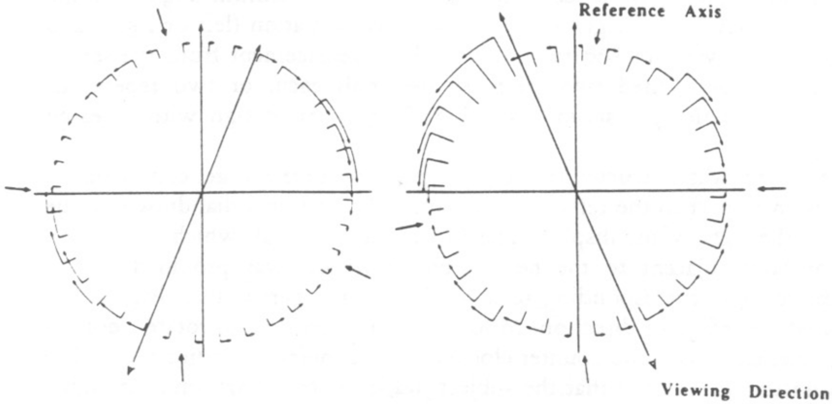


Figure 2. Circular plot of mean azimuth error in Experiment 1. The eye symbols show the subject's view directions with respect to the reference and crossing axes. The length of each directed arc corresponds to the mean error (N = 8) in target azimuth at the position marked by the tail of the arc. Reversals of arc directions show target azimuths where azimuth errors were at local minimum.

rotation tend to be rotated towards a position perpendicular to the view direction. As in previous experiments, the largest direction errors are near ±45 and ±135 deg azimuth.

In order to investigate the existence of symmetry about the reference axis, the right station data in the set were reflected and replotted in ordinary cartesian form. As may be seen from Figure 3, the reflection largely superimposes the error data from both view stations confirming the expected symmetry in the error pattern. This observation of symmetrical response patterns provides a control for dial-specific response biases.[2] Had these been a dominant effect, the distinctive features of the error in the data from the left viewpoint would not have symmetrical counterparts in the data from the right viewpoint.

Discussion

The generally symmetrical pattern of mean error clearly shows a dependency on view direction and demonstrates a breakdown of position constancy in the error pattern. This result confirms the initial hypothesis that removal of the crossing axis should break down this constancy. The breakdown is particularly evident near ±90 target azimuths since these are generally not minimums near zero as they were for left and right viewing directions in previous

---

[2] Control calibration experiments with the adjustment dial have shown that the error in across subject means range +/−2 deg with less than a 2 deg clockwise bias and with a pattern uncorrelated with observations in experimental results described in this paper.

experiments with fully grided ground references (McGreevy and Ellis, 1986; Grunwald and Ellis, 1986).

