



PERGAMON

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Vision Research 43 (2003) 2233–2243

Vision  
Research

[www.elsevier.com/locate/visres](http://www.elsevier.com/locate/visres)

## “Phase capture” in the perception of interpolated shape: cue combination and the influence function

Dennis M. Levi<sup>\*</sup>, Roger Wing-Hong Li, Stanley A. Klein

*School of Optometry, University of California, Berkeley, CA 94720-2020, USA*

Received 3 October 2002; accepted 22 April 2003

### Abstract

This study was concerned with what stimulus information observers use to judge the shape of simple objects. We used a string of four Gabor patches to define a contour. A fifth, center patch served as a test pattern. The observers' task was to judge the location of the test pattern relative to the contour. The contour was either a straight line, or an arc with positive or negative curvature (the radius of curvature was either 2 or 6 deg). We asked whether phase shifts in the inner or outer pairs of patches distributed along the contour influence the perceived shape. That is, we measured the phase shift influence function.

We found that shifting the *inner* patches of the string by 0.25 cycle results in almost complete phase capture (attraction) at the smallest separation ( $2\lambda$ ), and the capture effect falls off rapidly with separation. A 0.25 cycle shift of the *outer* pair of patches has a much smaller effect, in the opposite direction (repulsion).

In our experiments, the contour is defined by two cues—the cue provided by the Gabor carrier (the ‘feature’ cue) and that defined by the Gaussian envelope (the ‘envelope’ cue). Our phase shift influence function can be thought of as a cue combination task. An ideal observer would weight the cues by the inverse variance of the two cues. The variance in each of these cues predicts the main features of our results quite accurately.

© 2003 Elsevier Ltd. All rights reserved.

*Keywords:* Shape perception; Psychophysics; Influence function; Cue combination; Phase perception; Classification image

### 1. Introduction

Humans have a highly acute ability to judge the shape of an object, and to identify and localize distortions in the shapes of smooth objects (e.g., Watt, Ward, & Casco, 1987; Whitaker & McGraw, 1998; Wilkinson, Wilson, & Habak, 1998; Zanker & Quenzer, 1999). We are concerned with what stimulus information observers use to judge the shape of simple objects. Consider the three simple shapes, each defined by five Gabor patches, illustrated in the left column of Fig. 1 (a straight line, a “smiley-face” and a “frowny-face”). In each of these shapes, the patch centers (the center of the Gaussian envelope), and the patch features (the carrier phase) are perfectly aligned along the contour. In the middle column of Fig. 1, the patch centers are again perfectly aligned along the contour, however, the features (carrier) of the 2nd and 4th patches have been shifted

downward by 90 deg. When these patterns are presented briefly, the center patch appears to no longer be on the contour—i.e., the phase shift in patches 2 and 4, induces a change in perceived position that alters the contour.

The question of which stimulus features influence perceived position, and/or the precision of position judgements, has been extensively studied, since Westheimer and Mckee (1977) first noted that the perceived position of a line is determined by the centroid of its luminance distribution (Akutsu, McGraw, & Levi, 1999; Hess, Dakin, & Badcock, 1994; Levi & Westheimer, 1987; Morgan & Aiba, 1985; Morgan & Glennerster, 1991; Toet, Smit, Nienhuis, & Koenderink, 1988; Watt, Morgan, & Ward, 1983; Whitaker, McGraw, & Levi, 1997; Whitaker & Walker, 1988). In many recent studies the emphasis has been on the relative importance of the global (envelope) vs. local (carrier) features that comprise Gabor patches like those shown in Fig. 1. For example, an elegant recent study (Whitaker, Bradley, Barrett, & McGraw, 2002) showed that the precision of judging relative position was dominated by carrier

<sup>\*</sup> Corresponding author.

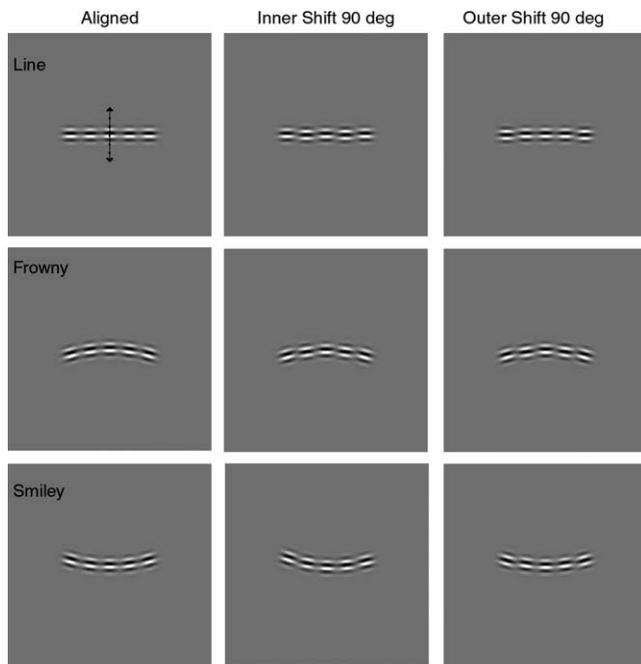


Fig. 1. Examples of our stimuli: a straight line (top row), a “frowny face” (middle row—i.e., an arc with positive curvature) or a “smiley face” (bottom row—i.e., an arc with negative curvature). The observers’ task was to judge whether the center ‘test’ patch (indicated by the arrow) was above or below the contour defined by the four outer patches. From trial to trial, the phase of the four outer patches was varied: either all four patches were phase aligned (left column); patches 2 and 4 were shifted by 90 deg (inner shift, middle column); or patches 1 and 5 were shifted by 90 deg (outer shift, right column) but the patch centers were perfectly aligned along the contour.

information when separation between patches was small, and by envelope information when separation between patches was large.

The question of which features (envelope vs. carrier) influence global shape perception has had less attention. Carrier orientations that match the contour may enhance contour detection either at detection threshold (Bonneh & Sagi, 1998; Kapadia, Ito, Gilbert, & Westheimer, 1995; Levi & Sharma, 1998; Polat & Sagi, 1993; Saarinen, Levi, & Shen, 1997), or when the target “contour” is embedded in distractors (Field, Hayes, & Hess, 1993; Kovacs & Julesz, 1993; Pettet, McKee, & Grzywacz, 1998). Detecting position jitter in a path is also not influenced by whether the local feature orientations are aligned with or orthogonal to the path (Keeble & Hess, 1999). On the other hand, the aspect ratio of a square can be judged more precisely when the local features are aligned with the global contour than when they are orthogonal to it (Saarinen & Levi, 1999). Similarly, the orientation of the gap in a C-like figure can be discerned at a lower contrast when the local features are aligned with the C-contour than when they are orthogonal to it (Saarinen & Levi, 2001). Thus, context (feature orientation) seems to matter for certain tasks and stimuli but not for others. Keeble and Hess

(1999) found little influence of the carrier orientation on the thresholds for detecting a distortion in a circle made up of Gabor patches except at the smallest separation where thresholds were better for gratings that were aligned with the circle than for gratings that were orthogonal to the circle. In a closely related study, Levi and Klein (2000) found that shape perception is determined by two factors: the primary determinant is the separation between the samples; however, the orientation of the samples can modulate performance. At small separations, performance is best when the samples are aligned with the global shape, poorer when they are orthogonal, and intermediate when they are all horizontal or vertical. At larger separations these contextual differences largely disappear (see also Levi, Klein, Sharma, & Nguyen, 2000).

