PERGAMON

Vision Research 40 (2000) 1969–1978

Vision Research

www.elsevier.com/locate/visres

Surface integration influences depth discrimination

Carol Yin [a,*], Philip J. Kellman [a], Thomas F. Shipley [b]

[a] *University of California, Los Angeles, CA 90024, USA*
[b] *Temple University, Philadelphia, PA 19122, USA*

Received 21 September 1999; received in revised form 14 September 2000

Abstract

Image fragments arising from partial occlusion may be perceptually unified by a surface integration process on the basis of similar color or texture. In a new objective measure pitting surface feature similarity against binocular disparity, observers discriminated whether a colored circle had either crossed or uncrossed disparity relative to a surrounding gray rectangle. Sensitivity to disparity was impaired only when (1) the configuration of the other surface fragments in the display supported the integration of a surface behind the rectangle and circle, and (2) matched the color of the central circle. Results were consistent with the hypothesis that a surface integration process integrated similarly-colored surface fragments into a smooth surface, even when those fragments were at different depths. Surface integration caused small and reliable effects on depth perception despite unambiguous disparity information. Perceived depth does not depend solely upon disparity, and may be determined after three-dimensional figural unity is established. © 2000 Published by Elsevier Science Ltd. All rights reserved.

Keywords: Surface perception; Form and shape perception; Amodal completion; Depth perception; Perceptual grouping; Stereopsis

1. Introduction

Visual perception in a three-dimensional (3D) environment is complicated by the fact that objects often partly occlude one another, blocking portions of some objects from an observer's line of sight. Variations in lighting add another twist to this situation, as portions of object contours may become invisible due to low contrast between the object and its background. In both these situations, an object's two-dimensional (2D) shape is not fully specified in its optical projection. Processes of visual completion or integration[1] allow observers to derive representations of these objects from incomplete or fragmented information.

Two sources of information that can be used in visual integration come from edges and surface features (e.g. color and texture) in the image. Although edge processes have received much attention, vision researchers have also hypothesized a complementary set of processes based on surface features (Yarbus, 1967; Grossberg & Mingolla, 1985; Livingstone & Hubel, 1987; Kellman & Shipley, 1991; Olzak & Thomas, 1992). There is some evidence that these two complementary surface- and edge-based systems may have their physiological substrates in the blob and interblob systems, respectively (Livingstone & Hubel, 1987, but see also Tootell, Silverman, Hamilton, De Valois & Switkes, 1988; Tootell, Silverman, Hamilton, Switkes & De Valois, 1988).

Of the two, the role of edges in visual completion has been better understood. For example, Kellman and

* Corresponding author. Present address: Computation and Neural Systems, Division of Biology, MC 139-74, California Institute of Technology, Pasadena, CA 91125, USA. Tel.: +1-626-3952359; fax: +1-626-8444514.

E-mail address: cyin@caltech.edu (C. Yin)

[1] The term 'visual completion' is problematic because it has been used by different researchers to mean different processes (Pessoa, Thompson & Noe, 1998; Yin, 1998) as well as to imply something about the completeness of the unspecified region (e.g. whether there is an analog representation of the occluded region in striate cortex). We have previously used 'visual completion' to mean that certain visible areas separated by gaps are represented as a unified object or surface, but have remained agnostic as to the nature of the occluded region in that representation. In this paper, we continue to use 'visual completion' to refer to the general class of phenomena, but we introduce the term 'surface integration' to more accurately describe a *surface-feature* based process that unifies image fragments without implying anything about the nature of the unified representation.

PII: S0042-6989(00)00047-X

Shipley's (1991) theory of edge interpolation provides a quantification of the Gestalt law of good continuation. Visual completion of edges is achieved when the interpolated path can be a smooth, monotonic curve. When such a curve is possible, the visible edges are *relatable*. Relatability is most likely between collinear edges, and breaks down for interpolation of acute angles (see Kellman & Shipley, 1991, for a formal definition of relatability). Field, Hayes and Hess (1993) found psychophysical evidence consistent with this same geometry, which they termed an 'association field'. Within this field, integration of collinear edge elements is preferred over less collinear elements.

Although considered in the Gestalt psychologist tradition (Koffka, 1935; Michotte, Thinés & Crabbé, 1964), the role of surface features in visual completion, in contrast, has remained largely neglected, perhaps because most studies of visual completion phenomena have focused on only one of several goals of visual completion, that of recovering the shape of the occluded region. But visual completion processes also must produce other outputs such as the figural unity of the object from disparate image fragments and the perceived lightness and color of the object (Yin, 1998). This paper addresses how we use surface features to determine figural unity: Which image fragments in the retinal projection belong to the same object?

