

## DA VINCI STEREOPSIS: DEPTH AND SUBJECTIVE OCCLUDING CONTOURS FROM UNPAIRED IMAGE POINTS

KEN NAKAYAMA\* and SHINSUKE SHIMOJO†

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, U.S.A.

(Received 10 August 1989; in revised form 17 January 1990)

**Abstract**—Distant surfaces are occluded by nearer surfaces to different extents in the two eyes, leading to the existence of unpaired image points visible in one eye and not the other. An ecological analysis of the real world situation that could have given rise to such unpaired points indicates the presence of a *depth constraint zone*, defined by visibility lines between which possible real world points must lie. The leading edge of this zone starts at the edge of a fused binocular occluding surface and recedes linearly with increases in horizontal distance to the unpaired point. Psychophysical evidence indicates that the human visual system makes use of this unpaired information in a remarkably adaptive manner, showing an increase in perceived depth for increasing horizontal separations between the unpaired target and fused edge, at least over a significant angular range (approx. 25–40 min arc). We also show that unpaired points in binocular images can lead to the formation of subjective occluding contours and surface having the qualitatively appropriate sign of depth. Furthermore, we show that the visual system could not recover depth of unpaired points camouflaged from the other eye against silhouettes. Our findings indicate that the visual system makes use of occlusive relations in the real world to recover depth, contour, and surface from unpaired points. The fact that such processes must utilize eye-of-origin information implies that they share this essential characteristic with classical or Wheatstone stereopsis. The necessity of eye-of-origin information also suggests that the processing may begin relatively early in cortical visual processing, possibly as early as V1. Finally, the novel emergence of subjective occluding contours from unpaired monocular stimuli raises the possibility that this process is mediated by visual experience, built up by the association of unpaired points and occluding contours.

Stereopsis	Binocularity	Depth perception	Subjective contour	Occlusion	Eye of origin
Silhouettes	Visual cognition				

### INTRODUCTION

Recent work on binocular vision suggests that depth from binocularly presented targets depends critically on the solution to the correspondence problem. According to this view, an important theoretical and experimental question is how points visible in one eye get correctly matched to corresponding points in the other eye. To limit the number of possible false matches, most theories have assumed that disparity varies only gradually over space (Sperling, 1970; Julesz, 1971; Nelson, 1975; Marr & Poggio, 1976). Thus, matches which require very large shifts in disparity over a restricted neighborhood are either forbidden or inhibited.

This preoccupation with the correspondence problem, however, glosses over the fact that surface discontinuities are common in daily life and that they contain rich information for object recognition. An inevitable consequence is that many points in everyday binocular scenes can never be matched. Ever since Leonardo da Vinci (see Wheatstone, 1838) at least some researchers have acknowledged, either implicitly or explicitly, that occluding surfaces, the edges of which have a vertically oriented component, can occlude distant surfaces to different extents in the two eyes (see Panum, 1858; Lawson & Gulick, 1967; Kaye, 1978; Gillam & Borsting, 1988; Westheimer, 1986). Thus, there is an interocularly unpaired zone along a surface discontinuity unless the edge is horizontally oriented.

Consider two very common examples where the visual system must deal with such unpaired points. First is the case where a central square is in front of a background (Fig. 1A). Second is

\*Present address: Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, U.S.A.

†Present address: Department of Psychology, College of Arts and Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153, Japan.

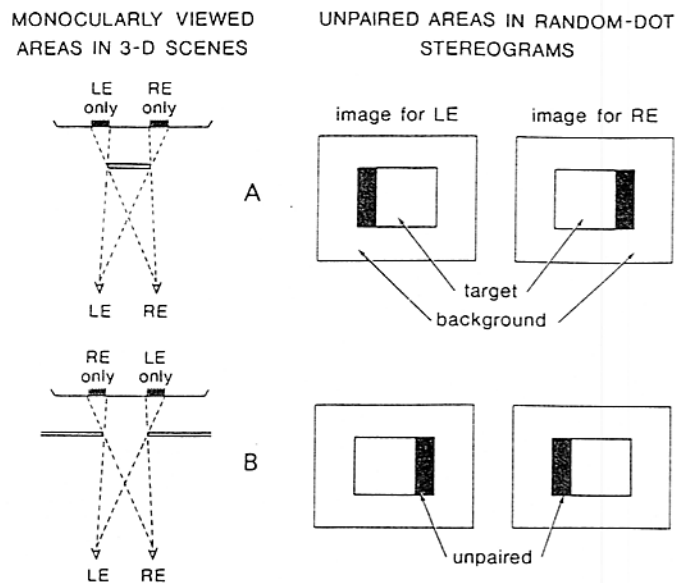


Fig. 1. Top view of two examples where background points in one eye are not visible in the other. (A) Looking at a central square in front of a background. (B) Looking at a distant surface through a central square aperture.

the case where there is a square window through which one sees a background (Fig. 1B). In both cases there are regions in the scene which are visible to one eye and not to the other. These unpaired zones are depicted as shaded areas in Fig. 1.

If we consider an exhaustive set of combinations of depth and monocular areas in Fig. 1, it should be clear that there are just two cases which simulate a real world situation and which are ecologically valid. Right-eye-only points are seen to the right of a closer occluding edge, and left-eye-only points are seen to the left of an occluding edge. The two other cases which are possible to create in stereograms cannot plausibly arise in natural scenes: right-eye-only points seen to the left of an occluding edge and left-eye-only points seen to the right of an occluding edge. Figure 2 summarizes the valid and invalid combinations of interocularly unpaired points in combination with paired regions.

Because unpaired areas have no counterpart in the other eye, it might be expected that they might lead to binocular rivalry and interocular suppression. Yet, given the ubiquity of such unpaired points in natural scenes and their potential usefulness as a rich source of information on three-dimensional layout of surfaces, it might also be expected that some classes of unpaired points would escape rivalry and suppression. Such a question motivated a recent study on binocular rivalry in our laboratory using random dot stereograms (Shimojo & Nakayama, 1990). As predicted from our

hypothesis regarding occlusion, we found that *valid* monocular unpaired regions escaped rivalry whereas *invalid* unpaired regions did not. Another important observation was that the *valid* regions always appeared as a part of the more distant surface, whereas *invalid* regions often appeared more in front and were more ambiguous in depth. Our success in understanding rivalry in terms of real-world occlusion constraints led us in the present paper to a more quantitative analysis of the perception of depth of such monocular points, comparing and measuring perceived depth of these points in stereograms which were either "valid" or "invalid".

Cases	Stimuli		Occlusion Constraints
	Left Eye	Right Eye	
(1)			Invalid
(2)			Valid (Back)
(3)			Valid (Back)
(4)			Invalid

Fig. 2. Occlusion constraint table, showing valid and invalid combinations of unpaired monocular points in relation to fused image regions.

