Pergamon

0042-6989(93)E0035-6

Vision Res. Vol. 34, No. 14, pp. 1863–1874, 1994

0042-6989/94 $7.00 + 0.00

Integration of Shading and Texture Cues: Testing the Linear Model

WILLIAM CURRAN,* ALAN JOHNSTON*

Received 2 February 1993; in revised form 3 June 1993

One of the first attempts to develop a formal model of depth cue integration is to be found in Maloney and Landy's [(1989) *Proceedings of the SPIE: Visual communications and image processing, Part 2* (pp. 1154–1163)] "human depth combination rule". They advocate that the combination of depth cues by the visual system is best described by a weighted linear model. The present experiments tested whether the linear combination rule applies to the integration of texture and shading. As would be predicted by a linear combination rule, the weight assigned to the shading cue did not vary as a function of its curvature value. However, the weight assigned to the texture cue varied systematically as a function of the curvature values of both cues. Here we describe a non-linear model which provides a better fit to the data. Redescribing the stimuli in terms of depth rather than curvature reduced the goodness of fit for all models tested. These results support the hypothesis that the locus of cue integration is a curvature map, rather than a depth map. We conclude that the linear combination rule does not generalize to the integration of shading and texture, and that for these cues it is likely that integration occurs after the recovery of surface curvature.

Depth Curvature Shape from shading

INTRODUCTION

The past few years have witnessed a growing interest in the problem of how the visual system integrates information from individual depth processing modules to produce a single percept of depth. Initial work in this area has indicated that the integration of depth cues by the human visual system may be best described by a weighted linear model in which the contributions of depth cues sum algebraically (Bruno & Cutting, 1988; Dosher, Sperling & Wurst, 1986; Johnston, Cumming & Parker, 1993; Landy, Maloney & Young, 1990; Maloney & Landy, 1989; Rogers & Collett, 1989; Young, Landy & Maloney, 1993).

Dosher *et al.* (1986) investigated the integration of stereo rotation disparity and proximity luminance covariance (PLC), PLC being a technique first adopted by Schwartz and Sperling (1983) in which line intensity co-varies with depth. They reported that a simple linear model accounted for the results of their experiment. It was found that more weighting was given to the stereo cue than to the PLC cue when subjects were presented with a still preview of the stimulus. However, in the absence of a still preview, more weighting was given to the PLC cue, thus emphasizing the role of context in the assignment of weights to depth cues.

Bruno and Cutting (1988) obtained statistical support for a linear combination rule encompassing four depth cues—relative size, height in the projection plane, occlusion and motion parallax. A magnitude estimation procedure, in which subjects judged the relative distance between three square panels, was employed to assess subjects' perceived exocentric distances. Each depth cue was either present or absent in a given stimulus, resulting in 16 different combinations. They found effects of relative size, height in plane and motion parallax. More importantly, the absence of significant interactions between the depth cues provided support for the argument that a linear combination strategy was adopted by the visual system in the processing of these depth cues.

Johnston *et al.* (1993) report that, when a texture depth cue is added to stereograms portraying a range of surfaces (ellipsoids, cylinders and "roofs"), subjects' perception of depth is increased (although not necessarily made more veridical). When both cues portrayed incongruent depth information, the data were well accounted for by a weighted linear rule, with stereo being more heavily weighted than texture. As in the experiments of Dosher *et al.* (1986) Johnston *et al.* reported that the weight assigned to the depth cues employed appeared to be somewhat dependent on context; as viewing distance was increased beyond 1 m, the weight assigned to texture increased and the weight assigned to stereo decreased.

Maloney and Landy (1989) propose a formal statistical framework as a model of how the human visual system may integrate depth estimates derived from

*Department of Psychology, University College London, Gower Street, London WC1E 6BT, England.

independent depth modules. The model is characterized by several assumptions, including the assumptions that depth cues are integrated in a linear fashion, and that this integration or "depth fusion" is carried out at each point in the visual scene. They suggest that the output following depth fusion is represented in a depth map of surface points in the scene. Maloney and Landy demonstrate that the weights assigned to depth cues can be assessed psychophysically using perturbation analysis. If a subject is presented with a stimulus composed of inconsistent depth cues, then the weight assigned to a given cue may be obtained by varying the cue in question while allocating a fixed depth value to the other cues. The weight assigned to the cue is simply the ratio between the change in the subject's estimate of overall depth and the change in the cue.

To test the validity of their model, Landy *et al.* (1990) estimated the weights given to two depth cues, texture and motion. The stimuli used were computer-generated, vertically orientated, textured cylinders rotating back and forth about a centrally placed horizontal axis. With the motion depth cue being assigned a constant value and the texture cue varying in depth, Landy *et al.* estimated the weight of the texture cue by measuring the slope of a line fitted to the plotted data of perceived depth as a function of the value of the texture cue. They estimated the weight of the motion cue by subtracting the weight of the texture cue from 1, rather than independently determining the weight of the motion cue by reversing the roles of the cues in a second experiment and determining whether the weights of the two cues did in fact sum to 1—a strategy recommended by Maloney and Landy (1989) as a rigorous test of the model. Although, we note that for a linear model this should be considered a redundant test.

Although Maloney and Landy's depth integration model has found support in several independent studies, the results are not unequivocal. Stevens, Lees and Brookes (1991) report that the integration of binocular disparity and surface contour (mono) cues does not appear to be a straightforward additive process. They point out that, if these cues were to be combined in a quantitatively additive manner, it should be possible to trade off the stereo amplitude of a surface feature, such as curvature, against that surface feature's mono amplitude while maintaining the same overall perceived amplitude. They report that, although a trade off between mono and stereo information can be made within limits, their additivity is not robust.

Several studies lend support to Maloney and Landy's depth integration model (Dosher *et al.*, 1986; Bruno & Cutting, 1988; Johnston *et al.*, 1993). However, none of these studies provide evidence to support the hypothesis that, prior to integration, depth information is represented in a range map. It is possible that the locus, or source, of integration may be found in an alternative representation to that proposed by Maloney and Landy; such as a representation based on local orientation, or a representation based on curvature.