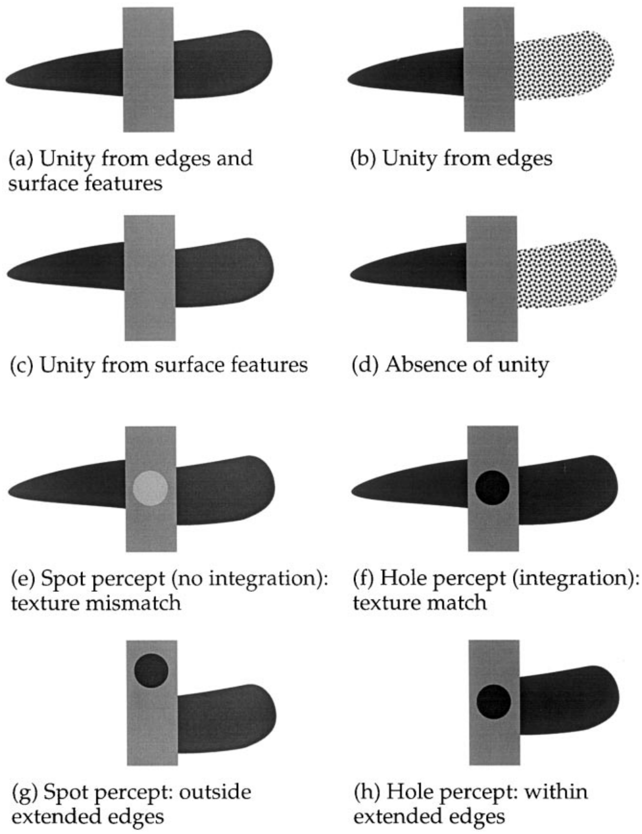

Fig. 1. The recovery of figural unity and shape results from the interaction of edge and surface features. (a) Depicts single black object that is partly occluded by a gray rectangle. Unity is established by both edge and surface feature information. In (b), edge relationships give some impression of unity despite mismatching surface features. In (c), edge relationships no longer provide figural unity, but matching surface features do. In (d), absence of matching surface features destroys any remaining sense of unity. In (e), the mismatching circle is not integrated with the black object and thus looks more like a 'spot'. In (f), the circle is successfully integrated and looks like a hole on the occluder that reveals a part of the occluded black object. In (g) and (h), successful integration of the circle with the partly-occluded black object depends upon whether the circle lies outside (g) or within (h) the extended edges of the object.

We use the term *surface integration* to name this process by which retinally separated image regions of a partly visible object are unified based solely on similarities of their surface qualities (e.g. color or texture). This process differs from those previously studied because prior studies involving surface features focused upon how surface features are 'filled-in' across a single region such as for stabilized images or across the blind spot (Yarbus, 1967; Ramachandran, 1992, 1993) to produce the percepts of color and lightness and texture. For example, in Grossberg and Mingolla's (1985) model of visual perception (see also Grossberg, 1994), surface features are responsible for providing a visible percept of a surface, but plays no role in visual integration. Kellman and Shipley (1991) proposed that surface features may play a role in unit formation, and that a surface integration process may complement the edge process by integrating image fragments of similar surface color or texture within interpolated or extended edges (see Fig. 1e–h).

Surface integration processes make a less obvious contribution to visual integration, compared with edge integration. Edge relationships can provide shape and unity information even when surface feature information does not support a unified percept (Kellman & Shipley, 1991). Surface features, however, can provide information about unity, but generally do not provide shape information about the occluded region, as illustrated in Fig. 1c (see also Yin, 1998).

Our previous work has provided some initial evidence for the surface integration process (Yin, Kellman & Shipley, 1997). In a paired-comparison task, observers were asked to judge which of two displays had a more 'realistic' hole. Displays always contained a circular area on a gray rectangle. Adjacent to the rectangle were one or two bars, appearing partly occluded by the rectangle. A realistic hole percept would result if the central circle was integrated onto the same surface as that of the occluded bar or bars (see Fig. 1f, h). Our data supported the hypothesis that two factors influenced successful surface integration: the surface features of the image fragments must match, and the edges of the surfaces must support the amodal extension or interpolation of the surface behind the occluder.

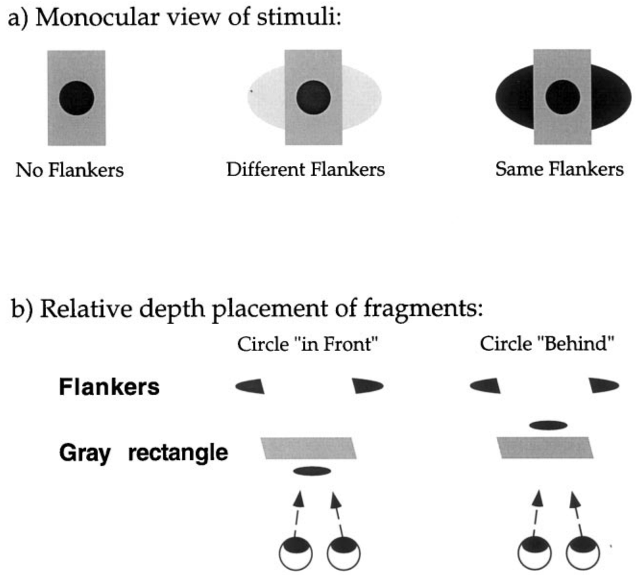

Fig. 2. These are monocular views of the no-flanker, different-flanker, and same-flanker stimuli used in Experiment 1. Observers in our experiments see these stimuli as stereograms. The bottom display (b) shows the depth stratification of the stimuli from a top view, with the observers at the front and the hemi-ellipses at the rear.

To provide objective evidence for this hypothesis, this paper describes a new measure of surface integration that employs a depth-discrimination task. We used a signal detection analysis to measure sensitivity (d') on a depth-discrimination task, where the primary depth information was given by binocular disparity. Lower sensitivity to disparity served as an indicator of successful surface integration. The rationale behind this paradigm comes from computational approaches to vision: surfaces of objects tend to have smoothly changing depths. Marr and Poggio's (1976) stereo-matching algorithm utilized this smoothness constraint by facilitating the activation of points along a smooth surface and inhibiting points that were not on that surface's depth plane (see also Pollard, Mayhew & Frisby, 1985). This constraint reflects a property of surfaces in the world, and the visual system is likely to utilize this constraint in surface integration. Specifically, a surface integration mechanism may assign similarly-colored surface fragments to a common depth plane. This process may interfere with the perception of that fragment's veridical depth, decreasing depth sensitivity (see also Mitchison & Westheimer, 1984).

