HYPERACUITY, SUPERRESOLUTION AND GAP RESOLUTION IN HUMAN STEREOPSIS

SCOTT B. STEVENSON, LAWRENCE K. CORMACK and CLIFTON M. SCHOR
School of Optometry, University of California, Berkeley, CA 94720, U.S.A.

(Received 4 January 1989; in revised form 3 March 1989)

Abstract—Different types of stereoscopic acuity were studied with tasks adapted from studies of visual direction acuity. Dynamic, random-element stereograms portraying multiple surfaces in depth and a temporal 2AFC procedure were used for all measurements. The three tasks required detection of a depth offset (Hyperacuity task), a depth-axis thickening (Superresolution task), and a depth-axis gap between surfaces (Gap Resolution task). Thresholds for the three tasks were on the order of 3 set arc, 30 set arc and 200 set arc of retinal disparity, respectively. These results are comparable to those for the analogous visual direction tasks on which they were patterned, suggesting that the underlying judgments involved are similar. Results are used to estimate the intrinsic noise of horizontal disparity processing.

Stereopsis Acuity Resolution Disparity averaging Intrinsic noise

INTRODUCTION

Traditional measures of stereoscopic acuity have universally asked subjects to detect or null a difference in horizontal retinal disparity between targets having different visual directions. In this study, we have used overlapped random-dot surfaces to measure stereocuity for targets having the same visual direction; that is, for every location in the field of view, targets at different depths are simultaneously present. Results from measurements with these stimuli show that there are different types of stereocuity, just as there are different types of acuity for visual direction. Furthermore, these different stereoscopic depth acuities are similar in relative sensitivity to the analogous visual direction acuities, suggesting a similarity in the underlying basis for decision-making in the tasks.

Studies of stereocuity have shown that the characteristics of the targets whose depths are compared have a great impact on the measured thresholds. Stimulus duration (Woo, 1974; Harwerth & Rawlings, 1977; Ogle & Weil, 1958; Uttal, Fitzgerald & Eskin, 1975), contrast (Schor & Heckman, 1989), spatial frequency (Schor, Wood & Ogawa, 1984; Westheimer & McKee, 1980b), velocity (Westheimer & McKee, 1978), distance from the horopter (Badcock & Schor, 1985; Schumer & Julesz, 1984) and proximity to other targets (Tyler, 1979; Weistheimer & McKee, 1980a; Westheimer & Truong, 1988) all have been shown to have a significant effect on stereocuity thresholds. In every case, however, essentially the same aspect of the stimulus configuration has been judged: the relative depth of targets in different visual directions. In our studies, we have held most stimulus characteristics constant at optimal levels (high contrast, spatially and temporally broad-band, dynamic random-element stereograms), and instead have varied the task demands to compare different types of stereocuity with essentially the same stimulus.

The tasks we have used were adapted from methods used in studies of visual direction acuity, as described by Westheimer (1979, 1987). They are referred to here as Hyperacuity, Superresolution and Gap Resolution acuities, and are outlined in Figs 1 and 2. In a Hyperacuity task, subjects detect a small relative displacement of targets which are separated sufficiently (in space or time) so as to make them independently visible. For example, a vernier offset detection task uses targets which are separated in space orthogonal to the direction of the displacement to be detected. In a Superresolution task, using Westheimer's (1987) characterization, subjects detect a small relative displacement of targets which have no separation except that produced by the displacement itself. For example, a two-point resolution task measures Superresolution...
if the subject is given a priori knowledge that the size and shape of the individual targets is held constant. In a Gap Resolution task, subjects detect a relative displacement of adjacent targets sufficient to produce a dip in the combined target profile. For example, measurement of acuity with a Snellen E test requires detection of the peaks and dips in luminance produced by the bars of the E.