While most studies of shape perception have focussed on thresholds, Keeble and Hess (1998) asked observers to judge the direction of displacement of a central Gabor patch (with vertical carrier) relative to two reference patches with orientations at  $\pm 45$  deg, and reported that in this condition, observers made large errors toward the contour formed by the three Gabor patches. Thus, contours defined by the carrier can strongly influence the point of subjective alignment.

Most previous studies (including our own) have focussed on determining which cue is used, implicitly assuming that the dominant cue determines performance. However, a more reasonable question may be how the different cues are combined to determine the shape. Recent work suggests a rather simple model: different cues are given weights based on how reliable they are (see Jacobs, 2002 for a recent review). This approach explains how haptic and visual cues are combined (Ernst & Banks, 2002; Hillis, Ernst, Banks, & Landy, 2002). Other work suggests similar cue combination rules operate in other domains, e.g. stereopsis (Landy, Maloney, Johnston, & Young, 1995; Young, Landy, & Maloney, 1993) and in selective attention (Murray, Sekuler, & Bennett, 2003).

In the present study we used a string of four Gabor patches to define a sampled contour. A fifth, center patch served as test pattern: we asked whether phase shifts in the inner or outer pairs of patches distributed along the contour, influence perceived shape. Our study bears a close relationship to the work of Hon, Maloney, and Landy (1997). Hon et al. measured the influence of eight sample points on their observers’ interpolation of parabolic sampled contours using a perturbation analysis. They found that the nearest points (equivalent to our “inner” patches) have the strongest influence; however, their results also provide some evidence in support of the hypothesis that each sample point is a “full” member of the sampled contour (i.e., that observers cannot exclude perturbed points from the contour and simply base their judgement on the other

points). Our study, like that of Hon et al. measured the influence of the sample points on the interpolation of a parabolic curve; however, rather than perturb the positions of the sample points, we “perturb” their phase. Our results, in combination with extant data, allow us to estimate the “influence function” for our sampled shapes, and to model how different cues to shape are combined.

## 2. Methods

### 2.1. Stimuli

The stimuli (illustrated in Figs. 1 and 2) consisted of strings of five circular Gabor patches. The patches were constructed to have 0.66 carrier cycles per Gaussian envelope standard deviation ( $\sigma$ ), corresponding to a spatial frequency bandwidth of 0.825 octaves. The carrier orientation was always aligned with the contour. Unless otherwise specified, the patches were briefly presented ( $\approx 200$  ms) on a Sony Trinitron F520, 21" flat screen monitor at a contrast of 80%, on a mean luminance background ( $\approx 80$  cd/m<sup>2</sup>).

The contours, as illustrated in Fig. 1, were either a straight line, or a circular arc. Two viewing distances were used. At the closer distance the radius of curvature was 6 deg and the spatial frequency of the Gabor carrier

was 3.33 c/deg. At the larger distance the radius of the circle was 2 deg and the spatial frequency was 10 c/deg. At both distances the radius of the circle was 20 periods ( $\lambda$ ) of the Gabor carrier. The observers were instructed (with the aid of a figure) to judge whether the center ‘test’ patch was above or below the contour defined by the four outer patches (which provided samples of the contour). They were told that the contour was either a straight line or a circle. From trial to trial, the phase of the four outer patches was varied: either (i) all four patches were phase aligned; (ii) patches 2 and 4 (the “inner patches”) were shifted downwards by 90 deg; or (iii) patches 1 and 5 (the “outer patches”) were shifted downwards by 90 deg. In all three cases the patch centers were perfectly aligned along the contour. At the start of each trial, a reticule (illustrated in the middle panel of Fig. 2) was presented to mark the location of the test patch. The reticule consisted of four diagonal lines, each 2' in width, 5' in length, and oriented at 45, 135, 225 and 315 deg from the tangent of the arc. The end of each line starts at 5 times the standard deviation of the patch (from the target patch center) and radiates outward. The reticule disappeared after 300 ms, and was followed immediately by the stimulus. In order to minimize bias, all 9 stimulus conditions [3 curvatures (positive, negative and zero, i.e., radius infinity) and 3 phases (all aligned; 2 and 4 shifted by 90 deg; 1 and 5 shifted 90 deg)] were randomly interleaved in a single run of 450 trials ( $\approx 50$  trials per condition). In order to avoid using edges or other absolute position cues, the orientation of the entire contour (all 5 patches) was randomly varied (by  $\pm 2.5$  deg). From run-to-run, we varied the separation between patches and the viewing distance. Fig. 2 illustrates examples of three of the separations tested. For the curved stimuli the separation between adjacent patches was constant as measured by the straight-line distance between the adjacent patch centers (see inset in Fig. 8).

In order to assess the perceived position of the central patch relative to the contour, we used a staircase to control the position of the center ‘test’ patch. The initial target patch position for each condition was a large step either above or below the ‘intended’ position (i.e. on the contour). The observer responded by pressing one of four buttons to indicate both the direction (high or low) and a confidence level. Staircase trials were determined by the observer’s prior response (both the direction and confidence). When the observer signaled high confidence, the position of the test patch was shifted two steps in the opposite direction. When the observer signaled low confidence, the target patch was shifted only one step in the opposite direction. When the observer’s previous response was in the opposite direction to that of the current response (regardless of the certainty level) the current position was recorded as a reversal. Following each reversal, the step size was halved until

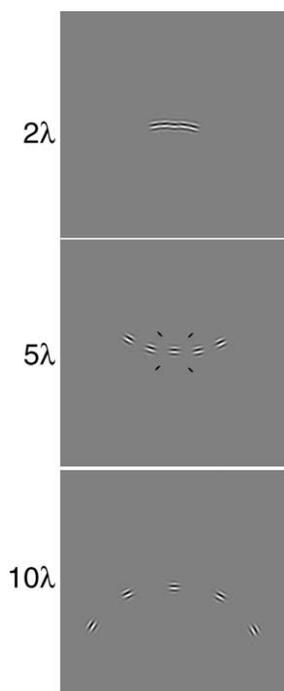


Fig. 2. Examples of our stimuli with different separations:  $3\lambda$  (top),  $5\lambda$  (middle) and  $10\lambda$  (bottom). The center panel also illustrates the “reticule” used for fixation. Note that the reticule was *not* presented at the same time as the stimulus (as illustrated here). It appeared at the start of a trial, disappeared after 300 ms, and was followed immediately by the stimulus.