Our results provide evidence for a surface integration process that impairs depth discrimination by integrating surface fragments with similar features to the same depth plane. The results have two major implications. First, our findings suggest that surface integration occurs via the coordinated action of both surface feature- and edge-based processes. Second, our findings show that in stereo vision, depth assignment may operate upon integrated visual representations and compete with disparity information of local image fragments. We elaborate upon these ideas in the general discussion.

2. Experiment 1

Observers were shown a bounded circular surface and asked to judge the depth of that circle relative to a surrounding bounded rectangular surface, as illustrated in Fig. 2. In some conditions, two half-ellipses ('flankers') appeared to form an oval surface behind the rectangle. The flankers' color either matched or mismatched the circle's. When it matched, surface integration was predicted to integrate the circle with the oval, resulting in an incorrect perception of the circle's depth.

2.1. Methods

2.1.1. Participants

Participants were 30 male and female undergraduates from the University of California, Los Angeles, who served in partial fulfillment of requirements for the introductory psychology course. All participants had normal or corrected vision. All participants also had normal stereopsis, as assessed by a depth discrimination test described below. All participants were naive as to the purpose of the experiment.

2.1.2. Design and stimuli

The stimuli consisted of a blue or yellow circle, placed either in front or behind a gray rectangle. Additionally, some stimuli appeared to have a blue or yellow oval placed behind the gray rectangle.

The circles subtended 1.5°, and were centered upon gray rectangles of $2.5 \times 4.18°$ angles. The circles were given near (crossed) or far (uncrossed) binocular disparity of 1.43 min arc. The flankers (half-ellipses) subtended 6° to either side of the rectangle and were always given 5.7 min arc far disparity relative to the gray rectangle. Fig. 1b illustrates the relative depth placements of these elements.

A block of 12 trials was formed by factorial combinations of the following three independent variables: Two levels of circle color (blue or yellow), two levels of circle depth (front or back), and three flanker conditions. In the *same flanker* condition the flanker color matched the circle; in the *different flanker* condition the flanker color differed from the circle; the *no flanker* condition was a control condition.

2.1.3. Procedure and apparatus

Observers were asked to determine whether the circle appeared in front of the gray square (e.g. more like a 'spot'), or whether it appeared behind (e.g. more like a 'hole'). Observers responded by pressing one of two keys on a keyboard. Observers were also instructed to not judge the depth of any other shapes that might appear in the trials, without explicit reference to the possible presence of ovals in the background.

In a depth perception pretest, observers were tested for stereoscopic vision. The test consisted of ten multiple-choice trials where Ss were asked to determine which disk in a row of five disks was at a different depth, and to report whether the disk was nearer or farther than the other disks. All observers exceeded a minimum accuracy of 70% on this test.

In the practice session, observers were shown no-flanker examples of a circle in front and a circle in back, and asked to decide which was which. Observers were then shown the 12 stimuli with depth disparities doubled to exaggerate the depth difference, and to make sure the observers understood what was required of them in the task. Each stimulus was displayed until the observer responded correctly. If the observer responded incorrectly, the computer beeped, and observers were instructed to look again to see why their first response was incorrect. In the practice session, observers went through two blocks of practice trials at exposures of 400, 307, 200 and 107 ms, in that order, with six trials at each exposure, with the auditory feedback.

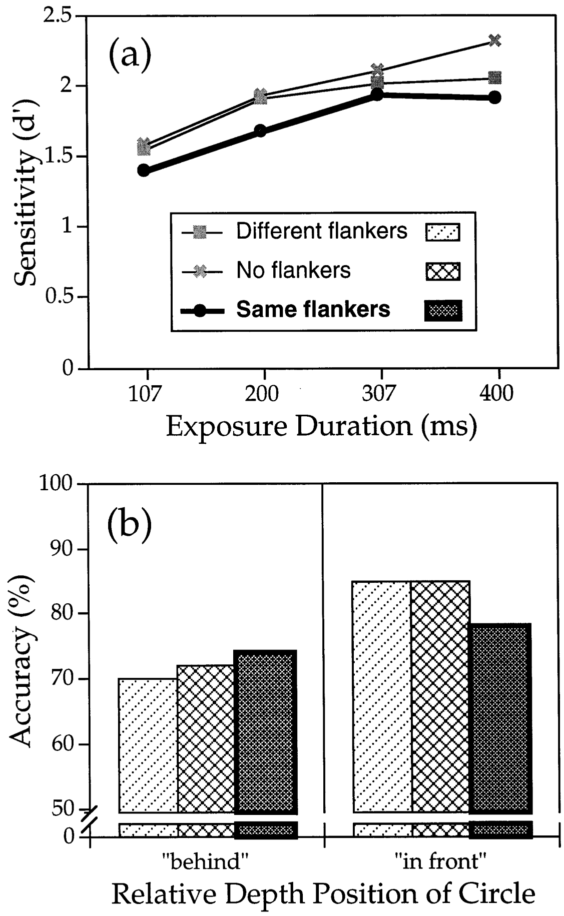

Fig. 3. Observers' sensitivities, as measured by d', are plotted against exposure duration (a). The bottom curve, accented by a thicker line width, shows that performance in the same-flanker condition is poorer than either of the other flanker conditions. The second graph (b) shows that the observers' accuracy's is lowest in the same-flanker condition, but only when the circle is in front. This supports our hypothesis that surface integration utilizes the surface smoothness constraint.

