Vision Res. Vol. 26, No. 6, pp. 973–990, 1986
Printed in Great Britain.

0042-6989/86 $3.00 + 0.00

TRADEOFFS BETWEEN STEREOPSIS AND PROXIMITY LUMINANCE COVARIANCE AS DETERMINANTS OF PERCEIVED 3D STRUCTURE

BARBARA ANNE DOSHER,[1]* GEORGE SPERLING[2] and STEPHEN A. WURST[2]
[1]Department of Psychology, Columbia University, New York, NY 10025, U.S.A. and
[2]Department of Psychology, New York University, New York, NY 10003, U.S.A.

(*Received* 19 *March* 1985; *in revised form* 6 *December* 1985)

Abstract—A 2D polar projection of a 3D wire cube (Necker cube) in clockwise rotation can be perceived either veridically as a clockwise-rotating cube (*rigid* percept) or as a counterclockwise-rotating rubbery, truncated pyramid (*nonrigid* percept). The 3D percept is influenced by various cues: linear perspective, stereo disparity, and proximity-luminance covariance (PLC, the intensification of edges in proportion to their proximity to the observer). Perspective, by itself or in combination, is a very weak cue whereas PLC is a powerful cue [Schwartz and Sperling (1983) *Bull. Psychon. Soc.* **21**, 456–458]. Here we determined psychometric functions for perceptual resolution in static displays and dynamic rotating displays (with and without a static preview) as determined by stereopsis and PLC in isolation and with both cues jointly, possibly in conflict. Stereopsis was the dominant cue in static displays and in most dynamic displays. When a static display preceded a dynamic display, it strongly influenced the subsequent dynamic percept. Perceptual resolution in all conditions was accurately described by a winner-take-all model in which the strength of evidence for each percept from different cues is simply *algebraically* added.

Structure from motion   Kinetic depth effect   Stereopsis   Luminance-depth covariance   Human visual perception

INTRODUCTION

Structure from motion

Several cues affect how observers assign a three-dimensional (3D) interpretation to two-dimensional (2D) line drawings in motion, among them perspective, stereopsis and luminance cues. Under rotation, the percept of an object in depth (3D) occurs even when a static "snapshot" of a line drawing—analogous to a shadow projection of a wire figure—appears flat, or is perceived incorrectly in depth. The emergence of a depth percept from a 2D projection of 3D *rotational* motion is called the kinetic depth effect (Wallach and O'Connell, 1953). In wire figures, the perception of 3D orientation and direction of rotation is geometrically ambiguous, the 2D projection being consistent with an infinite number of rigid and nonrigid (rubbery) 3D structures. Usually, however, only two perceptual states are observed; hence, the kinetic depth effect is a *bi-stable* phenomenon involving a *coupled* assignment of motion (direction of rotation) and form.

Early work on the kinetic depth effect (Attneave, 1971; Gregory, 1970; Hochberg, 1964; Wallach and O'Connell, 1953) utilized parallel projections of simple objects, such as Necker cubes [see Fig. 1(a)]. Parallel projections provide no linear perspective cues, so that the two alternative 3D form percepts are stereoisomers—structurally equivalent except for a mirror reflection. In the case of Necker cubes, the two states correspond to seeing a particular face as either forward or rear, but the structure in either case is a cube. The motion percept is coupled with the form percepts. When a particular face is seen forward, the object appears to rotate in one direction, while if that face is seen rear, the object appears to rotate in the opposite direction.

It is typical for computer displays to depict objects *with perspective*, that is, with polar projection simulated from fairly close viewing ranges. With significant cues from perspective, the two alternative form percepts correspond,

*The research was performed at NYU while B. A. Dosher was a visiting scholar there. Supported by AFOSR Grant 80-0279.

under rotation, to a rigid object (i.e. a cube rotating, say, clockwise) and a nonrigid object (i.e. a rubbery, distorting, truncated pyramid rotating, say, counter-clockwise). Previous experimenters (Braunstein, 1962, 1977; Braunstein *et al.*, 1982; Braunstein and Stern, 1980; Green, 1961; Lappin *et al.*, 1980; Petersik, 1979, 1980; Petersik and Pantle, 1979) emphasized the fact that subjects have a bias to perceive the rigid, rather than the nonrigid alternative in these circumstances. Theorists such as Ullman (1979a) modeled the perceptual process as a structure-from-motion algorithm that always selects the rigid 3D interpretation that is consistent with 2D evidence. However, the alternative, nonrigid percept actually occurs quite frequently. Under long-term continuous viewing, the two percepts will alternate. When a real, rigid, wire object such as a cube is twirled in the hand and viewed *monocularly*, the nonrigid illusion can persist despite the presence of conflicting tactile cues: the hand moves one way and the object appears to move the opposite way.

In this paper, we parametrically investigate two cues involved in the determination of 3D percepts from 2D projections and the tradeoffs between them. These cues are: proximity luminance covariance (PLC) and stereo disparity (binocular disparity). We propose and assess a weighting model for integrating information from these two cues; the model applies whether the cues are in conflict or in cooperation.

Proximity Luminance Covariance (PLC)

Schwartz and Sperling (1983) investigated the strength of two cues in controlling the depth percept: the amount of linear perspective and a novel manipulation that co-varied line-intensity with position in depth, termed proximity-luminance covariance (PLC). The amount of linear perspective is indexed here by the ratio of the projected size of a far and near face of a Necker cube in face-forward orientation. In *positive* PLC (PLC > 0), lines that were near the observer for the rigid (cube) stimulus were bright, and those that were far were dim [see Fig. 1(b)]. In *neutral* PLC (PLC $= 0$), all lines were equal in brightness [see Fig. 1(a)]. And in *negative* PLC (PLC < 0), lines toward the observer for a rigid percept were dim, and those in the rear were bright [see Fig. 1(c)].

Over all polar (perspective) projections, with neutral PLC, Schwartz and Sperling (1983, p. 457) found the probability of perceiving the rigid form in an extensive experiment was only 61.2%. The theoretically neglected "rubbery" percept occurred almost 40% of the time! Furthermore, the amount of linear perspective, which ranged from parallel (size ratio of 1.0) to extreme polar (size ratio of 2.33), had little effect on the form/rotation-direction percept. Linear perspective accounted for about 1% (0.5–1.4%) of the variance in direction of rotation judgments. In contrast, the intensity manipulation, PLC, had massive effects, accounting for almost 50% of the variance (37.6–54.0%) in perceived structure. (Percents-of-variance assume a linear-weighting model; ranges refer to the range across individual subjects.) Since only the luminance cue had any substantial effect, Schwartz and Sperling (1983) could not test any combination rules.

Schwartz and Sperling investigated only a relatively all-or-none version of a proximity-luminance covariance. A very high brightness value was assigned to the depth positions closest to the observer, and a very dim brightness value to the furthest depth position (or vice versa), with mid-depth lines somewhere in between. Because of the large brightness range, this was an extreme version of the PLC manipulation, and had a correspondingly large effect: the rigid percept occurred on more than 90% of PLC > 0 trials and on less than 10% of PLC < 0 trials.

In displays of several objects at different depths, if each object was allowed the maximum-to-minimum brightness variation, this would create contradictory information between objects. Therefore, it would be more natural to assign a more limited range of brightness covariation to each object, and to allow the degree of variation to depend, in part, on the depth of the object as a whole.

To discover the optimal arrangement of luminance cues to disambiguate depth in multi-object displays, requires first, the formulation of rational schemes for the assignment of luminance to objects as a function of their 3D depth, and second, the empirical investigation of the actual effectiveness of these schemes in influencing the selection of rigid vs "illusory" percepts.

Our method of assigning intensities to lines as a function of their 3D depth derives from the properties of infinitesimally thin self-luminous lines. Let the lines of a wire 3D object be self-luminous. The inverse square for points