*Ecological analysis of occlusion: depth constraint zone*

To gain a more detailed understanding of the natural constraints associated with occlusion it is important to assess the information available from unpaired points which are embedded in the binocular image. Consider the set of binocular half images seen in the bottom of Fig. 3A, showing a matched pair of squares plus a single right-eye-only bar to the right of this fused square. A first step is to determine the possible positional range of a target in the 3-dimensional real world that could have given rise to such a right-eye-only target, doing this for different horizontal positions. For the moment, let us act as if the binocularly fused surface *S* extends infinitely to the left and is bounded by the right edge shown in Fig. 3A. We identify lines of sight *a, b, c* appearing to radiate from the right eye (corresponding to the different angular positions of the monocular target in this eye) and the line (*Lx*), the left eye visibility line which defines the limits of visibility determined by the occluder. Points to the left of *Lx* are invisible to the left eye. As should be obvious from the figure, a given right-eye-only point representing a given angular direction, say *b*, can only have been created by points lying on *b* which are behind *Lx*. Thus for each horizontal angular position in the right eye, there is a segment of the line of sight which could have given rise to

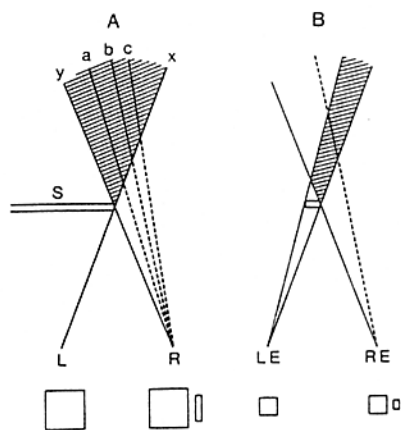


Fig. 3. Occlusion geometry. (A) Bottom: binocular images for right eye only stimulus in relation to fused binocular square. Top: top view of occlusion configuration with visibility lines for each eye (*Lx* and *Ry*) which delineate the regions of visibility of each eye occasioned by the presence of the binocularly viewed occluder (*S*). Depth ambiguity line segments for the right eye (see *a, b, c*) define a depth constraint zone (shaded region). (B) Special case for occluders of limited horizontal extent. Here the depth constraint zone (shaded region) is also bounded in the rear by the left eye visibility line passing along the left edge of the fused surface.

the unpaired point at a given distance from the occluding edge. In other words, each of these line segments represents the degree of depth ambiguity at each distance. If we consider all such segments corresponding to positions in the right eye, they sweep out a depth constraint zone, the locus of possible points which could have given rise to right eye only targets. This is shown as the shaded region in Fig. 3A. The leading edge of this zone recedes linearly with increasing distance from the edge of the binocularly viewed "occluder". As should be clear from this geometrical reconstruction, the recovery of depth is not exact as it is for conventional stereopsis because precise positional information regarding the locus of the unseen image (in the other eye) is lacking.

From this analysis, it should be evident that "valid" monocular points can arise from only limited regions of space, thus providing important one-sided constraints on the depth interpretation for an unpaired monocular point. Technically speaking, however, the constraint is actually two sided because the analysis above assumed that the binocular occluder extended infinitely to the left. If we assume that the occluder is also bounded on the left, then there exists another left-eye visibility line just passing along the left edge of the occluder. This delineates a new region to the left of which contain points which are again visible to the left eye. This further constrains the depth of unpaired targets, so that they cannot arise from even more distant points. Figure 3B shows a view of this more complete occlusion situation in its entirety, indicating that the depth constraint zone is bounded in back as well as in front. It should be noted that a reduction of the horizontal extent of the occluder so that it subtends the same angular width of the unpaired bar, is the well-known Panum's limiting case—two identical targets in one eye, one in the other (Panum, 1858; Westheimer, 1986). By restricting our analysis to extended occluding surfaces we are examining an occlusion situation which is thus related to the Panum case but is more frequently encountered in our natural viewing environment.

An immediate question to ask is whether the human visual system could make use of this implicit information provided by unpaired targets as outlined in Fig. 3. Because the depth constraint zone only provides a range of possible depths, not an exact value of depth, we began our study with a *qualitative* experiment.

We asked whether "valid" monocular points would be more consistently seen as more distant than "invalid" points, for varying horizontal distances from the occluding edge.

#### EXPERIMENT 1: PAIRED DEPTH COMPARISON OF VALID AND INVALID MONOCULAR POINTS

##### Method

*Subjects.* Four subjects, the authors (KN and SS) plus two who were naive (AR and MB), served as observers in the first experiment. Because we were concerned about the possible ill effects of strong eye preference, we began with an eye-dominance test for each subject. In this test, the subject was presented with a random-dot correlogram in which an interocularly-unpaired stimulus was surrounded by an interocularly-paired region, and asked to make a two-alternative, forced-choice judgment as to whether the unpaired region had faded or was suppressed. The results of these series of tests indicated that none of the four subjects showed a statistically significant eye dominance ( $P > 0.5$ ), and therefore, none were eliminated from the main experiment. (For more details of the eye-dominance test, see Shimojo & Nakayama, 1990.)

##### Stimuli and procedures

The stereogram consisted of paired white, homogeneous rectangular regions at the top and the bottom of the screen, and a white, vertical bar unpaired target on the left or right side of each rectangle. Figure 4 illustrates an example of top and bottom half-images. The large

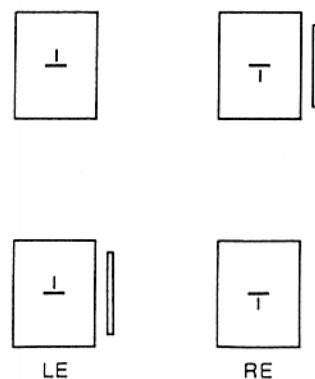


Fig. 4. An example of the stimulus used for forced choice experiment. Valid and invalid monocular bars are placed to the same side of a fused rectangle. Subjects task is to say which bar (upper or lower) is more in front. Nonius targets (to address the issue of oculomotor errors) are in the center of each large rectangle.

rectangular regions were interocularly paired and fusible, whereas the vertical bars were unpaired. A target presented to the valid eye and a target presented to the invalid eye were always presented at the same horizontal distance on a given trial. The position of the valid or invalid bar (either at the top or bottom) was randomly assigned. To avoid possible artifacts related to oculomotor convergence errors, a fusible horizontal line segment was added in the center of each rectangle flanked by a nonius stimulus consisting of two vertical line segments, each presented to one of the two eyes. When they appeared to align vertically, binocular convergence error would be less than 1.9 min arc.

The side on which the target was shown (left or right of the rectangle), the position of valid and invalid pairs (the valid at the top, or the invalid at the top), the distance of the target from the edge of the rectangular were all randomized across trials, and 5 trials were obtained for each position at each distance. Thus, 160 trials (=2 left-right sides  $\times$  2 top-bottom positions  $\times$  8 distances  $\times$  5 trials) were obtained for each subject.

The rectangular region, the target, the fusible horizontal line segment, and each of the vertical nonius segments extended  $137 \times 417$  min arc,  $7.6 \times 331$  min arc,  $42 \times 4.5$  min arc and  $1.9 \times 22.7$  min arc respectively at an observation distance of 70 cm. The luminance of the rectangular region and the vertical target was approx.  $100 \text{ cd/m}^2$  seen against a darker background. The stimulus pattern for each eye was generated on a CRT screen (Commodore AMIGA 1000), and viewed by a prism haploscope.

The subject's task was simply to make a two-alternative, forced-choice judgment as to whether the top or the bottom target appeared more distant from the fusible rectangle, and to push one of the two buttons accordingly, while monitoring the vertical alignment of nonius. No feedback was given.

##### Results

The results are shown in Fig. 5 where the percentage of seeing the valid stimulus as more distant is plotted against the horizontal position relative to the right and left edge of the square. As is obvious from the figure, the subjects had a tendency to perceive the valid target as more distant within a certain distance from the edge of the square, although this upper limit of distance for the depth effect varied considerably across eyes and subjects (from 30 to 60 min arc).