During the experiment, observers received no feedback about their performance. Eight blocks were presented at each of the following exposure durations: 107; 200; 307; and 400 ms, for a total of 384 trials. The exposure durations were randomized between blocks with the constraint that no exposure duration was used again until all other durations were used. Within a block, exposure duration remained constant. Each block of 12 trials began with a fixation cross that disappeared when the observer initiated the first trial with a keypress. Each stimulus was displayed at the center of the screen and removed without masking. Once the observer responded, the next stimulus was presented on the screen. Observers were told to take short breaks if they needed to, but told to wait to do so until the fixation cross indicated the end of the block had been reached.

The presentation of stimuli and the collection of participants' responses were controlled by a program written using MacProbe (Hunt, 1994), run on an Apple Power Macintosh 7100/66. The stimulus images were presented on a Mitsubishi Diamond Plus 20 in. monitor. Responses were entered onto a Macintosh Extended Keyboard. Stereo images were displayed with the CrystalEyes StereoGraphics system (see Lipton, 1991) with LCD shutter goggles.

Before the experiment, we established a baseline performance criterion[2] for including participants' data in the final results. According to this criterion, we excluded ten observers' data from analysis because of low accuracy. This exclusion did not change the overall pattern of results. The data we report are from the 30 observers who were 60% accurate in at least one exposure duration.

2.2. Results

Discrimination sensitivity was measured with d' by arbitrarily designating 'circle-in-back' as the signal and

[2] To be included, an observer had to show accuracy of at least 60% in the no flanker control condition at any exposure duration. Observers who performed at ceiling (100% at all exposure durations) in the no flanker condition were also excluded. Because the stimuli were not masked, we also established a reaction time criterion that excluded any observers whose reaction times exceeded three standard deviations of the mean reaction time. These criteria were necessary because the disparities available to us, given the resolution of our computer monitors, proved to be slightly difficult for our inexperienced observers. However, it is important to note that the exclusion was based on baseline ability to perceive disparity given the constraints of our equipment, not on surface integration abilities.

'circle-in-front' as noise. Sensitivity was lowest in the same-flanker condition, while sensitivity did not differ between the different-flanker and no-flanker conditions.

These results were supported by a 3×4 ANOVA, showing an effect of flanker type [$F(2,58) = 5.86$, $MSe = 0.35$, $P = 0.005$]. A planned comparison found that d' in same-flanker conditions were significantly different from the different-flanker and from the no-flanker conditions [$F(1,29) = 6.19$, $MSe = 0.23$, $P = 0.019$, and $F(1,29) = 8.00$, $MSe = 0.5$, $P = 0.008$, respectively]. The different-flanker and no-flanker conditions did not differ from one another, $F(1,29) = 2.12$, $MSe = 0.31$, $P = 0.156$. The ANOVA also revealed an effect of exposure duration $F(3,87) = 23.26$, $MSe = 0.26$, $P < 0.0001$, but no interaction, $F(6,174) < 1$. Fig. 3a shows these results.

2.2.1. Bias analyses

Bias did not differ with configuration and declined with longer exposure durations. Bias was measured by calculating $\log \beta$ (Wickens, 1998). Observers had a slight bias toward responding 'in front'. A 4×3 ANOVA found an effect of exposure duration [$F(3,87) = 3.09$, $MS = 0.30$, $P = 0.03$], no effect of flanker type [$F(2,58) < 1$], and no interaction [$F(6,174) < 1$]. The mean biases for the 107, 200, 307 and 400 ms conditions were 0.41, 0.31, 0.22 and 0.19, respectively. Post hoc comparisons showed that bias was higher in the 107 ms condition, compared with the 307 ms condition [$F(1,29) = 5$, $MSe = 0.33$, $P = 0.03$] and the 400 ms condition [$F(1,29) = 8.75$, $MSe = 0.27$, $P = 0.006$]. No other significant differences were found at the 0.05 alpha level.

2.2.2. Accuracy analyses

For the displays in which the circle was in front of the gray rectangle, accuracy for the same-flanker displays was significantly lower than accuracy for both the different-flanker and no-flanker displays (see Fig. 3b), as shown by a planned comparison $F(1,29) = 5.05$, $MSe = 0.08$, $P = 0.03$. There was no such difference between flanker types when the circle was 'in back'.

2.3. Discussion

The primary finding of Experiment 1 was that sensitivity was reliably reduced when flankers were present and the circle matched their surface color. This effect did not occur for differently colored flankers or in the absence of flankers. Decreased sensitivity in the same flanker condition may indicate an effect on depth discrimination from a surface integration process. Specifically, sensitivity to disparity was impaired when the circle was in front of the rectangle and matched the color[3] of the ovals, suggesting that this front circle was pulled back toward the depth of the ovals (see Fig. 3b).

The surface integration process seems to be strong enough to interfere with disparity information. These data provide evidence that converges with findings from the perceptual report tasks for a surface integration process (Yin et al., 1997). The effects of surface integration on depth discrimination found here, although modest, are particularly notable given that disparity has a powerful influence on perceived depth. Surface integration does not completely override disparity information, but can exerted a measurable and reliable influence.

However, there is an alternative account for these findings that does not require any surface integration. The surface integration process we have proposed produces an integrated representation of an object when its optical projection is fragmented. Some other grouping process might be possible — one that groups similarly-colored fragments into a similar depth plane, but does not indicate whether the fragments originate from one object or many. This distinction is critical in allowing us to determine whether the image fragments are from one object or are themselves each an object.