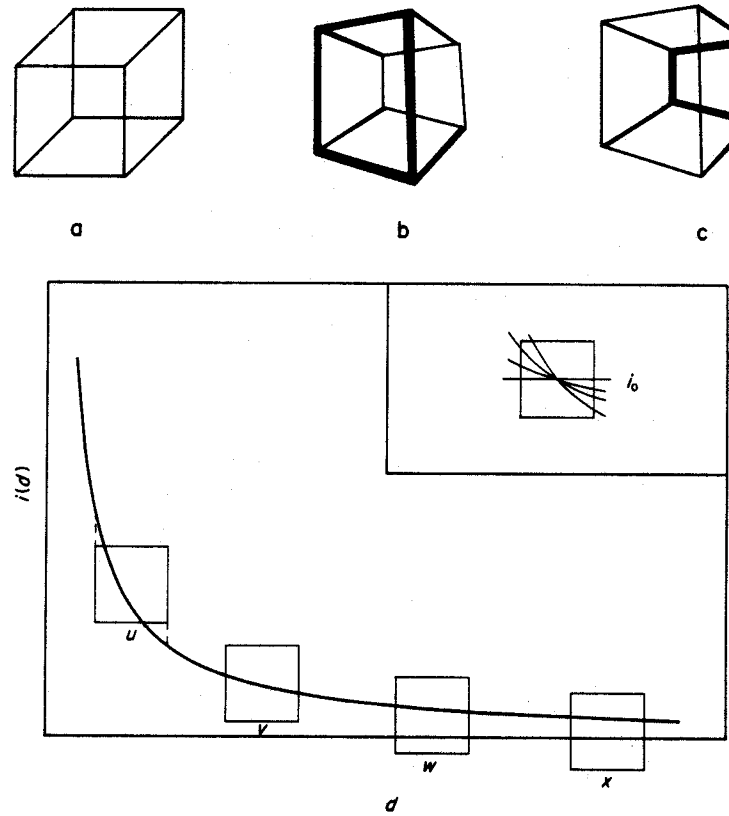

Fig. 1. Illustrations of proximity luminance covariance (PLC). (a) Necker cube, parallel projection, with no luminance cues to depth (equal-intensity of all lines). (b) Cube, polar projection, with positive PLC (front face intensified). Heavy lines represent more intense lines in a CRT display. (c) Cube, polar projection, with negative PLC (rear face intensified). (d) Diagram indicating the basis of computing the exact form of PLC as the inverse linear law for distances of self luminous lines. Abcissa: d, distance from observer to cube; ordinate: $i(d)$, intensity of intensified lines as function of d. The range of d and $i(d)$ is shown for a cube of constant size at four distances (u, v, w, x) from the observer. Displayed values of $i(d)$ are scaled multiplicitively so that all images have the same luminance at the distance corresponding to the center of the cube. Similarly, the sizes of the displayed images are scaled so that a cross section through the center of the cube would have the same size in all conditions.

translates into an inverse *linear* law for lines. This means that, at the observer's viewpoint, the observed intensity of a self-luminous line decreases in direct inverse proportion to the distance of the line. For a wire object of a given physical size, the observed intensity difference between front and rear lines depends on the physical distance between the object and the viewpoint; a self-luminous object that is distant from the observer varies little in the observed intensity ratio of front-to-rear edges, while the same object when near to the observer produces large front-to-rear observed intensity ratios. Choosing a 3D distance for a self-luminous wire object determines line intensities in its 2D projection, and hence the amount of PLC, in a physically meaningful way. For ease of comparison of objects at various distances, intensity is scaled so that a line at the 3D midpoint of the object apears equally intense in all cases [see Fig. 1(d)].

Given a lawful method of varying the amount of PLC, we can now empirically measure observers' psychometric functions for PLC. To avoid confounding the retinal size of the object or its overall brightness with the amount of PLC, we construct stimuli as follows. First, the target front-to-rear intensity ratios for 2D images of 3D Necker cubes at various effective distances are derived and cubes constructed. These intensity ratios determine the levels of PLC for determining the psychometric function. Cubes (all of the same retinal size) are constructed to embody the target front/rear intensity ratios. For each cube, its intensities are multiplied by a positive constant so that the intensity level assigned to the center point of each of the test figures is equal. This scheme amounts to exaggerating or contracting the depth-axis with respect to the assignment of luminance, and dissociates retinal size and neutral luminance (which are kept constant) from the manipulation of PLC.

Stereo disparity

In the real world, observers receive a 2D projection to *each eye*, and the disparity between the two views can be used to infer depth information. The degree of stereo disparity depends on the distance from the object to the viewer. The view to the two eyes is nearly identical for sufficiently distal objects, and the disparity between the eyes' views is large when objects are near [see Fig. 2(a)]. A 2D display system that mimics the binocular (stereo) disparity information in real viewing requires the generation of two separate images and a system to project each image to each eye independently. In our experiment, a computer creates a Left-eye version of a Necker cube on the left half of a display screen, and the Right-eye version on the right. The two versions are then displayed to the separate eyes by a system of mirrors and baffles [see Fig. 2(b)]. By systematically varying the amount by which the Left-eye and Right-eye views differ, and measuring the observers' responses, their psychometric functions for stereo disparity can be determined.

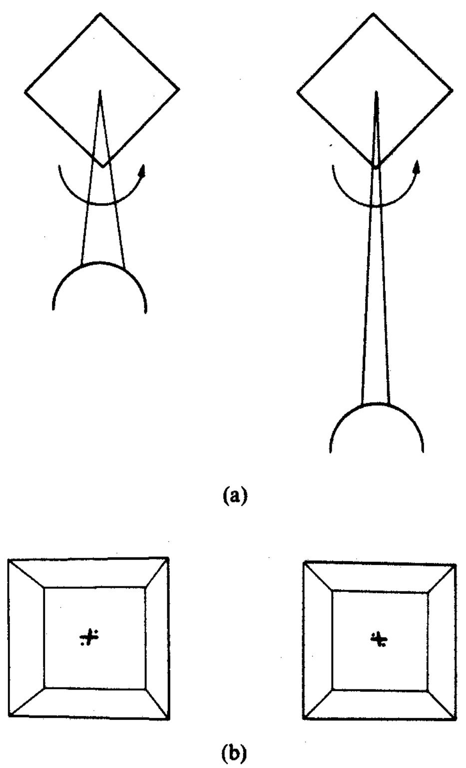

Fig. 2. Rotational stereo disparity. (a) Schematic illustration of how the angular difference between left and right eye views of objects naturally depends on vergence angle α which, for an object rotating about a vertical axis, is equivalent to α degrees of rotational stereo disparity. Under normal binocular viewing, α is inversely proportional to viewing distance. (b) The mirror system which directed the two (different) angular views, appearing on the left and right areas of a CRT screen, to the left and right eyes, respectively. The images shown illustrate the extreme conditions of rotational stereo disparity used in the experiments (-7 deg). Reversing the L and R images yields $+7$ deg of rotational disparity.

Amount of Stereo disparity is manipulated by varying the *view angle* of the object between the Left image and the Right image. This is equivalent to varying the interpupillary distance and scaling it in terms of object size. It would be better, of course, if the amount of stereo disparity were describable in terms of retinal disparity. In these displays, however, retinal disparity varies over a continuous range, different retinal disparities being produced by different parts of the object, the disparity for any particular part changing during rotation. Thus viewing angle is the better descriptor, but we also indicate the maximum and minimum retinal disparities that occur during a full object rotation.

A useful property of object rotation around a vertical axis in generating Left-and Right-eye images, is that these images differ only in temporal phase. For a clockwise rotating object (viewed from the top), Right-eye images precede Left-eye images, and for counterclockwise rotation, the reverse is true. Although angular distance between views translates into a time difference, varying rotation speed shows that only angle, and not time, is critical for the displays under consideration, justifying our description of the stereo cue as *stereo rotation disparity*, the angle between the two eyes' images. Given a particular objective direction of object rotation (e.g. clockwise), the display may be perceived veridically as a rigid object rotating with the front moving to the left, or non-veridically as a *non*rigid object rotating with front moving to the right.

In the real world of self-luminous 3D objects, the amount of stereo disparity would co-vary both with retinal size and with luminance. In our displays, retinal size is held constant, and the amounts of stereo rotation disparity and of PLC are varied independently.

Cues in opposition

Independent psychometric functions for PLC and for stereo disparity can be generated by manipulating each of these cues independently, while holding the other at the neutral, or zero, level. In the experiments reported here, these independent psychometric functions are tested in the context of a large crossed design which yields a family of joint psychometric functions. Almost half of the conditions in the crossed design involve presenting the two cues in opposition. For example, presenting a stimulus where stereo rotation disparity favors a non-rigid percept and proximity-luminance covariance favors the rigid percept. The focus of this study is how the visual system integrates multiple cues, especially conflicting cues. Each cue, varied independently, alters the *dominance* or probability of observing the rigid versus non-rigid percept, but does not, to a first approximation, alter the nature of the two percepts. The hypothesis tested here is that the effects of two independent cues are first weighted and then combined additively to determine the relative dominance of the two alternative percepts.

METHODS

Subjects

Four subjects with normal, or corrected-to-normal vision participated. Two subjects, B.D. and S.W. were authors, and had foreknowledge of the design, two subjects, S.W. (same as above) and D.P., were graduate students, and one subject, L.K. was an undergraduate and paid for her services. However, since the main factors of the design are easily perceivable, all subjects had general knowledge of the experimental design by the end of a single practice block.

After the formal experiments, three of the subjects were tested for their ability to perceive various kinds of simple and complex stereo displays (e.g. dynamic random-dot stereograms, Julesz, 1971) and they appeared normal in all respects. A fourth subject (L.K.), who showed very strong effects of stereo manipulations in the experiments, was unavailable for subsequent testing, but is believed to have been normal.

Apparatus

Each stimulus was displayed as illuminated vectors on a CRT (Hewlett-Packard 1310A High Speed Graphic display scope) with a P4 fast, white phosphor. The CRT was controlled by a PDP-11/34 and a special-purpose visual interface (Kropfl, 1975), and controlled by special purpose software (Melchner *et al.*, 1985), which allowed a resolution of 1024 × 1024 addressable pixels, each of which could be programmed for one of 256 intensity levels. Pixel spacing is approximately 0.193 mm vertically and horizontally. Vectors or lines were painted to the nearest pixel on this 1024 × 1024 grid.