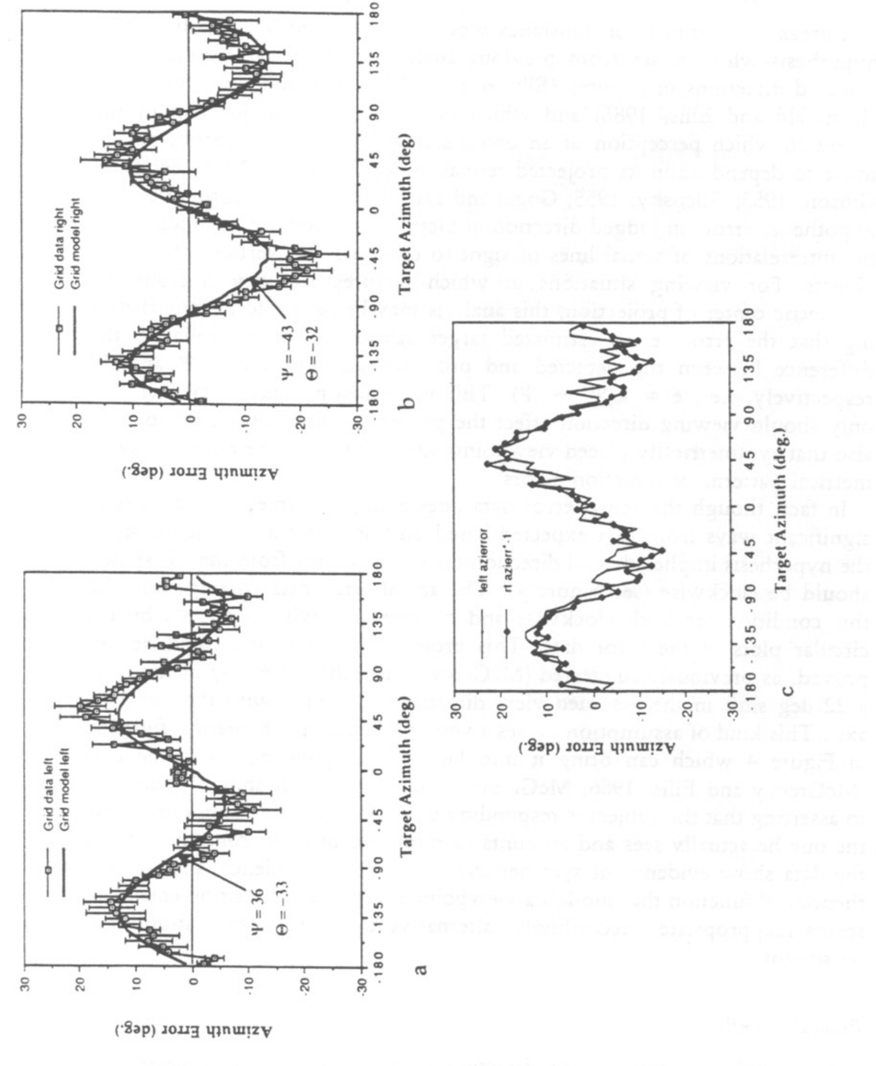
Alternative hypotheses

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 *et al.*, 1987; Gogel and Da Silva, 1987; Grunwald and Ellis, 1986) and which 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 contours and vertices of 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, e, in estimated target azimuth is proportional to the difference between the depicted and projected azimuth angles Ψ and Ψ' respectively, i.e., $e = k(\Psi' - \Psi)$. This formulation makes clear that not only should viewing direction affect the pattern of direction estimation but also that symmetrically placed viewpoints should produce the observed symmetrical patterns of direction errors.

In fact, though the actual error data does exhibit symmetry, it 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 (see Figure 4). The actual error data corresponding to this condition are both clockwise and counter-clockwise as shown by the circular plots of the error data. This projected angle model could be improved, as previously suggested (McGreevy and Ellis, 1986), by introducing a 22 deg shift in the assumed view direction to align it with the reference axis. This kind of assumption causes a vertical shift in the theoretical function in Figure 4 which can bring it into better correspondence with the data (McGreevy and Ellis, 1986; McGreevy *et al.*, 1985). This shift is equivalent to asserting that the subject is responding to a potential projection rather than the one he actually sees and amounts to modeling position constancy. Since the data show evidence of symmetrical viewpoint dependence, the use of a theoretical function that models a viewpoint-independent, position constancy 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


Grid data left
Grid model left
Ψ = 36
Θ = −33
Azimuth Error (deg.)
Target Azimuth (deg)
30
20
10
0
-10
-20
-30
-180
-135
-90
-45
0
45
90
135
180
a
Grid data right
Grid model right
Ψ = −43
Θ = −32
Azimuth Error (deg.)
Target Azimuth (deg)
b
left azierror
rt azierr*-1
Azimuth Error (deg.)
Target Azimuth (deg.)
c

and vergence demands as well as disparity and motion parallax cues to its physical distance, since no head restraint was used. 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 as would be the case if the pitch of view vector were overestimated to be −90 deg. 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 68 or −112 deg. Some evidence for this is found in Figure 2 which shows for both view stations that the apparent azimuth of targets located on the side of the direction of view rotation is rotated towards a plane orthogonal to the view direction.

The binocular information possibly causing this apparent rotation is at odds with the monocular information that is drawn on the display, e.g., the decreasing projected size of the cube as its depicted distance increases. 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 a visually proximate, monocularly viewed component of the visual field.

Accordingly, it is reasonable to suspect a similar process could influence the judgements in this experiment. In this case 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 and induce an overestimate of the view vector pitch. This process provides a hypothetical mechanism for the equidistance tendency observed in the first experiment. Its effects could be expected to be dominating were it not for the opposing influence of the numerous monocular depth cues provided by familiar shapes in the image. Since the monocular cues are well developed from familiar shapes in these images, the binocular cues would not be expected to determine totally the

◀ *Figure 3. Cartesian plots of mean azimuth errors in Experiment 1 are plotted for both left (a) and right (b) viewing stations. Error bars are ±1 standard error, $N = 8$. The heavy traces are theoretical functions derived from the assumption that the subjects misjudge the view vector. The estimated viewing parameters for the theoretical trace are at the tail of the arrow and show the expected overestimation of the true values ($\Psi = \pm 22$, $\theta = -22$). These functions have been fitted to the data as described in the discussion of Experiment 2 and correlate fairly well with the data (left station: $r = 0.896$, $p < 0.001$, $df = 72$; right station: $r = -0.925$, $p < 0.001$, $df = 70$). Part (c) illustrates the symmetry between the pattern of error from the left and right view stations by reflecting the data from the right station and replotting it with that from the left station.*

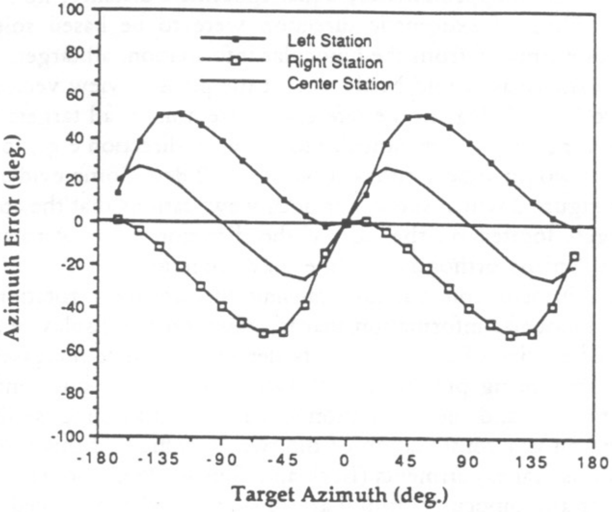


Figure 4. Predicted azimuth errors. If the subject's direction errors were entirely determined by the difference between the true depicted value of a target's azimuth angle and its projection, errors like those shown in this figure would be expected. The three traces show the expected error pattern can be vertically shifted if the depicted targets are assumed to be observed from a left (22.5 deg), right (−22.5 deg), or center (0 deg) viewing station.