Evidence has emerged in recent years undermining the notion that the locus of integration is at the level of a depth and/or orientation representation. Todd and Akerstrom (1987) reported that, when making depth judgements about textured elliptical surfaces, there was no impairment of subjects' performance when a regularly textured surface was replaced with an irregularly textured surface in which the texture elements varied randomly in both size and shape. This was despite the fact that manipulating the texture in this way resulted in a significant reduction in the correlations between the lengths of the individual optic elements and the depths of their corresponding surface elements, and between optic element compression and surface element orientation. This led Todd and Akerstrom (1987) to conclude that, rather than perceiving surfaces by assigning local depth or orientation values, observers' judgements were based on a more global level of image structure. In other words, they reject the idea that the locus of integration is at the level of point-by-point depth or orientation representations. Similarly, the recent work of Cumming, Johnston and Parker (1993) suggests that curvature is not calculated by first extracting local surface orientation from the compression of single texture elements. Stevens *et al.* (1991) suggest that surface curvature features are detected separately by stereo and mono processes. They also argue that it is these surface curvature features, rather than local, pointwise values such as depth or surface orientation, which constitute a "common language" or representation, and that the integration of cues occurs at this level. Johnston and Passmore (1992, 1994) reported low curvature discrimination thresholds in a surface alignment task (Weber fractions of around 0.1 were reported), demonstrating considerable precision in the task. At threshold, the changes in the surface normals were around a factor of 10 less than the threshold for detecting a change in slant for shaded spherical patches. This is taken as evidence that curvature information may be represented by the visual system in the form of an explicit description, rather than being represented implicitly as changes in surface orientation. The fact that the Weber fraction reported for the curvature discrimination task was stable across the range of surface curvatures tested further suggests that curvature is a primary representation (Foster, Simmons & Cook, 1993). Thus there is a growing body of evidence which questions the notion that the locus of integration is either a depth or orientation representation. Furthermore, for shading and texture, recent research suggests a curvature representation as an alternative locus of integration.

The present experiments were designed to test whether the visual system integrates shading and texture in a linear manner, or whether some other, as yet unspecified, combination rule better describes their integration. The experiments were designed to measure the weight of the two cues independently, rather than inferring the weight of one cue from that of the other. These experiments also provide the opportunity to investigate the locus of integration, for example by showing that a simple model, like the linear model, holds for one kind of three-dimensional shape representation but does not hold for another. Thus by simply comparing the ease with which a simple model is fitted to data described in terms of depth and curvature descriptors, we can infer which of the two representations was more likely to have been employed by the visual system.

METHOD

Subjects

The author, WC, and two other subjects, PP and CF, participated in the experiment. PP had extensive experience in curvature discrimination tasks, and CF was a naive subject with no prior experience of psychophysical experiments. All subjects had normal or corrected vision.

Stimulus generation and display

An image of a pair of spherical surface patches was constructed by ray casting (Foley, van Dam, Feiner & Hughes, 1990). The stimulus generation software allowed control over the curvature of the patches, their location in the modelling space, the viewpoint and the location of a single point light source for each patch. The surfaces were rendered using a Phong illumination model,

$$p = sI_a + sI_p(\mathbf{N} \cdot \mathbf{L}) + gI_p(\mathbf{H} \cdot \mathbf{N})^n,$$

where P is the computed brightness, s is the albedo, I_a is the intensity of ambient illumination, I_p is the intensity of direct illumination, and g is the proportion of light reflected specularly. **N** and **L** are the surface normal and light source direction unit vectors and **H** is the unit vector which bisects **L** and the line of sight. The spread of specular reflection is controlled by the parameter n.

Texture was added to the spherical patches using a texture mapping technique. The plane cannot be mapped onto a doubly curved surface without distortion. The nature of the distortion depends upon the mapping function. An equidistant azimuthal mapping, which preserves radial distances, was chosen. We can think of the equidistant azimuthal mapping as the result of positioning the north pole of the sphere from which the surface patch is derived on the origin of the texture map and then transferring the texture onto the sphere by rolling along lines passing through the origin. In this projection there is no distortion of the texture along meridians of longitude, although there is some shrinkage of the texture along parallels of latitude. Shrinkage is minimal near the central point of the display and maximal at the occluding boundary. For the majority of curvature values used, texture distortion at stimuli edges was $<10\%$. The texture is scaled in accordance with the following expression

$$\sin(\pi/2 - \phi)/(\pi/2 - \phi),$$

where ϕ is the elevation, in radians, and the texture shrinkage, expressed as a percentage, is given by

$$(1 - \sin(\pi/2 - \phi)/(\pi/2 - \phi))*100.$$

The texture map provides the albedo value for any point on the visible hemisphere. For a given ray through a pixel on the screen the surface normal at the intersection point of the ray and the sphere was computed and specified in terms of elevation and azimuth. Those parameters were used to index the texture map. Since in general the specified location in the texture map would lie between grid points, the albedo values were calculated using bilinear grey-level interpolation. The advantage of the texture mapping technique is that for radial directions there are equal amounts of texture for equal amounts of distance along the surface. In the alternative technique in which the texture is carved from a solid block the size of each texture element depends upon the angle of cut and position relative to the voxels in the solid.

The texture map chosen for the experiment was a grey-level checkerboard pattern. The surface patches were displayed against a background pattern composed of random grey-level noise. Thus the generated display gave the impression of an opaque screen with two apertures, each measuring 2 cm in diameter, through which two spherical surface patches appeared to protrude. Pairs of spheres of various diameters were generated and viewed through the apertures. The diameter of a generated sphere was never less than the diameter of the aperture through which it was viewed.

The ray casting was computed from a 75 cm distant viewpoint. The stimuli were displayed under polar projection on a 19 in. Sony Trinitron monitor screen under the control of a SUN Sparcstation 330. The grey-level display provided 8-bit resolution per pixel. In order to linearize the display a lookup table of luminance values was determined with a micro-photometer and used to control stimulus brightness. The subject viewed the screen monocularly from a distance of 75 cm. The position and direction of the light source are specified with reference to a coordinate frame centred on the patch. The z-axis extends out from the centre of the patch. The light source slant describes the angle between the light source vector and the z-axis and the light source tilt specifies the direction of slant, with 0 deg of tilt referring to the upper quadrant of the yz-plane. A given sphere was lit by a single light source positioned above it at a slant of 45 deg.

Procedure

Pairs of stimuli were presented side by side (see Fig. 1). One stimulus had texture and shading cues with identical curvature values and the second stimulus was comprised of texture and shading cues with different curvature values, with curvature being defined as the inverse of the radius. These stimuli were labelled the consistent-cue and inconsistent-cue stimuli, respectively. The inconsistent-cue stimuli were generated as follows. The shading was calculated on the basis of each surface normal for the shading cue surface. As stated earlier, the texture map provides the albedo value for each point on the surface. For discrepant surfaces, those stimuli for which the shading and texture curvatures differed, the texture map was indexed by the azimuth and elevation of the intersection of each ray and the texture cue surface. This procedure is equivalent to painting the texture for one curved surface onto another, appropriately shaded, curved surface. The texture was scaled with curvature to ensure that the texture element size, as defined on the texture cue surface, remained constant.

One cue in the inconsistent-cues stimulus (the "anchor cue") remained fixed. The perceived curvature of the inconsistent-cues stimulus was measured as a function of the curvature of the second cue, which was assigned one of eight curvature values in each block of trials; these values were 0.25, 0.33, 0.50, 0.58, 0.67, 0.75, 0.83, and 1.0 cm^{-1}. In each experiment the curvature value of the anchor cue of the inconsistent-cues stimulus was fixed at one of three settings: 0.4, 0.7, or 0.9 cm^{-1}. In Expt 1 the shading cue of the inconsistent-cues stimulus was anchored and the curvature value of the texture cue was allowed to vary. In Expt 2 the roles of the two curvature cues in the inconsistent-cues stimulus were reversed. Note that although the curvature values sampled are uniformly spaced, they become less uniformly spaced when expressed in terms of depth.