Stimuli

Each stimulus consisted of a Left and Right part, corresponding to the Left- and Right-eye views of a 3D skeletal cube displayed in 2D with linear perspective plus a fixation cross (nonius), parts of which were visible to each eye and parts to both. The fixation nonius consisted of a plus sign displayed to both eyes plus diagonal radii, negative diagonals in the left eye and positive diagonals in the right eye. When the eyes were properly verged, the nonius appeared as a cross with all diagonals: it served as a vergence and binocularity check on every trial.

The skeletal cubes were projected with a constant degree of linear perspective—a projection ratio (length of the near face/length of the far face) of 1.57/1. This corresponds to viewing a cube from a distance of 2.25 W (W = cube width = 200 pixel units under parallel projection) from the center of the cube, and produces a noticeable degree of perspective foreshortening. Since this perspective information is consistent with the cube stimulus, it should favor a rigid percept, although Schwartz and Sperling (1983) found that perspective was a weak cue.

Each eye's view (half stimulus) was displayed within a 300 × 300 pixel (58 × 58 mm) region of the CRT screen, with a center-to-center offset between the Left and Right portions of 500 pixels (97 mm). The Left- and Right-eye views were projected to the corresponding eyes via an optical system [see Fig. 2(b)], resulting in a central stereo image, and flanking monocular images at a viewing distance of 82 cm.*

*Subjects report largely being unaware of the configuration or direction of rotation of the flanking images during these experiments. When specifically asked to attend to these flanking images, subjects report directions of rotation that are relatively decoupled both from each other and from the stereo image, unless the PLC condition has an equal and large effect on all. Interestingly, the monocular images appear neither more nor less solid than binocular images.

The beginning position was always a cube with face forward. Each stimulus was displayed in rotation around an imaginary vertical axis. Motion was simulated by a series of picture frames corresponding to discrete positions in the rotation of a 3D cube. A sequence of 72 different still frames produced a full 360 deg, each frame corresponding to a 5 deg rotation. Each distinct frame was painted on a CRT screen 7 times at 10 msec intervals, corresponding to a frame-duration of 70 msec, or 5.04 sec per full rotation. Under these display conditions, the motion appeared smooth and continuous, approximating continuous motion.

The stereo rotation disparity conditions were produced by lagging (or leading) the left view relative to the right view by a constant number of degrees throughout the rotation. The PLC manipulation was an approximation to the ideal continuous manipulation of intensity along each line of the stimulus. Vertical lines (with a constant depth coordinate z) were assigned a single intensity value; oblique lines were also assigned a uniform brightness which was the intensity appropriate to the average depth coordinate for the two endpoints. Schwartz and Sperling (1983) and Schwartz (1983) report that this uniform-line approximation to a fully continuous PLC, while occasionally discriminable from fully continuous PLC in a still frame, is indiscriminable from fully continuous PLC under rotation, and has essentially the same effect in determining percept dominance.

The PLC manipulations are defined in terms of luminous intensity. Since luminance is generally nonlinear with the selected CRT intensity level (0 to 255), an empirical linearization table based on luminance measurements of CRT test patches was used to select appropriate luminous intensity levels. The neutral luminous intensity level (level for points at the depth level of the axis of rotation, or for the entire stimulus in a neutral PLC condition) was 3.41×10^{-5} cd/cm of line on the CRT screen (Sperling, 1970). The maximum intensity was 59.1×10^{-5} cd/cm; the minimum was 1.15×10^{-5} cd/cm; the viewing distance was 82 cm.

Procedure

Subjects sat in a room that was dark except for the projection of the stimulus on the CRT. They were seated immediately behind the system of mirrors for stereo vergence and viewed the display screen without head restraint. The mirrors were adjusted for each subject to produce optimal vergence by centering a vertical line seen only by the right eye on a horizontal line seen only by the left eye. This display also served for an occasional vergence check. Vergence and binocularity were insured at the beginning of each trial by fusion of the fixation nonius as described above. Subjects were instructed not to initiate the rotation of the stimulus unless the fixation nonius was fused and intact.

For sessions consisting of a still frame (plus fixation) preview, subjects first responded concerning the nature of the still view of the stimulus. The response was either "cube" or "truncated pyramid", and subjects were asked to pick the dominant percept even when the figure gave little stereo information. Subjects pressed a key with the right hand to indicate "cube" and with the left to indicate "pyramid". For sessions where the still frame consisted only of the fixation nonius, subjects simply pressed a key when they had verified stereo fusion.

The stimulus began to rotate approximately 1 sec after the initial response (either for stereo vergence or for still frame). Subjects were instructed to continue fixating the nonius located at the center of the cube, and to report only the first direction of rotation they saw, either "front face moving to the right" or "front face moving to the left". Subjects pressed a key with the right hand to indicate "right" and with the left hand to indicate "left". (Which of these responses corresponded to a rigid percept depended on the objective direction of the rigid figure, a balanced factor.)

Each session lasted 45–70 min depending on scheduling constraints. Each trial took somewhere between 3 and 5 sec, depending on the self-paced speed of the subject.

Design

The two main experiments yield three families of psychometric functions that derive from all combinations of stereo rotation disparity (SRD) and proximity-luminance covariance (PLC). The two main experiments differ only in their presentation history. Experiment 1 begins each trial with a fixation nonius for 0.5 sec, followed by a still frame depicting the cube-stimulus plus nonius. The subject first verifies fixation and responds to indicate the still percept (cube or truncated pyramid). This initiates rotation of the display, and the subject then responds again to indicate the direction of rotation for the *first percept* from the moving stimulus. Experiment 2 begins each trial with a still frame containing only a fixation nonius (no cube stimulus). After the subject has verified fixation, a rotating cube stimulus is presented. The subject responds to indicate the first percept of the moving stimulus. These experiments lead to three sets of data, labeled: Still (S). Rotating after Still (sR), and Rotating (R), respectively. (Note that the S and sR data represent yoked responses.)

Seven levels of stereo rotation disparity (SRD) were tested: -7, -3, -1, 0, 1, 3 and 7 deg of rotational difference between the two eye views. Positive rotation differences favor the rigid 3D perceptual interpretation, while negative ones favor the nonrigid 3D interpretation. The absolute maximum pixel differences (at any point of the rotation) between corresponding points in Left- and Right-eye images are 0 pixels at 0 deg, 4 pixels at 1 deg, 12 pixels at 3 deg and 24 pixels at 7 deg. At the viewing distance of 82 cm, these values produce approximately 0, 3, 9, and 19 min of retinal stereo disparity. (Note that these are the extreme values of retinal stereo disparity; stereo disparity differs across the figure and during the rotation.)

Seven levels of proximity-luminance covariance (PLC) are tested: log F/R of -2.68, -0.68, -0.14, 0, 0.14, 0.68, and 2.68. F/R stands for the Forward-to-Rear luminance ratio, where the values for Forward reflect the maximum luminous intensity (e.g. when a corner is forward), and the Rear reflects the minimum luminous intensity. These values correspond to F/R ratios of 1/474, 1/4.82, 1/1.39, 1/1, 1.39/1, 4.82/1, and 474/1. Negative PLC variations are computed by applying the corresponding positive conditions as a function of $-z$, where z refers to the depth dimension, with zero at the center of the cube (also the center of rotation).

In the basic crossed design, there are 7 stereo $\times$ 7 PLC = 49 conditions. Each of these 49 crossed conditions appeared in a random order in an intermixed (mixed-list) fashion. The sample size per cell is 20 per subject; the two experiments yielded 1960 experimental trials and 2940 responses per subject. Ten of the 20 trials in each cell were presented with the rigid cube in objective rightward rotation, and 10 presented leftward rotation. The direction of objective rotation did not interact with any other variables, so it is collapsed as a factor and not considered further.

Collateral experiments intermixed only stereo disparity conditions for neutral PLC, and only PLC conditions for neutral stereo disparity. In these experiments, 9 levels of stereo disparity (-7, -5, -3, -1, 0, 1, 3, 5, and 7 rotational degrees) and 9 levels of PLC (-2.68, -1.56, -0.68, -0.14, 0, 0.14, 0.68, 1.56, and 2.68 log F/R luminance ratio) were tested. Trials were randomized among the 9 stimuli and two objective rotation directions, and the sample size per cell (collapsing over rotation direction) was 48.

RESULTS

Three families of psychometric functions

Figure 3 contains observed psychometric functions for each of the four subjects and for the three display conditions: still view (S), rotating after still (sR), and rotating (no still preview) (R). (Recall that the S and sR responses represent consecutive judgements on the same trial.) Each data point is based on 20 observations: each panel contains 49 data points; the 12 panels represent 11,760 observations. Figure 3 graphs the percentage of cube percepts (out of a possible 20) versus the level of stereo rotational disparity, with the proximity-luminance covariance (PLC) as the curve parameter. To further illustrate intersubject differences in the relative importance of stereo and PLC, two data panels are regraphed in Fig. 4 to show the percentage of cube percepts versus PLC, with stereo rotational disparity as the curve parameter.

The general form of the psychometric functions in Figs 3 and 4 is consistent with data from two collateral experiments in which only stereo rotational disparity or PLC were varied with the other cue at its neutral level. Figure 5 shows psychometric functions for each of the four subjects for PLC alone and for stereo alone at the neutral value of the alternate variable. Judgments on a still view (S) and subsequent rotating judgments (sR) are shown. Each data point in Figure 5 is based on 48 observations.

