What prior uniocular processing is necessary for stereopsis?

Gerald Westheimer and Suzanne P. McKee

In order to induce stereoscopic depth, retinal images in the right and left eyes must differ. Our experiments demonstrate that these uniocular differences for simple three-line configurations at stereo-threshold cannot usually be distinguished one from the other. We confirm that depth values are associated with individual features by virtue of their disparity, rather than by a comparison of their image separations in the two eyes. In ordinary stereoacuity patterns, no prior identification of monocular relationships is needed to produce discriminable depth differences. For certain patterns, disparity of the positions of the averages of the targets' internal light distributions in the two eyes can substitute for disparity of pattern contours in stereoscopic depth discrimination.

Key words: stereoscopy, disparity, hyperacuity

Stereoscopic depth resolution has a threshold of only a few seconds of arc, much smaller than the size of retinal elements and comparable to values for hyperacuity thresholds, e.g., vernier alignment. When a pattern has components at different depths, the geometry of the situation entails that the configurations seen by the two eyes are somewhat different. The simplest case to consider is a target consisting of two horizontally separated points or lines. When there is a stereoscopic depth difference, the separations on the two retinas between corresponding target points will differ. If, for example, there are two vertical rods and the left is nearer than the right, the separation of the retinal images of the rods in the right eye will be wider than that in the left.

We here wish to distinguish between two possible mechanisms for the recognition of a stereoscopic depth difference, mechanisms which differ in the sequence in which information is extracted from the retinal images. We ask: Are the separations of the retinal images of features determined individually in each eye first, and is this operation followed by the identification of the nearer feature through comparison of corresponding retinal image separations in the two eyes? Or, alternatively, is a depth assigned to each feature, and are the depth values of the individual features then compared to make a decision of which is nearer? In other words, is the detection of difference in feature separation in the two eyes (hyperacuity) a necessary precondition to stereopsis or not? The facts that hyperacuity thresholds are only a few seconds of arc and that such small distance differences can be detected for a wide variety of target configurations beyond the traditional vernier patterns\textsuperscript{1} make the prior uniocular processing of separations a viable proposition. And since there is continued movement of the two eyes, even relative to each other, it would seem reasonable to arrive at a stereoscopic depth determination via central
comparison of retinal image separations in the two eyes, because these would remain largely invariant with changes in convergence.

Stratton, who first raised this question, later concluded that the basis of stereacuity was the fine localization ability found in vernier acuity. This is also the view proposed by Julesz, who regarded some of his results as demanding as a prior condition to stereopsis the uniocular detection of small distance differences—which he called superresolution. (He defined process A as being prior to process B if B utilizes A.) Hering made the relevant qualitative observation that monocular differences in distance just above threshold give rise to a striking difference in depth when viewed stereoscopically. An explicit phrasing of the question by Walls led to a detailed experimental analysis by Berry, whose data support the conclusion that stereoscopic depth resolution thresholds are
made in the face of retinal image separation differences in the two eyes that could not be discriminated if presented to one eye alone; the data of Stigmar are also in line with such a view. Several related experiments are either consonant with or at least do not contradict the proposition that "depth discrimination could not be interpreted as merely the combination of vernier discriminations performed by each separate eye."

We here present experimental evidence leaving no doubt that the differences between the two retinal images for a target configuration at stereoscopic depth threshold cannot always be detected by themselves, and we are therefore led to the view that the assignment of depth difference to a pair of features does not require monocular separation discrimination as a precondition. Just the same, accurate localization of each feature on each retina is essential to the performance of a stereo-task; we are therefore reporting on some experiments relating to the process by which this is achieved.

Methods

Patterns were made up of white vertical lines against a dark background, created by computer control on a pair of Tektronix 602 display units with P4 phosphor. By means of a beamsplitting pellicle and Polaroid filters suitably placed before the scopes and the eyes, each eye could see only one scope. Subjects sat 2.5 m from screens; observation was always binocular with natural pupils, with optimal correction of any refractive error. During an experimental run, a target was presented every 3 sec for a duration of 500 msec (experiments in sections A and B) or 620 msec (section C). In the intervening period, there appeared a fixation square, outlined by four dots or brackets 0.75° apart, which assured good binocular and foveal fixation. A threshold was obtained by giving the middle line one of seven different locations during each presentation and requiring the subject to make a binary decision based on error feedback. Two different experiments were performed. In one, the center line was presented symmetrically displaced in opposite directions in the two eyes, and a stereoscopic depth threshold was obtained. It was expressed by the disparity (in seconds of arc) at which the subject could identify with 75% accuracy whether the middle line appeared in front of or behind the frontal plane established by the other two lines. This is a measure of the minimum detectable disparity; the results are plotted against mean target separation in Fig. 2.

If this detection is accomplished by the scheme on the lower left side of Fig. 1, the two retinal images of the patterns with threshold disparity should also be discriminable from each other. The companion experiment, carried out with identical method and procedure, involved the binocular presentation of a three-line target whose middle line had been displaced small amounts either way.
from the center, but now in the same direction in the two eyes. The subject had to identify whether in a given presentation the separation between the left line and the center line was larger than that between the center and the right. A threshold was obtained, outlining the pattern configuration whose direction of asymmetry would be correctly identified on 75% of occasions. It is expressed in identical measure as the stereoscopic threshold and is also shown in Fig. 2. For all but the smallest target separations, the detection of binocular disparity was much better than that of the underlying inequalities of spacing in the two eyes.