Fig. 1. Schematic diagram of targets and corresponding spread functions used to measure (a) Hyperacuity, (b) Superresolution and (c) Gap Resolution. The left and right panels show alternatives given in a forced-choice procedure. A comparison of (a) single unbroken target and broken target measures offset Hyperacuity; (b) single target and two parallel targets measures width or thickness Superresolution acuity; (c) two parallel, separated targets and four targets of equal overall separation (approximating a filled area) measures Gap Resolution acuity. For luminance domain acuity measures, the targets could be thin lines; in these stereoscopic acuity measures, the targets were random-dot planes, with all offsets and separations occurring along the depth, or Z-axis. The spread functions under each target represent hypothetical distributions produced in the nervous system on viewing the targets: presumably, the information on which the subject bases his/her choice in the task. The axes at lower left refer to these spread functions: for our stereoscopic acuity experiments, X is Retinal Disparity and Y could be some measure of Interocular Correlation or Matching Probability.

Fig. 2. Schematic illustration of the supra-threshold appearance of stimuli used in these experiments. Each panel depicts a perspective view of the targets in the right column of the panels in Fig. 1. (a) Hyperacuity stimulus appeared as a fronto-parallel plane of dynamic random-dots split across the middle, with the bottom half closer in depth than the top. (b) Superresolution stimulus appeared as a thick random-dot slab, yielding "Pykno-stereopsis". (c) Gap resolution stimulus appeared as two distinct, overlapped random-dot surfaces with empty space between them, yielding "Dia-stereopsis".

Figure 1 represents these visual direction acuity tasks in schematic form, and describes our experiments on depth acuity as well. Each panel contains a configuration of targets above a hypothetical spread function which is produced in the visual system on viewing them. The exact shape of the spread functions is not known, but we have used luminance point-spread functions to illustrate our points. The target configurations can be thought of as a side-view representation of our random-dot surfaces. Figure 2 shows a perspective view of the target configurations used in our experiments, corresponding to the right-hand column of Fig. 1.

The Hyperacuity, Superresolution and Gap Resolution acuity tasks are represented in Fig. 1a–c, respectively. In each panel, the target
configurations are arranged in pairs, indicating the alternatives which might be offered in a 2AFC paradigm to measure each type of acuity. The axes at the lower left refer to the hypothetical spread functions, representing interocular correlation or matching strength of our random-dot stimuli (Y) vs horizontal retinal disparity (X). Presumably, some representation of this function in the visual system limits performance in each task.

In the Hyperacuity task, subjects must report which interval contained an offset in depth between the upper and lower half-fields of the display, presumably by comparing the means or peaks of the individual disparity-correlation distributions produced by each. In the Super-resolution task, subjects must report which interval contained a thicker depth surface. This judgement might be made by comparing the variance or spread of the alternative distributions. In the Gap Resolution task, subjects must report which interval contained an unfilled region or clear separation in depth between two overlapped surfaces, as compared to two surfaces of the same separation with the depth interval between them filled by additional surfaces. (For separations near the gap threshold, we assume that the spread functions in the 'filled' condition overlap sufficiently that they fill the gap uniformly, as shown in Fig. 1c.) The subject's judgement in the Gap Resolution task might be made by comparing the kurtosis, or specific shape of the two alternative distributions.

Thus, the measurement of these stereoacuities can extend our knowledge of how effectively the visual system can use retinal disparity to make a variety of depth judgements. Additionally, if underlying spread functions limit performance in these tasks as we have suggested, then measurement of these acuity thresholds provide an estimate of their extent, indicating the intrinsic uncertainty of stereoscopic depth localization.

**GENERAL METHODS**

The subjects in all our experiments were corrected to optimal in visual acuity, had good stereopsis and were practiced psychophysical observers.

The majority of our measurements were made using dynamic, random-dot stereograms displayed on 60 Hz non-interlaced video monitors. Some additional measurements (experiment IV) were made using dynamic, random-line stereograms displayed on oscilloscopes at a 200 Hz frame rate.