In addition to the data shown in Figs 3, 4, and 5, ρ correlations (Hays, 1981, pp. 555–558) were computed between responses in Still and in Rotating After Still. It is empirically impossible to determine correlations between categorical responses when the probability of either response is near 0 or 1. However, conditions exhibiting midrange probabilities of a cube percept, showed large, statistically significant ρ correlations for all subjects.

A qualitative examination of the psychometric functions in Figs 3 and 4, indicates that

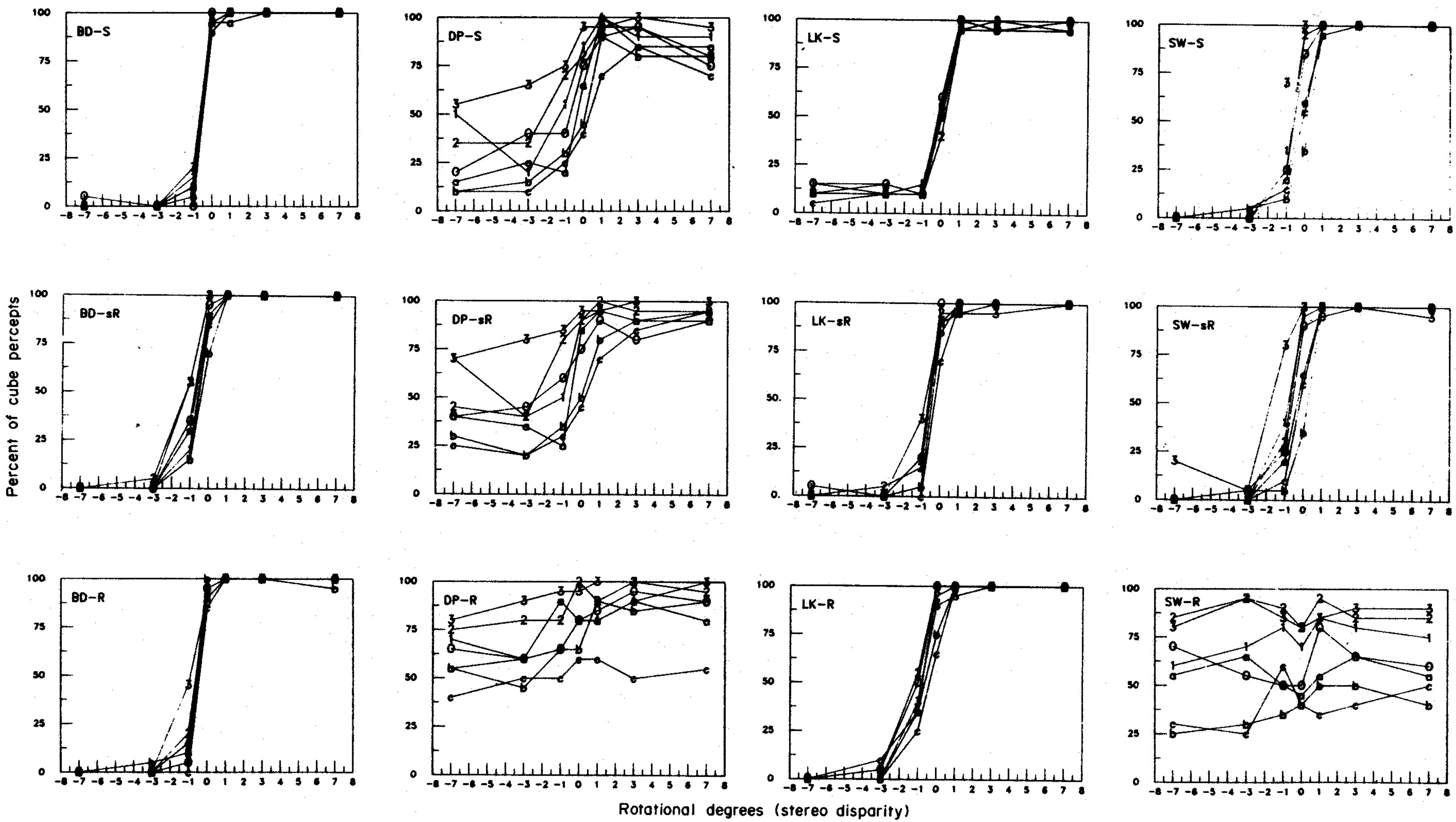
BD-S
DP-S
LK-S
SW-S
BD-sR
DP-sR
LK-sR
SW-sR
BD-R
DP-R
LK-R
SW-R
Percent of cube percepts
Rotational degrees (stereo disparity)

Fig. 3

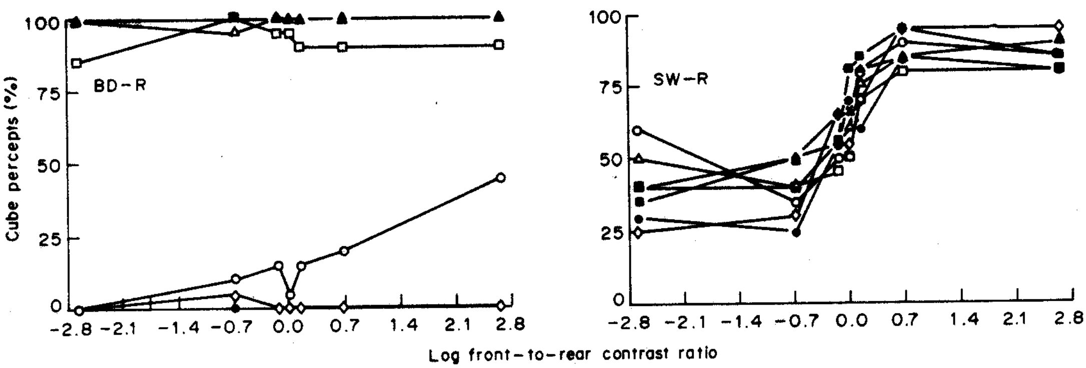

Fig. 4. Two data panels of Fig. 3 graphed with PLC as the independent variable, and STEREO as the curve parameter. Data for "rotating, no preview" (R); subjects B.D. and S.W. The open squares indicate neutral stereo disparity. Positive (favoring rigid) rotational disparities are represented by the solid squares (+1 deg), solid triangles (+3 deg), and open triangles (+7 deg). Negative (favoring nonrigid) rotational disparities are represented by the open circles (−1 deg), open diamonds (−3 deg), and filled circles (−7 deg.

stereo rotational disparity is a strong to overwhelming cue in determining the form-percept (rigid cube vs nonrigid truncated pyramid) for all subjects in the Still condition, and consequently in the Rotating After Still condition, since these responses were highly correlated. In the Rotating condition, stereo disparity is a strong cue only for two of four subjects.

The four subjects differ substantially from one another and across presentation conditions with respect to the degree of importance placed on the proximity-luminance cue. At one extreme, subjects B.D. and L.K. always places relatively low weight on PLC. At the other extreme, subject S.W. is entirely controlled by PLC when a rotating stimulus is seen with no still preview (S.W.-R, Figs 3 and 4) but S.W.'s perceptions are overwhelmingly controlled by stereo information when the rotating figure is preceded by a still preview (S.W.-sR, Fig. 3). The difference between the sR and R responses illustrates a radical difference in the perception of the dynamic rotating form that is entirely dependent upon prior exposure to a static image.

Context also plays a role as may be best seen by comparing Fig. 5 (right column) to Fig. 4. The sR condition of Fig. 5 is physically identical to the sR condition of Fig. 4 with zero stereo disparity. Only the context of other trials is different—there are many trials with nonzero stereo in Fig. 4, none in Fig. 5. (The effect of PLC is greater for all subjects in the procedure of Fig. 5; only two subjects are shown in Fig. 4.) In the context of only zero-disparity trials, the effect of PLC is quite comparable to that observed by Schwartz and Sperling (1983).

It is impossible to assess the complex internal relations in the data by eye. The next section presents and tests a quantitative model of cue combination. The model, when fitted to the data, allows the estimation of weighting factors for the two cues, which can be compared across conditions and subjects.

A linear combination model

The model consists of two parts: first, a computation of the relative strength of the cube percept (vs the nonrigid percepts) as a function of experimental parameters, and second, a conversion of this strength into a probability of perceiving the cube on a particular trial. The reason for proceeding in two steps is that a simple additive description of the cue effects is possible in the strength domain but not in the

Fig. 3. *Facing page*. Observed psychometric functions for four subjects and three viewing conditions. Observed probability of a cube percept (inferred from direction of rotation judgments for conditions R and sR) is graphed against Rotational Stereo Disparity as the independent variable. Luminance Covariance condition is the curve parameter, with the neutral condition labeled 0, increasing positive PLC conditions labeled 1, 2, and 3, and increasing negative PLC conditions labeled *a*, *b*, and *c*. Each panel is labeled with the subject's initials, and a code for the viewing condition: S for still preview, sR for rotating view following still preview, and R for rotating view with no preview.