reaching the minimum step size (1/100 of the interpatch separation). In order to keep the observers engaged, approximately 10% of the trials were presented at a large step (2 times the current step size). These “large” offsets were not used in the staircase, or in calculating reversals; however, they provided data at offsets larger than the 50% point, which is useful for constructing psychometric functions from the data, and for estimating thresholds. This staircase converged rapidly and each staircase resulted in, on average, 18 reversals per run. The staircase is quite efficient, giving a “sweat” (see Eq. (42) of Klein, 2001) of about 1.35. The ideal sweat for this task is 1.25. Since reversals that occurred early in the staircase may not be as reliable as those at the end of the staircase, we excluded the first two reversals from the analysis, and averaged the rest. This average can serve as an estimate of the perceived location of the test contour (PSE), but to quantify our results we constructed psychometric functions, and performed probit analysis. Each condition was repeated at least four times (giving approximately 200 trials per condition). The results reported in the figures represent the weighted means of at least four individual PSE estimates (50% point of the psychometric functions) obtained from the probit analyses. The probit PSEs are close to those obtained by averaging reversals (compare open and filled symbols in the top panels of Fig. 4). We tested 4 observers (including one of the authors), each with normal or corrected to normal vision. Viewing was monocular.

### 3. Results

#### 3.1. Perceptual errors in shape perception

Observers make perceptual errors even in the absence of a phase shift of the neighboring patches (Fig. 1, left column) and we refer to these as “standing errors”. Fig. 3 shows the PSE for each of the four observers plotted as a function of separation, specified in units of the carrier spatial period ( $\lambda$ ). The three curvatures are coded by symbol, and the two viewing distances by symbol size (larger symbols for the closer viewing distance). We used an unconventional sign convention, such that the frowny curve has positive curvature. Although these results are quite idiosyncratic several common points emerge. First, perceptual errors tend to increase with separation. Second, they are, in general, considerably larger with curved contours than with aligned contours (squares). Third, at large separations, the errors are often substantial, approaching (or, for NS, larger than) the carrier period. Observer NS’s data in Fig. 3 represent oblate errors (a “squashed” circle for which the width of the ellipsoid is greater than the height). SC tended toward prolate errors (a stretched or “pointy” circle, i.e., one for which the height is greater than the width). If the

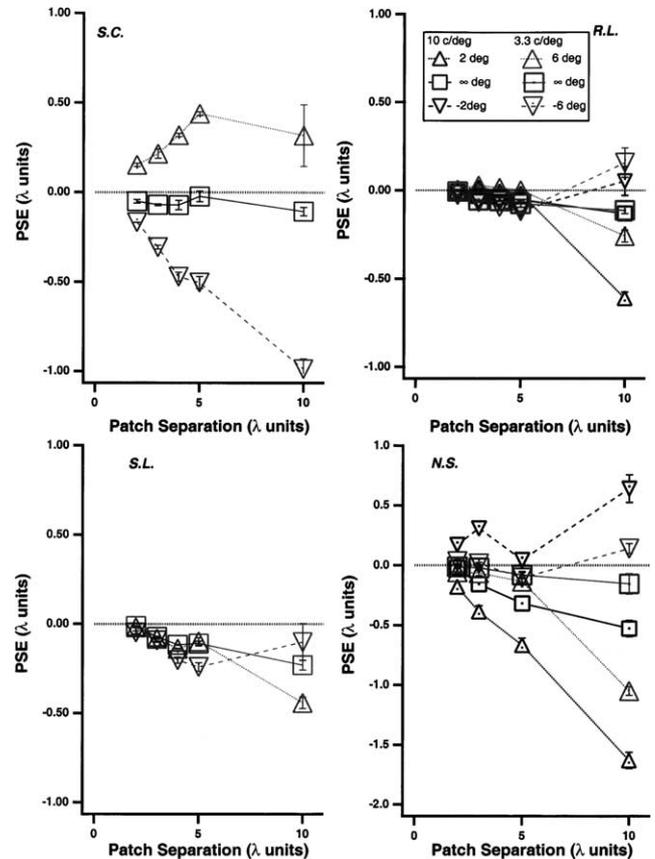


Fig. 3. The “standing error” or PSE (specified in units of the carrier wavelength) is plotted as a function of the patch separation (also specified in units of the carrier wavelength) for each observer when all 5 patches are phase aligned.

PSE in Fig. 3 were zero then the distance from the middle patch to the center of the circle would be  $20\lambda$ , identical to the distance from the inner and outer patches to the center of the circle. For a prolate PSE the distance to the center patch is larger than  $20\lambda$ . For example, if the PSE were  $0.5\lambda$ , the distance from the middle patch to the center would be  $20.5\lambda$ . If that error had been oblate the distance would have been  $19.5\lambda$ .

#### 3.2. Phase-capture in shape perception

We are interested in the effect of phase shifting the carrier of either the inner or outer patches defining the contour. Fig. 4 shows two examples of the PSEs for observer RL with the inner or outer panels phase shifted, or with no phase shift at all (abscissa—these “no shift” data correspond to RL’s  $3\lambda$  data in Fig. 3). The circles show independent estimates of the PSE for each of the 4 replications; each open symbol is an estimate based on the mean of the staircase reversals (for one run) and each filled symbol is the probit least squares estimate of PSE (for one run). The thin dotted line at PSE = 0 indicates placement of the central patch on the

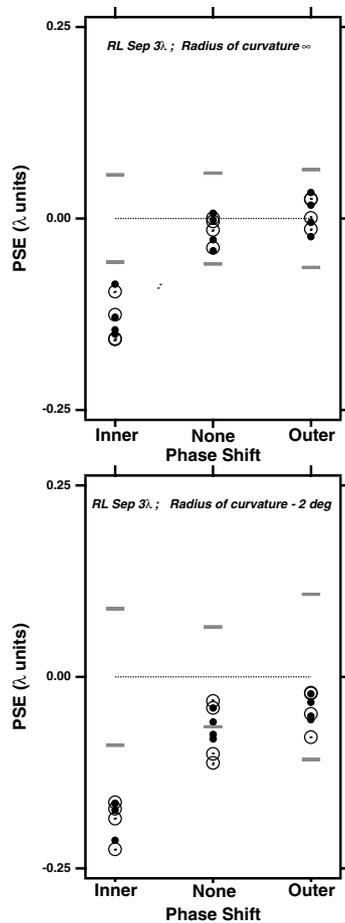


Fig. 4. PSEs (ordinate) for observer RL with the inner or outer patches phase shifted, or with no phase shift at all (abscissa). The circles show independent estimates of the PSE for each of the four replications; each open symbol is an estimate based on the mean of the staircase reversals (for one run) and each filled symbol is the probit PSE (for one run). The thin dotted line at zero indicates veridical perception (PSE = 0). The small horizontal line segments show  $\pm Th/2$ . Note that shifting the inner patches produces a substantial increase in PSE (i.e., the PSE becomes more negative). The top and bottom panels are for zero curvature and a curvature of  $-2$  deg respectively.