Figure 3 shows a diagram of the haploscope arrangement used for both display systems. The viewing distance was 57.3 cm and the haploscope mirrors were adjusted so that subjects converged to that distance. The subject's head position was stabilized by a chin cup and forehead rest. The edges of the display image were hidden behind frosted plastic sheets with circles cut out to form circular apertures subtending 11° in diameter. These were placed between the mirror and display on each side so that the edges of the apertures were defocused and the surround region had approximately the same mean luminance as the random-dot display. A microcomputer controlled the stimulus and took responses from the subject via key presses.

The video system used TSD video monitors (model HRM 15-c, P4 phosphor, 60 Hz non-interlaced) driven by Revolution no. 9 video boards in an IBM AT microcomputer. In-house software was used to drift a random grey-level lookup table across random pixel values, making dynamic random two-dimensional noise with 256 gray levels and 2.3 min arc pixel size. Contrast was set to 80% for a static pattern and mean luminance was 40 cd/m².

Disparity was produced by delaying the horizontal sync signal to one monitor, using a Data Delay Devices (model PDU-13256) 8-bit programmable delay chip interfaced with the IBM-AT. Disparity could be changed in units of 1.9 sec arc. Disparity control was calibrated using a Moiré interference technique which yielded a 30:1 magnification of image position, allowing for precise measurement of the offsets produced.
When an experiment required presentation of two or more surfaces in depth, disparity was changed from frame to frame so that when averaged over time, multiple disparity planes appeared superimposed. If one assumes that the rate of presentation is sufficiently high that all depth planes are presented within the integration time of stereopsis, then the presentation of multiple surfaces is effectively simultaneous, and in this paper they are referred to as such. In our experiments, no more than four surfaces were ever presented in this way. In this extreme case, each of the four surfaces was repeated at 15 Hz on our video system and this exceeds the maximum modulation rate for perceiving motion in depth in random-element stereograms (White & Odom, 1985; Norcia & Stevenson, 1981).

The sequence of stimulus events in each trial is schematized in Figs 4a–c for the three acuity tasks. The two alternatives in the forced choice procedure were signalled by beeps from the computer and feedback was given after each trial. The intervals were each 2 sec in duration in every condition. Subjects responded by pressing a key on the computer keyboard, which initiated the next trial. On rare instances in which the subject was distracted during a trial, the trial was repeated with a new random order of the two intervals. Seven levels of stimulus magnitude were presented in random order, comprising a block of trials. Thirty blocks were run in a session, which lasted usually about 30 min. Subjects were allowed to rest as often as needed during the session.

The method used to determine threshold is illustrated in Fig. 5. Psychometric functions were fit to a cumulative normal distribution by converting proportion correct for each stimulus level to its Z-score equivalent, thus linearizing the psychometric function. This function was fit with linear regression to determine the threshold, defined as the disparity level yielding 75% correct performance. Correlation values for the linear model were generally 0.9 or higher, and the residuals showed no clear trend that would have implied that a different model would fit any better.

**Experiment I: hyperacuity**

In our first experiment we measured sensitivity to disparity offsets in a single surface of
Resolution in human stereopsis

dynamic random-dots, with the split-field configuration shown in Fig. 2a. The disparity of the upper half of the random-dot display was fixed, while that of the lower half was varied in magnitude of crossed disparity from trial to trial according to the method of constant stimuli. The disparity was introduced after a warning tone and remained fixed for the duration of the interval. During the blank interval and at all other times the random dot surface was flat (Fig. 4a).

Results. Results from the disparity offset hyperacuity threshold measurements for three subjects are shown in the left-hand side of Fig. 6. Thresholds for this task vary from 1.7 to 3.6 sec arc of retinal disparity. These values are as low as any reported for stereoacuity with dynamic or static, “local” or “global” targets (e.g. McKee, 1983), indicating that the stimulus and methods used introduce no artificial limitations on disparity resolution.