Such an alternative account for our findings is suggested by models of surface smoothing in stereopsis (e.g. Mitchison & McKee, 1987a). A surface smoothing process that groups similarly-colored fragments to the same plane, but is not sensitive to shape, could also cause the reduction in sensitivity found in Experiment 1. In this case, however, this process would not produce the percept of figural unity, that these fragments belong to one object. The distinction between surface integration and a space-based grouping mechanism is similar to the one used by Vecera and Farah (1994) who distinguished object-based attention from attention to spatially-grouped elements. In Experiment 2 we test for color grouping by including control displays where surface integration cannot occur due to unrelatable edges.

3. Experiment 2

In order to test for surface *integration* rather than grouping by color, Experiment 2 included the three conditions from Experiment 1 as well as two more control conditions: same reflected-flankers and different reflected-flankers (see Fig. 4). These non-integratable

[3] Because the two colors we chose also differed in achromatic color (i.e. lightness) to exaggerate their differences, the surface integration process may use either chromatic color, achromatic color, or, more likely, a combination of both as the basis for integration.

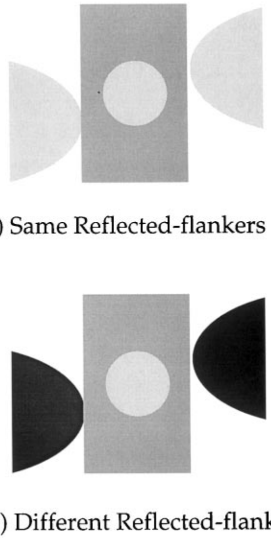

Fig. 4. The additional reflected-flanker stimuli was created for Experiment 2 were created by reflecting the flankers of the same-flanker and different-flanker stimuli from Experiment 1.

flankers were constructed by reflecting each half-ellipse about a vertical axis so that edges were no longer relatable. If the decreased sensitivity is object-based, supporting a surface integration process, then the same reflected-flanker stimuli should not impair depth discrimination because the flankers cannot form a single object. The different reflected-flankers were included to balance the design and to keep the observers from inferring that the current hypothesis involved the same reflected-flankers.

A wider range of exposure durations was also used in Experiment 2, in order to see if the pattern of responses in Experiment 1 continued to develop with increased exposure. Five exposures were used, including one halving the shortest exposure used previously and one doubling the longest exposure used previously.

3.1. Method

3.1.1. Participants

Thirty-seven new male and female participants were recruited from the same pool described above.[4]

3.1.2. Stimuli and design

The original set of 12 stimuli from Experiment 1 were augmented by eight new reflected-flanker stimuli (see Fig. 4). The reflected-flanker stimuli were adapted from the original stimuli by taking the visible portions of the occluded oval (half-ellipses) and reflecting them across the vertical axis. Two other measures were taken to emphasize the distinctness of each half-ellipse: the half-ellipses were vertically shifted in opposite directions (the right half-ellipse upwards and the left half-ellipse downwards) approximately 8 min arc so that they were no longer horizontally aligned. Half-occluded gaps (1.43 min arc) were also added between the half-ellipse and the occluder to enhance their separation. Half-occluded features are unpaired features that appear in one eye's view but not in the other, and have been shown to be strong cues to surface discontinuities (Nakayama & Shimojo, 1990; Shimojo & Nakayama, 1994; Anderson & Nakayama, 1994). These new stimuli were constructed by manipulating three factors: Circle color (yellow and blue), flanker color (yellow and blue), circle depth placement (near or far disparity). In all other respects, the physical characteristics of these reflected-flanker stimuli were the same as the original stimuli.

3.1.3. Procedure and apparatus

The procedure differed from that used in Experiment 1 in a few minor respects: The stimuli in Experiment 1 had been centered upon the screen based on the stimuli's horizontal dimensions. As a consequence the gray occluders in the no-flanker conditions appeared at a slightly farther depth than did the gray occluders in the same- and different-flanker conditions. In Experiment 2 this was corrected so that the gray occluder always appeared at the same convergence angle. A reduction screen was also used to prevent the edges of the monitor from being seen.

We also avoided mentioning 'holes' or 'spots' to remove any possible influence due to cognitive processing. Observers were simply instructed to decide whether the circle was in front of the gray rectangle or behind, and to respond 'front' or 'back'. The mapping of 'front' and 'back' to the left hand or right hand was also counterbalanced across subjects. During the experiment, a small 'F' or 'B', indicating 'front' or 'back' appeared at the appropriate bottom corner of the screen to serve as a reminder to the observers.

All participants had stereo vision, as assessed by their ability to see a random-dot stereogram using the stereo goggles. The display session included the reflected-flanker stimuli, again with the disparity exaggerated. The practice trials consisted of one block of trials, with four stimuli presented at each of the exposure durations as before.

For the experiment, four blocks of 20 trials were presented at each of the following exposure durations: 53; 107; 200; 400; and 800 ms, for a total of 400 trials. The participants' response times were also collected by the computer. All other aspects of the procedure and apparatus were identical to those used previously.

[4] Forty observers participated, but three observers' data were excluded according to the a priori criterion mentioned earlier. Two observers performed at ceiling and another observer was excluded because her mean RT across all conditions was over three standard deviations slower than the other observers. These exclusions did not change the overall pattern of results. The data of the remaining 37 contributed to the results reported here.

Because the predicted finding was that out of the five conditions, the same-flankers alone would decrease sensitivity, it was possible that an overall effect of flankers would not be found. In order to test for the effects found in the previous experiment, and to test the new hypotheses, pairwise comparisons were also planned with the critical comparison being between the same-flanker and the same reflected-flanker conditions.

3.2. Results

An omnibus 5 × 5 ANOVA showed that the flanker effect was marginal, $F(4,144) = 2.14$, MSe = 0.31, $P = 0.079$. There was also an effect of exposure duration, $F(4,144) = 51.29$, MSe = 0.73, $P < 0.0001$, but no interaction between exposure duration and flanker type, $F(16,576) < 1$.