apparent distances to the objects in the images. Furthermore, the rotated viewing direction would introduce asymmetries into these monocular features of the image such as the texture gradient that could be expected to introduce corresponding asymmetries into their interaction with binocular cues to distance. Examples of this kind of feature in the direction judgement data are the mismatches between the theoretical functions and those data that occur at symmetrical target positions for data from the left and right view stations (see Figure 3).

The overestimation of pitch discussed above is a form of classical error called "slant overestimation" (Sedgwick, 1986)[3] and may provide a mechanism for the incorrect estimate or use of the viewing parameters. Figure 5 shows a family of theoretical azimuth error curves for different overestimates of the viewing vector pitch together with the data from Experiment 1. These curves are constructed on the assumption that the viewer correctly measures the line of sight angles to all contours and vertices but makes an error in the

---

[3] 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, really is a special case of this kind of slant overestimation. It is equivalent to asserting that the overestimation is equal to the complement of the actual pitch angle.

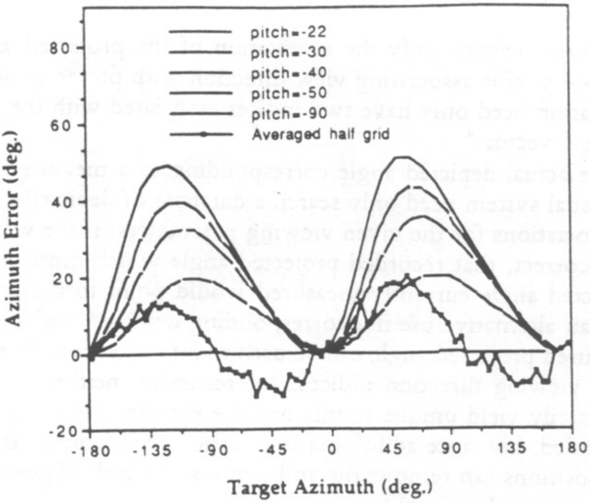


Figure 5. Plot of expected azimuth error if a subject misjudged the depression angle of the viewing direction. Errors are calculated for a left viewing station (azimuth = 22.5 deg) with a true depression angle of −22.5 deg assuming that the subject misjudged the depression by the parameter of each of the curves. Average error data from Experiment 1 are also plotted for comparison. These data are the average of the left station with those from the right station which have first been reflected to correct for symmetry differences.

interpretation of the projected target angle by in a sense looking up its 3D characteristics in the wrong table. For example, the trace labeled "pitch = −40" shows the expected azimuth errors from a subject who when looking at a scene from a left viewing station ($\Psi = 22.5°$, $\theta = -22.5°$) incorrectly assumes that the actual pitch is −40 deg. He then looks up the 3D interpretation of the projected angles that he does see in the wrong table, i.e., the one for a −40 degree pitch, and finds incorrect corresponding depicted angle values. These curves show that errors in pitch estimation alone can not account for the data from Experiment 1.

In addition to the pitch, the yaw and roll of the view vector may be incorrectly estimated and cause a similar kind of error in a more generalized look-up table that associates depicted angles with their projections. Though the roll of the view vector also can influence the geometric relationship between depicted and projected angles; the kinematics of the head constrain the amount of roll associated with head rotations in pitch and yaw. For constrained ranges of rotation, the cross-coupling of yaw onto roll has been estimated, for example, at only about 2 per cent (Chouet and Young, 1974; Larsen and Stark, 1988), but in any case, any roll around the view direction after a pitch and yaw will not influence the size of projected angles with vertices along the view vector, only the orientation of the projected angle. Accordingly, a look-up table associating view direction with projected angles as a first approximation need only have two indices associated with the pitch and yaw of the view vector.[4]

To determine the actual depicted angle corresponding to a measured projected angle, the visual system need only search a database of depicted-angle-projected-angle associations for the given viewing parameters. If the viewing assumptions were correct, that recorded projected angle which most closely matched the projected angle currently measured would point to the correct depicted angle. In an alternative use if a corresponding depicted angle can be assumed, the measured projected angle can be used to infer the view direction by recovering the viewing direction indices. Interestingly, neither of these searches will necessarily yield unique results because the viewing constraints may be under specified, but since additional assumptions concerning size and habitual viewing positions can remove the ambiguities, the lack of geometric uniqueness is not necessarily a problem.