Subjects' curvature discrimination thresholds and points of subjective equality were measured using an adaptive method of constant stimuli, APE (Watt & Andrews, 1981). Discrimination threshold is defined as the standard deviation of the error distribution and corresponds to the 84% point on the psychometric function. The point of subjective equality (PSE) is defined as the 50% point on the psychometric function. The subjects' task was to indicate, with the press of a

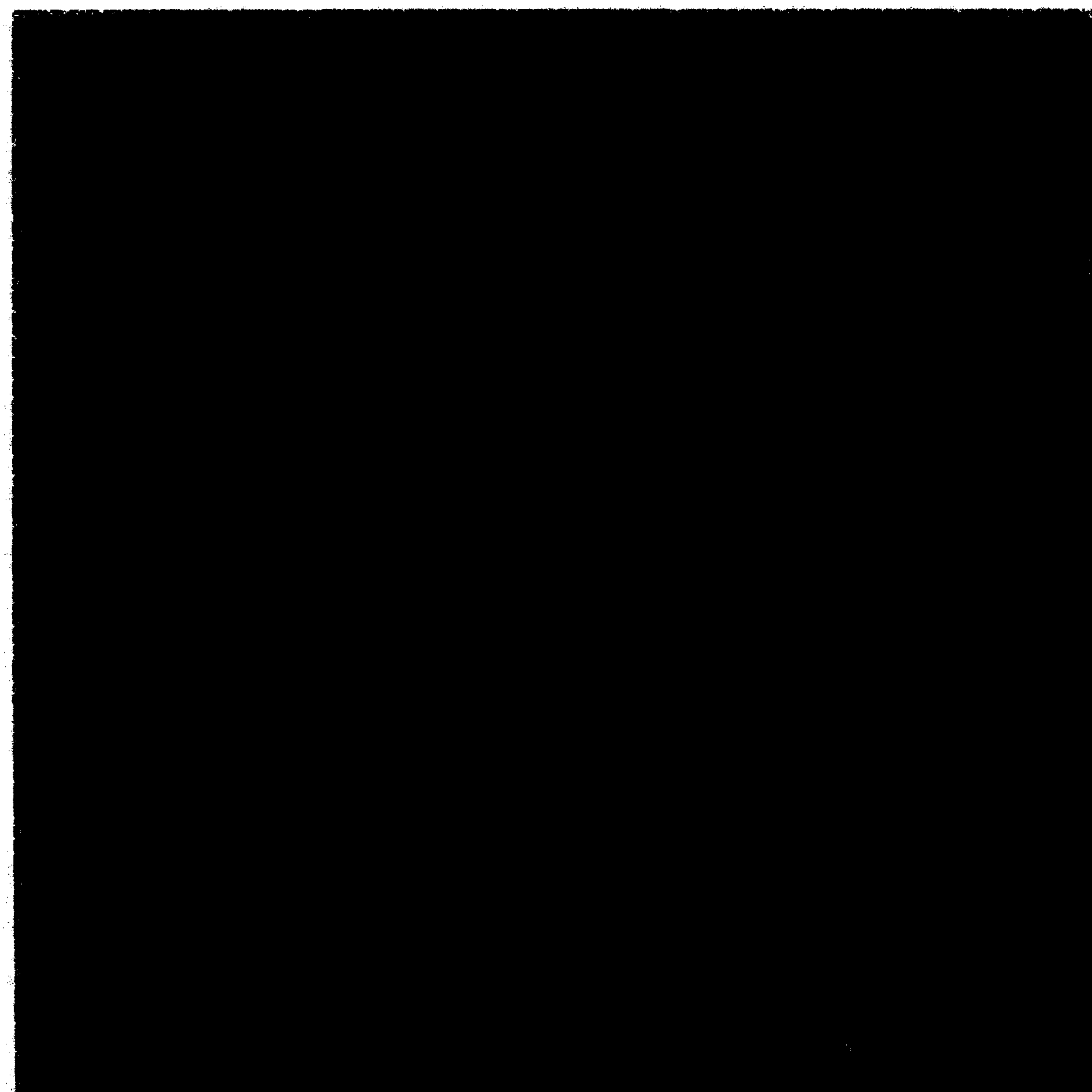


FIGURE 1. An example pair of stimuli employed. Subjects were told to treat each display as though viewing two spheres protruding through two apertures. The diameters of the spheres were never smaller than the diameters of the apertures. For the pair of stimuli shown the inconsistent-cues stimulus (left) has a shading curvature value of 0.83 cm^{-1} and a texture curvature of 0.7 cm^{-1}; the consistent cues stimulus has a curvature value of 0.7 cm^{-1}.

button, which of the two surfaces appeared more curved. The inconsistent-cues stimulus appeared equally often on each side of the display across trials. Pairs of stimuli were displayed until a response was given. Following a subject's response the display was replaced by a random grey-level noise pattern. The time lapse between terminating one display and presenting the next pair of stimuli was approx. 3 sec. Each psychometric function was calculated on the basis of the subjects' response to 64 pairs of stimuli. Each data point is the average of at least three separate measurements.

If the integration of texture and shading is best accounted for by a simple weighted linear model, as proposed by Maloney and Landy (1989), then the data obtained from the above experimental conditions should fall on a plane defined by the formula:

$$c_p = w_t c_t + (1 - w_t) c_s$$

where c_p is perceived curvature, w_t and $1 - w_t$ are the texture and shading weights, and c_t and c_s are the portrayed texture and shading curvature values. (Note that the above model has only one parameter. This is because there are only two cues and the linear model, as proposed by Maloney and Landy, is constrained by the assumption that the weights of the available depth cues in a given scene must add to 1; thus, the appropriate linear model under the conditions of this experiment has only one free parameter.) The weight of the given cue, say texture, can then be estimated by calculating the slope of the line fitted to the data obtained from the experimental conditions in which a given shading curvature is combined with a range of texture curvatures.

RESULTS

Figure 2(a–f) plots subjects' results for both experiments. Figure 2(a, b, c) plots subjects' perceived curvature as a function of the value of the texture curvature cue, with the shading cue remaining fixed at one of three curvature values—0.4, 0.7, and 0.9. Figure 2(d, e, f) plots subjects' perceived curvature as a function of the value of the shading curvature cue, with the texture cue remaining fixed at one of the above three curvatures. The data were similar across subjects; thus the data from each experiment were summed and averaged to produce one set of results per experiment [see Fig. 3(a, b)].

