MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 737

August, 1983

HIDDEN CUES IN RANDOM LINE STEREOGRAMS

H. K. Nishihara & T. Poggio

Abstract: Successful fusion of random-line stereograms with breaks in the vernier acuity range has been previously interpreted to suggest that the interpolation process underlying hyperacuity is parallel and preliminary to stereomatching. In this paper (a) we demonstrate with computer experiments that vernier cues are not needed to solve the stereomatching problem posed by these stereograms and (b) we provide psychophysical evidence that human stereopsis probably does not use vernier cues alone to achieve fusion of these random-line stereograms.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support was provided under Office in part by National Science Foundation Grant 79-23110MCS. The research described in this report originally appeared in *Nature*, 300, 347-349, 1982.

Vernier acuity shows that our visual system can interpolate the relative position of a feature with astonishing precision, on the order of a fraction of the distance between neighboring photoreceptors in the fovea. The remarkable properties of this interpolation process have been the subject of several recent studies.[1—9] A major question still unanswered is whether this process is parallel—operating automatically at all times over a large part of the visual field—or is serial—operating selectively at isolated locations only when judgments in the hyperacuity range are required (as in a forced choice task).

Julesz and Spivack's[10] claim that stereo fusion of their random line stereograms is based on vernier cues alone is apparently the best evidence for the former possibility. An example of Julesz and Spivack's ingenious random line stereograms is shown in figure 1a. Figure 2 illustrates, by a small 4 by 5 array of picture elements, the way figure 1a was generated. Each picture element contains a thin vertical line segment at one of two possible horizontal positions—separated by 1/5 the picture element diameter. The random selection of the two positions to make the 4 by 5 array produces 5 vertical lines through the array with frequent small horizontal breaks. Once the left array is generated this way, it is copied into the right array where a rectangular region of picture elements at the centre (dotted lines) is shifted horizontally an integral number of picture element diameters (here the shift is one picture element). Thus the pattern of breaks is correlated in the left and right arrays with zero disparity around the boarders and a fixed horizontal disparity in the central region. Figure 1a is the same except that the array has 100 by 100 picture elements and the horizontal shift is two picture elements. Similar patterns can also be made with horizontal line grids..

The only obvious monocular information is the pattern of minute horizontal breaks occurring at random in the thin vertical line grids. The stereograms clearly yield stereopsis—the centre square is seen in front of the surround—even with monocular breaks of as little as 16 seconds of arc, which is below the threshold for resolving two lines. Note that the stereo matching problem posed by this type of stereogram is not solvable by simply measuring the position disparity between nearest line segments in the two arrays of the stereo pair. The correct matches in the central region are shifted horizontally by one or more picture elements so the nearest matches are all false targets. Thus the matching problem can only be solved by taking into account, in some manner, the pattern of breaks in an area. For this reason Julesz and Spivack argued that it is necessary to detect the vernier breaks and match corresponding ones prior to stereo fusion.

The key question is whether vernier cues are indeed the only monocular information present in this type of stereogram. H. Barlow and, independently, W. Richards suggested to us that receptive fields—for instance, centre-surround ganglion cells—each integrating over angular extents of several minutes of arc might detect coarse monocular structures in such stereograms. If so, this structure alone may be sufficient for driving binocular stereo-matching without need of vernier precision. Indeed, a recent theory of human stereo vision[11—14] (see also ref. 15) proposes that the two images are filtered through several, roughly bandpass, channels, each with a different centre-surround receptive field size; and that stereo-matching is based mainly on zero-crossings—boundaries between regions of positive and negative response—in each of these channels. Essentially all images—for instance random dot stereograms—when processed in this way, reveal a clear coarse structure (see for example figure 3.6 in ref. 13).

When the random line stereogram of figure 1a is filtered in the same way, coarse monocular features do appear, as shown in figure 1b. The underlying cause of this coarse structure is the variation in spacing between neighboring lines. While the absolute changes in spacing are small, the percentage change in the line spacing is appreciable. For the patterns described by Julesz and Spivack, the ratio of smallest to largest line spacing is 2/3. Figure 1b shows the locations where a centre-surround operator would give positive (white) and negative (black) responses to the random line pattern shown in 1a. The coarse structure here is well preserved by the photoreceptor sampling at 30" intervals. Interpolation between these points to increase spatial precision is not necessary.

Many matching algorithms can successfully use the monocular features provided by the larger operators to solve the stereo correspondence problem. As an example, figure 1c shows the output obtained with the cooperative algorithm of Marr and Poggio[16,17] when applied to the binary array shown in figure 1b. It correctly identifies the different disparities of the centre square and the surrounding background. Essentially the same results are obtained with the other stereo matching algorithms developed in our laboratory when they are applied to the same filtered images. Thus these random-line stereograms can be solved without need of vernier precision.