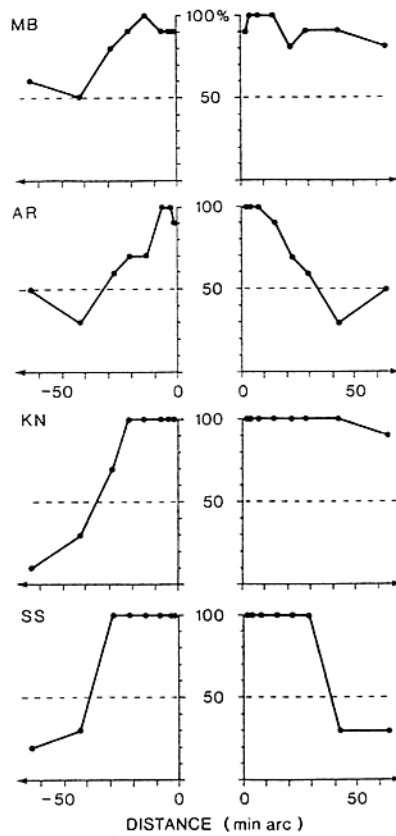


Fig. 5. Data from qualitative depth experiment depicted in Fig. 4. Percentage response of seeing the valid points as farther behind, plotted as a function of distance from the left and right edge of the fusible rectangular occluder. Data from four observers are shown separately, two experienced (KN, SS), two naive (MB, AR).

It may be worth noting again that a valid target on the right side of the rectangle was presented to the right eye, whereas a valid target on the left side was actually presented to the left eye. Thus, the symmetrical depth functions on the left and the right side cannot be attributed to simple eye dominance or any other possible factors specific to one eye or the other.

These results on qualitative depth were highly consistent with our predictions based on the occlusion constraints, showing that valid unpaired points were seen more distant than "invalid" points.

**EXPERIMENT 2: DISPARITY MATCHING**

The results obtained in expt 1 suggest that the visual system actually implements occlusion-related constraints to determine depth of unpaired stimuli. However, the results were only qualitative and relative in the sense that it revealed a systematic depth relationship between the valid and invalid stimuli. It does not

reflect the subjects' judgments as to perceived depth relative to the fusible stimulus rectangle. Our more detailed analysis of the occlusion constraints (as illustrated in Fig. 3A) indicates the presence of a depth constraint zone, confining a "valid" target to a range of depths which increases with increasing horizontal separation from an occluding edge. Because of the linearly receding nature of the theoretically closest possible depth interpretation, our interest was to determine whether the encoding of depth quantitatively would reflect this relation.

As such, our approach to examine the perceived depth of these monocular targets was to use a psychophysical disparity matching procedure, varying the disparity of a binocularly fused target so that it was seen as the same distance as the unpaired target.

*Method*

Two subjects (the authors) participated in this experiment. The set of stimuli used in expt 2 was similar to that used in expt 1 except that the bar-shaped target at the bottom is fusible (interocularly paired) and its disparity relative to the bottom rectangle was adjustable by the subject (shown in Fig. 6). Under free fixation, the subject's task was to adjust the disparity of the matching target at the bottom of screen until its perceived depth matched the unpaired target at the top of screen while monitoring the vertical alignment of nonius. The measurement of perceived depth of an unpaired monocular point can be complicated by the fact that the depth of the unpaired point often "takes on" the depth of the fused matching target. Thus, as the observer moves the binocular probe so it appears as farther and farther away, the

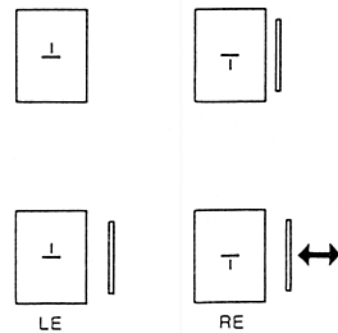


Fig. 6. An example of the stimulus used in the depth matching expt 2. Similar to expt 1, except observer sees one of the vertical bars binocularly (at the bottom) and sets its disparity to match the perceived depth of the unpaired bar at the top.

monocular target also can appear to move back. This happens even when the two targets are quite far apart on the display. To overcome this problem, the observer moved the binocular target over as wide a range as possible, going forward and backward. Under these circumstances, it was possible to find a disparity setting where depth appeared to be most stable and was most easily matched by the disparity target. A session consisted of 48 trials (2 sides  $\times$  8 steps  $\times$  3 trials), and each subject performed 3 sessions with a break of at least 5 min between them. Thus, 144 trials were obtained from each subject.

### Results

The results are shown in Fig. 7 where the matched disparities to both valid and invalid unpaired targets are plotted as a function of distance for each side in each subject. The leading edge of the depth constraint zone (as illustrated in Fig. 3) is represented by the dotted oblique lines.

The results were similar for the two observers: the disparity matching functions for invalid

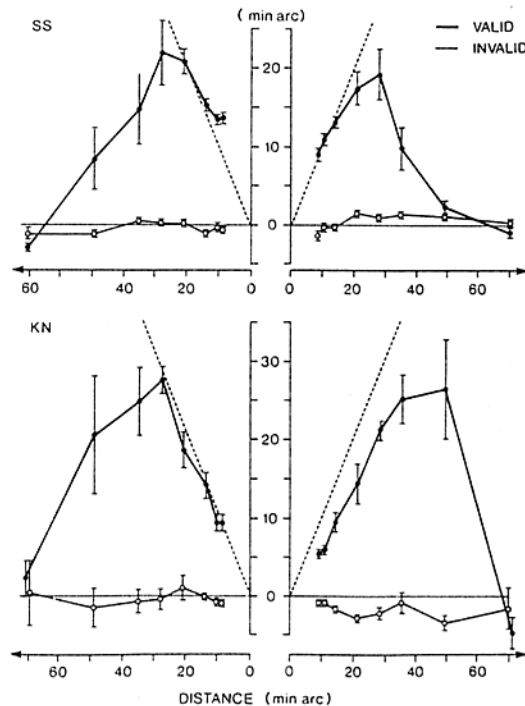


Fig. 7. Disparity matching results (expt 2). Matched disparity to unpaired points is plotted as a function of horizontal distance from the left and right edge of binocularly fused occluder. Leading edge of depth constraint zone illustrated as  $L_x$  in Fig. 3 is shown as dotted oblique lines. Open circles represent valid monocular points, and solid circles represent invalid monocular points.

stimuli were flat and essentially at zero; whereas, the perceived distance of the valid stimuli increased, following the nearer limit of the constraint zone up to 30–40 min arc of distance, after which they returned to the zero disparity line.

### Discussion of experiments 1 and 2

Results so far can be quickly summarized. First, valid monocular points are perceived as more distant than invalid monocular points, and second, the perceived distance of these unpaired points is roughly comparable with the closest possible theoretical distance compatible with the occlusion constraint, at least for some range of horizontal angular separations.

As such, these results taken together, indicate that the visual system utilizes information as to which eye is stimulated in relation to fusible edges to make an appropriate interpretation of depth.

### EXPERIMENT 3: SILHOUETTE CONSTRAINTS?

So far we have concentrated on the problem of differential binocular occlusion, the hiding of some image points behind visible occluders. Although we may have given the impression that occlusion is the only possible way in which unpaired points can arise in scenes, this is not strictly true. We have neglected one rare yet clear possibility, that of one-eye-camouflage arising from silhouettes (see Kaye, 1978).

This situation can occur when objects are illuminated from behind and where one cannot see any internal details of object surfaces so that the target is camouflaged against the fused surface for one eye yet appears in sharp contrast with the background through the other. This one-eye-camouflage case is illustrated in Fig. 8A. Contrary to the occlusion situation, where such a case could not arise, it would be possible to see a target in the left eye only, even if it is to the right of fused target. This means that an unpaired stimulus in one eye may be perfectly consistent with this "silhouette" situation in the real world, and therefore, a "front" perception of unpaired target may be justified as "valid". This can occur only, however, if the luminance and color of the object giving rise to the unpaired target has the same color and luminance of the fused binocular target so that it cannot be visible against the fused surface.