The data in Fig. 2(a–f) demonstrates that the texture cue had practically no effect on curvature judgements for images with low curvature values, whereas the shading cue had a large effect. This is well illustrated by the 0.4 condition, in which shading curvature remained constant at 0.4 and texture curvature was varied. When the texture cue was assigned a curvature value of between 0.25 and 0.58, subjects' curvature judgements are equivalent to a surface curvature value of 0.4. Thus, varying the texture cue in the range 0.25–0.58 had no noticeable effect on subject's perceived curvature. However, when the texture curvature value remained constant and shading was varied across the above range (Expt 2), there was a clear effect on curvature perception; changes in perceived curvature were in the same direction as the changes in shading curvature. In contrast, as the curvature value of either cue increased beyond 0.58 in the 0.4 condition, the perceived curvature of the inconsistent-cues stimulus also increased. Furthermore, as the texture anchor cue curvature value is increased perceived curvature appears gradually to become more influenced by the texture cue than by the shading. This is demonstrated in Fig. 2(d–f) with the data suggesting that varying the shading curvature value has little or no effect on curvature judgements but for a texture curvature of 0.9.

Figure 2(a–f) can be thought of as representing cross-sectional views of the three-dimensional surface generated for each subject by the data combined from all the experimental conditions [see Fig. 3(b)]. Because a consistent-cues stimulus with $c_t = c_s = c_p$ must match an "inconsistent-cues" stimulus with $c_t = c_s = c_p$, the surfaces generated by each subject's data are constrained to rest on the axis describing this relationship. Landy and Young (personal communication) point out that such surfaces are well described by a general quadratic model incorporating the above constraint,

$$c_p = ac_t^2 + bc_s^2 - (a + b)c_t c_s + dc_t + (1 - d)c_s.$$

This reduces the number of free parameters to three.

To test how much variance was accounted for by a linear model the following procedure was applied to each subject's set of data. The six sets of data generated by each subject (three sets each from the texture and shading experiments) were treated as estimates of points on a continuous surface. If the integration of the texture and shading conforms to a simple linear rule, we should expect that the slope of each of the three sections of the surface generated by the texture conditions will be both constant and equal. Similarly, the slopes of the three surface sections generated by the three shading conditions should also be both constant and equal. If this is the case, then the resulting surface will be best described by a planar function of the form

$$c_p = dc_t + (1 - d)c_s$$

where c_t and c_s are texture and shading curvature values, and d is a parameter. Note that this model meets the theoretical constraint discussed earlier, and is a subset of the above three-parameter quadratic model. It is clear from Fig. 2(a–f) that, for most of the experimental conditions, the obtained data do not fall on a straight line, suggesting that the shading and texture cues were not combined in a linear fashion.

The planar function accounted for 90.4%, 81.6% and 90.9% of the variance in WC's, PP's and CF's data, respectively. However, fitting a six-parameter general quadratic function to the subjects' data and removing coefficients close to zero suggested the following one-parameter non-linear model may provide a better fit to the data

$$c_p = ac_t^2 - ac_t c_s + c_s.$$

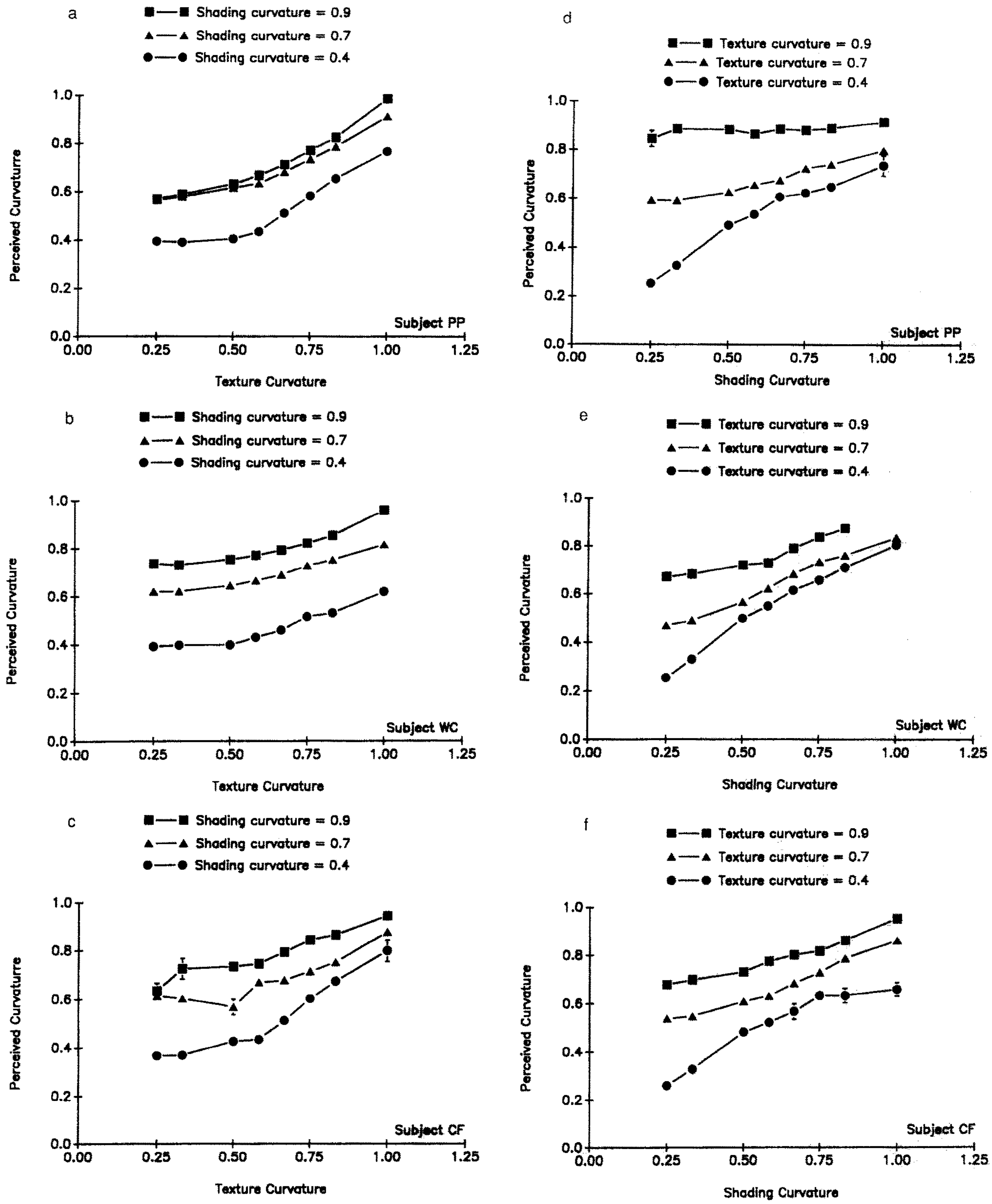
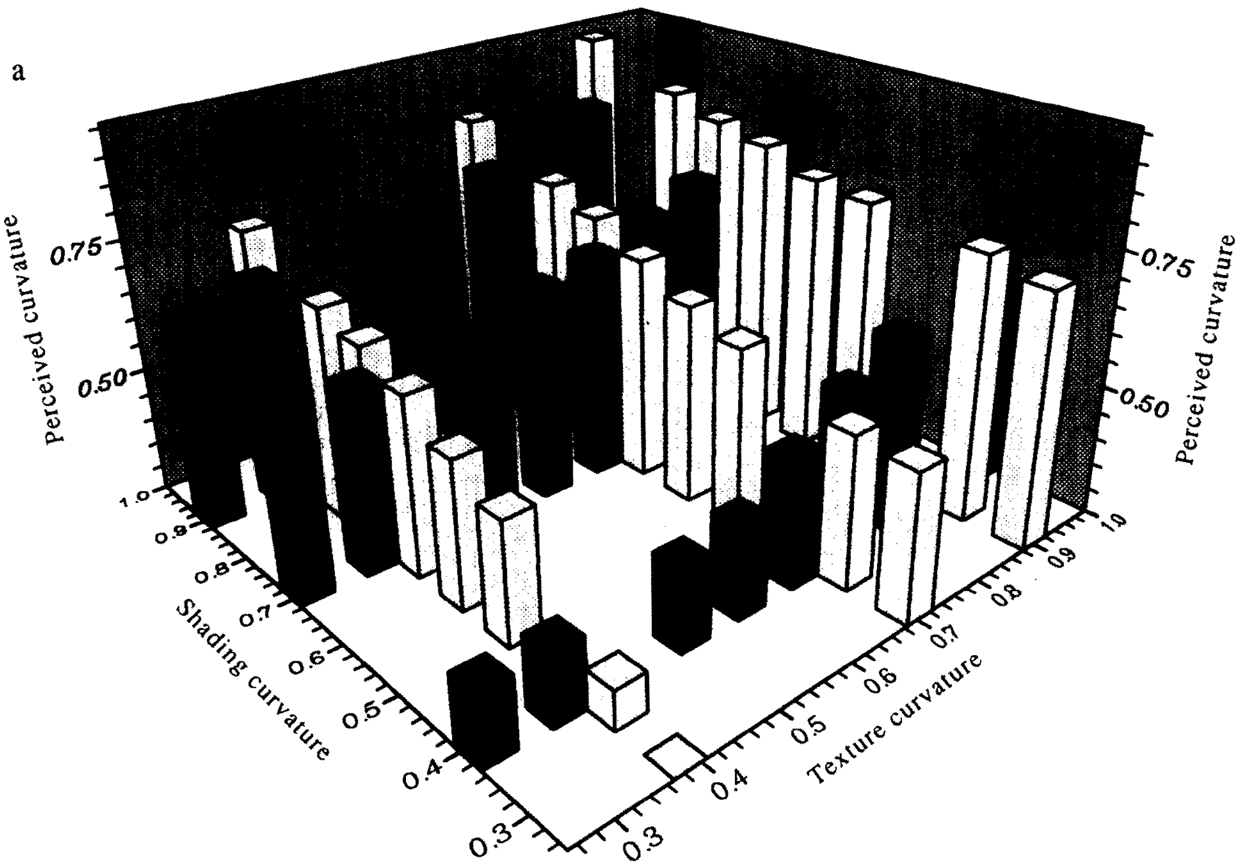