General theories of errors in determining depicted target directions may then be expressed in terms of errors the viewer makes in determining the view direction. A model of this sort (Tharp, 1989; Tharp and Ellis, 1990) was fitted to the data from Experiment 1 by conducting a grid search of a range of possible errors in pitch and yaw estimates to find a pair that minimized the RMS error between the inferred target azimuth errors that such erroneous viewing assumptions would produce and the actual observed azimuth errors. These fitted models plotted in Figure 3 show that this two parameter theory provides a pretty good fit to the data of Experiment 1.

As shown in Figure 3, the view vector parameters estimated from the data in Experiment 1 indicate that not only is pitch overestimated by 10–11 deg, as conjectured in the discussion of the experiment, but so is yaw, in fact yaw has a greater error of 14–21 deg. This difference makes sense in retrospect since proprioception would have been a good cue to view direction and the subject's heads were correctly pitched to view the display surface. They were not rotated in yaw to the left or right. Though the view vector model fits fairly well for a two parameter model, i.e., Pearson r is around 0.9, systematic deviations from it near the obliques suggest that further refinement may be necessary. Interestingly, an alternative model which was based on a "telephoto" error in measuring the divergence of lines of sight into a picture also had the same systematic difficulty matching the azimuth error data along the oblique axes (Grunwald *et al.*, 1988).

---

[4] If the head kinematics mimicked eye kinematics and obeyed Listing's and Donder's Laws, the look-up table would not in principle require more than two indices. We have made some rough measurements of the amount of roll associated with head clockwise and counterclockwise yaws of 30 deg and downward pitches of −15 deg and found the amount of associated roll to be of the order of 1 deg. It can, however, be up to about 30 per cent of the pitch or yaw for more extreme positions such as a 60 deg yaw associated with a 60 deg pitch up. But the amount of roll produced under these circumstances is very variable, depending upon the constraints placed on the torso, and is under the voluntary control of the subject.

We have reanalysed previously reported data in which the ground reference was a full grid, in order to determine if errors in assumed direction of the view vector could also model these. In fact the model works on the older data about as well as that in the current experiment with a correlation with the observed data of 0.89. In this case the best fit corresponds to an overestimate of pitch of only 4 deg and yaw of 9 deg and, as is the case with the present data, there is a tendency for the model to underestimate the errors near the oblique axes. The error in yaw compares roughly with independent direct measurements of perceived yaw of fully grided ground references. These data predict about a 6 deg overestimation of yaw for approximately comparable viewing conditions and image content (Ellis and Grunwald, 1988).

Cause of error in estimated viewing direction

Assumptions regarding the physical properties of objects or elements of pictures are necessary for picture perception because of the inherent ambiguity of the monocular pictorial information. These assumptions provide the means for quantitative interpretation of that information. Examples would be: that the reference lines dropped from the cube markers are parallel, equal and themselves perpendicular to the ground reference; that the marker cubes remain equal in depicted size and that the lines in the ground reference are all parallel and coplanar. Other examples would be assumptions regarding the regularities of background textures that would allow geometric interpretations of the texture gradients present in perspective projections.

It is probably correct to argue that for monocular perspective displays these shape assumptions are the principal basis for the construction of a perceived space from the provided line of sight information. The properties of this inferred virtual space are opposed, however, by the properties of the physical space of the picture surface which provide a mechanism to produce the pattern of direction errors that have been recorded. A simple test of this hypothetical distorting mechanism would be to repeat the previous experiment in a real scene, a situation where there is no binocular conflict and in which there are an abundance of cues to the correct view direction. Experiment 2 investigates this possibility.

Experiment 2

Methods

Subjects

Four nonpilots and four pilots with normal or corrected to normal visual acuity participated in the study. Five of the subjects were graduate psychology students one of whom was a pilot. Three of the nonpilots were experienced psychophysical observers.

Apparatus and stimuli

The geometric conditions in Experiment 1 were generally duplicated although electronically produced apertures and dials were replaced by actual equivalent sized objects with similar functions. The stimuli and viewing geometry used in Experiment 1 were physically reproduced in a parking lot adjacent to the Life Science Building at the Ames Research Center which was viewed from observation stations on top of this building. Two large cubes (91 × 91 × 91 cm) were constructed from 1.27 cm white polyvinyl chloride pipe. Each cube was suspended from its center 3.6 m above the ground by a single aluminum pole and allowed to tumble irregularly in the breeze. One of the two cubes marked the position of the reference cube, and the other marked the location of the target cube. An assistant, in radio contact with the experimenter, moved the target cube. During the experiment the subjects viewed the stimuli through two different sized windows in the two observation stations. Each station was 104 cm square × 61 cm deep. The angular sizes of the windows measured from a distance equal to the 61 cm depth of the station were 30 and 60 deg.

Immediately adjacent to the windows in each observation station was a circular, clear plastic angle indicator dial for collecting angular data geometrically equivalent to that used in Experiment 1. The face of the dial was normal to the subjects line-of-sight to the reference cube and had two lines on it, one fixed and one moveable. The fixed one was parallel to the vertical axis of the window. Subjects in the experiment rotated the dial to adjust the moveable line to match the angle on the face of the dial with a specified azimuth angle of the target cube. Response recording and stimuli selection were controlled by a microcomputer.