circle defined by the inner and outer envelopes. Probit analysis also provides an estimate of the precision (the threshold,  $Th$ , or slope of the psychometric function) of the judgement, and this is represented by the small horizontal line segments, which show  $\pm Th/2$ . There are several clear points: First, the PSEs estimated from probit analysis and from the staircase are closely similar and quite reproducible. Second, this observer shows, on average, a small negative error in perception for a straight line (top panel) with no carrier shift, and a larger negative error for a curve (radius of curvature  $-2$  deg, bottom panel, corresponding to the smiley curve). Third, shifting the inner patches downward produces a substantial increase in PSE (i.e., the PSE becomes more negative), indicating that the phase shift of the inner patches has “captured” the target. This shift is substantial, exceeding the threshold. Shifting the phase of

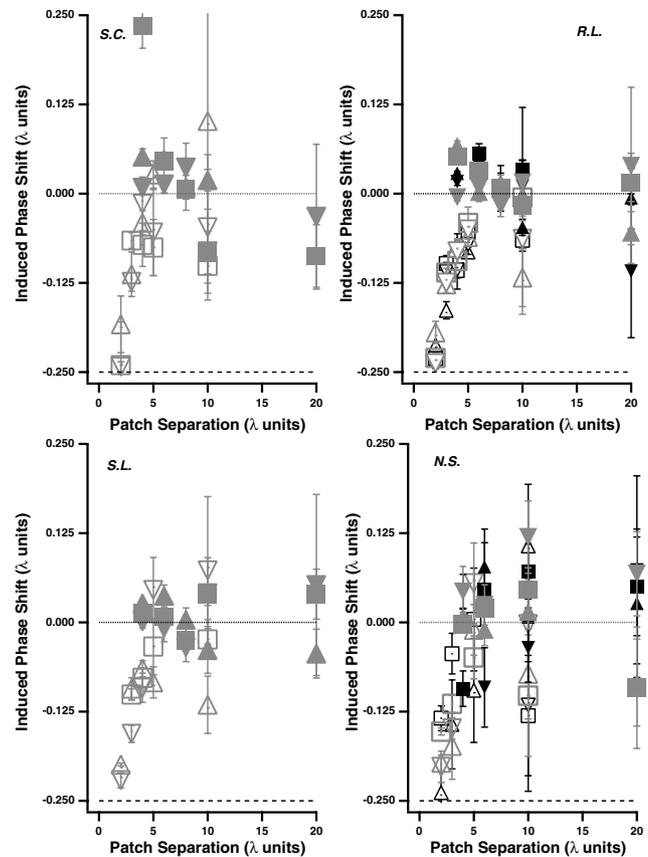


Fig. 5. The induced phase shift vs. patch separation. i.e., the *change* in PSE induced by the phase shift (i.e., we subtract out the “standing” errors shown in Fig. 2). Data are shown for two different patch sizes/curvatures (10 and 3.33 c/deg corresponding to curvatures of 2 and 6 deg—obtained by a factor of three change in viewing distance and coded by symbol size). The dotted line at 0.0 indicates no induced phase shift, the dashed line at  $-0.25\lambda$  indicates complete phase capture.

the outer patches produces a negligible shift in PSE in the opposite direction as the PSE shift caused by the phase shift of the inner patches. For an ideal observer who places the middle patch on a circle, the PSE shift caused by the outer patches, should be about  $-1/4$  of the shift caused by the inner patches.

In order to reveal the influence of carrier phase shifts, we subtract out the “standing” errors (i.e., the PSE when the carrier is aligned with the contour) and plot the *change* in PSE induced by the phase shift (Fig. 5). Although there are clearly individual differences, the overall trends are quite similar across observers. First, regardless of curvature, shifting the *inner* patches by 0.25 cycle (open symbols) results in almost ideal capture<sup>1</sup> ( $0.25\lambda$ —dotted line) at the smallest separation ( $2\lambda$ ), and the effect falls off more or less linearly with separation until a near zero effect is reached. Surprisingly, even at  $10\lambda$ , the shift of

<sup>1</sup> By “ideal capture”, we mean the capture predicted by an ideal observer model, to be discussed later. The model predicts an induced shift of  $-0.24\lambda$  at a separation of  $2\lambda$ .

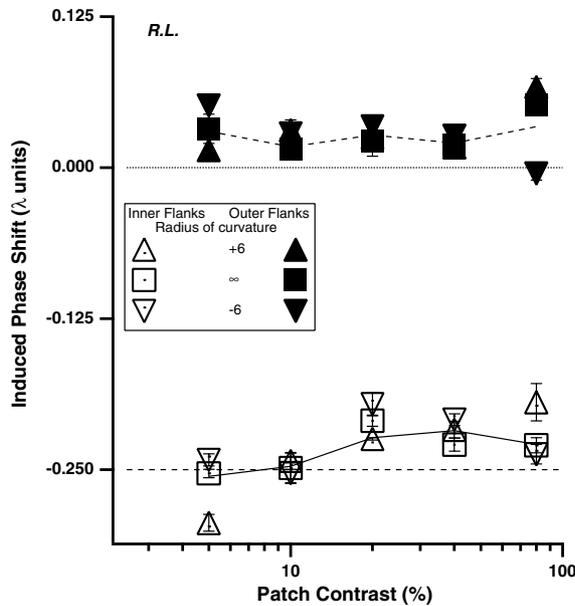


Fig. 6. The effect of contrast. Complete capture by the inner patches evident at the smallest separation ( $2\lambda$ ), is almost independent of patch contrast (data are shown for observer RL).

the inner patches results in a small residual change in perceived position. The nearly ideal capture by the inner patches evident at the smallest separation ( $2\lambda$ ), is almost independent of patch contrast (Fig. 6).

A 0.25 cycle shift of the *outer* pair of patches (solid symbols in Figs. 5 and 6) has a much smaller effect, in the opposite direction (as expected). These results can be seen more clearly in the mean data (averaged across observers and viewing distances—black symbols in Fig. 7).