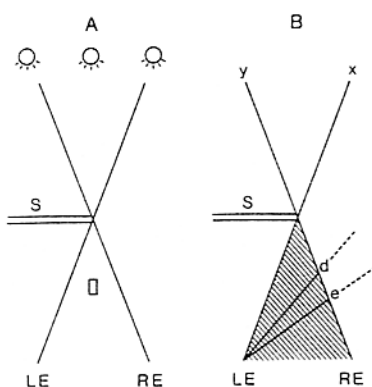


Fig. 8. One-eye camouflage. (A) Top view of silhouette situation with back lighting. Target is camouflaged in the right eye when seen against the binocular square but is visible in the left eye when seen against the background. (B) Silhouette geometry. For given horizontal positions of unpaired points in the left eye (corresponding to *d* and *e*), real world points must arise from the line segment that is within the depth constraint zone (shaded area).

Figure 8B shows a corresponding depth constraint zone for this silhouette situation (see shaded area), swept out by possible depth line segments (lines *d* and *e*). This shows that the only possible real-world interpretation of such a binocular image is for the target to be seen in front of the fused target, between the two visibility lines (shaded area). Note that all points must be closer to the observer relative to the fusible surface, and that the farthest edge of this silhouette zone becomes progressively closer as the horizontal distance of the unpaired target from the edge is increased. Figure 9 shows the occlusion and silhouette constraints in terms of unpaired points presented to the right side of an

Cases	Left Eye	Right Eye	Occlusion Constraints	Silhouette Constraints
A (1)			Invalid	Valid (Front)
A (2)			Valid (Back)	Invalid
B (1)			Invalid	Invalid
B (2)			Valid (Back)	Invalid

Fig. 9. Occlusion and silhouette constraint tables. Occlusion and silhouette are compared in terms of eye-of-origin and luminance polarity of the stimulus. (A) The same luminance condition. (B) The opposite luminance condition for a critical test of silhouette constraints. For simplicity, analysis is restricted to the right side of the fused edge.

occluder, comparing situations where luminance between occluder and monocular points are matched (Fig. 9A) and when they are not (9B). As is clear from the analysis illustrated in the figure, the conditions under which a silhouette interpretation might be plausible are very rare because the color and luminance of the various targets would have to be the same.

To determine whether silhouette constraints are actually implemented in the interpretation of the depth of unpaired points, we conducted an experiment where the luminance conditions for the silhouette assumption were clearly violated, comparing it to the case where they were not. Based on the analysis shown in Fig. 9, our approach was to use opposing luminance polarity for the unpaired target and the fused square (see top left inset of Fig. 10). Against a gray background, a monocular target was black and the fused square was white. Under these conditions there is no plausible way in which a black unpaired target could be rendered invisible when seen against a white square. In addition to this incompatible condition, we also ran the situation where both the fused square and the monocular targets were black and thus consistent with a silhouette interpretation. Stimuli and procedures were otherwise identical to those in expt 2. It should be noted that one naive subject (MB) used a slightly different strategy of observation compared to the two authors (see Methods section of expt 2). In her case, she moved the binocular target as far back as possible, going back and forth to find the point beyond which the unpaired could not appear to be pulled farther back. This was the only way she could make judgments with confidence.

The results for both conditions for three observers (including the new naive observer) are seen in Fig. 10. First, it should be noticed that the pattern of results for the valid occlusion case replicates expt 2 both for the same and opposite polarity luminance conditions. There is an increasing perception of "far" (following the constraint line dashed and labeled O in the figure) as the valid occluded target becomes more separated from the fused edge. This effect decays when the distance goes beyond 30–40 min arc. Most relevant for the issue at hand is the perceived depth of the left-eye-only point, comparing the case where the silhouette conditions are fulfilled (left hand graphs) vs the case where they are not (right hand graphs). There is no systematic difference in the perceived depth for

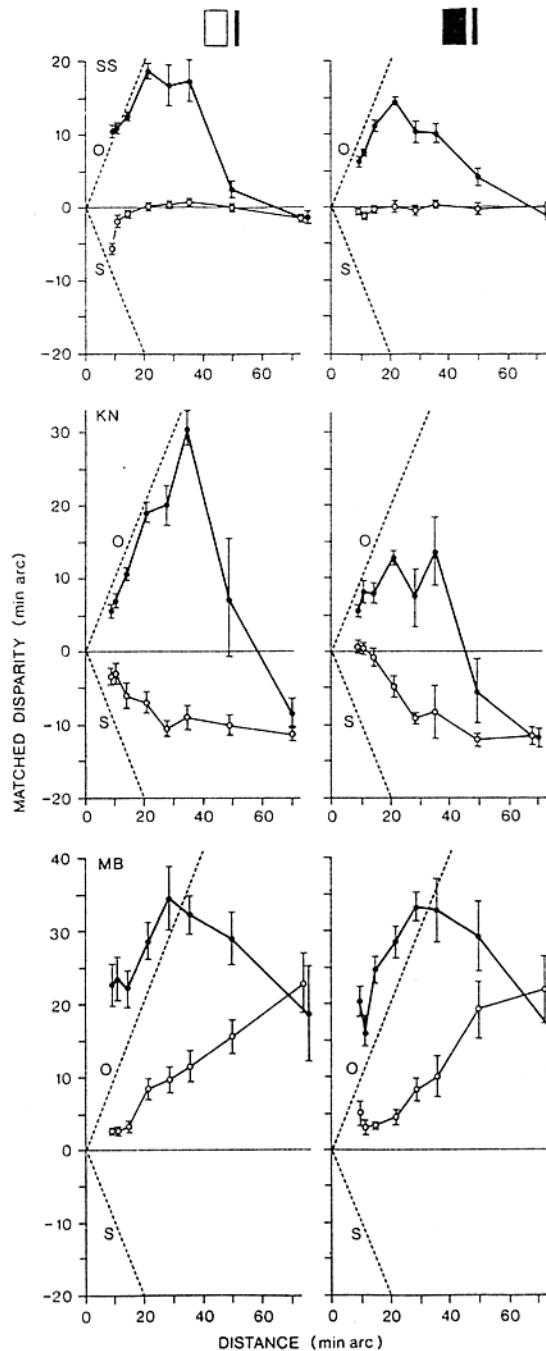


Fig. 10. Results from silhouette experiment comparing situation where binocular image could have arisen from a silhouette (right graphs) vs the cases where this is not possible because of reversed luminance polarity of fused square and unpaired bar (left graphs). Matched disparity is plotted against horizontal distance from the right edge. Solid circles represent stimuli given to the right eye (valid for occlusion, invalid for silhouette), open circles represent stimuli presented to the left eye which are invalid for occlusion and valid for silhouette only in the matched luminance condition (see right graphs).

the silhouette valid vs silhouette invalid case (compare the open circles in the left and right

graphs). Second, there was no systematic tendency for the perceived depth curve to hug the trailing edge of the silhouette constraint zone (dashed and labeled S in the figure) as would be predicted from a silhouette interpretation. Both of these findings indicate that the visual system is unlikely to take account of the conditions necessary to achieve silhouettes in the processing of depth from unpaired images.

The results, however, reinforce the importance and robustness of occlusion constraints in the processing of unpaired points.

#### *The inverse case: subjective occluding contours from unpaired dots*

So far, our discussion has dealt with the perceived depth of unpaired points—at what distance do these points appear in relation to fused binocular targets? Our results indicate that human observers see appropriate depth from unpaired targets if they simulate the real world situation of a fused surface occluding background points in one eye and not the other. It seems that the visual system has implicit knowledge about the facts of occlusion in the determination of perceived depth.