Section B. Disparity discrimination of a moving target with respect to a stationary comparison target. A particularly telling demonstration of the independence of stereos-discrimination from monocular hyperacuity judgments is afforded by utilizing the fact that stereoscopic acuity is robust to target movements of up to several degrees per second. The target here consists of a pair of vertical lines in each eye, all 12 min arc long, the upper of which had one of a set of seven disparities with respect to the lower. The pattern was shown both in a stationary configuration and, in a separate series of experiments, with the upper line moving horizontally at a rate of 1.0°/sec. During the 500 msec exposure it moved through a distance of 30 min arc either to the right or left. To ensure that the upper line remained symmetrically disposed, it started its trajectory with a

Fig. 2. Stereoscopic depth discrimination and hyperacuity threshold for three subjects. Three white vertical lines, 15 min long, were presented for 500 msec. The separation indicated on the abscissa is half the distance between the outer two reference lines (lines A and C in Fig. 1). Detection involved angle δ (see Fig. 1) in each case. For the stereoscopic experiment, displacement of line B from exact center of the distance AC was in opposite directions in the two eyes, and subject had to respond whether line B appeared behind the plane created by lines A and C or not. In the hyperacuity experiment, the displacement of line B was in the same direction in the two eyes, and the subject’s response was whether AB was larger than BC or not. For one pair of comparable points in each graph the standard errors of the thresholds are given by a vertical line. For most target separations, stereoscopic depth discrimination is possible with uniconal target configurations that cannot be distinguished.
mean binocular position equal to 15 min arc to the right of the bottom target if it moved leftward and 15 min arc to the left if it moved rightward. Fig. 3 gives the stereoscopic acuity threshold under these conditions; it is not greatly different from that in the stationary situation. Also shown in the figure are comparable hyperacuity thresholds. They were obtained by making both eyes view the pattern which was seen by only the right eye during the stereo-task. The subject’s task here was to make a judgment of vernier offset or, in a separate experiment, to decide whether the moving upper line’s trajectory was displaced to the right or to the left with respect to symmetry around the position of the steadily exposed lower line. As can be seen, the subject’s performance was poor in the latter task; that is, she could identify the direction of asymmetry if the upper line’s whole trajectory was laterally displaced by 90 sec arc, whereas as little as $\frac{1}{20}$ of this value could be detected if it was a constituent part of the monocular view of a stereoscopic acuity pattern. A second subject demonstrated comparable effects.

We are therefore forced to abandon the scheme on the left side of Fig. 1 and conclude that stereoscopic acuity was achieved by means of direct disparity processing without prior hyperacuity discrimination.

**Fig. 3.** Thresholds (sec of arc) for detection of disparity (stereoaucuity), or lateral displacement (vernier acuity). For vernier configuration, both eyes were shown the right eye’s view of the stereo-configuration. Target consists of two bright lines, 12 min arc long, separated by a 4 min arc gap. In moving condition, velocity of upper target line was 1°/sec. Duration, 500 msec.

<table>
<thead>
<tr>
<th></th>
<th>STATIONARY</th>
<th>MOVING</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEREO</td>
<td>VERNIER</td>
<td>STEREO</td>
</tr>
<tr>
<td>JR</td>
<td>12.8 ± 1.4</td>
<td>7.9 ± 0.8</td>
</tr>
<tr>
<td>SM</td>
<td>5.3 ± 0.3</td>
<td>4.8 ± 0.5</td>
</tr>
</tbody>
</table>

**Section C. Redistribution of light within targets of fixed contour.** The findings described so far emphasize that stereoscopic depth discrimination does not have a primary basis in monocular separation discrimination; that is, depth is not assigned stereoscopically by discrimination of corresponding image separation in the two eyes. Very precise information must therefore be obtained for the position in each eye of the images of features belonging to each component of the stereo-targets.

Since the differences in position are exceedingly small, the mechanism by which position is associated with a feature is of more than passing interest. We have addressed the corresponding question in the course of a previous investigation into visual hyperacuity. We demonstrated there that spatial localization of a feature can proceed via the identification of the place occupied by the mean or “center of gravity” of its light distribution. We here report on a similar experiment in stereoaucity.

The target consisted of a pair of side-by-side features in each eye, 15 min arc high and separated by a mean distance of 12 min arc (see Fig. 4). All of these were always made up of seven equiluminous vertical lines, six of which formed a uniform vertical ribbon of light of total width 2.8 min arc (the lines were...
Fig. 4. Upper, Schematic diagram, not to scale, of the two stimulus conditions. Each target for each eye is a vertical ribbon made of six bright vertical lines, 15 min arc long, equally spaced 30 sec arc apart. There is also a seventh line, either placed in center of ribbon or eccentrically displaced to produce a disparity of the centroid of the left ribbon. Targets appear as two wide bars of light separated by 12 min arc. Centroid condition, Threshold (seconds of arc) for detection of disparity between right and left ribbons. Threshold disparity calculated as the difference between the average positions of the seven lines making up left ribbon seen by the two eyes. Note that there is no disparity in the position of the outer edges (contours). Contour conditions, Threshold (seconds of arc) calculated as for centroid condition. Note that the whole left ribbon, with the seventh line remaining centered, is shown in a position bodily displaced in the right and left eyes to produce disparity.