### 3.3. Weighting of visual cues for shape perception: the “Influence Function” for sampled shape

We are interested in how each cue to the sample position is weighted to provide an estimate of contour shape. Consider the circular curve shown as an inset in Fig. 8A. The four patches represent sample points along the curve, each separated by a distance “sep” between patch centers. The X represents the intended position of the “test” patch (i.e., on the curve defined by the four sample points). The offset (indicated by the arrow) represents the observers’ setting of the test patch position relative to the horizontal line between the centers of the Gaussian envelope of inner patches. The offset (Off) is related to our previously defined PSE by  $Off = PSE + sag$ ; where sag is the vertical distance from the center of the envelope of the inner patches to the central sample point of a circle that goes through the center of the envelopes of the inner and outer patches. Using linear multiple regression, we can compute the regression coefficients, which correspond to the weights, of all cues ( $C_b$ ,  $C_{Env_{out}}$ ,  $C_{P_{in}}$  and  $C_{P_{out}}$ ) as follows:

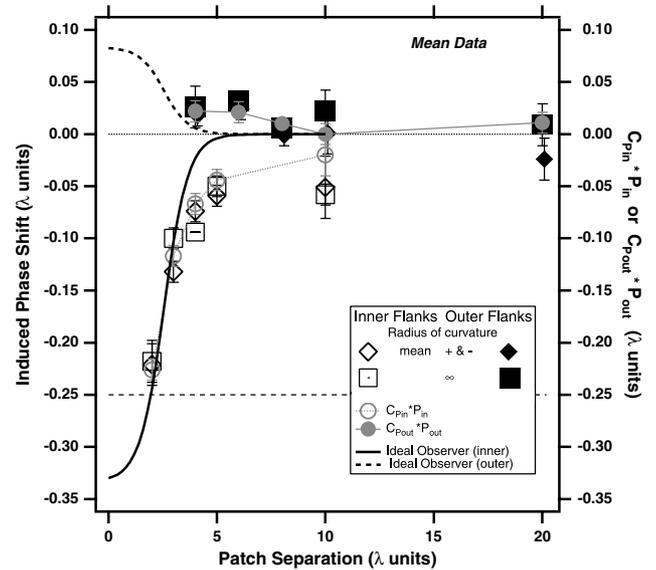


Fig. 7. The induced phase shift (left ordinate) vs. patch separation. The data of Fig. 4 have been averaged across observers and viewing distances. The thick black curves show the prediction obtained by weighting each cue by its inverse variance (see text). The right ordinate and gray circles show the coefficients obtained from our regression analysis multiplied by the phase shift (Fig. 8B), plotted in  $\lambda$  units for the inner (open circles) and outer (solid circles) patch carriers.

$$Off = C_b + C_{Env_{out}}H + C_{P_{in}}P_{in} + C_{P_{out}}P_{out} + \epsilon \quad (1)$$

$C_{Env_{out}}$  is the coefficient of the envelope of the outer patches;  $C_{P_{in}}$  is the coefficient of the carrier phase of the inner patches;  $C_{P_{out}}$  is the coefficient of the carrier phase of the outer patches;  $C_b$  is a constant, representing observer bias;  $H$  is the envelope location of the outer patches (the vertical distance in min of arc between the inner and outer patches),  $P_{in}$  and  $P_{out}$  are the phase shifts of the inner and outer carrier in min of arc, and  $\epsilon$  is the residual error. For the 3.33 c/deg carrier  $\lambda = 18$  min so the quarter cycle phase shift would be  $P_{in}$  or  $P_{out} = \pm 4.5$  min. Recall that in our experiments, for a given separation and curvature, we have three values of  $H$  ( $H = 0$  corresponding to a curvature of zero [radius of curvature = infinity], and a positive and negative  $H$ , corresponding to the negative and positive curvatures). For each separation, we computed the weights (coefficients) for each of the three cues. These weights are unitless, and correspond to the gain or influence function of each cue. This influence function may be thought of as the classification image for shape (Levi & Klein, 2002; Murray et al., 2003).

Fig. 8A shows the values of Off (specified in  $\lambda$  units) from the linear regression fits to the mean data ( $N = 4$  observers) for the 3.33 c/deg data with  $P_{in}$  and  $P_{out}$  set to zero (no phase shift). The squares, which hover around  $Off = 0$ , are for radius of curvature = infinity. Triangles are for radius of curvature =  $\pm 6$  deg, and they follow closely (but not exactly) the prediction for a circle (i.e.,