In this section, we show a number of demonstrations which provide very different and perhaps even stronger evidence for this hypothesis. We show that the visual system can “construct” an occluding contour when faced with unpaired monocular points. The existence of such “constructed” contours can be seen by considering the stereograms seen in the top portion of Fig. 11. These three part stereograms are specifically designed so that both crossed and uncrossed fusers can observe a stereogram with the appropriate sign of disparity. Cross fusers should fuse the left and center images, whereas uncrossed fusers should fuse the center and right images. In the middle of one of the stereo half-images we have placed a colinear vertical set of three dots just above the arrow-shaped fixation aid. Each dot in this vertical column does not have a counterpart in the other eye. When fused, these stereograms show differential depth such that dots in the right half of the binocularly fused configuration are seen at a different depth from those on the left. Furthermore, the single unpaired column of dots generally takes on the same depth value as those sets of dots which are seen as behind.

Most important to observe, however, is the existence of a sharp “knife edge” border, a vertical subjective contour or illusory edge



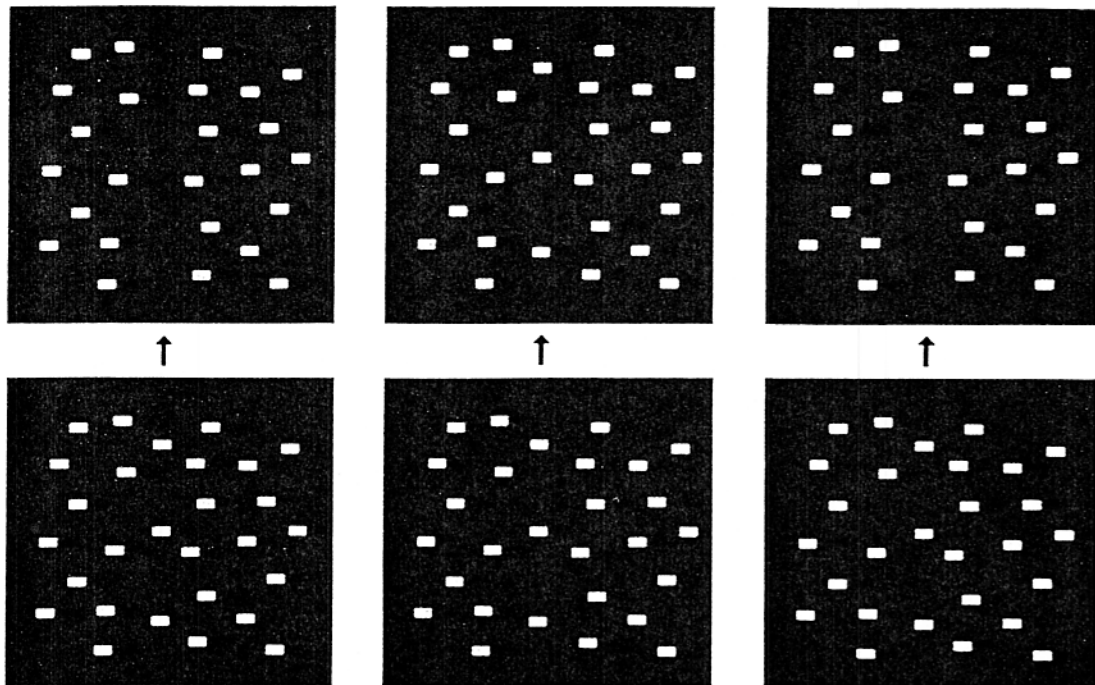


Fig. 11. Three part stereograms. Cross-fusers should fuse the left and center images and uncrossed fusers should fuse the center and right images. Top: "knife-edge" subjective occluder induced by the existence of unpaired points indicated by arrows. When viewed as instructed observers see a subjective contour to the immediate right of left-eye-only points. Bottom: control stereogram where all points are fusible and there is no "knife-edge".

running down the middle of each stereogram. The position of this border is lawfully related to the eye of origin of the unpaired monocular dot. If the stereograms are viewed correctly (see above) then the unpaired dots are seen in the left eye and the subjective contour is seen immediately to the right of these dots. If the stereogram on the top is viewed in the opposite way (cross fusers combining the center and right image or uncrossed fusers combining the left and center image), then the monocularly visible dots will be seen by the right eye only and the illusory edge is seen to the immediate left of these dots. Another observation to note is that the blank region between the front dots are perceptually filled in so that they appear as a flat, opaque surface which partially occludes the surface behind.

To underscore the importance of the unpaired monocular points in the formation of the subjective contours, we also generated a pair of control stereograms (see bottom stereogram of Fig. 11). Here everything is the same as on the top of this figure, except that three unpaired points in each stereogram have been paired or "matched" with their identical counterpart in the other eye. Viewing these stereograms, which have no unpaired points, one can see that the same difference in depth is present. The right

half of the binocularly fused image has a different depth than the left. Supporting our main point, however, is the fact that the subjective contour present in the top of Fig. 11 is not visible here. No subjective contour can be seen which runs along either side of the vertical column of dots. So it is the differential interocular visibility of the three points that must generate the illusory occluding edge. More specifically, and in accord with our discussion of occlusion, the location of the subjective contours is highly specific and consistent with occlusion constraints. They appear to the immediate right of left-eye-only points and to the immediate left of right-eye-only points.

Also to be noted is the fact that the opaque filling-in effect (surface extrapolation from the fusible front dots to the vertically lined unpaired dots) is far less compelling in this control condition where all points are binocularly fused. Depth differences generated by differences in the disparity of matched binocular points cannot account for the illusory occluding edges and surfaces.

#### *Subjective contours define depth*

So far we have shown that subjective contours can be generated by unpaired binocular

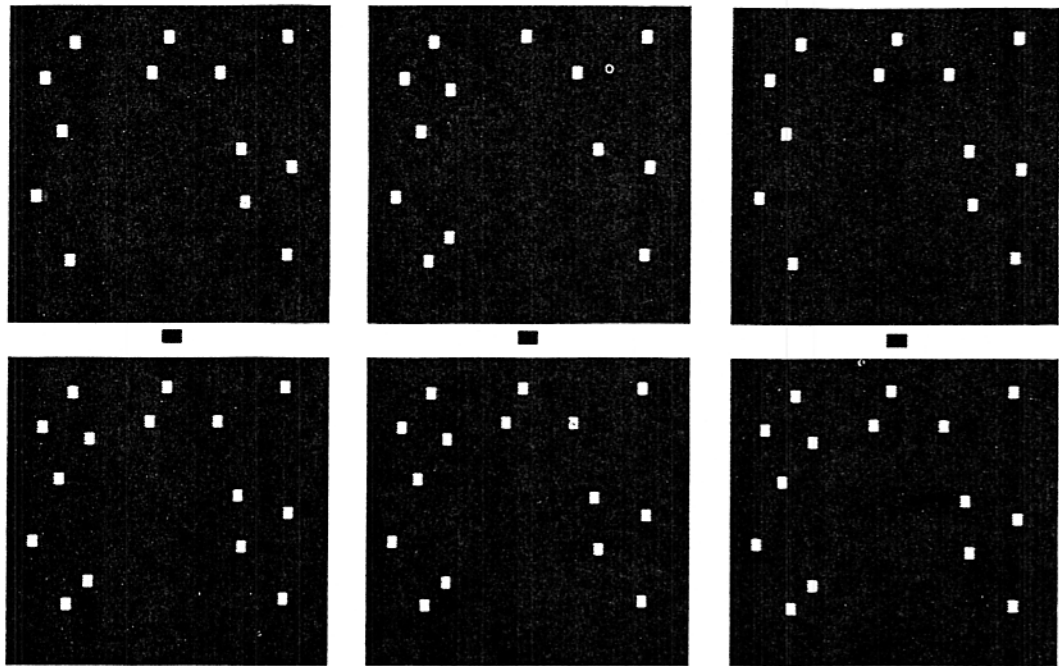


Fig. 12. Top stereogram: emergence of subjective occluding triangle, seen in front of fused random dots. The triangle shaped surface is induced by two vertically oriented dots visible only to the left eye and two obliquely oriented dots visible only to the right eye. See Fig. 13 for details regarding the location of these unpaired dots. Bottom stereogram: same as above, except that all points are fused and in the same depth plane. In this set of stereograms, crossed fusers should fuse the left and center images, and uncrossed fusers should fuse the center and right images.