Experiment II: superresolution

In our second experiment we measured sensitivity to disparity difference between two overlapped surfaces of dynamic random-dots (a thickness judgment), as indicated in Fig. 2b. As in the previous experiment, the disparity of the upper half of the random-dot display was fixed. The thickness of the lower half was varied by adding equal amounts of crossed and uncrossed disparity to the two overlapped surfaces, according to the method of constant stimuli. The thick disparity surface was introduced after a warning tone and remained fixed for the duration of the signal interval. During the blank interval and at all other times the random dot surface was flat (Fig. 4b).

Results. Thresholds for thickness detection, shown by the middle bars in Fig. 6, ranged from 22 to 38 sec arc for the three subjects tested. While these thresholds are low in an absolute sense, they are considerably higher than thresholds measured in the offset task, indicating a higher degree of difficulty for thickness judgments. Comments from the subjects supported the characterization of this task as a thickness judgment.

Experiment III: gap resolution

The third experiment measured sensitivity to a gap, or unfilled region between two random-dot surfaces, as indicated in Fig. 2c. The signal, or “gap” intervals in the 2AFC procedure contained two full-field surfaces whose disparity difference varied from trial to trial according to the method of constant stimuli. The blank, or “filled” intervals contained four full-field surfaces which spanned the same disparity range as the two surfaces in the signal intervals, with the two intermediate surfaces acting to fill the gap between the two outer surfaces.

Each interval was preceded by a tone, after which the disparity was introduced for the duration of the signal interval. Briefly between intervals and at all other times the random dot surface was flat (Fig. 4c).

Results. Thresholds for gap detection, shown by the right-hand bars in Fig. 6, ranged from 136 to 351 sec arc for the three subjects tested. Thus, the gap detection task proved to be relatively difficult for our subjects, compared to the offset and thickness detection tasks. Comments from subjects made after having run in this experiment generally support its characterization as a gap detection. In addition, none of the subjects reported seeing motion in depth during the blank intervals when four disparity levels were presented in succession, supporting our assumption that they are effectively simultaneous.

Experiment IV: comparison of gap detection with video and oscilloscope systems

Our experiments comparing these three types of acuity actually began with an oscilloscope-based apparatus, and measurements of gap detection thresholds made using this system are reported here for comparison to those made with the video monitors. This system has...
limitations (lower mean luminance, lower video bandwidth, inability to split the field) which prompted us to move to the video system in our quest for an optimal stereo stimulus, but it has the advantage of a much higher frame rate (200 Hz). A comparison of results measured with the two systems under similar task conditions provides a check on our assumption that multiple depth planes were effectively simultaneous despite the relatively low frame rate of the video system. In addition, a comparison of the two systems indicates the degree of generality of our results across changes in stimulus characteristics, since the oscilloscope system displayed vertical dynamic random lines, instead of random dots.

Methods. The oscilloscope system (Tektronix model 5110, P31 phosphor) was driven by a 200 kHz vertical raster and a 100 Hz horizontal triangle-wave raster, yielding a frame rate of 200 Hz. The Z-axis was modulated with an analog noise source to produce vertical one-dimensional low-pass random noise with continuous gray levels and high-frequency corner at 4 c/deg. Contrast was set to 85% for a static pattern of equal peak-peak voltage as the noise.

Disparity was introduced by adding a d.c. shift to the horizontal raster signals in opposite directions on each scope. Disparity was calibrated with a Moiré interference technique as for the video system above. Disparity was controlled by an Apple II+ microcomputer and associated in-house electronics. The same computer determined trial sequence and tabulated responses from the subjects.

A two-alternative forced-choice procedure identical to that used in Experiment III was used.

Results. Gap detection thresholds for two subjects, measured with the oscilloscope system, are shown in Fig. 7, along with comparison data measured with the video system. Thresholds measured with the oscilloscope system were 130 and 438 sec arc, while those measured with the video system were 136 and 171 sec arc for the same subjects. These results show that the gap task can be performed with the random-line stereogram stimulus nearly as well as with the random-dot stimulus. The difference found for subject LKC can be partially attributed to experience, since although he had considerable experience with random-dot stereograms, he had relatively little experience with the random line stimulus. Subject SBS had considerable experience with both types of display.