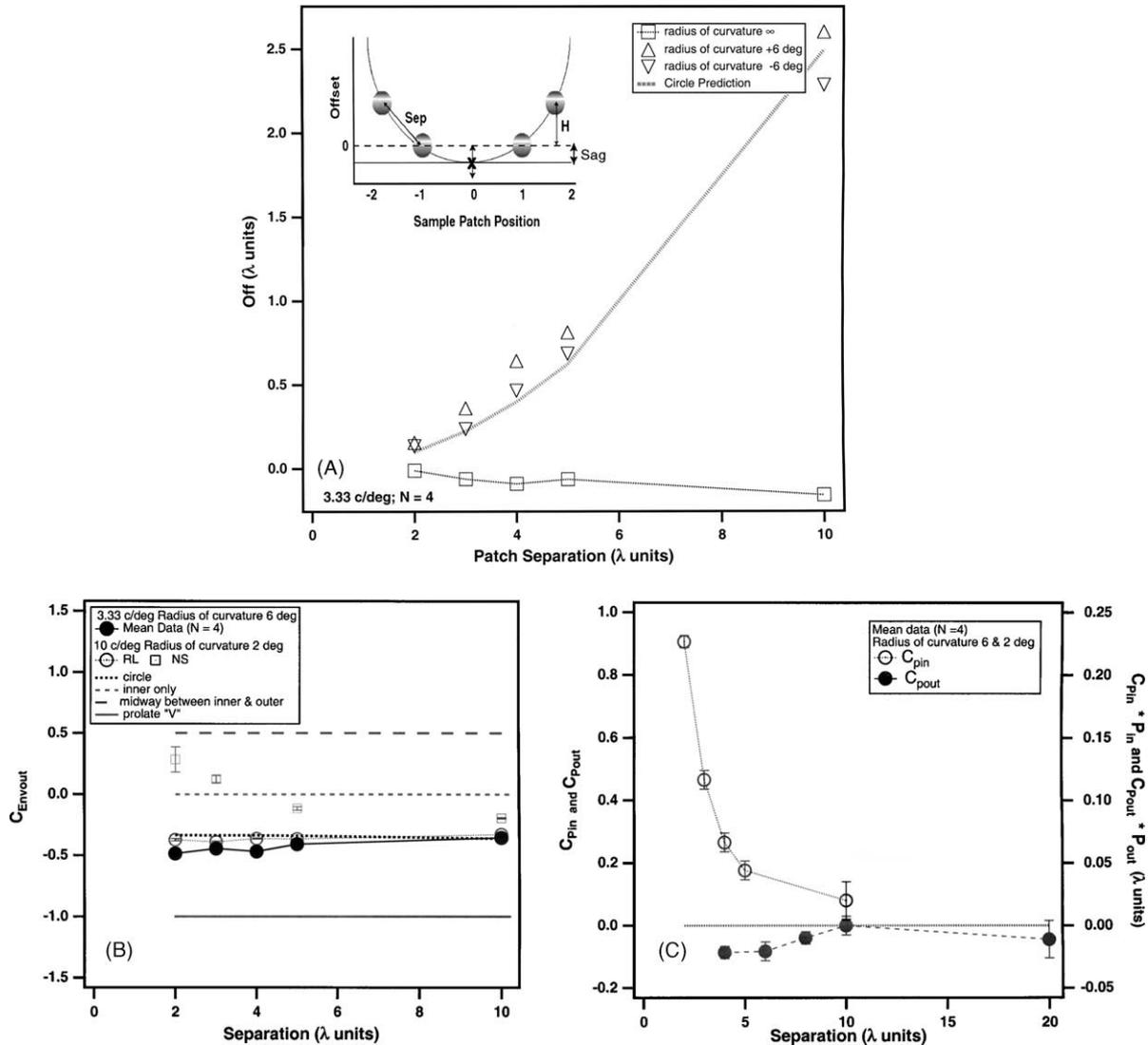


Fig. 8. Inset: A parabolic curve with four sample patches, each separated by distance “sep”. The X represents the intended position of the “test” patch (i.e., on the curve defined by the four sample points). The offset (indicated by the arrow) represents the observers’ setting of the test patch position. We define Off=0 at the height of center of the Gaussian envelope of the two inner sample points, in order to compute the weightings of the remaining cues (see text for details). (A) Shows the values of Off (specified in  $\lambda$  units) from the linear regression fits to the mean data ( $N = 4$  observers) for the 3.33 c/deg data with  $P_{in}$  and  $P_{out}$  set to zero (no phase shift). The squares are for radius of curvature = infinity. Triangles are for radius of curvature =  $\pm 6$  deg. The thick dotted line is the prediction for a circle (i.e., the sag). (B) Filled circles show the weighting of the envelope of the outer patches ( $C_{Envout}$ ) for the 6 deg curvature (3.33 c/deg), averaged across observers and plotted as a function of separation. The dashed line at  $C_{Envout} = 0$  indicates performance if the observer completely ignored the outer patches (i.e., if the setting of the test patch were completely determined by the envelope of the inner patches). The solid line at  $C_{Envout} = -1$  corresponds to the extreme prolate case of a “V”. The large-dashed line at  $C_{Envout} = 0.5$  indicates performance if the observer set the test patch to line up with the average of the inner and outer envelope positions. The black dotted line at  $C_{Envout}$  approximately  $-0.33$  shows the predicted weighting of the envelope of the outer patches if the observer set the test patch on the contour. (C) Left ordinate: The regression coefficients of the inner ( $C_{Pin}$ —open circles) and outer ( $C_{Pout}$ —solid circles) patch carriers, averaged across both observers and viewing distances. For the inner patch carrier, the coefficients (gain) approaches 1 at the smallest separation, and falls off rapidly with increasing separation, although it is not yet zero at  $10\lambda$ . The outer patch carrier has a small but significant weighting in the opposite direction, which also decreases with separation. The right ordinate shows the inner and outer carrier coefficients multiplied by the phase shift and plotted in  $\lambda$  units.

the sag, shown by the thick dotted line). The small deviations from the prediction represent the standing errors described earlier.

The filled circles in Fig. 8B show  $C_{Envout}$  for the 6 deg radius of curvature (3.33 c/deg), averaged across observers and plotted as a function of separation. The in-

dividual data are closely similar to the mean ( $\approx -0.43$  across observers and separations). The horizontal lines in Fig. 8B show four predictions: (1) The dashed line at  $C_{Envout} = 0$  indicates performance if the observer completely ignored the outer patches (i.e., if the setting of the test patch were lined up with the envelope of the

inner patches); (2) The large-dashed line at  $C_{\text{Env}_{\text{out}}} = 0.5$  indicates performance if the observer set the test patch to line up with the average of the inner and outer envelope positions; (3) The black dotted line shows the predicted weighting of the envelope of the outer patches if the observer set the test patch on a circular contour (i.e.,  $C_{\text{Env}_{\text{out}}} = \text{sag}/H$ ). This predicted weighting is approximately  $-1/3$ . Settings above the dotted line reflect oblate errors (squashed circle,  $C_{\text{Env}_{\text{out}}} > \text{sag}/H$ ), while settings below the dotted line reflect prolate errors (pointy circle,  $C_{\text{Env}_{\text{out}}} < \text{sag}/H$ ); (4) The solid line at  $-1$  corresponds to the extreme prolate case of a “V”.

Our results clearly rule out predictions one, two, and four and show that observers indeed weight the positions of the envelope of the outer patches in order to judge the contour shape nearly veridically on a circle, but on average, with an error in the prolate direction. These results, in agreement with Hon et al. (1997) show that observers are quite good at interpolating a parabolic or circular sampled contour. Hon et al. reported that the “influence field” was scale invariant in three of their four observers. We tested scale invariance in two observers by making the measurements at two viewing distances. For RL (open circles in Fig. 8B),  $C_{\text{Env}_{\text{out}}}$  was approximately the same, and nearly veridical at both distances. Observer NS (open squares), showed much less reliance on the outer patch envelope at the larger viewing distance, and for the smaller separations, she seems to have relied almost exclusively on the inner patches.

Fig. 8C shows the coefficients of the inner ( $C_{P_{\text{in}}}$ —open circles) and outer ( $C_{P_{\text{out}}}$ —solid circles) patch carriers, averaged across both observers and viewing distances. For the inner patch carrier, the coefficients (gain) approaches 1 at the smallest separation, and falls off rapidly with increasing separation, although it is not yet zero at  $10\lambda$ . Interestingly, the outer patch carrier has a small but significant weighting ( $\approx -0.1$ ), in the opposite direction, which also decreases with separation. These weightings are also shown in Fig. 7 (gray symbols—right-hand ordinate), where they have been multiplied by the induced phase shift in order to show the gain in  $\lambda$  units. The ratio of weightings of the outer patch to inner patches at  $4\lambda$  is approximately  $-1/3$ , just as would be expected for a circular shape. These values, derived from our regression analysis are similar to the induced phase shifts.

#### 4. Discussion

Our results, in agreement with Hon et al. (1997) show that observers are good at interpolating the shape of a sampled curved contour. Further, our regression analysis shows that observers use both the envelope and phase of all of the samples on the contour in making

their judgements. Our “phase” capture results show the influence function of the carrier information when the envelope is aligned with the contour. A key question in sensory science is how cues are combined to form a perceptual decision. Recent work suggests a rather simple model. For example, Ernst and Banks (2002) suggest that haptic and visual cues are combined by a maximum-likelihood integrator that weights the inputs by their inverse variances. Other work suggests similar cue combination rules operate in other domains, e.g. stereopsis (Young et al., 1993). Here we are interested in how cues to shape are combined.