points. Implicit in the discussion of subjective contours is the notion that it lies in a different depth plane from the monocular point, that it must be of necessity in front because it partially "covers" the monocular point. In fact these stereograms contain disparity information defining a depth of paired dots which is also in front. Thus, one might conclude that the formation of the subjective contour results directly or indirectly from the disparity-defined depth of these fused points.

In this section, we will demonstrate that this is not the case. We show that stereograms containing no differences in disparity, thus having no stereoscopic depth, can lead to vivid perceived depth of a surface partially bounded by subjective occluding contours which is seen in front of all fused points.

The stereogram seen at the top of Fig. 12 is almost identical to that seen at the bottom of this figure except that two very specific dots have been simply *removed* from each half image. Thus two vertically oriented sets of dots have been removed from the left portion of the right eye image and two obliquely oriented dots have been removed from the left eye image. So the combined stereogram contains mostly paired

dots, except for two monocular pairs in each eye. Such paired and unpaired dots are schematized in Fig. 13. Viewing these stereograms, our observers noted that subjective contours were seen to the right of the left monocular dots and to the left of the right monocular dots (see white lines in Fig. 13) such that a "triangle" was discernible.

Most important to note, however, is the fact that this fragmentary figure of a triangle is seen as clearly *in front* of the field of dots. This demonstration shows that interocularly unpaired dots, in the absence of any differential retinal disparity of matched binocular points, is

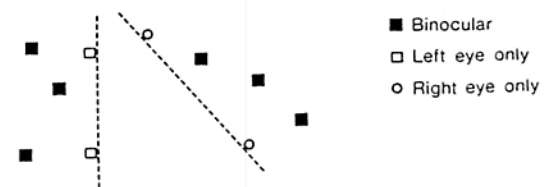


Fig. 13. Schematic description of the stereogram seen in the top of Fig. 12 showing the location of unpaired monocular points. Binocular points labeled by solid squares, left eye points by open squares, right eye points by open circles. Dashed lines denote location of the subjective edges which partially delineate the triangle shaped surface.

sufficient to create the perception of a surface in depth. The control case is seen in the lower stereogram of Fig. 12, where all points are matched. As expected, only a single background plane is seen, no subjective contours or closer surface is discernible.

A subsidiary observation is also important to note. The subjective contour formed by the unpaired dots appears as right next to these dots, but not necessary as exactly vertical. The contour orientation can have a horizontally oriented component such that it provides the linkage between several unpaired points as schematized by the oblique dashed lines in Fig. 13 and seen in the top stereogram of Fig. 12. As such, the contour formation mechanism seems to be adaptive to variations of visual stimuli in the real environment.

#### EXPERIMENT 4: SUBJECTIVE CONTOURS AND SURFACES IN NAIVE OBSERVERS

These observations were made by experienced observers, having familiarity with stereoscopic experiments, though not necessarily knowledgeable about the aims and goals of these demonstrations. To see if these observations could be confirmed by naive observers, we ran a phenomenological experiment in which 11 naive subjects participated. The subject observed four stereograms through a phase haploscope, a pair of shuttered spectacles, opening at the appropriate time for the presentation to each eye of its CRT half-image (stereoscopic display controller of Stereo-optic Systems Inc. was used with Commodore Amiga 1000).

The stimulus pattern occupying top portion of Fig. 11 was paired with the control pattern illustrated at the bottom of this figure. They were placed, one above the other, and viewed through the phase haploscope in the first stereogram. The top and bottom positions were simply exchanged for the second stereogram. The stimulus extended  $4.7 \times 10.8$  deg of visual angle at the observation distance of 170 cm. The element's size was  $12 \times 6$  min arc, and the relative disparity between the front surface (right side) and the behind surface (left side) was 15 min arc. Free observations were allowed without eye fixation. The experimenter first asked the subject to describe any depth seen in the display. The subject was categorized as "spontaneously saw it" when he reported (a) sharper or more straight edges with monocular dots than with binocular dots, or (b) the lateral

position of edges pushed leftwards relative to the binocular control, being right next to the right side of monocular dots. If the subject's spontaneous report did not meet any of these criteria, then the subject was asked to concentrate on the sharpness, straightness and lateral position of the edge, and make an explicit comparison between the top and the bottom patterns in this regard. If the subject met those criteria after these instructions, then the subject was categorized as "saw it when suggested", otherwise he was categorized as "did not see it".

The stimulus pattern illustrated in Fig. 12 was employed for the third stereogram, and the positions were simply exchanged for the fourth stereogram. The stimulus extended  $3.7 \times 7.1$  deg at the observation distance of 170 cm. The element's size was  $3 \times 3$  min arc. Again, free observation was allowed. The subject was simply asked whether he saw any contours, edges or a shape besides the dots, either in the top or the bottom stimulus patterns. If the answer was positive, the subject was further asked to describe them, to designate which part of the stimulus looked in front, and also to tell the difference between the top and the bottom stimulus patterns. The subject was categorized as "spontaneously saw it" when reported a subjective shape floating in front of the background and identified it correctly as either top or bottom. If the subject did not spontaneously report this, then the subject was explicitly asked whether any shape or depth was visible in the center part of either the top or the bottom pattern, and whether there was any difference between the top and the bottom patterns in these aspects. Depending on their protocols, the subject was either categorized as "saw it when suggested", or "did not see it".

Table 1 shows the results. As is clear from the table, all 11 naive subjects saw the subjective contours created by unpaired dots at least in one of the two types of stereograms, even though there were considerable individual differences. To give an example of a more phenomenological description of the percept, one subject (A.C.) reported for the first stereogram, "It is easy to tell the difference between the top and the bottom. At the top, the edge is wavy and dots look isolated and not interpolated. At the bottom, the edge is sharper, clearer and more straight. Blanks among front dots are interpolated, so it looks like an occluding surface." Another subject (C.D.) reported as for the third

Table 1. Table of subjects who saw the subjective contours defined by unpaired monocular dots. See the text for the criteria of categorization

Subject no.	Spontaneously saw it	Saw it when suggested	Did not see it
1	a b		
2		b	a
3	a b		
4	b		a
5		a b	
6	a		b
7	a b		
8		a	b
9	b	a	
10	a b		
11	b	a	

(a) "Cliff" stereograms (adaptation of Fig. 11).

(b) "Triangle" stereograms (adaptation of Fig. 12).

stereogram, "There is a magical triangle which occludes in front, but it's only in the top pattern."