Planned comparisons showed that sensitivity (d') was lower in the same-flanker condition compared with the same reflected-flanker condition, $F(1,36) = 4.62$, MSe = 0.42, $P = 0.038$ (see Fig. 5a). There was a significant difference between the same-flanker condition and the average of the other flanker conditions, collapsed across all exposure durations [$F(1,36) = 7.08$, MSe = 0.33, $P = 0.01$]. The same reflected-flanker condition was also not different from the no-flanker, different-flanker, and different reflected-flanker conditions, all $F(1,36)$'s < 1. These findings suggest that surface integration was responsible for the poorer depth discrimination found in Experiment 1.

For the three conditions that were replicated from Experiment 1, planned comparisons showed that sensitivity differed between the same-flanker and different-flanker conditions $F(1,36) = 5.78$, MSe = 0.23, $P = 0.02$. However, the previously significant difference between the same-flankers and no-flankers became non-significant, $F(1,36) = 2.21$, MSe = 0.37, $P = 0.15$, although the trend is in the direction of the previous finding. No difference was again found between the no-flanker and different-flanker conditions, $F(1,36) < 1$.

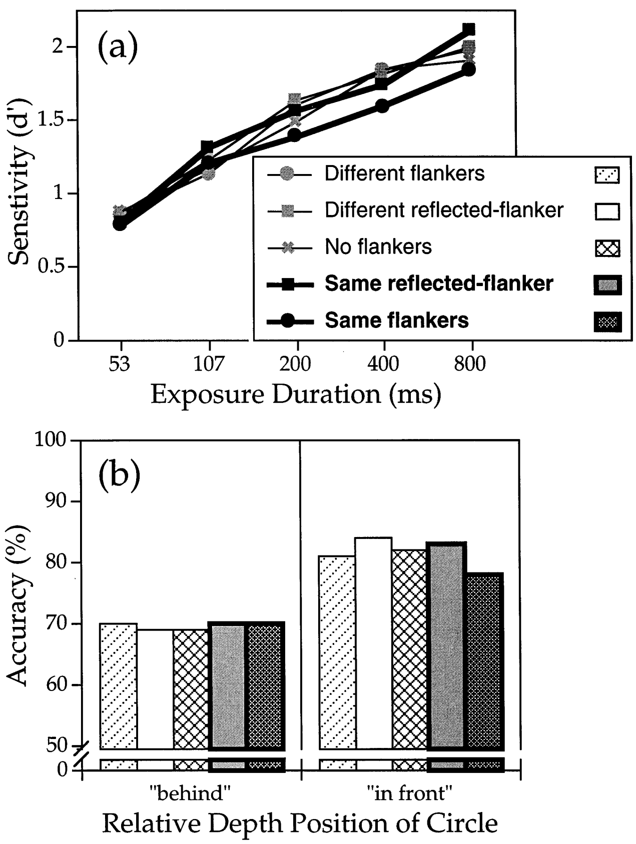

Fig. 5. For both the sensitivity (a) and accuracy (b) graphs, both the same flanker condition and the same reflected-flanker condition are depicted in a thicker line width for ease of comparison. Observers' sensitivities were poorest in the same flanker condition (a). Again, the observers' accuracy's were lowest in the same flanker condition, but only when the circle was in front (b).

3.2.1. Bias analyses

There was also a slight bias toward reporting the circle in front, although no systematic effects were found to be significant at the 0.05 alpha level. The mean biases were 0.18, 0.24, 0.22, 0.18 and 0.21, for the 53, 107, 200, 400 and 800 ms exposure durations, respectively.

3.2.2. Accuracy analyses

For the displays in which the circle was in front of the gray rectangle, accuracy for the same-flanker displays was significantly lower than accuracy for the average of the four other conditions, as shown by a planned comparison $F(1,36) = 4.16$, MSe = 0.07, $P = 0.049$ (see Fig. 5b). As in Experiment 1, there was no such difference between conditions when the circle was in back.

3.3. Discussion

The difference between same flanker and same reflected-flanker conditions suggests that it is indeed an object-based, surface integration mechanism that is responsible for decreased sensitivities, and not grouping of similarly-colored proximal elements.

The difference in sensitivities between the same-flanker and different-flanker conditions found in Experiment 1 was replicated. The difference found earlier between the no-flanker and same-flanker conditions was not replicated, but the trend is in the same direction.

Although we did not systematically query our observers, a number of them, during the debriefing session at the conclusion of the second experiment, voluntarily reported that the circle looked like a hole when the circle and oval colors were identical. This suggests that the circles may have been perceptually as well as behaviorally integrated with the oval.

4. General discussion

These experiments furnish objective support for the existence of a surface integration process that operates on similarly-colored surface fragments. This is an object-based integration process and not a space-based grouping of separate elements on the basis of color similarity. These findings show that surface feature similarity contributes to visual completion and confirm our earlier findings with magnitude estimation and paired comparison procedures (Yin et al., 1997).

In this discussion, we will explore the implications of our findings for theories of stereo vision which suggest that stereo processes operate on visual representations that are higher order than previously thought. We will discuss the implications of our findings for theories of visual completion, specifically the interdependency between surface feature and edge processes. We will also mention some possible neural mechanisms that may mediate surface integration.