##### 4.1. How are carrier and envelope information combined in spatial vision?

In our experiments, the contour is defined by two cues—the cue provided by the Gabor carrier (we call this the ‘feature’ cue) and that defined by the Gaussian envelope (the ‘envelope’ cue). The most straightforward way for an ideal (or human) observer to perform the task is to first combine the two cues for patch location, to obtain an estimate of the shift of the inner and outer patches,  $s_{\text{in}}$  and  $s_{\text{out}}$ . Based on the values of  $s_{\text{in}}$  and  $s_{\text{out}}$  the observer adjusts the offset of the middle patch to place it on the circle. Since the envelopes are always fixed on the circle with zero shift the locations of each patch following cue combination are:

$$s_{\text{in}} = w_{P_{\text{in}}} P_{\text{in}} \quad (2)$$

and

$$s_{\text{out}} = w_{P_{\text{out}}} P_{\text{out}} \quad (3)$$

where  $w$  is the cue combination weighting of the phase cue. For the aligned case where the patches are all very close to being on a straight line, a parabolic approximation to the circle allows the offset of the middle patch to have a simple relationship to the locations of the inner and outer patches:

$$\text{Off} = s_{\text{in}} + 1/3(s_{\text{in}} - s_{\text{out}}) \quad (4)$$

$$= 4/3s_{\text{in}} - 1/3s_{\text{out}} \quad (5)$$

By comparing Eqs. (2), (3) and (5) with Eq. (1) we see that:

$$C_{P_{\text{in}}} = 4/3w_{\text{in}} \quad (6)$$

and

$$C_{P_{\text{out}}} = -1/3w_{\text{out}} \quad (7)$$

We can generalize this to the curved case by making a small perturbation,  $p$ , of the inner or outer position in the direction of the phase shift. A new circle is fit to the new inner and outer patches. The required shift of the central patch to put it onto the new circle is calculated. The influence function coefficients,  $C$ , are given by the

ratio of the central patch shift to the perturbation,  $p$ . To a good approximation the influence coefficients are:

$$C_{P_{out}} \approx -1/3 - (\text{sep}/3r)^2 \quad (8)$$

and

$$C_{P_{in}} = -1 - C_{P_{out}}$$

where  $\text{sep}/r$  is the patch separation divided by the radius of curvature. For a separation of  $\text{sep} = 10\lambda$  and a radius of curvature of  $r = 20\lambda$  the exact values are  $C_{P_{out}} = -0.3636$  and  $C_{P_{in}} = 1.3636$ . The approximation given by Eq. (8) works well in this case which is the most extreme case in our data (the largest value of  $\text{sep}/r$ ).

Can we use the variance in each of these cues to predict the cue combination weights,  $w$ ? An ideal observer would weight the cues by their inverse variance, so that the more reliable cues are given a stronger weight. Fig. 9 shows the threshold ( $\sqrt{\text{variance}}$ ) for carrier sensitivity obtained from sine-wave Vernier thresholds of Whitaker (1993) and our own unpublished data using sinusoidal ribbons (see Levi et al., 2000 for stimulus details). The dotted diagonal line shows the best fitting straight line to the (log) threshold vs. separation data (with each specified in  $\lambda$  units). To obtain an estimate of the variance, we squared the threshold values estimated from the linear regression fit. Since these data were obtained with sinusoidal gratings or ribbons (with

no envelope) they reflect the variance of the carrier information alone. Fig. 9 shows that thresholds rise rapidly with separation of the ribbons, exhibiting a 10-fold increase as separation increases from the abutting case to a separation of three grating periods ( $3\lambda$ ). The results of Whitaker et al. (2002) provide a rich data source for computing the variance for the envelope cue over a wide range of separations. As seen in their Fig. 3, the envelope cue has a constant threshold of approximately 0.15 times the Gaussian standard deviation ( $\sigma$ ) for separations up to  $\approx 15\sigma$ . For our stimuli with  $\sigma = 2\lambda/3$ , this translates to a threshold of  $\approx 0.1\lambda$ , and a variance of  $\approx 0.01\lambda^2$  (i.e.,  $0.1^2$ ) for separations up to about  $10\lambda$ . The thick black curves labeled “ideal observer” in Fig. 7 shows the PSE prediction ( $\text{PSE}_p$ ) obtained by weighting each cue by its inverse variance:

$$w = (1/V_c + 0 * 1/V_{Env}) / (1/V_c + 1/V_{Env}) \quad (9)$$

$$\text{PSE}_{P_{in}} = 4/3 * -0.25w_{in} \quad (10)$$

$$\text{PSE}_{P_{out}} = -1/3 * -0.25w_{out} \quad (11)$$

where  $V_c$  and  $V_{Env}$  are the variances of the carrier and envelope cues respectively and where  $-0.25$  is the phase shift of the carrier. The envelope cue has a zero offset because when calculating PSE we have subtracted off the influence of the envelope shift. The prediction provides a reasonable description of our results up to a separation of about  $3.5\text{--}4\lambda$ .

Interestingly, however, our observers continue to show some sensitivity to the phase shift at larger separations. We suspect that the low sensitivity of the carrier prediction at large separations is based on the use of sinusoids with many cycles (Whitaker, 1993; our unpublished data), rather than on localized patches with just a few visible cycles. For localized patches the phase shift is discriminable even with large separations, an impossibility for Gabor functions with many cycles.

## 5. Summary

- Human observers (with normal foveal vision) often demonstrate standing errors in the perception of shape. These perceptual errors tend to increase with separation, and are larger with curved contours than with straight lines.
- Our results, in agreement with Hon et al. (1997), show that observers are good at interpolating the shape of a sampled curved contour. Further, our regression analysis shows that observers use both the envelope and phase of all of the samples on the contour in making their judgements.
- Shifting the *inner* pair of patches by 0.25 cycle results in almost complete phase capture (weight near 1) at the smallest separation ( $2\lambda$ ), and the effect falls off

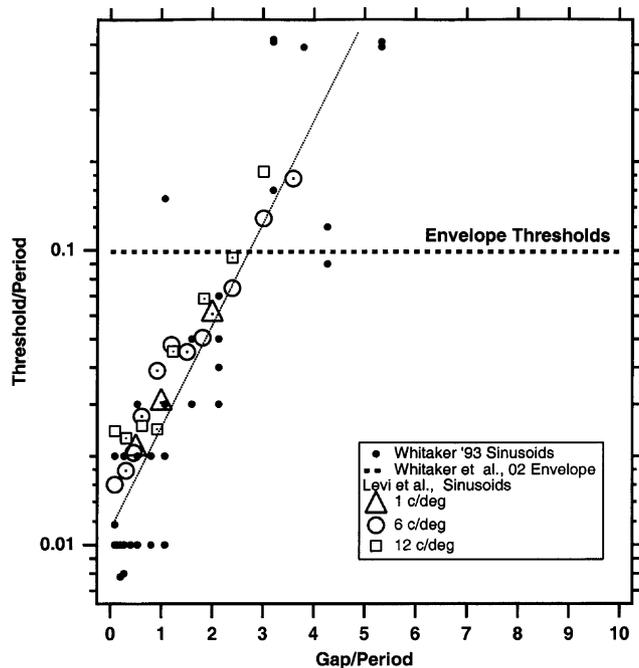


Fig. 9. Threshold for carrier sensitivity obtained from sine-wave Vernier thresholds of Whitaker (1993) and our own unpublished data using sinusoidal ribbons (see Levi, Klein, & Carney, 2000 for methods). The dotted line shows the best fitting straight line to the (log) threshold vs. separation data (with each specified in  $\lambda$  units). The thick dashed line is the envelope threshold from Whitaker et al. (2002). For our stimuli, this translates to a threshold of  $\approx 0.1\lambda$ .