#### *Discussion: demonstration 1 and 2, experiment 4*

The results presented in the demonstration, plus our experiments with naive observers indicate that the simple existence of unpaired points in binocular images can call forth something qualitatively different from simple depth or rivalry. It appears that the visual system generates the perception of an occluding edge, seemingly out of nowhere. This subjective contour is seen "in front" of the unpaired dot in a very lawful manner. First, it is seen to the right of left-eye-only points and second, it is seen to the left of right-eye-only points. Third, it always appears in front of the unpaired point. Fourth, this depth is so compelling that a whole scene containing no binocular disparity cues at all can generate a vivid surface in front, bounded only by occluding contours generated as a consequence of these unpaired points.

All of these facts, which would otherwise be puzzling, conform to the everyday facts of occlusion. In normal scenes, nearer surfaces occlude background image points to differing extents in the two eyes.

#### GENERAL DISCUSSION

##### *Occlusion and stereopsis: a broader view of stereoscopic vision and depth processing*

In this paper we show two phenomena relating the presence of unpaired points and depth. Unpaired points can be perceived at the appropriate depth if they are near fused targets and the eye-of-origin is appropriate. In addition unpaired points can lead to the emergence or

formation of subjective occluding contours, seen in front.

Most important to note is the fact that eye-of-origin information is critical in all of these experiments and demonstrations. Perception varies markedly if one exchanges right and left eye stimulation. Of course this is as it should be if we think of the real-world case of occlusion. The visibility of right-eye-only points and left-eye-only points are lawfully related to the angular positions of fused portions of ecologically valid binocular images.

This dependence on eye-of-origin information is also evident in two recent studies in our laboratory. First, we found that interocular rivalry and suppression were absent only when unpaired monocular areas simulated a real world scene (Shimojo & Nakayama, 1990), a finding mentioned earlier and one which motivated the present study. Second, Shimojo, Silverman and Nakayama (1988) simulated the case where an observer sees distant moving targets through a narrow vertical slit such that there was no temporal overlap in the visibility of the moving target in each eye. Yet, appropriate depth perception was obtained in this occlusion simulation, using direction of motion and interocular order information.

Thus, we have four separate findings which can be linked to rivalry, depth, subjective contours, and motion, respectively—all of which are related to occlusive conditions in the real world and require explicit eye-of-origin information. Because these phenomena are closely related to notions put forth by Leonardo da Vinci (Wheatstone, 1838), we group them under his name, thus making the distinction between da Vinci and Wheatstone stereopsis.

At first glance it might appear that there is no obvious relation between these two types of stereopsis. Yet on closer examination, it is evident that there may be an important link. Again, it revolves around the issue of eye-of-origin information. Classical stereopsis relies on the fact that two physical points in space, lying at different distances lead to differential retinal disparity of the two pairs of binocular image points. But just noting a difference in retinal disparity is not enough to identify depth. The sign of this difference is critical, determining whether one sees the one target in front or in back. The basis for this signed disparity, of course, is eye-of-origin information. Only by specifying which eye is which can it be determined whether disparity is crossed or uncrossed.

In other words, "the maintenance of labeling of the eyes for disparity selective combination would be the fundamental requirement for stereopsis" (Shimojo et al., 1986).

Such considerations argue for a more broadened conceptualization regarding binocular depth processing to incorporate both Wheatstone and da Vinci stereopsis. Both reflect the horizontal difference in station point (Gibson, 1950) when sampling the optical array. Each, however, takes advantage of a different consequence of this fact: Wheatstone stereopsis is based on the signed difference in retinal disparity; da Vinci stereopsis is based on the differential occlusion of image points. Thus, against the apparent distinction between the two kinds of stereopsis, we consider them as two ways of utilizing eye-of-origin information for depth, edge and surface perception.

#### *Neurophysiological locus and mechanism?*

Studies on the functional anatomy of primate extrastriate cortex have indicated the existence of many visual areas. Starting with the striate cortex, which is the entry point of visual information from the LGN, recent studies reveal as many as 10–20 visual areas arranged in hierarchical fashion (Maunsell & Newsome, 1987). Can we identify one of these areas where the processing of unpaired points might begin?

Eye-of-origin information is explicitly represented in the ocular dominance structure of V1 (Hubel & Wiesel, 1968). Thus, many cells in this structure respond preferentially to either right eye or left eye stimulation. Beyond V1, however, it appears that eye-labeling is less explicit. For example, cells in V2, V3, V4, MT and VP are all strongly binocular, with very little difference in the response to right or left eye stimulation (Burkhalter & Van Essen, 1986; Maunsell & Newsome, 1987). Because eye-of-origin information is critical for the tasks described in this study, and because it is apparently lost after V1, it suggests that da Vinci stereopsis begins surprisingly early in visual cortical processing, possibly as early as V1 or its immediate projection targets.

#### *Implications for stereo algorithms*

Because of the classical Wheatstone view of stereopsis, the problem of stereopsis has been considered as that of finding corresponding points or resolving ambiguity, particularly in the context of stereo algorithms. This is generally thought to be a complex task because of the potentially large number of false matches that

can be made. One of the main approaches to the problem has been to exploit the assumption of smoothness in the real world to constrain the number of potential matches. By constructing appropriate parallel algorithms to capitalize on this constraint, considerable progress has been made (Julesz, 1971; Sperling, 1970; Nelson, 1975; Marr & Poggio, 1976). Yet, it appears that no existing algorithm can correctly recover the depth of all stereoscopic scenes nor can they account for the properties of human or animal stereopsis (Mitchison, 1988). Much of the difficulty arises at the edges of real surfaces, points in binocular image where smoothness cannot be assumed and where the system is faced with unpaired points.

In treatments so far, these unpairable points have been considered as "noise", creating undesirable signals to be suppressed or neglected. Our findings, however, indicate just the opposite. Unpaired points contain rich information as to where discontinuity boundaries are located. They also indicate the sign of this discontinuity—the eye-of-origin information can specify which side should be in front and which side should be in back. Instead of neglecting or discarding unpaired points, we suggest that future algorithms incorporate specific constraints relating eye-of-origin information to surface boundaries. This would be consistent with psychophysical findings which show that the human system can perceive depth in random dot stereograms more quickly when appropriate unpaired points are added (Gillam & Borsting, 1988).

#### *Neural inference in visual processing and the issue of learning*

Unlike stereopsis which has been considered as an innately hard-wired depth cue, occlusion and other pictorial cues have been considered as acquired through visual experience. Thus there has been an implicit dichotomy between "physiological" and "cognitive" cues to depth. These attitudes may have stemmed in part from the existence of units selective to binocular disparity (Barlow, Blakemore & Pettigrew, 1967) and the lack of evidence for comparable neural units for pictorial cues.

Against this dichotomy, our current findings indicate that occlusion related constraints are implemented very early in visual processing. Remembering that both Wheatstone and da Vinci stereopsis are mediated by early eye-of-origin processing, we would like to argue that



both kinds of depth processing can be regarded as inferences implemented early in the visual pathway. Both exploit real-world geometry, particularly the difference in the horizontal position of the two eyes. da Vinci stereopsis, relying on the property of differential occlusion. Wheatstone stereopsis, relying on the different angular positions of eye-labeled points.

Lending support to the view that such occlusion processing is closely related to classical stereopsis is the fact that such neural "inferences" rely only on local distance information and are thus very similar in kind to disparity information. Both can operate only over a limited spatial range and process only limited sets of information. With respect to occlusion processing, the perceived depth of unpaired points follows the leading edge of the occlusion constraint for only a limited horizontal distance. Beyond a horizontal separation of about 30–40 min arc depth reverts back to the front-parallel plane (see Figs 7 and 10). In addition, the perceived depth of monocular points is not obviously affected by whether or not the scene could be interpreted as a silhouette, nor is there any evidence that the perceived depth has any relation to luminance constraints dictated by the properties of silhouettes (see Fig. 10). These limitations make it more likely that the depth effects are mediated by relatively local eye-labeled positional mechanisms, analogous to mechanisms responsible for Wheatstone stereopsis.