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 on average about 5 deg; its suspension allowed it to rotate irregularly. The cubes markers provided a significant stereoscopic stimulus since the binocular disparity of the target varied between 6.6′ to 9.8′ as it was positioned around the reference cue. This maximum disparity difference of 3.2′ is at least about thirty times a typical stereo threshold but within normal values of fusion area for the range of retinal eccentricities experienced by the subjects.

Subjects made the same exocentric direction judgments used in Experiment 1. Positions of the target cube were unobtrusively marked on the ground to allow an assistant to position the target cube accurately at the planned azimuth angles. The zero-degree-azimuth reference axis always pointed away from the center of the reference cube and was parallel to the white lines painted on the black asphalt parking lot to indicate parking places.

Prior to participation, each subject was shown the actual location of this axis. In conformance to earlier experiments the reference cube was not centered between nor positioned on any of the painted lines. Except for the immediate vicinity of the cubes, the parking lot was often filled with cars.

A 24 × 2 × 2 × 2 × 2 (Target Position × Window Size × Direction of Viewing × Replications × Flight Experience) mixed factorial design was used in the experiment, with repeated measures on the first four variables. The between subjects variable, flight experience, referred to membership in either the pilot or nonpilot group.

Subjects were required to make azimuth judgments of 24 equally spaced (15 deg) target positions for two directions of viewing (±22 deg) and two window sizes (30 and 60 deg FOV). Each subject proceeded through the design twice for a total of 192 judgments of target azimuth (24 Target Azimuths × 2 Window Sizes × 2 Directions of Viewing × 2 Replications). As in Experiment 1, the dependent variable was the subjects' error in judging target direction, azimuth error. Azimuth error was computed as in Experiment 1. Decision time was also recorded but will not be discussed in this paper.

Procedure

The subjects were seated so that their heads were centered in the windows. Subjects were discouraged from moving their heads toward or away from the stimulus scene but no head restraint was used in order to preserve naturalistic viewing conditions.

A method of adjustment was used in which the subjects manually adjusted the display angle indicator to match accurately the depicted horizontal angle shown by the position of the target cube, reference cube, and the reference axis. They signaled the computer to take the data by pressing a button adjacent to the dial. No premium was placed on rapid judgements but the subjects were told not to take more than about 30 seconds per judgement. Each subject was given written instructions describing the task, was shown how to manipulate the apparatus, and was then allowed up to 10 practice trials to become familiar with both the equipment and the task. No feedback was given concerning the accuracy of his judgments.

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 three hours to complete. Short rest periods (about 2–5 min) were provided when the subject was required to change direction of viewing (about every 1 or 2 blocks). For each subject, directions of viewing and the order in which blocks were presented were random. To balance possible hardware biases in data collection the data collection equipment at one station was switched with the equipment at the other station halfway through the experiment (i.e., after four subjects).

Results

The azimuth error data were analysed by analysis of 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 nor did it 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 ($F(23,138) = 3.861$, $p < 0.001$).

The nature of the statistical interaction that was observed between viewpoint and target azimuth is again clarified by circular plots in Figure 6. This figure illustrates the underlying symmetry in the error data, which is similar to that in Experiment 1 but it also shows the expected generally smaller size of the errors and the absence of the "equidistance tendency" since there is no pronounced tendency for the errors to be towards the crossing axis.

Discussion

The absence of the equidistance tendency in Experiment 2 confirms the supposition that the full set of spatial cues in a natural viewing situation would remove the bias in Experiment 1 hypothetically introduced by the binocular conflict or other related picture surface cues. In that experiment the azimuth errors were generally away from the reference axes and toward the crossing axis. In contrast to the relatively large bias in Experiment 1, the expected smaller errors in Experiment 2 are less consistent, and frequently 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 observation in Experiment 1 that the error pattern still does not exhibit position constancy.

As in Experiment 1, the observed error pattern in Experiment 2 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 4. As in Experiment 1, the results do not closely resemble these curves so new alternatives need to be considered to explain both the smaller average size of the error and the particular pattern itself.

Though fitting the view vector error model to the data from Experiment 2 yields the expected nearly correct estimate of the view vector parameters (Figure 7) the error patterns are not markedly sinusoidal and do not fit the