FIGURE 2. Perceived curvature as a function of the curvature represented by the texture cue with the value of the shading cue fixed at one of three levels. Curvature is expressed as the inverse of the radius. (a) subject PP; (b) subject WC; (c) subject CF. Perceived curvature as a function of the curvature represented by the shading cue with the texture cue held constant at one of three values. (d) Subject PP; (e) subject WC; (f) subject CF. Perceived curvatures are measured using an adaptive method of constants (Watt & Andrews, 1981). Each data point is the average of at least three separate PSE determinations and is therefore based on at least 192 trials. The bars show $\pm$1SE of the measurement.

Note that this model is a subset of the three-parameter quadratic model, and also incorporates the diagonal constraint discussed above. The non-linear model accounted for more variance in the data than the linear model, accounting for 94.3%, 88.8% and 93.4% of WC's, PP's and CF's data, respectively.

Although the two models clearly accounted for different amounts of variance, it is not possible to test directly a



b

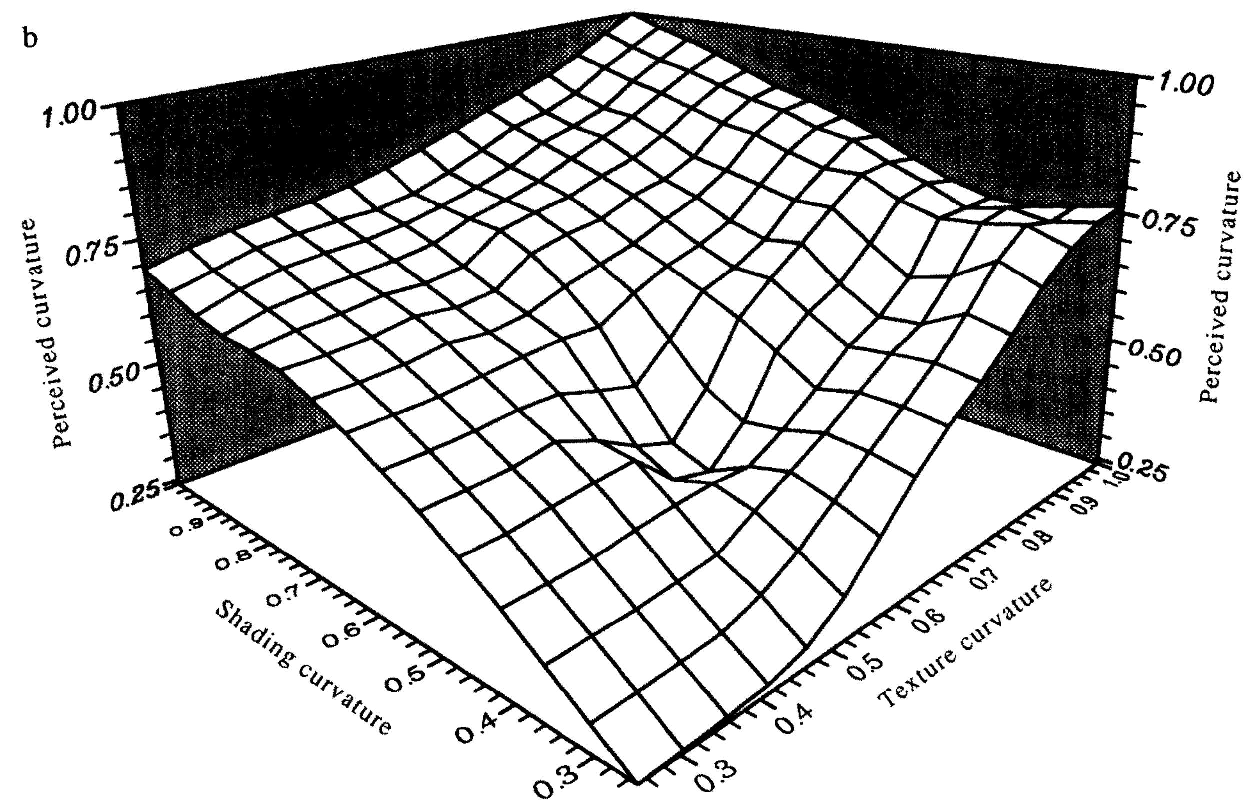


FIGURE 3. Because of the similarity of performance between subjects, the data for all subjects were summed and averaged. These mean responses are represented in a three-dimensional bar plot (a), and surface plot (b), of perceived curvature as a function of shading and texture curvature. The dark bars in (a) indicate perceived curvature when the shading cue was anchored at curvature values of 0.4, 0.7 and 0.9 cm^{-1} and the texture curvature was varied between 0.25 and 1 cm^{-1}. The light bars indicate perceived curvature when the texture cue was anchored at the above three curvature values and the shading cue's curvature was varied between 0.25 and 1 cm^{-1}.