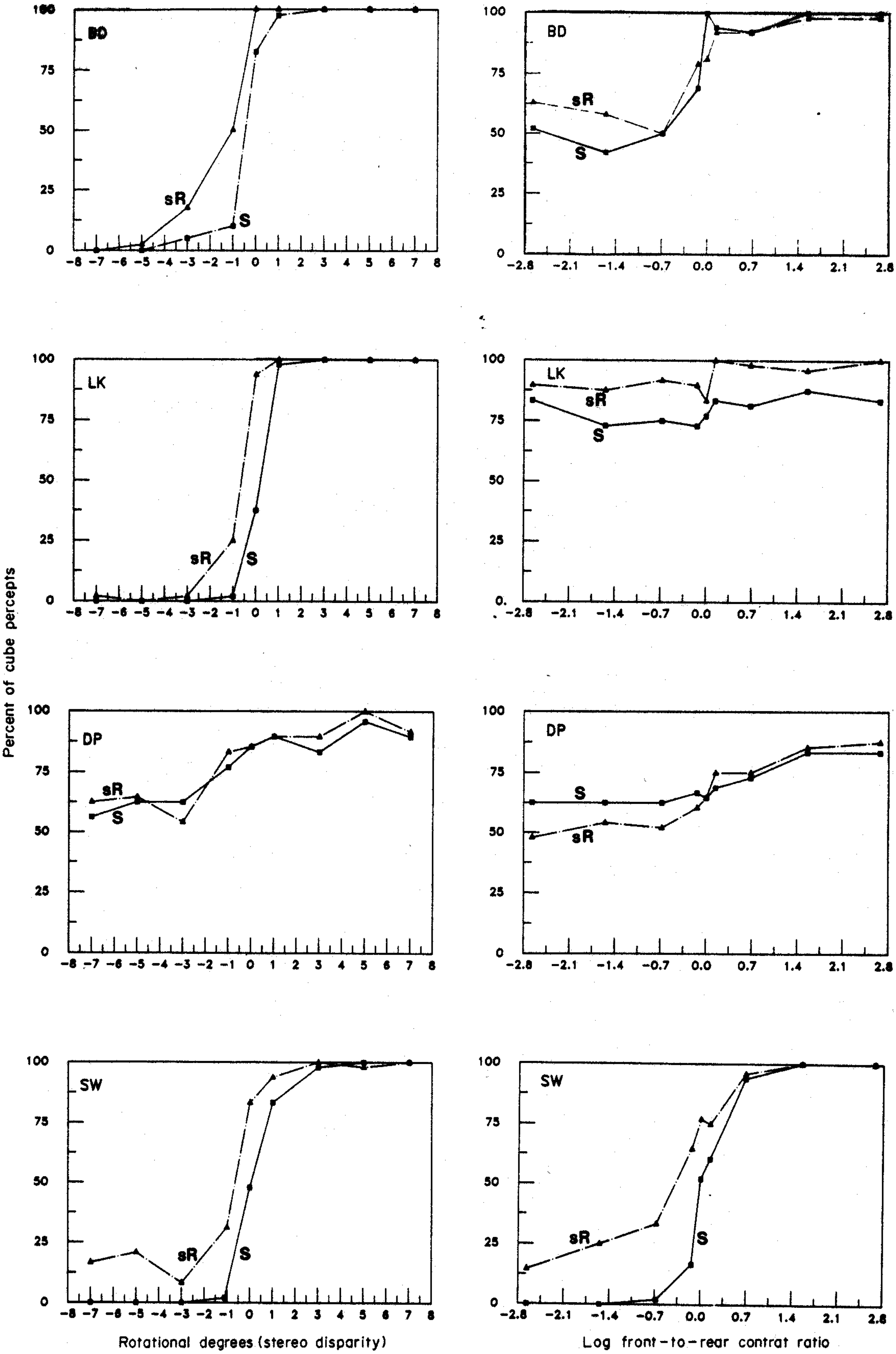

Fig. 5. Psychometric functions for PLC and STEREO cues obtained in isolated experimental blocks. The alternative cue took on its neutral value. Individual panels are labeled with the subject's initials, with the curve parameter corresponding to the S (still) and yoked sR (rotating after still) judgments in these experiments. Ordinate: observed probability of a rigid percept; abscissa: cue level.

probability domain. The strength computation is given by equation (1)

$$S(i,j,k,n) = W(k)\cdot S_{\mathrm{STEREO}}(i) + V(k)\cdot S_{\mathrm{PLC}}(j) + B(k) + \varepsilon(n). \quad (1)$$

In Equation (1), $S(i,j,k,n)$ represents the *strength* of the cube percept (relative to the nonrigid percept) as a function of the display conditions (indexed by i,j,k) and the trial number, indexed by n. Specifically, i, $i \in (-7, -3, -1, 0, 1, 3, 7)$, is an index representing the condition of STEREO rotational disparity: j, $j \in (-2.68, -0.68, -0.14, 0, 0.14, 0.68, 2.68)$, is an index of the PLC condition: k, $k \in$ (S, sR, R), is an index on the display condition.

$S_{\mathrm{STEREO}}(i)$ is a measure of the strength contributed to a cube percept by Stereo condition i. $S_{\mathrm{STEREO}}(i)$ is constrained to be zero (0) for the neutral stereo condition, $i = 0$, and the $S_{\mathrm{STEREO}}(i)$'s are range-scaled so that the max $[S_{\mathrm{STEREO}}(i)$, for all $i] - \min[S_{\mathrm{STEREO}}(i)$, for all $i] = 2$. Range scaling constrains the maximum and minimum values of the STEREO strength component to be approximately $+1$ and -1, respectively. $S_{\mathrm{PLC}}(j)$'s are similarly constrained strength values representing the contribution of PLC component to the combined strength.

The $W(k)$'s are condition-dependent weighting factors for the Stereo cue, while the $V(k)$'s are condition-dependent weighting factors for the PLC cue. The reason for range-scaling the STEREO and PLC strengths is that the weighting parameters then capture all the information about the relative importance of the two sets of cues.

The $B(k)$'s are bias factors that depend on viewing condition, k. In a condition k, when STEREO disparity and PLC are zero, any tendency of the subject to perceive stimuli as rigid (or nonrigid) is represented by $B(k)$. This single parameter (in each viewing condition) incorporates the net tendency towards rigid perception attributable to polar perspective and all other possible sources of bias inherent in the viewing condition itself.

Finally, $\varepsilon(n)$ is an additive error (noise) term: ε is assumed to be distributed as a standard normal density function with mean $\mu = 0$ and standard deviation $\sigma = 1$. Subjects are assumed to respond "cube" or "rigid" whenever the composite strength value $S(i,j,k,n) > 0$. Since identical noise is assumed for all conditions, the distributional assumptions are equivalent to a Thurstone Case V (Thurstone, 1947) model with, however, many fewer free parameters.

The core of the strength model embodies assumptions similar to, but more specific than, an additive conjoint measurement model (Krantz *et al.*, 1971; Falmagne, 1976; Falmagne *et al.*, 1979). Because the much stronger Thurstone Case V models works extremely well, the conjoint measurement model was not developed for this case.

The functioning of the additive strength-combination model is illustrated for a single subject in schematic form in Fig. 6. The top two portions of this figure show the range-scaled strength values for each of the seven levels of STEREO and of PLC. The central portion Fig. 6 illustrates the multiplication of each of the range-scaled strengths by the corresponding weighting factor, resulting in component strength values in units of z-scores relative to a unit normal distribution. The STEREO and PLC component strengths add with a bias factor, and noise to yield composite strength. For each viewing condition shown in Fig. 3 (Still, Rotating after Still, and Rotating), this produces 49 (7 STEREO $\times$ 7 PLC) unit-variance normal distributions, appropriately shifted relative to the zero-point, or criterion value, along a composite strength axis. The area of the normal distribution to the right of zero, yields an estimated probability of "rigid" object perception for each of the 49 cue combinations in each viewing condition.

Since there are three viewing conditions with 49 observed points per condition in the main experiments, there are 147 observed data points per subject. The model allows 19 free parameters (5 range-scaled Stereo strengths, 5 range-scaled PLC strengths,* 6 weights, and 3 biases). An iterative hill-climbing algorithm was used to estimate the 19 parameters of this model for each of the four subjects independently.

The additive strength model was used to generate a predicted probability $\hat{p}(i,j,k)$ of a rigid (cube) response for each of the 147 independent data points shown in Fig. 3. Since 20 trials were used to estimate each observed

*There are actually seven strength values each for Stereo and PLC. However, the strengths for neutral Stereo and neutral PLC are defined as zero, eliminating two free parameters, and the strength of one of each of the Stereo and PLC conditions is derivable from the others by virtue of the requirement for range scaling, eliminating two more free parameters. For example, since the Stereo strengths and PLC strengths were constrained so that the max $-$ min $= 2$, knowing the minimum determines the maximum, or vice versa.