At this point, our computer demonstration only shows that non-vernier, monocular cues are present and could be used for successful stereo matching, but it leaves open the question of whether our visual system actually uses them instead of the vernier cues. Psychophysical experiments can be used to test these alternative hypotheses. In particular, we examine the differences one would expect between systems relying entirely on one or the other approach. As noted above, the coarse structure shown in figure 1b is due to the large line spacing variation in the patterns. Increasing the line spacing while maintaining a fixed break size reduces the detectability of coarse structure present in the patterns but does not affect the detectability of the vernier breaks (larger spacings should actually improve vernier acuity[6]).

Figure 3 shows a random line stereogram with twice the line spacing as shown in figure 1, but with the same break size and disparity for the central region. This stereo pair is difficult to fuse when viewed at a distance which gives a vernier break size of about 15" while the pair of figure 1a can be fused easily at that distance. These patterns were computer generated on a high resolution CRT and viewed stereoscopically from a distance of 4 meters at which the size of the vernier breaks was about 15". We tested 4 subjects under these conditions and for all of them the stereogram of figure 3a was much more difficult to fuse than figure 1a. This result is consistent with the computer simulations (compare figure 1c with figure 3c).

Other psychophysical observations, though less critical, are also consistent with the large channel hypothesis. For instance, we expect that (a) random line stereograms could be fused for very small break sizes if line spacings were made proportionally smaller—fusion would eventually be limited by the loss of contrast as the receptive field extends across more than several lines; and (b) the disparity limit should be larger for horizontal random line patterns. The regions of positive response in figure 1b are elongated slightly in the direction of the lines. This gives a larger average horizontal distance between vertical zero-crossings for the case of horizontal random-line patterns, and thus larger disparities can be viewed without confusion from false targets. Julesz and Spivack[10] report both phenomena—fusion continuing weakly with break sizes beyond the threshold for vernier acuity, and larger disparity range for horizontal patterns.

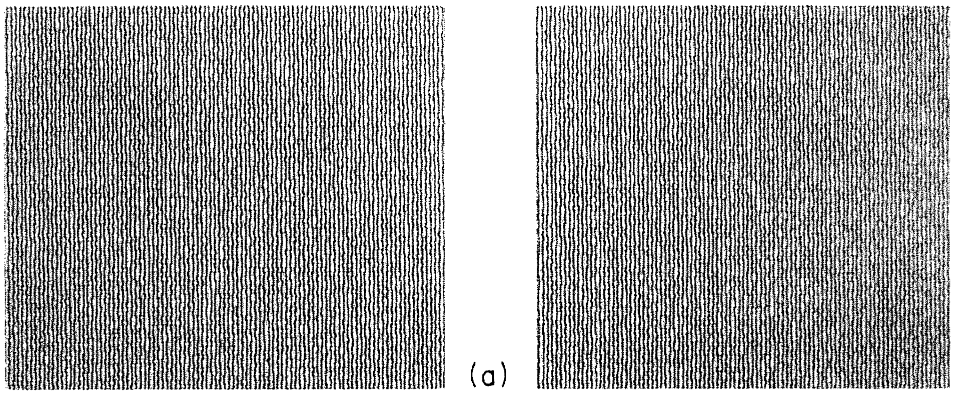
In summary, our computational experiments establish that the stereo matching problem posed by the random line stereograms *can* be solved without vernier interpolation, while the psychophysical data suggest that human perception of these stereograms does not rely on vernier acuity cues. The reason for this is that the elegant random dot and random line stereograms of Julesz[18] and Julesz and Spivack[10], when appropriately processed, exhibit otherwise hidden cues that simplify the stereo matching problem. (The same cues may have a significant role in human stereo vision of natural images.) The stereograms therefore cannot be used to support the hypothesis that the detection of vernier breaks is computed in parallel and made available to other later processes.

FIGURES

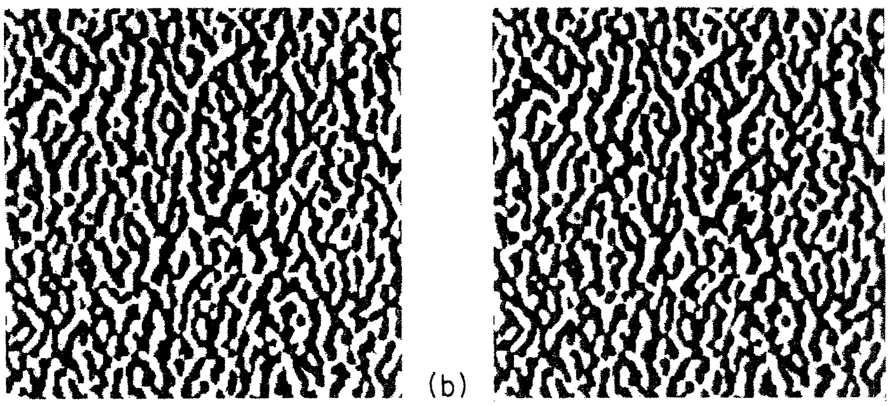
Figure 1 (a). Random-line stereograms of thin vertical line segments with small breaks, portraying a centre square in front of the surround. When the patterns as printed here are viewed at 1.4 meters, the vernier breaks are about 15" and the minimum line spacing is 1' (the spacing and the diameter of the cones in the human fovea is about 30"). The central square has a disparity relative to the background of 10 times the break size.