TABLE 1. Model parameters

	a	b	d
Subject WC	$P < 0.05$	NS	NS
Subject CF	$P < 0.05$	NS	$P < 0.05$
Subject PP	$P < 0.05$	NS	NS

Results of testing whether the parameters dropped from the three-parameter model, to derive the one-parameter non-linear and linear models, are significantly different from zero. From the table we can see that the parameter a, which was dropped to derive the linear model, is significantly different from zero for all subjects; this suggests the fit of the linear model could be improved upon.

whether this difference is statistically significant. Instead, we compared both models against the three-parameter quadratic model

$$c_p = ac_t^2 + bc_s^2 - (a+b)c_t c_s + dc_t + (1-d)c_s$$

of which both models are subsets. Thus for the linear model we set $a = b = 0$ with d being fit to the data, and for the non-linear model we set $b = d = 0$ with a being fit to the data. Our strategy was to attempt to show that the three-parameter model was not significantly different from the non-linear model, but was significantly different from the linear model. Preliminary exploration indicated that the proposed non-linear model was appropriate for these data, whereas the linear model was not. t-Test analysis (Table 1) demonstrates that b and d are not significantly different from zero, although d was just significantly different for CF, and that a is significantly different from zero for all subjects. This indicates that the linear model is inappropriate for these data. Furthermore, when the standardized residuals are plotted against texture curvature for the linear model fit, there is evidence of systematic curvature for all subjects, as opposed to the random scatter that is typical of good model specification (see Fig. 4), indicating a quadratic term is needed. In the case of the non-linear model there is no apparent systematic pattern when standardized residuals are plotted against the squared texture curvature. Regression analysis (Chatterjee & Price, 1977)* shows that, for all subjects, the non-linear model accounted for significantly less variance in the data than the three-parameter model [WC: $F(2,43) = 5.4$, $P < 0.05$; CF: $F(2,44) = 19.48$, $P < 0.001$; PP: $F(2,44) = 8.048$, $P < 0.001$]; as did the linear model [WC: $F(2,43) = 21.5$, $P < 0.001$; CF: $F(2,44) = 30.68$, $P < 0.001$; PP: $F(2,44) = 22.98$, $P < 0.001$]. Although the results of the regression analysis are inconclusive, in each case the F-values are considerably higher for the linear model than the non-linear model. Furthermore, the pattern of residuals and the results of the t-tests suggest that the non-linear model provides a more parsimonious representation of the data than the Maloney–Landy model.

The data suggest that the weight of the texture cue varies as a function of that cue's curvature value. Although the linear model does not provide the best fit to the data, we can consider the use of a local linear approximation. Since, from elementary calculus,

$$dc_p = \frac{\partial c_p}{\partial c_t} dc_t + \frac{\partial c_p}{\partial c_s} dc_s$$

we can see that the change in perceived curvature dc_p is a linear function of increments in shading and texture cues weighted by the values of the partial derivatives of the surface. For the Maloney and Landy model the weight of a given cue, defined by the slope of the function relating perceived curvature and the value of the cue, is equal to the value of the parameter which scales that cue,

$$w_t = \frac{\partial c_p}{\partial c_t} = d.$$

However, this is not true of the non-linear model. In order to make explicit the changes in the perceived curvature for a change in a given cue, the local slope values were estimated by means of the partial derivatives of the perceived curvature with respect to shading curvature and texture curvature. Thus the weight of the texture cue, that is the influence that changing the texture cue has on perceived curvature, is given by the equation

$$w_t = \frac{\partial c_p}{\partial c_t} = 2ac_t - ac_s$$

and the weight of the shading cue was estimated from the equation

$$w_s = \frac{\partial c_p}{\partial c_s} = 1 - ac_t.$$

In the Maloney and Landy model (1989) the weights scale the absolute values of the depth cues to derive a value for perceived depth. However, in the differential model

$$dc_p = w_t dc_t + w_s dc_s$$

the weights scale increments in three-dimensional cues to derive a value for the change in perceived three-dimensional shape. In other words, the weights describe how *changes* in shading or texture determine the *change* in perceived curvature.

Figures 5(a, b) is a plot of the weights assigned to texture and shading, respectively, for the 0.4, 0.7 and 0.9 conditions. We can see from Fig. 5(a) that as the curvature value of the texture cue increases there is a proportional increase in the influence of texture on subjects' perception of surface curvature. However, in the case of the shading cue, Fig. 5(b) shows that, as the

*Chatterjee and Price show that this test can be expressed directly in terms of the sample multiple correlation coefficient. Let R_p denote the sample multiple correlation coefficient obtained by fitting the full model with all the p variables to a set of data, and let R_q denote the sample multiple correlation coefficient obtained by fitting the reduced model with q number of variables to the data. The F-statistic for testing the null hypothesis that $(p-q)$ specified variables have zero regression coefficient is

$$F = \frac{\left(R_p^2 - R_q^2\right) / (p-q)}{\left(1 - R_p^2\right) / (n-p-1)}$$

with d.f. $= p-q$, $n-p-1$.

value of the shading cue increases, the influence that shading has on subjects' perception of surface curvature remains constant.

To assess the likelihood of subjects using depth measurements rather than curvature measurements in making their judgements, the data were transformed into depth values and tested with both models described plus the three-parameter model. Scaling the axes in this way rather than in terms of curvature results in a reduced fit for all three models. The amount of variance accounted for by the linear model was reduced from 90.4%, 81.6% and 90.9% to 76.6%, 78.8% and 82.9% for subjects WC, PP and CF, respectively. The amount of variance accounted for by the three-parameter model was reduced from 95.4%, 91.8% and 96.5% to 89.4%, 82.3% and 91.9% for subjects WC, PP and CF, respectively. Thus the non-linear model in the curvature domain was a better fit than the three parameter model in the depth domain for all subjects (WC, PP and CF: 94.3%, 88.8% and 93.4% for the non-linear model in the curvature domain as compared to 89.4%, 82.3% and 91.9% for the three-parameter model in the depth domain). The better fit of the non-linear model in the curvature domain when contrasted with the three-parameter model in the depth domain suggests that it is likely subjects were deriving information about surface form from a curvature map rather than from a range map.