In line with this discussion of neural inference, it might be asked whether such processes are based on built-in or learned mechanisms. As far as most of the phenomena under discussion are concerned, it is conceivable that natural selection would favor the pre-wiring of neural connections to interpret depth correctly before any kind of visual experience. This is also consistent with our analyses that a small set of straightforward rules or constraints can predict these depth- and contour-related phenomena.

However, our analysis also indicate that these constraints are, in fact, well reflected in the prior probabilities of our visual experience. For instance, the visual system encounters combinations of visual inputs such as "right-eye-only stimulus on the right side of occluding edge" (a valid combination) at *almost every moment* in the normal environment, whereas it essentially *never* experiences visual combinations such as "right-eye-only stimulus on the left side of occluding edge" (an invalid combination). Fur-

thermore, all the information necessary for neural inference (eye-of-origin and relative position to the edge) are locally available at very early levels of cortical processing. Thus, it may be even more conceivable that synaptic modification based on stimulus-driven neuronal activity is responsible for the apparently inferential processes.

Even for the formation of the more fundamental and apparently "hard" structures such as ocular dominance columns, the contribution of Hebb-like learning rule has never been ruled out (see Willshaw & Von der Malsburg, 1976; Reiter, Waitzman & Stryker, 1986; Miller, Keller & Stryker, 1989). Thus at least theoretically, it is possible for the visual system to encode appropriate depth and contour perception through normal visual experience.

The possibility of learning becomes more attractive, however, when we consider the emergence of the subjective occluding contour and surface arising from the unpaired points shown in Fig. 12. It is difficult to understand how a built-in mechanism under Darwinian selective pressure could have responded appropriately to such a situation which is hardly encountered in daily life. On the other hand, the continuous association between unpaired points and real occluding contours certainly provides the basis for such neural connections. Theoretical neural networks, for example, can evoke associative responses to stimuli which have never been presented alone. Thus, it is possible that the single unpaired dot could call forth the perception of the subjective contour just as a small fraction of a stimulus pattern is sufficient to retrieve the original input pattern in associative memories (Anderson, 1972; Kohonen, 1972, 1977).

#### *Neurophysiological representation of subjective contours*

Well described by Schumann (1900) and studied more recently by Kanizsa (1955), subjective contours have aroused much interest, particularly because they are not present in the gray level image, yet they are particularly vivid under certain circumstances. Some early workers have called them cognitive contours suggesting a higher level of representation in the visual system and it has been assumed that the perception of the subjective contour reflects cognitive or inferential processes for perceiving the layout of occluding surfaces (Coren, 1972; Rock & Anson, 1979).



Yet, recent evidence also suggests an early physiological representation. Most compelling are neurophysiological studies on alert primates which show responses to such contours in primate area V2 (Von der Heydt, Peterhans & Baumgartner, 1984; Von der Heydt & Peterhans, 1989; Peterhans & Von der Heydt, 1989). Supporting these findings is the fact that such contours elicit orientation-specific adaptation with comparable tuning characteristics to real contours (Paradiso, Shimojo & Nakayama, 1989). Both kinds of findings are consistent with an early representation of subjective contours.

Although these two positions, cognitive vs physiological, may appear to be irreconcilable, plainly they are not. In particular, we agree with the cognitive theorists who argue that the perception of subjective contours is tightly related to occlusive edges and depth. Yet our critical assertion is that some of the processes responsible for the perception of subjective contours must begin sufficiently early to have direct access to explicit eye-of-origin information. What reconciles these two apparently conflicting positions is our analysis which indicates that so-called cognitive processes can be implemented very early, as early as cortical area V1 or its immediate projection targets. Thus, there is no mutual exclusivity between cognition and early cortical physiology. Neuronal activity in the earliest areas of cortical visual processing may provide the required substrate for essential aspects of visual cognition.

#### REFERENCES

- Anderson, J. (1972). A simple neural network generating an interactive memory. *Mathematical Biosciences*, *14*, 197-220.
- Barlow, H. B., Blakemore, C. & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *Journal of Physiology, London*, *193*, 327-342.
- Burkhalter, A. & Van Essen, D. C. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *Journal of Neuroscience*, *6*, 2327-2351.
- Coren, S. (1972). Subjective contour and apparent depth. *Psychological Review*, *79*, 359-367.
- Gillam, B. & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, *17*, 603-608.
- Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, London*, *195*, 215-243.
- Julesz, B. (1971). *The foundations of cyclopean perception*. Chicago: University of Chicago Press.
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, *49*, 7-30.
- Kaye, M. (1978). Stereopsis without binocular correlation. *Vision Research*, *18*, 1013-1022.
- Kohonen, T. (1972). Correlation matrix memories. *IEEE Transactions on Computers*, *C-21*, 353-359.
- Kohonen, T. (1977). *Associative memory—A system theoretic approach*. Berlin: Springer.
- Lawson, R. B. & Gulick, W. L. (1967). Stereopsis and anomalous contour. *Vision Research*, *7*, 271-297.
- Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*, 283-287.
- Maunsell, J. H. R. & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Reviews in Neuroscience*, *10*, 363-401.
- Miller, K. D., Keller, J. B. & Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation. *Science*, *245*, 605-615.
- Mitchison, G. (1988). Planarity and segmentation in stereoscopic matching. *Perception*, *17*, 753-782.
- Nelson, J. I. (1975). Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, *49*, 1-88.
- Panum, P. L. (1858). *Physiologische Untersuchungen ueber das Sehen mit zwei Augen*. Kiel: Homann.
- Paradiso, M. A., Shimojo, S. & Nakayama, K. (1989). Subjective contours, tilt aftereffects and visual cortical organization. *Vision Research*, *29*, 1205-1213.
- Peterhans, E. & Von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience*, *9*, 1749-1763.
- Reiter, H. O., Waitzman, O. M. & Stryker, M. P. (1986). Cortical activity blockade prevents ocular dominance plasticity in the kitten visual cortex. *Experimental Brain Research*, *65*, 182-188.
- Rock, I. & Anson, R. (1979). Illusory contours as the solution to a problem. *Perception*, *8*, 665-681.
- Schumann, F. (1900). Beitrage zur Analyse der Gesichtswahrnehmungen. Erste Abhandlung. Einige Beobachtungen ueber die Zusammenfassund von Gesichtseindrueken zu Einheiten. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *23*, 1-32.
- Shimojo, S. & Nakayama, K. (1990). Real world occlusion constraints and binocular rivalry. *Vision Research*, *30*, 69-80.
- Shimojo, S., Bauer, J., O'Connell, K. M. & Held, R. (1986). Pre-stereoptic binocular vision in infants. *Vision Research*, *26*, 501-510.
- Shimojo, S., Silverman, G. H. & Nakayama, K. (1988). An occlusion-related mechanism of depth perception based on motion and interocular sequence. *Nature*, *333*, 265-268.
- Sperling, G. (1970). Binocular vision: A physical and a neural theory. *American Journal of Psychology*, *83*, 461-534.
- Von der Heydt, R. & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience*, *9*, 1731-1748.
- Von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260-1262.
- Westheimer, G. (1986). Panum's phenomenon and the confluence of signals from the two eyes in stereoscopy. *Proceedings of the Royal Society, London B*, *228*, 289-305.
- Wheatstone, C. (1838). On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society, London*, *128*, 371-394.
- Willshaw, D. J. & Von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society, London*, *194*, 431-445.