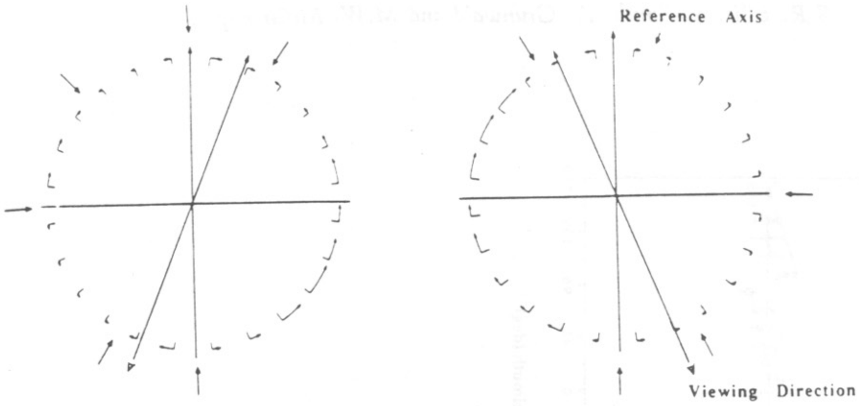
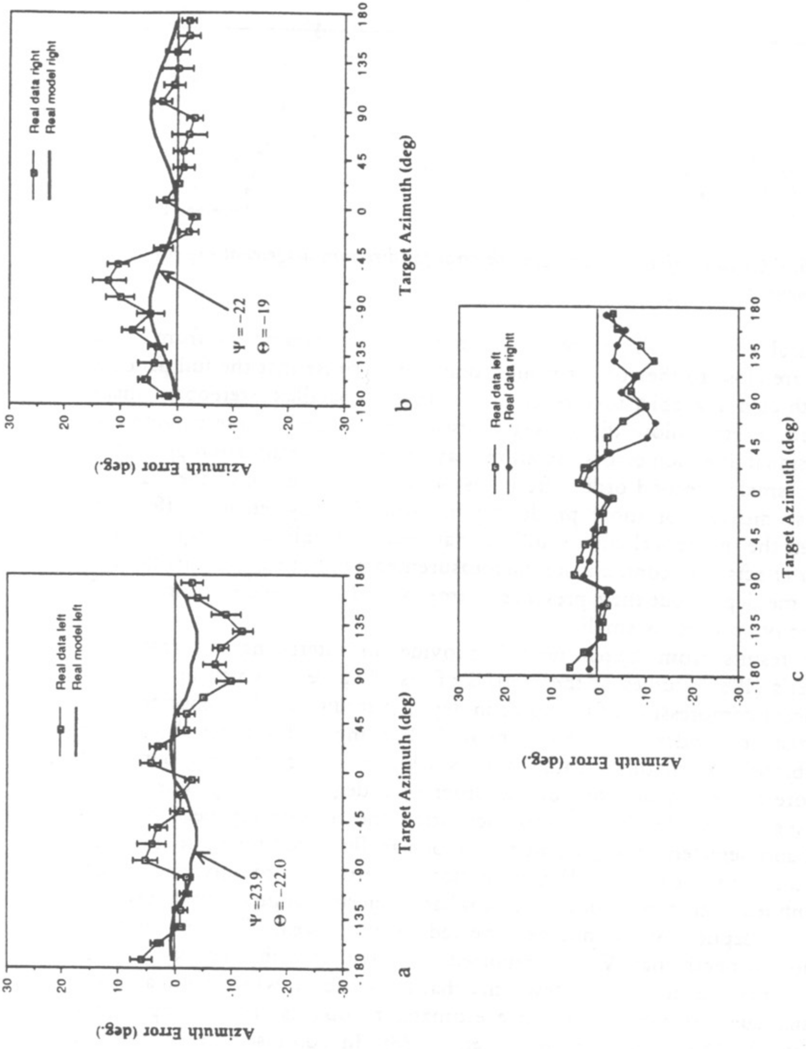


Figure 6. Circular plot of mean azimuth error for direction judgement experiment in Experiment 2.

theoretical curves well. Since the estimated view parameters from these curves are close to the correct values, one may surmise that the full panoply of depth cues available, texture gradients, motion parallax, stereopsis, shape assumption, etc., allow the viewer to estimate accurately these parameters. The residual direction errors therefore may reflect other estimation processes, perhaps smaller second order effects, associated with the use of the dial as a response measure or those producing the small but systematic differences between the theoretical curves and the data near the obliques. These kind of smaller effects may contaminate the measurements in both Experiments 1 and 2 to some degree but their presence is emphasized when the error in the 3D interpretive process is small.

The results from Experiment 2 provide an interesting contrast with Wagner's (1985) studies of the metrics of visual space in which he reported significant compression of the space in depth, a result like that reported as an "equidistance tendency" in Experiment 1. Reading of his paper shows that his subjects viewed their spatial layouts while standing on the ground and therefore at low inclination, of the order of 3 deg or less depending upon their specific eye height. At this inclination the relationship between projected and depicted angles begins to explode as illustrated for right angles by the third term in equation (1) getting close to zero. Accordingly, the corrective influence that perception of familiar projected angles could have on spatial perception would not be expected to work well. The kind of compression in depth that Wagner reported is consistent with Gilensky's (1955) earlier work, but it is also noteworthy that for scenes in which familiar shape and size cues are present, distance estimates to objects can be surprisingly accurate (IATSS, 1983; Loomis *et al.*, 1988). In contrast, the judgmental errors that we have measured in pictorial viewing situations in which the viewing direction may be misjudged may be fairly well modeled by assuming that the error in estimating the view direction causes subjects to look up


a
Real data left
Real model left
Ψ = 23.9
Θ = −22.0
Azimuth Error (deg.)
Target Azimuth (deg)
30
20
10
0
-10
-20
-30
-180
-135
-90
-45
0
45
90
135
180
b
Real data right
Real model right
Ψ = −22
Θ = −19
Azimuth Error (deg.)
Target Azimuth (deg)
c
Real data left
Real data right
Azimuth Error (deg.)
Target Azimuth (deg)

external shapes in incorrectly indexed look-up tables associating distal shapes with their proximal projections.