DISCUSSION

In the Maloney and Landy (1989) model of depth cue integration the percept results from the process of

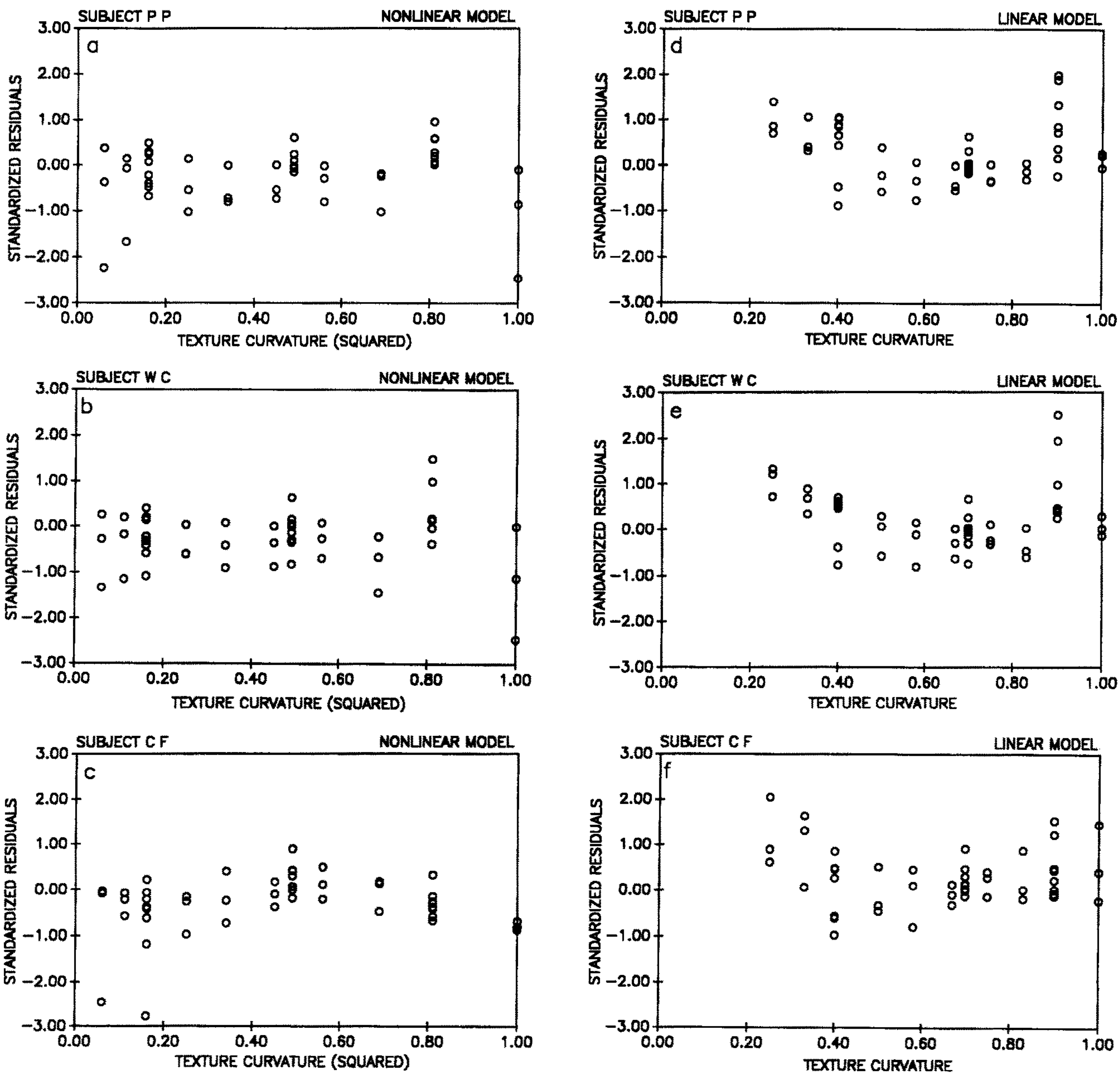


FIGURE 4. (a, b, c) The standardized residuals plotted against squared texture curvature values for the non-linear model for each subject. (d, e, f) The standardized residuals plotted against texture curvature values for the linear model for each subject. While there is random scatter of the points in (a)–(c), there is clearly systematic curvature in (d)–(f) suggesting that a quadratic term is needed.