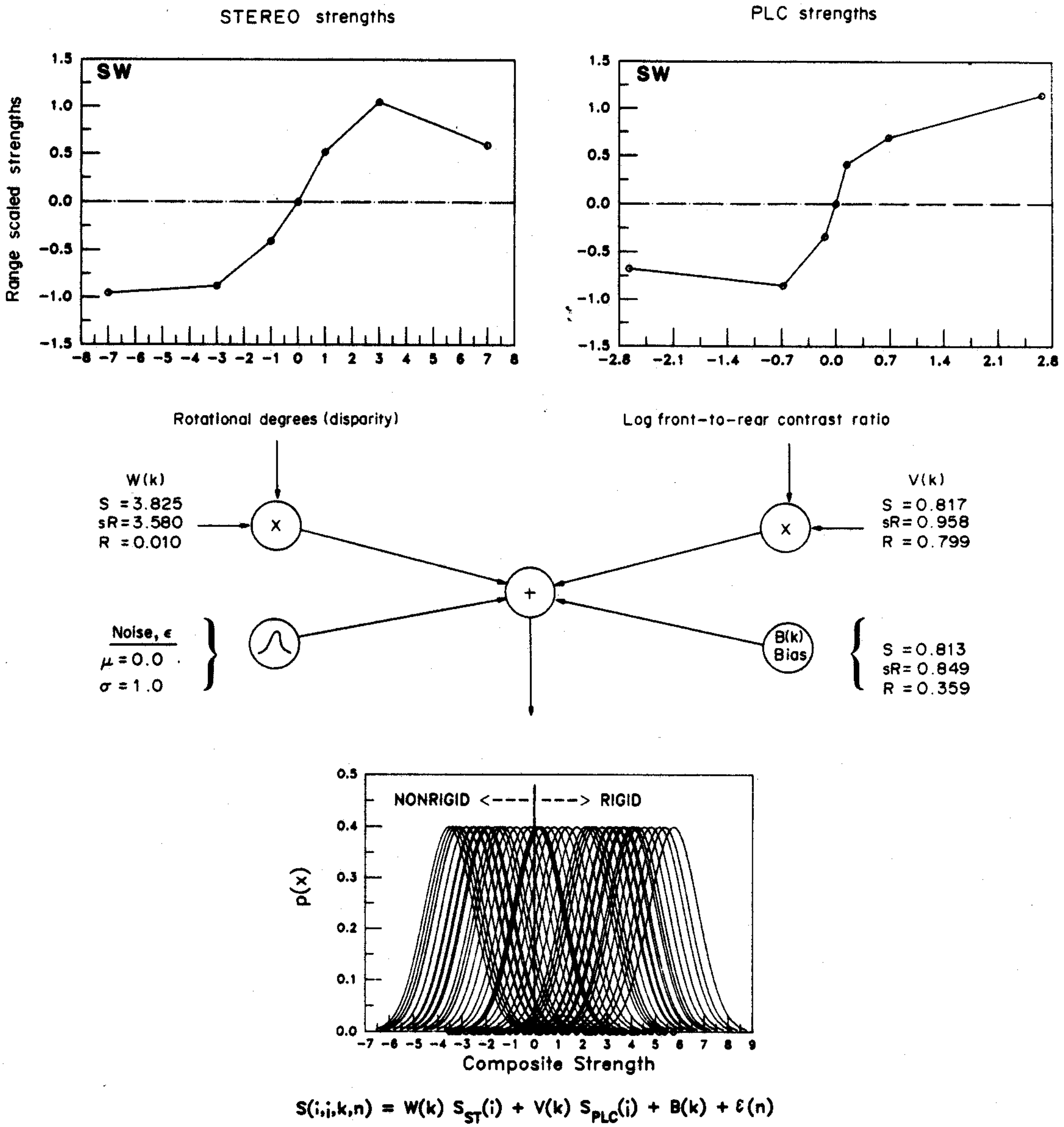

Fig. 6. Schematic illustration of the operation of the additive model of cue combination for one subject (S.W.) and one viewing condition. The top panels show S.W.'s range-scaled estimated PLC and Stereo strengths. The middle section illustrates the weighting and combination rules, and the bottom section shows the resulting family of density functions representing the composite strength (favoring a rigid percept) corresponding to the data for the Still preview condition.

probability, $p(i,j,k)$, the obtainable values are quantized (0.00, 0.05, . . . , 1.00) corresponding to the actually observed number $r(i,j,k)$ of (0, 1, . . . , 20) "cube" responses. In order to account for this quantization error, the log of the binomial likelihood of observing r "cube" responses given n trials ($n = 20$) and given a predicted probability $\hat{p}$ was computed for each observation [e.g. binom($r,n,\hat{p}$)], and the sum of the log-likelihoods was maximized using an iterative hill-climbing algorithm (Reed, 1976) which converges in a manner similar to STEPIT (Chandler, 1965). (In order to maintain estimates of probabilities corresponding to observed frequencies of 0 or 20 within reasonable bounds, *weighted* strengths on individual factors were constrained to be between -4.0 and 4.0. This constraint affected only the stereo weighting estimates of subject B.D. Estimates were also run with constraints between -5.0 and

Table 1. Weights and bias estimates for main crossed-factor experiment

	Parameter								
		Stereo			PLC			Bias	
Subject	W(S)	W(sR)	W(R)	V(S)	V(sR)	V(R)	B(S)	B(sR)	B(R)
S.W.	3.825	3.580	0.010	0.817	0.958	0.799	0.813	0.849	0.359
D.P.	0.984	0.904	0.422	0.759	0.772	0.772	0.488	0.785	0.838
L.K.	1.813	3.215	3.378	0.028	0.303	0.570	0.297	1.105	1.807
B.D.	3.991	3.927	4.110	0.309	0.618	0.262	1.365	1.700	1.432

Note: S, sR, and R are abbreviations for Still preview, Rotating after Still preview and Rotating (no preview) conditions. Stereo weight estimates for B.D. were partially constrained during model fitting and may be underestimates.

+5.0, with essentially identical results except, of course, for the few points which are estimated as equal to the constraints.)

Estimated goodness-of-fit, strengths and weights

Goodness of fit. Because of the large number of cells in which the expected frequency is below 5 or above 15 of 20. χ^2 is not an appropriate test of the quality of the fit of the model to the data. In order to give some quantitative measure of goodness of fit, the root mean squared error (r.m.s.$_{obs}$) between observed (quantized) and expected (predicted) probabilities (non-quantized) was computed. For comparison, the r.m.s.$_{binom}$ that would be expected given binomial variability around the estimated $\hat{p}s$ was computed, i.e. r.m.s.$_{binom}$ represents the expected binomial error of the perfect-fitting model. The values are as follows: For S.W., r.m.s.$_{obs}$ = 0.064 (1.277 observations out of 20) and r.m.s.$_{binom}$ = 0.071 (1.419); for D.P., r.m.s.$_{obs}$ = 0.071 (1.420) and r.m.s.$_{binom}$ = 0.087 (1.742); for L.K. r.m.s.$_{obs}$ = 0.062 (1.232) and r.m.s.$_{binom}$ = 0.056 (1.126); and for BD, r.m.s.$_{obs}$ = 0.063 (1.259) and r.m.s.$_{binom}$ = 0.042 (0.832). The fact that r.m.s.$_{obs}$ compares favorably to r.m.s.$_{binom}$ suggests not only that the model gives an extremely good account of the data, but that fewer parameters might have sufficed to describe all of the data in the three viewing conditions. For example, it might have been possible to assume a generalized form of psychometric function on range-scaled strengths, or to assume that $W(S) = W(sR)$. However, by using the parameters described above, we can be fairly sure that the model captures all of the main features of the data.

Strengths. The estimated range-scaled strengths for the seven levels of stereo rotation disparity and of proximity-luminance covariance are shown for each of the four subjects in Fig. 7. These are essentially psychometric functions with scaled strengths as the dependent variable. All of the subjects show typical scaled psychometric functions for rotational stereo disparity. The scaled psychometric functions for PLC are somewhat abnormal for subjects L.K. and B.D. This may be because, for both the subjects, the effect of PLC on the performance is extremely small, and hence the data do not adequately constrain the estimates. The weights and bias estimates for each subject and condition are listed in Table 1.

Weights. Examination of the weights and biases in Table 1 indicates that subjects S.W., D.P., and L.K. all show a greater relative emphasis on the proximity-luminance covariance (PLC) cue in determining the structure-from-motion percept when the rotating stimulus is not preceded by a still view. Subjects S.W. and D.P. have lower stereo weights in no-preview than preview conditions [W(R) $\ll$ W(sR)], and about equal weights on PLC. Subject L.K. has about equal weights on stereo, but a higher PLC weight for no-preview [V(R) > V(sR)]. Subject B.D. shows a constant, fairly low level of weighting on PLC, with a small reversal of the general pattern.* One way to estimate the *relative* dominance of stereo and PLC in the three response conditions is to take the ratio of the stereo and PLC weights for these conditions. The $\log_{10}$ of these three ratios [W(S)/V(S), W(sR)/V(sR), and W(R)/V(R)] are shown in Fig. 8. The greater relative importance of the stereo cue is indicated by log ratios above 0, and the trend toward lower stereo weighting in the no-preview condition is apparent as well.

Biases. All biases are positive (e.g. toward a cube percept). The "bias" toward the cube percept at least partially reflects the objective information toward the rigid form given by linear perspective. Is there any additional bias

*However, the stereo weights for B.D. may be slight underestimates due to the constraint on estimated strength parameters.