Experiment V: vertical disparities

In order to test the hypothesis that these experiments test properties of the stereoscopic depth processing mechanisms we measured thresholds for one subject in the three tasks using vertical disparity, rather than horizontal disparity. While stereopsis is specific to horizontal disparity, any artifacts in the stimulus which might be used to perform the tasks should be equally visible for horizontal and vertical orientations of the monitors. Thus, if the thresholds for vertical and horizontal disparity stimuli were found to be similar, it would suggest that something other than stereopsis was mediating the subjects’ responses.

Methods. The methods used were identical to those used in the first three experiments, except that the video monitors were rotated 90 deg.

Results. Figure 8 shows results from one subject in the three tasks with vertical disparity stimuli. Data from the same subject with horizontal disparity stimuli are replotted for comparison. For all three tasks, sensitivity to horizontal disparity is higher than sensitivity to vertical disparity. The largest difference occurs for the offset task, in which this subject showed more than a factor of ten difference in sensitivity between vertical and horizontal disparity.

The subjective impressions of the subject differ greatly for the two types of disparity. Horizontal disparity manipulations produce apparent changes in the depth, thickness or density of random-dot surfaces in the three
types of acuity tasks. Vertical disparity manipulations produce what is best described as an apparent change in coherence or solidness of the random-dot surface, and this effect is the same for all three tasks. Presumably, this occurs due to spurious matches between vertically adjacent rows of the display as the disparity magnitude approaches the angular size of the dots. Since adjacent rows are uncorrelated, they match to yield a percept of “Confusion”, rather than fusion; that is, the field appears to contain dots at many depths in constant motion, rather than dots at one depth forming a stable surface. The subject thus might have simply chosen the interval which appeared more confused. Regardless of the strategy used to perform the vertical disparity tasks, it is clear that our subjects used a unique process such as stereopsis to perform the horizontal disparity tasks.

**GENERAL DISCUSSION**

The results presented in this series of experiments demonstrate several things about the processing of horizontal disparity.

The Hyperacuity measurements (Experiment I) show that conventional stereoacuity measured with random-dot stereograms can be under 3 sec arc. This value is unusually low for stereoacuity measurements with random-dot patterns. Some previous measures of stereoacuity have indicated that stereothreshold is considerably higher with random-dot patterns than with simple, local targets (e.g. Harworth & Rawlings, 1977). Westheimer and McKee (1980a) showed that crowding of elements, as may occur with random-dot patterns, can cause a reduction of stereoacuity. It appears that our stimulus does not produce crowding effects of the type they describe, despite the high density of pixel elements.

While we have not reported measurements of stereoacuity with local targets for our subjects, informal measurements with a Howard-Dolman apparatus and a Bausch & Lomb Orthorator indicate that stereoacuity measures do not differ significantly between the “local” and “global” stimulus conditions. A more likely explanation for the sensitivity levels we report involves a combination of factors including the experience of our subjects and the broad spatial bandwidth and high contrast of our display. (The dynamic character of our displays does not appear to be essential to these levels of performance: we measured offset hyperacuity with static dots and obtained only slightly higher stereoacuity thresholds.)

The Superresolution measurements (Experiment II) show that one can use the apparent depth–axis thickness of a surface to discriminate a single flat plane from two overlapped planes, just as one can use the apparent width of a bar to discriminate a single thin line from two parallel lines (Westheimer, 1987). This perception of a thick, single surface for small depth separations has been described previously by Tyler (1983) as “Pykno-stereopsis” and our Superresolution thresholds represent the lower disparity limit for this percept.