4.1. Surface smoothness constraint operates within integrated object representations

Mitchison and his colleagues (Mitchison & Westheimer, 1984; Mitchison & McKee, 1987a,b; Mitchison, 1988) provided psychophysical evidence for a surface smoothness constraint (Marr & Poggio, 1976). They showed that when a stereoscopic display had more than one possible match, the visual system determined depth by interpolating disparity information from the edges of the stereogram. However, their data were obtained with a dot array stimulus depicting a planar surface, which did not allow them to make a distinction between space-based smoothing and object-based smoothing. Our stimuli did allow us dissociate the two: A space-based smoothness constraint would have allowed the grouping of all similarly-colored stimuli to a common depth plane, yielding no difference between our integratable and our non-integratable, reflected-flanker stimuli. Instead, we did find a difference, suggesting that the smoothness constraint can be object-based. Previously, it appeared that stereoscopic vision interpolated depth within 'segmented regions' (Mitchison, 1988), implying that segmentation based on contrast precedes depth interpolation. Our results[5] suggest that surface integration also precedes depth interpolation. This extends Mitchison's work by suggesting depth is interpolated within the boundaries of a single, perceptually unified surface, whether a single homogeneous region, or a set of regions unified by surface integration.

[5] The disparities used in our experiment fall within the range for the depth interpolation found by Mitchison and McKee (1987a,b).

Our data suggest that when a higher order visual representation such as a unified object is available, its information competes with local visual information such as the disparity and boundary ownership of image fragments. A similar relationship between perceptual organization and local information has been found in a variety of experiments. Enns and Rensink (1991) showed that the percept of a cube *preempts* access to the 2D lines depicting that cube (see also Rensink & Enns, 1995). He and Nakayama (1992, 1994) showed that surfaces preempt lower order visual features such as luminance boundaries in visual search and texture segmentation tasks. However, our data differ from prior evidence for preemption in that those studies showed the presence of a higher order representation dominates an ambiguous stimulus that may be accurately represented in two different levels of representation. In our current study, surface integration caused unambiguous disparity information to be misperceived.

4.2. Object unity and shape

The results of the studies reported here also have implications for the sequencing of edge and surface integration in models of vision. For many models, the two integration processes occur in a rigid sequence. For example, in Grossberg's (1994) FACADE model, boundaries are formed first at different disparities, and then surface features are filled in. The filling-in of the surface features subsequently allows the surfaces to be perceived. Our results suggest that the integration process is more interactive and not unidirectional. Boundary perception can be influenced by surface features, and surface feature integration can be influenced by the location of boundaries.

A key component that needs to be incorporated into future models of visual perception is the integrative role of surface features. Surface-feature and edge processes make different contributions in determining an object's unity and shape. During visual completion, one must recover the shape of the illusory or partly-occluded surface, but also must determine which image fragments belong to which object. The distinction between determining unity versus shape was noted early, by Gestalt psychologists (c.f. Koffka, 1935), but most research on visual completion has focused on shape recovery (see Yin et al., 1997). We have previously shown that surface feature similarity can establish unity for 2D image fragments that may be otherwise treated as discontinuous, based on information from edge interpolation. The unity provided by surface integration influenced subsequent 2D edge interpolation processes (Yin et al., 1997). The results of the current studies suggest that we can extend our findings on surface integration to 3D stimuli as well.

4.3. Neural implementation of surface integration

The neural mechanisms of surface integration are not yet understood, and this problem is particularly complicated by the fact that the neural substrates for perception of a continuous surface are themselves not well understood. Perception of continuous surfaces is commonly held to be the output of 'filling-in' processes that spread the properties of the surface from its boundaries (Krauskopf, 1963; Yarbus, 1967). This propagation of signals was presumed to be mediated by the connections between neighboring cells in visual cortex. The problem with this conceptualization is that while there is some psychophysical evidence for this 'filling-in' (Paradiso & Nakayama, 1991; De Weerd, Desimone & Ungerleider, 1998; Maddess, Srinivasan & Davey, 1998; Pessoa & Neumann, 1998; Pessoa, Thompson & Noe, 1998), these processes require time, and may not be sufficiently fast to mediate everyday perception and action. More critically, this 'filling-in' process may not mediate surface integration: filling-in of surface features occurs on the time scale of seconds, not milliseconds. This relatively long time course has been attributed to the need first to segment and remove contours before the 'filling-in' process can operate (De Weerd, Desimone, & Ungerleider, 1998). In the perception of partly occluded surfaces, where intervening contours abound, this scheme may be too slow for normal perception, where humans make saccades every 300 ms or so. However, if we are able to find evidence that the presence of a T-junction along a contour can effectively signal the contour system to 'remove' an intervening (and occluding) contour, then analog propagation of signals might be feasible.

Another candidate for mediating the perception of both continuous and fragmented (illusory and partly-occluded) surfaces relies upon long-range horizontal connections between pyramidal cells in primary visual cortex. Ts'o and Gilbert (1988) found evidence that some color-selective cells in the cytochrome oxidase blobs of V1 had correlated firing rates with color-selective cells that were in other spatially separated blobs. These long range connections may allow gaps or occluders in the image to be spanned to some degree, lessening the problem of how to propagate signals across boundaries for partly-occluded surfaces. Additionally, Ts'o and Gilbert provided some evidence suggesting that color-selective cells may also be connected to color-selective cells that are also sensitive to oriented edges, suggesting that the interplay between edge and surface feature information may occur very early on. While much more research needs to be done to understand the function of the long-range connections, it appears that part of the surface integration problem may be resolved in primary visual cortex.

Most likely, surface integration is the result of several neural processes operating in concert. In addition to integration via long-range horizontal connections in early visual cortex, surface integration might also be achieved via the activity of cells with large receptive fields. Higher visual areas such as V2 and V4 remap the visual field with large, overlapping receptive fields which Lennie (1998) suggests may aid spatial integration. In such an implementation, similarly-colored image fragments might trigger the activity of an extrastriate cortical cell with a large receptive field. Even in such a case, other processes must verify whether or not similarly-colored fragments belong to a single object.