Summary

Two experiments have been conducted in which subjects indicate the apparent exocentric azimuth direction of a marker with respect to a reference position and direction. This judgement constitutes a precise, systematic version of the Piaget "decentering" task. The task was presented either as a perspective projection onto a binocularly viewed, flat computer display or as a geometrically equivalent physical space. Elimination of binocular conflict between picture surface cues and monocular cues to the display's virtual space, markedly reduced a judgement bias in Experiment 1 resembling a spatial compression in depth. The azimuth errors observed in Experiment 1 can be modeled by a generalization of classic slant overestimation in which the viewer is assumed to overestimate both the pitch and yaw of the viewing direction. When this model is applied to the data of Experiment 2 space, it correctly recovers the true pitch and yaw of the viewing direction thus indicating that in the physical space the subjects are able to use a much better estimate of how they view the scene to estimate exocentric direction than in perceptually degraded displays.

References

Attneave, F. and Frost, R. (1969). The determination of perceived tridimensional orientation by minimum criteria. *Perception and Psychophysics*, **6**, 391–396.

Beck, J. and Gibson, J.J. (1955). The relation of apparent shape to apparent slant in the perception of objects. *Journal of Experimental Psychology*, **50**, 125–133.

Borresen, C.R. and Lichte, W.H. (1962). Shape constancy : dependence upon stimulus familiarity. *Journal of Experimental Psychology*, **63**, 92–97.

◀*Figure 7. Cartesian plots of mean azimuth errors in Experiment 2 are plotted in ordinary cartesian form for both left (a) and right (b) viewing stations. Error bars are ±1 standard error, $N = 8$. The heavy traces are theoretical functions derived from the assumption that the subjects misjudge the view vector. The estimated viewing parameters for the theoretical trace are at the tail of the arrow. These functions have been fitted to the data as described in the discussion of Experiment 2 and show the fitting procedure almost exactly estimates the actual viewing parameters ($\Psi = \pm 22$, $\theta = -22$). Nevertheless, though the fits are significant the theoretical traces do not correlate well with the data (left station: $r = 0.408$, $p < 0.05$, $df = 22$; right station: $r = -0.394$, $p < 0.05$, $df = 22$). (c) illustrates the symmetry between the pattern of error from the left and right view stations by reflecting the data from the right station and replotting it with that from the left station.*

Chouet, B.A. and Young, L.R. (1974). Tracking with head position using an electro-optical monitor. *IEEE Transactions in Systems, Man and Cybernetics. SMC-4*, 192–204.

Ellis, S.R. and Grunwald, A.J. (1988). Egocentric direction estimates in pictorial displays. *Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics*. Peking, China. IEEE CAT 88CH2556-9, 8–11.

Ellis, S.R., Smith, S. and McGreevy, M.W. (1987). Distortions of perceived visual directions out of pictures. *Perception and Psychophysics*, **42**, 535–544.

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

Gilensky, A.S. (1955). The effect of attitude upon perception of size. *American Journal of Psychology*, **68**, 173–192.

Gogel, W.C. (1977). The metric of visual space. In W.F. (Ed.) *Stability and Constancy in Visual Perception*. New York: Wiley.

Gogel, W.C. and Da Silva, J.A. (1987). Familiar size and the theory of off-sized perceptions. *Perception and Psychophysics*, **41**, 318–428.

Goldstein, E.B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology*, **13**, 256–266.

Gregory, R.L. (1966). *Eye and Brain: The Psychology of Seeing*. New York: World University Library.

Grunwald, A.J. and Ellis, S.R. (1986). Spatial orientation by familiarity cues. In Patrick, J. and Duncan, K.D. (Eds) *Training, Human Decision Making and Control*, 257–279. Amsterdam: North-Holland.

Grunwald, A.J., Ellis, S.R. and Smith, S. (1988). Spatial orientation in pictorial displays. *IEEE Transactions on Systems, Man and Cybernetics*, **18**, 425–436.

Hochberg, J. (1987). Machines should not see as people do, but must know how people see. *Computer Vision, Graphics, and Image Processing*, **37**, 221–237.

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

IATSS 527 project team (1983). Characteristics of visual reaction to dynamic environment. *International Association of Traffic and Safety Sciences*, **9**, 162–172.

Ittelson, W.H. (1951). Size as a cue to distance. *American Journal of Psychology*, **64**, 54–67.

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

Larsen, J. and Stark, L. (1988). Difficulties in calibration of eye movements. *Proceedings of the 1988 IEEE International Conference on Systems, Man and Cybernetics*, CH2556-9, 297–298.

Levine, M. (1984). The placement and misplacement of you-are-here map. *Environment and Behavior*, **16**, 139–157.