Schumer (1979), and more recently, Parker and Yang (1987) have studied the effect of overlapping random-dot surfaces in the context of “disparity averaging.” Schumer found that two surfaces appear as a single surface at intermediate depth at disparity separations as large as 2–3 min arc. Thus, the depth of the thick surface produced by two planes appears to be at the “average” of the individual planes’ disparities. Schumer (1979) suggests that the resulting average surface is indiscriminable from a
single surface. However, our results indicate that it is still discriminable from a single plane based on thickness difference when the separations are greater than about 15–30 sec arc. The disparity information carried by the individual surfaces is not lost simply because they overlap.

The Gap Resolution measurements (Experiment III) show that a relatively large separation between surfaces is required before subjects can detect the gap between them. This perception of two distinct overlapping surfaces is described by Tyler (1983) as “Dia-stereopsis.” Our Gap Resolution thresholds represent the lower disparity limit for this percept and the point of transition from Pykno- to Dia-stereopsis.

Assuming that the Gap Resolution task is limited by something like a Rayleigh criterion for gap resolution, this provides an estimate of the effective depth–axis thickness of a random-dot surface. By the Rayleigh criterion, a gap between two overlapping spread functions is just visible when the separation between them is about equal to the width of each one individually. The gap threshold value itself is therefore the best estimate of the thickness of a fronto-parallel random-dot surface: for our subjects, on the order of 2 min arc. This thickness value can also be thought of as the limiting noise level for disparity judgments of this kind.

The measurements with the oscilloscope system show that these results generalize to one-dimensional noise stimuli and higher frame rates, supporting our assumption that the gap resolution measurements were not affected by the relatively low frame rate of the video system. The vertical disparity measurements support the assumption that stereoscopic depth perception, based on horizontal disparity, mediated subject responses in our first three experiments, since one would expect that any other source of information would be used just as easily in the vertical disparity stimuli as in the horizontal disparity stimuli. Responses in the vertical disparity conditions were probably mediated by a fusional mechanism or by detection of an overall change in correlation due to random matches between neighboring rows of dots.

Considering all three types of acuity measures together, it is clear that stereacuity is not just a hyperacuity, but rather that stereoresolution depends on the task used to measure it, just as is true for visual direction acuities. Furthermore, the pattern of results parallels that observed for visual direction acuities measured with luminous points or lines. Thresholds for offset hyperacuities, like vernier acuity, are extremely low compared to gap thresholds measured with the same point or line stimuli and the actual values obtained are similar to those we present here for stereacuities. This similarity in results follows from the similarity in the task demands on the subject, specifically in the aspect of the distribution which must be judged in order for the task to be performed: e.g. the mean, variance or kurtosis of the spread function. While this does not explain the particular ratio of values found for these three tasks, it provides a framework for describing the information extracted in each.

We have described our stimuli in terms of hypothetical spread functions which limit performance in these tasks. In an ideal case, a perfectly correlated surface in depth would be a true plane with zero thickness. In the real case, it is probable that a combination of optical filtering, oculomotor noise and neural filtering introduce uncertainty or “noise” into the localization of this plane, producing the spread functions implied by our results. Assuming something like a Rayleigh criterion is operating in the gap judgments, the threshold for seeing a gap is itself an estimate of the sum of these noise sources. Since our subjects were corrected to optimal acuity, and since vergence movements would not affect the relative disparities of random-dot planes in the overlapped or split-field conditions, it is likely that the nervous system is contributing significantly to the limiting noise distribution.

In summary, our demonstrations that stereopsis can subserve different types of acuity tasks and that performance on these tasks follows performance on analogous visual direction acuity tasks suggests a unified view of acuity judgments. For every point in visual space, one can describe a three-dimensional ellipsoid of positional uncertainty formed by the spread functions on each spatial dimension. This ellipsoid of uncertainty then represents the limit of our ability to resolve points (or lines, or surfaces) as separate in space, whatever the axis of separation.

Acknowledgements—The authors would like to thank subject GSEG for her time and cooperation and the Brown University Psychology Department for generously loaning some of the equipment used. This work was supported by National Eye Institute NRSA grant EY00045 to Scott B. Stevenson.
REFERENCES