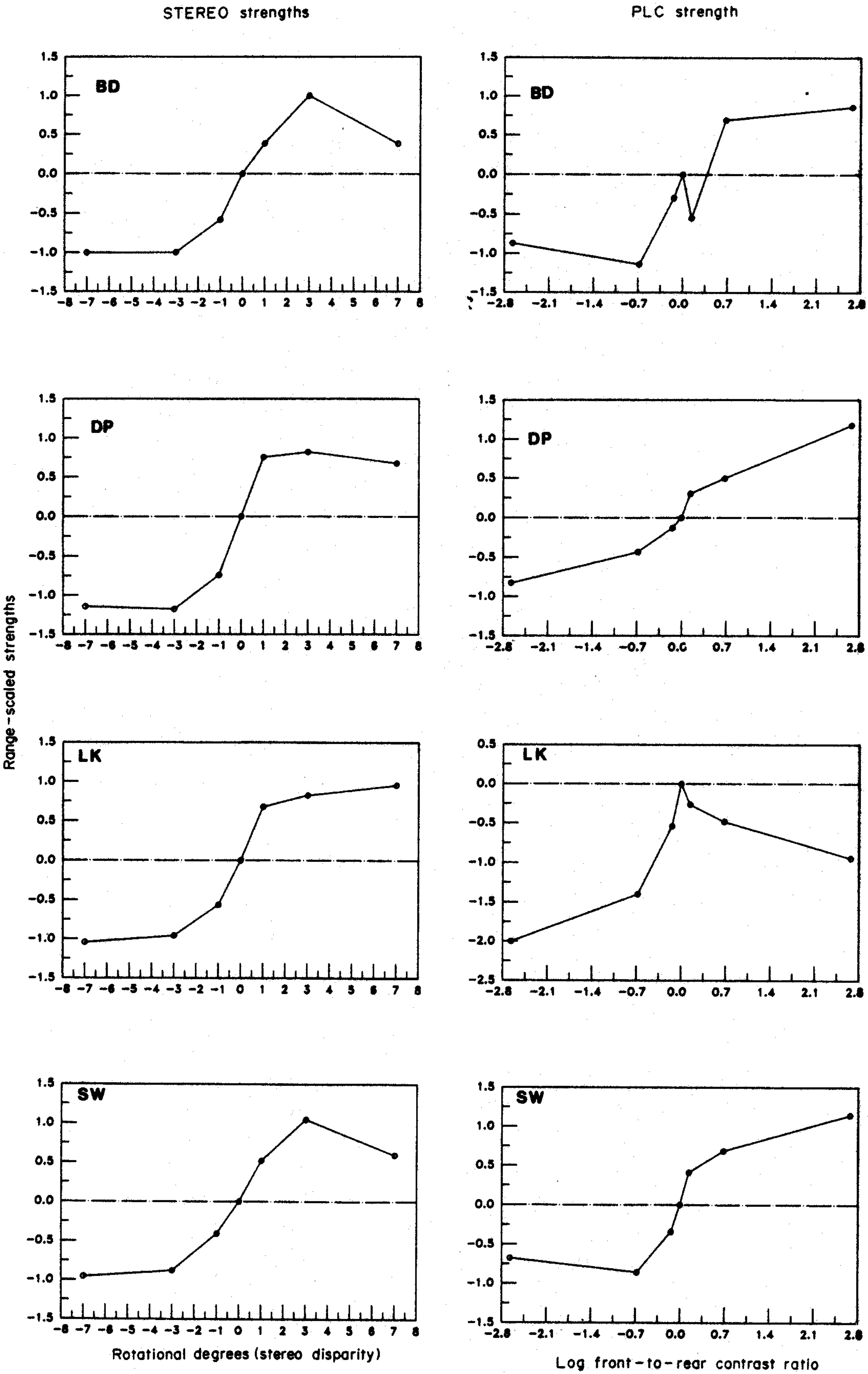

Fig. 7. Range-scaled model strengths for PLC and STEREO as a function of cue level, estimated for each subject.

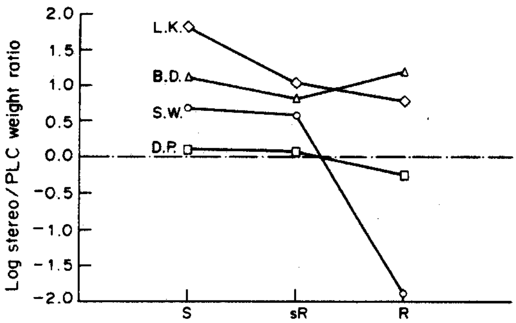

Fig. 8. The relative importance of proximity luminance covariance and rotational stereo in determining the depth percept for each subject and viewing condition. Ordinate: $\log_{10}$ of the ratio of the estimated model weights for component strengths of STEREO over PLC; abscissa: viewing condition (see caption, Fig. 3); curve parameter: subject.

towards "cube" that can be attributed to motion? Take the bias for "cube" in the Still preview, $B(\mathrm{S})$, as a baseline that includes the perspective information and any preference for the more-familiar cube and against the less-familiar truncated pyramid. The additional bias attributable purely to rotational mode is $B(\mathrm{R}) - B(\mathrm{S})$. This additional bias is utterly insignificant for two of the four subjects. Rotation is the only mode where "rigidity" (Ullman, 1979a, 1979b, 1983) is relevant, but subjects are inconsistent in the strength of their tendency to perceive rotating cubes as rigid cubes rather than as nonrigid truncated pyramids—some have it strongly, others do not.

Path dependence. Responses in Still Preview judgments (S) and in the consecutive (*yoked*) Rotating After Still judgments (sR) are not independent, that is, there is path dependence in the perception of structure in the moving displays. The assertion of path dependence (perceptual perseveration from the still preview to the rotating view) may be assessed in two ways. First, a correlation between responses in consecutive conditions: Still Preview and Rotating After Still (S, sR), indicates path dependence—even when the mean response in two consecutive conditions is similar. Second, path dependence is assessed directly by comparing the effect of different initial conditions (static preview, no preview) on the mean response in the Rotating condition.

First, consider the correlation between the yoked responses to S and to sR stimuli. In the midrange of conditions where it is feasible to observe correlations (i.e. when the probability of a "cube" response is not nearly zero or one), correlations between S and sR responses are consistently large and positive for all subjects, indicating the influence of the prior static perception upon the subsequent perception of movement. Second, the response to the moving stimulus (R, sR) depends enormously on the preview, the prior history. Consider rotation following no preview versus rotation following static preview: [R vs sR]. Witness the performance of subject S.W. His stereo weights are $W(\mathrm{sR}) = 3.580$ and $W(\mathrm{R}) \approx 0$, indicating overwhelming dominance of stereo disparity in viewing a rotating stimulus preceded by a still preview, and complete irrelevance of stereo disparity in the same rotating stimulus viewed without still preview. At the other extreme, subject B.D. always shows such strong dominance of stereo over PLC that path dependence is demonstrable only by correlation.

DISCUSSION: RELATION TO DYNAMIC COMPUTATIONAL MODELS

The linear combination model used in this paper to capture the nature of cue interaction in the derivation of structure from motion is essentially a descriptive, not process model. However, this descriptive model is consistent with some processes and may help to place strong constraints on any computational model of depth processing.

The algebraic addition of the effects of cues is consistent with, and suggestive of, a dynamic energy map description (Sperling, 1970). Energy maps associate with every perceptual state (stable and unstable states) a real-valued energy. Various sources of evidence (e.g. internal and external constraints) add to yield a net energy function, often called an energy surface. A marble "rolls" on the surface; its position represents the perceptual state. Relative minima of the surface (energy wells) represent stable perceptual states. The force that moves the marble from one state to another is represented by the slope of the energy surface. Since the slope is the derivative of the energy, algebraic addition of energies is equivalent to the algebraic addition of forces. Initially, at the instant the stimuli become visible, the marble is poised on the saddle hump between the minima that represent the competing perceptual states. At a saddle point, the forces are balanced. The STEREO and PLC cues add their maps to the configuration-determined map; at the saddle they add a slope that represents a force tending to move the marble towards one state or the other. The various forces (cue strengths) add algebraically without respect to their origin to determine the marble's initial movement and thus its eventual perceptual state.

On one hand, certain kinds of cooperation–competition interactions which are represented computationally as symmetric relaxation labeling algorithms have been shown to have equivalent energy map formulations (Hummel and Zucker, 1983; Mohammed and Zucker, 1983). On the other hand, the energy map is the natural companion to the linear weighting model in which all the evidence in favor and against perceptual alternatives, independently of origin of the evidence, is simply algebraically summed to determine the outcome (Sperling *et al.*, 1983). Energy map models offer a natural explanation for perceptual path dependency (Sperling, 1970), and cooperation–competition has been proposed as a mechanism for a large set of closely related phenomena involving depth computations (Dev, 1975; Julesz, 1971; Marr and Poggio, 1976; Nelson, 1975; Sperling, 1970, 1981). Thus, there are strong *a priori* reasons to expect that a linear weighting model might hold not only for the resolution of perceptual ambiguity but for a far wider range of perceptual, cognitive, and motivational processes that are describable by energy maps (Sperling *et al.*, 1983).