Loomis, J., Da Silva, J.A, Marque, S.L. and Fukusima, S.S. (1989). Visual matching and visually directed walking: a comparison involving exocentric intervals. *29th Annual Meeting of the Psychonomics Society*, November 10–12, 1988 Chicago.

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

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

McGreevy, M.W., Ratzlaff, C. and Ellis, S.R. (1985). Virtual space and two-dimensional effects in perspective displays. *Proceedings of the 21st Annual Conference on Manual Control*. June, 1985.

Piaget, J. (1932). *The Origins of Intelligence in Children*. London: Kegan Paul.

Pirenne, M. (1970). *Optics, Painting, and Photography*. Cambridge: Cambridge University Press.

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

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

Tharp, G. (1989). "Modeling human perception errors using perspective displays and the response to error display." Master's Report, Department of Mechanical Engineering, University of California, Berkeley.

Tharp, G. and Ellis, S.R. (1990). Training effects on exocentric direction judgement errors made with perspective displays. *NASA TM 102792*, NASA Ames Research Center, Moffett Field CA.

Thouless, R.H. (1931). Phenomenal regression to the real object. *British Journal of Psychology*, **21**, 338–59.

Wagner, M. (1985). The metric of visual space. *Perception and Psychophysics*, **33**, 483–495.

Wallach, H. (1985). Perceiving a stable environment. *Scientific American*, **252**, 118–124.

Appendix

Figure 8 shows the detailed geometry of a view of a right angle AOB in the *XZ* plane with a viewing elevation of Θ. The right angle is rotated in the plane through an azimuth of Ψ. Each of its arms of unit length are projected onto the axes of the *XZ*

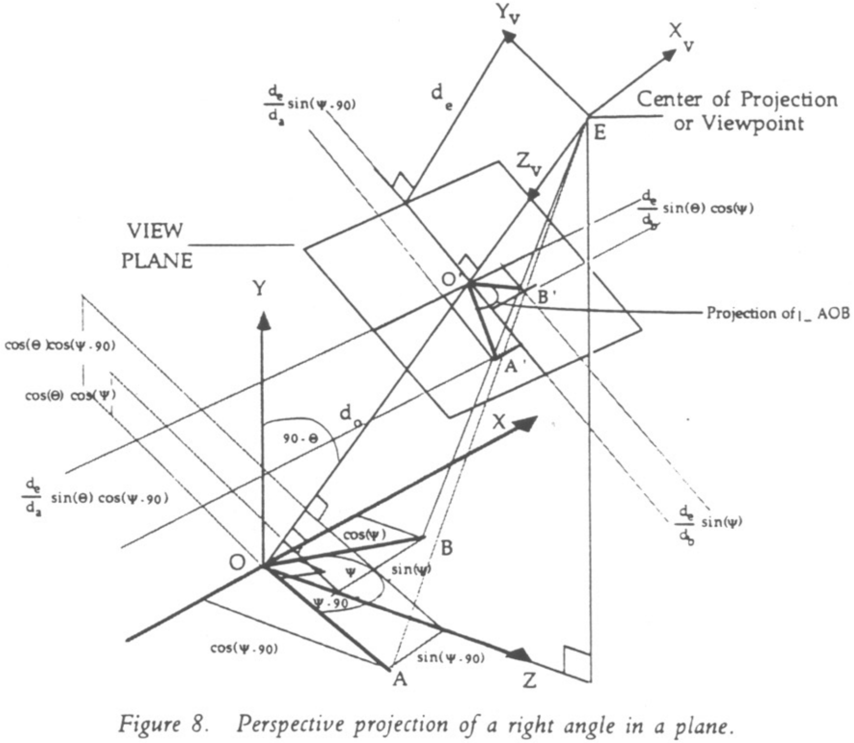


Figure 8. Perspective projection of a right angle in a plane.

plane and these components in turn projected onto the view plane d_o from 0. Thus, the projected vectors corresponding to vectors **A** and **B** in are $\mathbf{A}_v$ and $\mathbf{B}_v$ in the view plane X_vY_v.

$$\mathbf{A}_v = \left[\frac{d_e}{d_a}\sin(\Psi - 90), \frac{d_e}{d_a}\sin(\Theta)\ \cos(\Psi - 90)\right]$$
$$\mathbf{B}_v = \left[\frac{d_e}{d_b}\sin(\Psi), \frac{d_e}{d_b}\sin(\Theta)\ \cos(\Psi)\right] \tag{2a}$$

The projected angle A′O′B′ in the view plane can be directly calculated from:

$$\angle A'O'B' = \cos^{-1}\left[\frac{\mathbf{A}_v \cdot \mathbf{B}_v}{|\mathbf{A}_v||\mathbf{B}_v|}\right] \tag{2b}$$

The distances drop out and trigonometric identities allow reduction of the right side of equation (2a) to:

$$\sin(2\Psi)\tfrac{1}{2}(\sin^2\Theta - 1)((\cos^2\Psi + \sin^2\Theta\sin^2\Psi)(\sin^2\Psi + \sin^2\Theta\cos^2\Psi))^{-1/2} \tag{2c}$$