30 sec arc apart and each had a width of 15 sec arc). The seventh line was superimposed on this ribbon in a manner depending on the particular experiment. In the principal experiment, the seventh line was placed accurately in the middle of the ribbon for the right and left eye's view of the right member of the target pair. In the left member of the target, the seventh line was presented in displaced positions with respect to the center of the ribbon, and the displacement was in opposite directions for the two eyes. By suitable arrangement of the parameters of the displacement, it was possible to present the left target of the stereo-pair with a disparity which had been created by a shift in the mean of its monocular light distributions, without a shift of disparity of its contours. Stereoacuity was now determined by the usual method of constant stimuli (seven equally spaced steps of disparity ranging from crossed to uncrossed disparity through no disparity), but the parameter now was disparity of the mean or centroid of the light distribution whose outside dimensions remained fixed. As a companion experiment, stereoacuity was determined for contour information. This was achieved by leaving the seventh line in the center of the ribbon throughout for all four individual features and creating disparity by presenting the left-side targets in each eye displaced bodily to assume the required disparities. The two experimental conditions were run randomly interspersed so that the subjects could not anticipate the type of target (centroid or contour) to be presented.

The data from three subjects are given in Fig. 4. It is clear that the identification of the retinal location of a feature for the purposes

<table>
<thead>
<tr>
<th></th>
<th>CENTROID CONDITION</th>
<th>CONTOUR CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>GW</td>
<td>5.9±0.5</td>
<td>6.2±0.7</td>
</tr>
<tr>
<td>JR</td>
<td>7.8±0.9</td>
<td>8.0±1.0</td>
</tr>
<tr>
<td>SM</td>
<td>3.5±0.3</td>
<td>3.0±0.3</td>
</tr>
</tbody>
</table>
of disparity detection proceeds in a manner analogous to its identification for the purpose of relative localization in a hyperacuity task. A mechanism is called into play which determines the location of the first moment of the retinal light distribution of the feature, and it is this parameter, rather than contour location, which influences the location assigned to the whole feature under our test conditions. There are natural limitations, of course, to the operation of this mechanism, and they concern principally the width of the feature and possible nonlinearities of light summation.

Discussion

Considerable simplification has been achieved in this study by restricting the stimuli so that no confusion can possibly arise as to which uniocular patterns belong together to form a binocular feature for the purposes of attribution of depth. This is by no means universal in this kind of research, and in fact, starting with the Panum-Wheatstone Grenzfall, there has been a trend to introduce deliberate ambiguity. The best-known example is the random-dot stereogram (Julesz). It is sometimes said that they allow stereoscopic vision in the absence of monocular cues, but this is factually incorrect. Clearly identifiable picture elements are always visible monocularly in them; with appropriate connectivity established through stereoscopy, a global pattern emerges. As pointed out by Julesz and in particularly telling fashion by Marr and Poggio, ambiguity of binocular pairing is resolved in the interest of building a depth framework that is locally homogeneous. In our study, there is never any ambiguity, and this has allowed us to concentrate on a more sensory and less perceptual, if we may be allowed this old-fashioned distinction, level of processing.

The remarkable differences between hyperacuity and stereoscopic acuity as a function of target separation, shown in Fig. 2, can be interpreted in a variety of ways. For example, it can be argued that uniocular separation detection is really quite good and does in fact form the basis for stereoscopic depth discrimination, but for some reason it is unavailable in hyperacuity judgments. The shapes of the curves and their crossing over speak against such a view. Why would the uniocular separation discrimination be better than stereoscopic acuity for some target separations and not for others? We prefer to think of two distinct mechanisms, both, however, founded on an excellent capability to localize the image of a feature on the retina.

On the basis of our experiments, the answer to the question posed in the title is the following.

Retinal positions of the uniocular images of a feature have to be located very precisely. A narrow feature has its place allocated according to the location of the centroid of its light distribution. Depth is attributed directly to individual, binocularly seen features, and no prior identification or discrimination of monocular relationships is needed to induce discriminable depth differences.

REFERENCES


**Information for authors**

Most of the provisions of the Copyright Act of 1976 became effective on January 1, 1978. Therefore, all manuscripts must be accompanied by the following written statement, signed by one author: "The undersigned author transfers all copyright ownership of the manuscript (title of article) to The Association for Research in Vision and Ophthalmology, Inc., in the event the work is published. The undersigned author warrants that the article is original, is not under consideration by another journal, and has not been previously published. I sign for and accept responsibility for releasing this material on behalf of any and all co-authors." Authors will be consulted, when possible, regarding republication of their material.