Two computational models have been proposed by Ullman (1979b, 1983) for the computation of structure from motion. The first of these solves the structure from motion problem by assuming that the visual system uses linear algebra to compute the rigid structure that is compatible with the stimulus array. When there is error in the representation of the stimulus, or when there is no rigid alternative, no computation is proposed. There is overwhelming empirical evidence here and in Schwartz and Sperling (1983) that nonrigid perceptions occur almost as often as rigid ones, and that other, unrelated factors such as proximity luminance covariance can exert almost total control over perceptual state. Ullman (1983) himself recognized these and other problems with his original model, and proposed a second model in which a Euclidean constraint (rigidity) is applied from moment to moment, and the perceptual system drifts towards a solution satisfying this constraint. Possibly, such a computation may be a component of structure from motion. Possibly not. But such a computation would represent only one of the sources of evidence for structure from motion, and the computation of structure from motion has to be large enough to encompass all of the cues.

SUMMARY AND CONCLUSIONS

(1) Two cues to a structure-from-motion percept in the kinetic depth effect are stereo disparity (STEREO) and proximity luminance covariance (PLC). Variation in the strength of each cue, by itself, results in a normal psychometric function.

(2) STEREO and PLC—in collaboration or in opposition—both contribute to the determination of a form and motion percept. The combined effectiveness of the two cues is very well accounted for by *a weighted linear model* in which the contributions of STEREO and of PLC sum algebraically.

(3) STEREO usually is stronger than PLC; the ratio of the weighting parameters for STEREO to PLC is, with one exception, greater than or equal to 1 (see Fig. 8).

(4) In the main experiments, (Fig. 3), PLC has a smaller effect than the enormous effect reported by Schwartz and Sperling (1983). This difference in the significance of PLC as a cue appears to depend on the overall experimental context, since PLC had a much larger impact in small designs in which subjects viewed only positive, neutral and negative PLC exemplars, all with neutral STEREO (Fig. 5). The exact nature of these context effects is of some interest, and is being investigated further. But the fact that contextual factors can determine the overal weight assigned to any particular cue is not inconsistent with the analysis we offer here.

(5) The relative weights of cues are affected by viewing condition, such as viewing static or dynamic displays. However, because of the persistence of perception from one viewing condition to the next, there are large path-dependent effects (hysteresis) so that the apparent effectiveness of a cue in dynamic viewing depends on its effectiveness in a static preview. The cue weights reflect a large degree of situation-dependent variation in the degree of importance of cues to 3D structure and motion.

(6) There was greater bias towards rigid cube percepts under rotation than in static displays for two of the four subjects, and no greater tendency for "rigid" percepts under rotation than in static displays for the other two subjects. Similarly, Schwartz and Sperling (1983) found

only a small effect of *differential* levels of linear perspective on the fraction of rigid percepts, and only a small tendency to perceive objectively rigid cubes as rigid more frequently than as nonrigid. A strong "cube" bias in perception of a still preview is not a rigidity bias because both the cube and the truncated pyramid alternatives are rigid in a static view; a cube bias in static displays would be interpreted as due to familiarity. The tendency to perceive rotating Necker cubes as rigid *because* of the motion cues is not large and is certainly not universal.

(7) The linear weighting model is compatible with energy map models of perceptual states and consequently with cooperation-competition networks and with relaxation labeling algorithms for depth computation that have been shown to be equivalent to energy maps. In particular, energy map models naturally exhibit path dependence and they naturally provide for the algebraic addition of evidence from many sources without regard to their origin. Thus, while a linear weighting model may be regarded as merely an efficient and powerful descriptive tool, it is suggested by, and suggests, the form of neural and computational theories for the basic perceptual processes in structure from motion.

REFERENCES

Attneave F. (1971) Multistability in perception. *Scient. Am.* **225**, 62–71.

Braunstein M. L. (1962) Depth perception in rotating dot patterns: Effects of numerosity and perspective. *J. exp. Psychol.* **64**, 415–420.

Braunstein M. L. (1977) Perceived direction of rotation of simulated three-dimensional patterns. *Percept. Psychophys.* **21**, 553–557.

Braunstein M. L., Anderson G. J. and Reifer D. M. (1982) The use of occlusion to resolve ambiguity in parallel projection. *Percept. Psychophys.* **31**, 261–267.

Braunstein M. L. and Stern K. R. (1980) Static and dynamic factors in the perception of rotary motion. *Percept. Psychophys.* **27**, 313–320.

Chandler J. P. (1965) *Subroutine STEPIT.* Program QCPE 66, Quantum Chemistry Program Exchange. Indiana Univ., Bloomington.

Dev P. (1975) Perception of depth surfaces in random-dot stereograms: a neural model. *Int. J. Man–Machine Studies* **7**, 511–528 (1975).

Falmagne J. C. (1976) Random conjoint measurement and loudness summation. *Psychol. Rev.* **83**, 65–79.

Falmagne J. C., Iverson G. and Marcovici S. (1979) Binaural "loudness" summation: Probabilistic theory and data. *Psychol. Rev.* **86**, 25–43.

Green B. F. (1961) Figure coherence in the kinetic depth effect. *J. exp. Psychol.* **62**, 272–282.

Gregory R. L. (1970) *The Intelligent Eye.* Weidenfeld & Nicholson, London.

Hayes W. L. (1981) *Statistics*, 3rd edn. Holt, Rinehart & Winston, New York.

Hochberg J. E. (1964) *Perception.* Prentice Hall, Englewood Cliffs, N.J.

Hummel R. A. and Zucker S. W. (1983) On the foundations of relaxation labeling processes. *IEEE Trans. Pattern Analysis Machine Intell.* **PAMI-5**, 267–287.

Julesz B. (1971) *Foundations of Cyclopean Perception.* Univ. of Chicago Press.

Krantz D. H., Luce R. D., Suppes P. and Tversky A. (1971) *Foundations of Measurement*, Vol. 1. Academic Press, New York.

Kropfl W. J. (1975) Variable raster and vector display processor. *Bell Telephone TM* 75-1223-3.

Lappin J. F., Doner J. and Kottas B. (1980) Minimal conditions for the visual detection of structure and motion in three dimensions. *Science* **209**, 717–719.

Marr D. and Poggio T. (1976) Cooperative computation of stereo disparity. *Science* **194**, 283–287.

Mohammed J. L. Hummel R. A. and Zucker S. W. (1983) A feasible direction operator for relaxation methods. *IEEE Trans. Pattern Analysis Machine Intell.* **PAMI-5**, 330–332.

Melchner M. J., Yangarber R., Sperling G. and Dosher Barbara A. (1980) REDS: A *RE*al-Time *D*isplay *S*ystem for C/Unix environments. User Guide and Manual. *Mathematical Studies in Perception and Cognition, 85-6.* New York Univ.

Nelson J. I. (1975) Globality and stereoscopic fusion in binocular vision. *J. theor. Biol.* **49**, 1–88.

Petersik J. T. (1979) Three-dimensional object constancy: Coherence of a simulated rotating sphere in noise. *Percept. Psychophys.* **25**, 328–335.

Petersik J. T. (1980) Rotation judgments and depth judgments: Separate or dependent processes. *Percept. Psychophys.* **27**, 588–590.

Petersik J. T. and Pantle A. (1979) Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Res.* **19**, 143–154.

Reed A. V. (1976) MODPAC: A modular package of programs for fitting model parameters to data and plotting fitted curves. *Behav. Res. Meth. Instrum.* **8**, 375–377.

Schwartz B. J. (1983) Nonrigid perceptions consistently elicited by rigid object stimuli in the kinetic depth effect. Doctoral dissertation, New York Univ.

Schwartz B. J. and Sperling G. (1983) *Bull. Psychon. Soc.* **21, 456–458.**

Sperling G. (1970) Binocular vision: a physical and a neural theory. *Am. J. Psychol.* **83**, 461–534.

Sperling G. (1981) Mathematical models of binocular vision. In *Mathematical Psychology and Psychophysiology* (Edited by Grossberg S.). *Soc. Industrial Appl. Math–Am. Math. Assoc.* (*SIAM-AMS*) *Proc.* Vol. 13, pp. 281–300.

Sperling G., Pavel M., Cohen Y., Landy M. and Schwartz B. (1983) Image processing in perception and cognition. In *Physical and Biological Processing of Images* (Edited by Braddick O. T. and Sleigh A. C.), Vol. 11, pp. 359–378. Springer Ser. Information Sci. Springer, Berlin.

Thurstone L. L. (1947) *Multiple-factor Analysis: Development and Expansion of the Vectors of Mind.* Chicago University Press.

Ullman S. (1979a) The interpretation of structure from motion. *Proc. R. Soc. Lond. B* **30**, 405–426.

Ullman S. (1979b) *Interpretation of Visual Motion*. MIT Press, Cambridge Mass.

Ullman S. (1983) Maximizing rigidity: The incremental recovery of 3-D structure from rigid and rubbery motion. MIT A.I. Memo No. 721.

Wallach H. and O'Connell D. N. (1953) The kinetic depth effect. *J. exp. Psychol.* **45**, 205–217.