more or less linearly with separation until the effect is nearly zero. Surprisingly, even at  $10\lambda$ , the shift of the inner patches results in a small residual change in perceived position.

- Shifting the *outer* pair of patches by a 0.25 cycle has a much smaller effect in the opposite direction (repulsion), as predicted by an ideal observer.
- In our experiments, the contour is defined by two cues—the cue provided by the Gabor carrier (the ‘feature’ cue) and that defined by the Gaussian envelope (the ‘envelope’ cue). The variance in each of these cues predicts our results for normal observers quite accurately.

### Acknowledgements

We are grateful to Hope Queener for programming the experiments. This research was supported by Research grants R01EY01728 and R01 EY04776, and by a Core Center Grant P30EY07551, from the National Eye Institute, NIH.

### References

- Akutsu, H., McGraw, P. V., & Levi, D. M. (1999). Alignment of well-separated patches: multiple location tags. *Vision Research*, *39*, 789–801.
- Bonneh, Y., & Sagi, D. (1998). Effects of spatial configuration on contrast detection. *Vision Research*, *38*, 3541–3553.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local “association field”. *Vision Research*, *33*, 173–193.
- Hess, R. F., Dakin, S. R., & Badcock, D. (1994). Localization of element clusters by the human visual system. *Vision Research*, *34*, 2439–2451.
- Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science*, *298*, 1627–1630.
- Hon, A. K., Maloney, L. T., & Landy, M. S. (1997). The influence function for visual interpolation. *SPIE*, *3016*, 409–419.
- Jacobs, R. A. (2002). What determines visual cue reliability. *Trends in Cognitive Sciences*, *6*, 345–350.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*, 843–856.
- Keeble, D. R., & Hess, R. F. (1998). Orientation masks 3-Gabor alignment performance. *Vision Research*, *38*, 827–840.
- Keeble, D. R. T., & Hess, R. F. (1999). Discriminating local continuity in curved figures. *Vision Research*, *39*, 3287–3299.
- Klein, S. A. (2001). Measuring, estimating and understanding the psychometric function. A commentary. *Perception & Psychophysics*, *63*, 1421–1455.
- Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Science USA*, *90*, 7495–7497.
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, *35*, 389–412.
- Levi, D. M., & Klein, S. A. (2000). Seeing circles. *Vision Research*, *40*, 2329–2339.
- Levi, D. M., & Klein, S. A. (2002). Classification images for detection and position discrimination in the fovea and parafovea. *Journal of Vision*, *2*, 46–65.
- Levi, D. M., Klein, S. A., & Carney, T. (2000). Unmasking the mechanisms for Vernier acuity: evidence for a template model for Vernier acuity. *Vision Research*, *40*, 951–972.
- Levi, D. M., Klein, S. A., Sharma, V., & Nguyen, L. (2000). Detecting disorder in spatial vision. *Vision Research*, *40*, 2307–2327.
- Levi, D. M., & Sharma, V. (1998). Integration of local orientation in strabismic amblyopia. *Vision Research*, *38*, 775–781.
- Levi, D. M., & Westheimer, G. (1987). Spatial-interval discrimination in the human fovea: What delimits the interval? *Journal of the Optical Society of America A*, *4*, 1304–1313.
- Morgan, M. J., & Aiba, T. S. (1985). Vernier acuity predicted from changes in the light distribution of the retinal image. *Spatial Vision*, *1*, 151–161.
- Morgan, M. J., & Glennerster, A. (1991). Efficiency of locating centres of dot-clusters by human observers. *Vision Research*, *31*, 2075–2083.
- Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2003). A linear cue combination framework for understanding selective attention. *Journal of Vision*, *3*, 116–145.
- Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research*, *38*, 865–879.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, *33*, 993–999.
- Saarinen, J., & Levi, D. M. (1999). The effect of contour closure on shape perception. *Spatial Vision*, *12*, 227–238.
- Saarinen, J., & Levi, D. M. (2001). Integration of local features into a global pattern in human vision. *Vision Research*, *41*, 1785–1790.
- Saarinen, J., Levi, D. M., & Shen, B. (1997). Integration of local pattern elements into a global shape in human vision. *Proceedings of the National Academy of Science USA*, *94*, 8267–8271.
- Toet, A., Smit, C. S., Nienhuis, B., & Koenderink, J. J. (1988). The visual assessment of the spatial location of a bright bar. *Vision Research*, *28*, 721–737.
- Watt, R. J., Morgan, M. J., & Ward, R. M. (1983). Stimulus features that determine the visual location of a bright bar. *Investigative Ophthalmology & Visual Science*, *24*, 66–71.
- Watt, R. J., Ward, R. M., & Casco, C. (1987). The detection of deviation from straightness in lines. *Vision Research*, *27*, 1659–1678.
- Westheimer, G., & Mckee, S. P. (1977). Integration regions for visual hyperacuity. *Vision Research*, *17*, 89–93.
- Whitaker, D. (1993). What part of a vernier stimulus determines performance? *Vision Research*, *33*, 27–32.
- Whitaker, D., Bradley, A., Barrett, B. T., & McGraw, P. V. (2002). Isolation of stimulus characteristics contributing to Weber’s law for position. *Vision Research*, *42*, 1137–1148.
- Whitaker, D., & McGraw, P. V. (1998). Geometric representation of the mechanisms underlying human curvature detection. *Vision Research*, *38*, 3843–3848.
- Whitaker, D., McGraw, P. V., & Levi, D. M. (1997). The influence of adaptation on perceived visual location. *Vision Research*, *37*, 2207–2216.
- Whitaker, D., & Walker, H. (1988). Centroid evaluation in the Vernier alignment of random dot clusters. *Vision Research*, *28*, 777–784.
- Wilkinson, F., Wilson, H., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, *38*, 3555–3568.

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.

Zanker, J. M., & Quenzer, T. (1999). How to tell circles from ellipses: perceiving the regularity of simple shapes. *Naturwissenschaften*, *86*, 492–495.