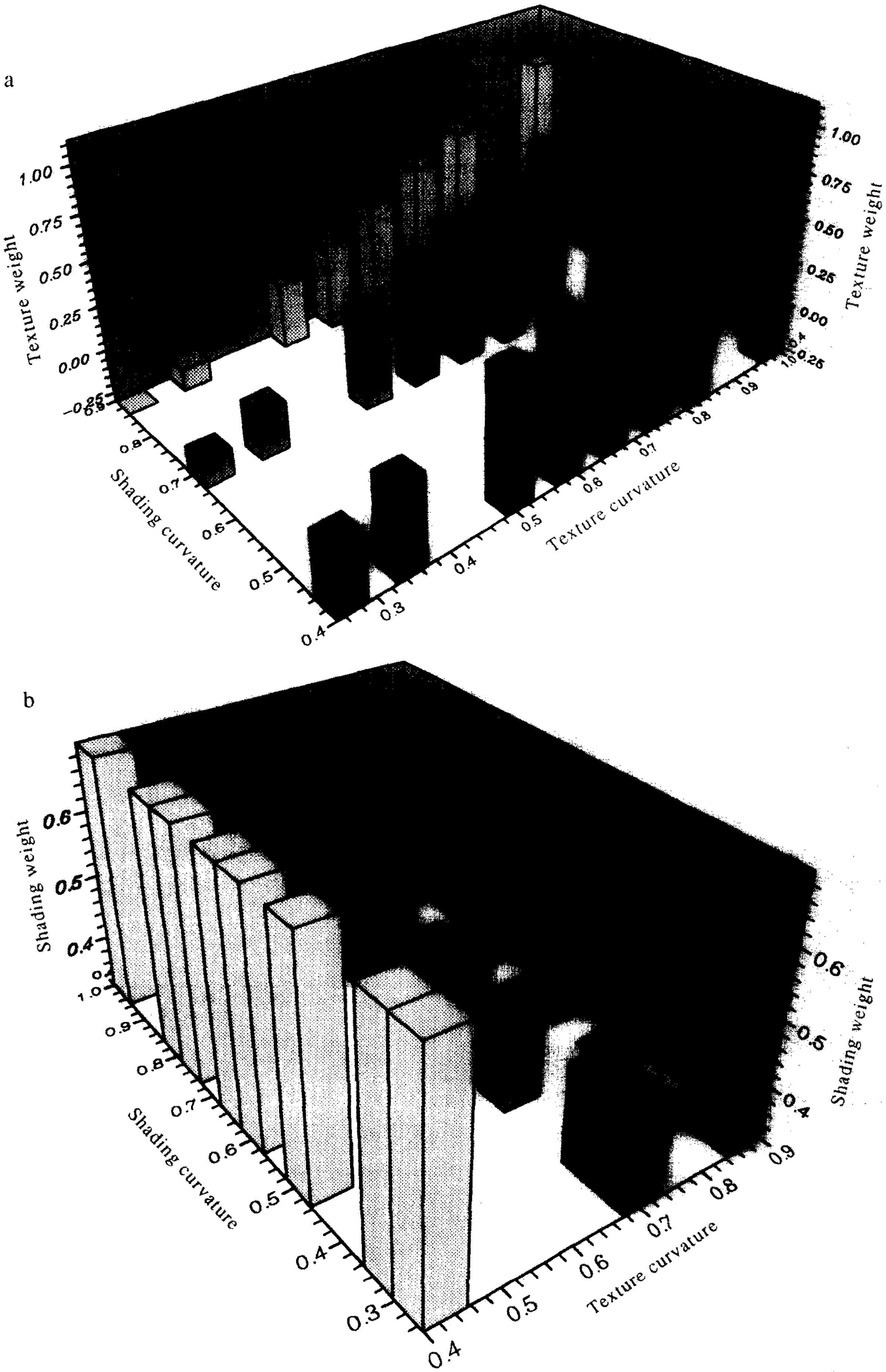


FIGURE 5. The weights of the texture and shading cues for each of the eight levels tested within each "anchor" condition. The weights scale increments in the texture and shading cues to derive a value for the change in perceived three-dimensional shape. The weights were calculated from the mean responses across subjects. (a) The weights assigned to the texture cue; (b) the weights assigned to the shading cue.

combining the outputs of independent depth modules in a specified proportion; a procedure which they refer to as weak fusion (Young *et al.*, 1993). The elegant technique of perturbation analysis allows one to estimate the relative weights of individual depth cues by varying one cue and measuring its influence on the combined percept. If the change in the combined percept is proportional to the change in the level of the cue for all increments in the cue then we can model the combined percept as a planar function of the values of the component cues. This is the situation we have referred to as conforming to a linear combination rule. However, Maloney and Landy accept that other influences might affect the combined percept, like the separation of the component cues. Under these conditions there will be a departure from linearity. Indeed one interpretation of the non-linear model presented here

$$c_p = c_s + a\,(c_t - c_s)c_t$$

is that perceived curvature depends upon shading curvature plus some influence from texture which is weighted by a proportion of the difference between the cues. Other interpretations may also be possible and we would wish to note that we do not assume an explicit implementation of this rule. It does however provide a useful description of the data.

Following Maloney and Landy, we have defined the weights as the values of the partial derivatives of the surface generated by the measured value of the combined percept, with the component cues taken as parameters, although the idea has been generalized to deal with more complex surfaces. This approach involves a local linear approximation and so we concur with Maloney and Landy that locally we can think of the cue combination rule as linear. However, it is clear that any curved surface will appear planar if investigated locally. So, in attempting to test Maloney and Landy's (1989) proposed model of depth cue integration in the context of shading and texture, we investigated the shape of the function for a wide range of cue combinations.

Our results show that the linear model does not generalize to the integration of these shading and texture cues, although for other pairs of cues the weights may not vary with component cue separation. If the integration of shading and texture was best described by a linear model then one would expect that the weight of a given cue would remain constant as the cue varied in its level of curvature. Although the weight assigned to shading was constant within experimental conditions, suggesting that the two cues may be integrated in a simple linear fashion, the weight assigned to texture was shown to vary systematically with texture curvature when the roles of the two cues were reversed. The value of the non-linear model described above is its role in highlighting the way in which the shading and texture weights change as a function of the values of the component cues.

The present experiments also provided the opportunity to test Maloney and Landy's proposal that the combination of cues results in an integrated representation of surface distance or depth. Our data suggest that, in the above experimental conditions, three-dimensional surface information was more likely to have been derived from a curvature map. This is supported by the poorer fit of all three models (the linear, the non-linear and the three-parameter) to the transformed depth data, and the better fit of the proposed non-linear model in the curvature domain when contrasted with the three-parameter model in the depth domain.

Thus, although several experimenters have reported empirical support for Maloney and Landy's depth cue integration model for the combination of a wide range of depth cues, the results of the experiments reported in this paper suggest that such a strategy may not be used for the integration of shading and texture cues. Instead, the results lend support to the hypothesis that the visual system may combine such cues in a non-linear fashion. It is also suggested that the locus of texture and shading cue integration is a curvature map, rather than a depth or orientation map.

REFERENCES

Bruno, N. & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General, 117*, 161–170.

Chatterjee, S. & Price, B. (1977). *Regression analysis by example*. New York: Wiley.

Cumming, B. G., Johnston, E. B. & Parker, A. J. (1993). Effects of texture cues on curved surfaces viewed stereoscopically. *Vision Research, 33*, 827–838.

Dosher, A. D., Sperling, G. & Wurst, A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research, 26*, 973–990.

Foley, J. D., van Dam, A., Feiner, S. K. & Hughes, J. F. (1990). *Computer graphics* (2nd edn). Reading, Mass.: Addison Wesley.

Foster, D. H., Simmons, D. R. & Cook, M. J. (1993). The cue for contour–curvature discrimination. *Vision Research, 33*, 329–341.

Johnston, A. & Passmore, P. J. (1992). Shape from shading: Psychophysical evidence for the primacy of surface curvature perception. *Perception, 21*, A54.

Johnston, A. & Passmore P. (1994). Shape from shading—I. Surface curvature and orientation. *Perception*. In press.

Johnston, E. B., Cumming, B. G. & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research, 33*, 813–826.

Landy, M. S., Maloney, L. T. & Young, M. J. (1990). Psychophysical estimation of the human depth estimation rule. In Schenker, P. S. (Ed.), *Proceedings of the SPIE: Sensor fusion III: 3-D perception and recognition* (pp. 247–257).

Maloney, L. T. & Landy, M. S. (1989). A statistical framework for robust fusion of depth information. *Proceedings of the SPIE: Visual communications and image processing, Part 2* (pp. 1154–1163).

Rogers, B. J. & Collett, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology, 41A*, 697–717.

Schwartz, B. J. & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society*, *21*, 456–458.

Stevens, K. A., Lees, M. & Brookes, A. (1991). Combining binocular and monocular curvature features. *Perception*, *20*, 425–440.

Todd, J. T. & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 242–255.

Watt, R. J. & Andrews, D. P. (1981). APE: Adaptive Probit Estimation of psychometric functions. *Current Psychology Review*, *1*, 205–214.

Young, M. J., Landy, M. S. & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*, 2685–2696.

Acknowledgements—This work was supported by a project grant awarded to Alan Johnston from the SERC Image Interpretation Initiative and a SERC Image Interpretation Initiative studentship awarded to William Curran. We would like to thank Dr Peter Passmore for programming assistance and Michael Landy, Mark Young and another, anonymous, referee for their comments on a previous draft.