The results from our current studies suggest that unity can be established on the basis of surface features that are available in the projected image. Edge processes that may be operating in parallel must take this unity into account when recovering shape. Our goal in these studies was to show that surface integration processes do contribute to visual completion. Our results pose serious problems for any theory of visual completion that postulates strict hierarchical stages in the recovery of unity, shape and perceptual quality: these processes appear to be interdependent.

Acknowledgements

This research is based on a doctoral dissertation submitted by CY to the University of California, Los Angeles, in partial requirements for a Ph.D. degree. Experiment 1 was presented at the Association for Research in Vision and Ophthalmology Annual Meeting in April of 1996. This work was supported by a UCLA Graduate Division Dissertation Year Fellowship to CY, SBR 9496112 to PJK, and SBR 9396309 to TFS. The authors thank Tom Mullen, Amy Milov, Janice Esguerra, and Ana Esilva for their assistance in data collection. The authors also wish to thank Alexander Grunewald, Ron Rensink and Shinsuke Shimojo for their helpful comments and suggestions during the writing of this manuscript.

References

Anderson, B. L., & Nakayama, K. (1994). Toward a general theory of stereopsis: Binocular matching, occluding contours, and fusion. *Psychological Review*, *101*, 414–445.

De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Response from De Weerd, Desimone and Ungerleider. *Trends in Cognitive Sciences*, *2*, 425–426.

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, *98*, 335–351.

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.

Grossberg, S. (1994). 3-D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, *55*, 48–120.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 173–211.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, *359*, 231–233.

He, Z. J., & Nakayama, K. (1994). Perceiving textures: beyond filtering. *Vision Research*, *34*, 151–162.

Hunt, S. M. J. (1994). MacProbe: a Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods*, *Instruments*, *Computers*, *26*, 345–351.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & Co.

Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterchromatic targets. *Journal of the Optical Society of America*, *53*, 741–744.

Lennie, P. (1998). Single units and visual cortical organization. *Perception*, *27*, 889–935

Lipton, L. (1991). *The crystaleyes handbook*. San Rafael: StereoGraphics Corporation.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*, 3416–3468.

Maddess, T., Srinivasan, M. V., & Davey, M. P. (1998). Response from Maddess, Srinivasa and Davey. *Trends in Cognitive Neurosciences*, *2*, 425.

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

Michotte, A., Thinés, G., & Crabbé G. (1964). Les complements amodaux des structures perspectives. Reprinted in 1991: G. Thinés, A. Costall, & G. Butterworth, *Michotte's experimental phenomenology of perception* (pp. 140–167). Hillsdale: Lawrence Erlbaum Associates Publishers.

Mitchison, G. (1988). Planarity and segmentation in stereoscopic matching. *Perception*, *17*, 753–782.

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research*, *24*, 1063–1073.

Mitchison, G. J., & McKee, S. P. (1987a). The resolution of ambiguous stereoscopic matches by interpolation. *Vision Research*, *27*, 285–294.

Mitchison, G. J., & McKee, S. P. (1987b). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research*, *27*(2), 295–302.

Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. Special issue: optics, physiology and vision. *Vision Research*, *30*, 1811–1825.

Olzak, L. A., & Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision Research*, *32*, 1885–1898.

Paradiso M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, *31* (7–8), 1221–1236.

Pessoa, L., Thompson, E., & Noe, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences*, *21*, 723–802.

Pessoa, L., & Neumann, H. (1998). Why does the brain fill in? *Trends in Cognitive Sciences*, *2*, 422–424.

Pollard, S. B., Mayhew, J. E., & Frisby, J. P. (1985). PMF: a stereo correspondence algorithm using a disparity gradient limit. *Perception*, *14*, 449–470.

Ramachandran, V. S. (1992). Filling in gaps in perception: I. *Current Directions in Psychological Science*, *1*, 199–205.

Ramachandran, V. S. (1993). Filling in gaps in perception: II. Scotomas and phantom limbs. *Current Directions in Psychological Science*, *2*, 56–65.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: evidence for low-level grouping. *Psychological Review*, *102*, 101–130.

Shimojo, S., & Nakayama, K. (1994). Interocularly unpaired zones escape local binocular matching. *Vision Research*, *34*, 1875–1881.

Tootell, R. B. H., Silverman, M. S., Hamilton, S. L., De Valois, R. L., & Switkes, E. (1988). Functional anatomy of macaque striate cortex: III. Color. *Journal of Neuroscience*, *8*(5), 1569–1593.

Tootell, R. B. H., Silverman, M. S., Hamilton, S. L., Switkes, E., & De Valois, R. L. (1988). Functional anatomy of macaque striate cortex: V. Spatial frequency. *Journal of Neuroscience*, *8*(5), 1610–1624.

Ts'o, D. Y., & Gilbert, C. D. (1988). The organization of chromatic and spatial interactions in the primate striate cortex. *Journal of Neuroscience*, *8*, 1712–1727.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology*: *General*, *123*, 146–160.

Wickens, T.D. (1998). *Elementary signal detection theory*. Department of Psychology, University of California, Los Angeles (Unpublished manuscript).

Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum Press (B. Haigh, Trans.).

Yin, C. (1998). A task-oriented taxonomy of visual completion. *Behavioral and Brain Sciences*, *21*, 780–781.

Yin, C., Kellman, P. J., & Shipley, T. F. (1997). Surface completion complements boundary interpolation in the visual integration of partly occluded objects. *Perception*, *26*, 1459–1479.