(b). The stereogram of (a) filtered through a centre-surround operator consisting of the difference of two gaussians (DOG).[19] The width of the operator's centre is 16 times the break size in (a). When (a) is viewed from a distance making the break size 15", this corresponds to what a filter with a 4' centre would compute—a medium sized channel as revealed by psychophysical experiments.[20,11,21] Positive values are shown white and negative black; white (black) values would then represent the activity of the corresponding on-(off) centre-surround ganglion cells. The patterns in (a) were first convolved with a gaussian ($\sigma = 20''$) to simulate the optics of the eye, this result was then sampled at 30" intervals prior to convolution with the difference of gaussian operator.

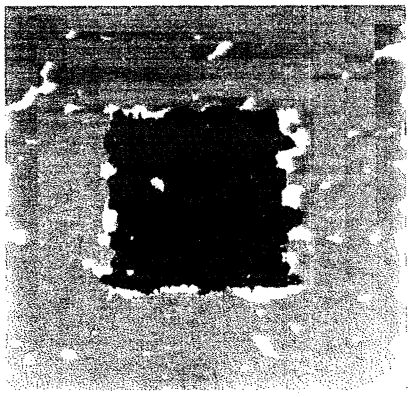
(c). The result of applying the stereo matching algorithm of Marr and Poggio (see legend of fig. 3c) to the (binary) array shown in (b). The grey levels here indicate the disparity of the matches obtained by the algorithm after 5 iterations, black corresponding to +6 pixels disparity (+3'), and zero disparity appears as grey. White indicates non-matched pixels. The algorithm successfully extracts the correct disparity information without need of vernier interpolation.



(a)



(b)



(c)

Figure 2 Method by which the stereogram of fig. 1a was generated (from Julesz and Spivack[10]).

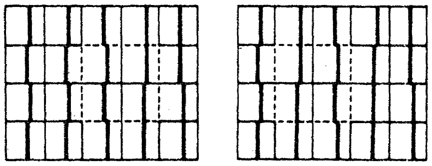
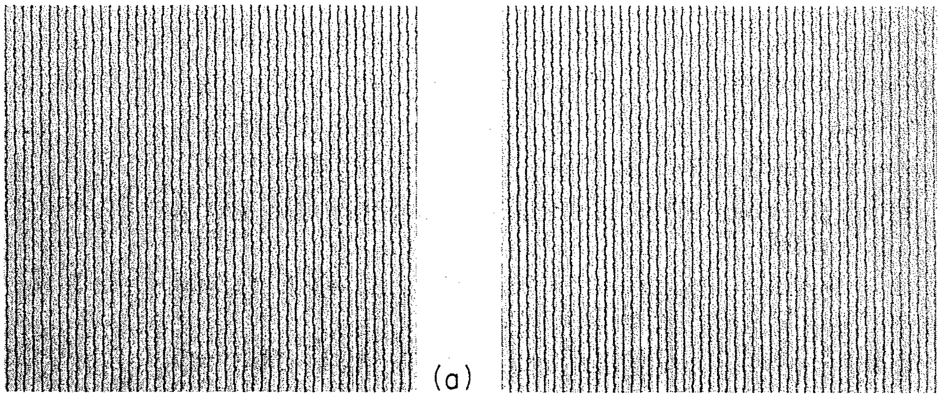


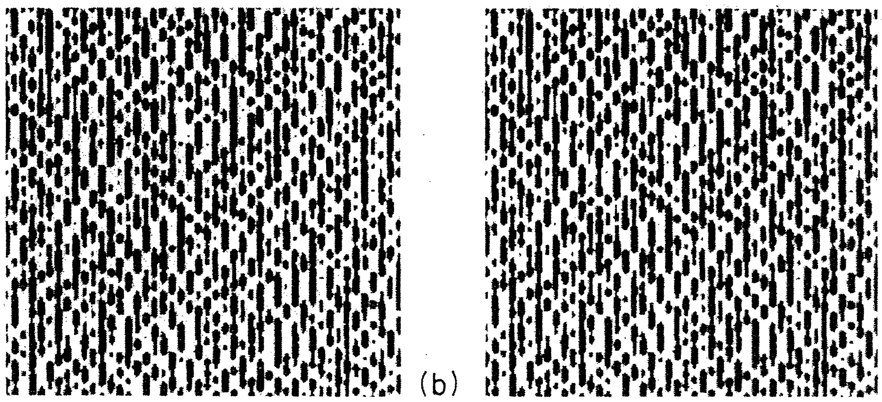
Figure 3 (a). A random line stereogram as in fig. 1a with the same break size but twice the line spacing. The disparity of the central square is the same as in fig. 1.

(b). The sign of the convolution of fig. 3a with a centre-surround operator as in fig. 1b. The vertical line grid in (a) dominates the coarse structure shown here; there is much less effect from the variation in line spacing due to the pattern of breaks.

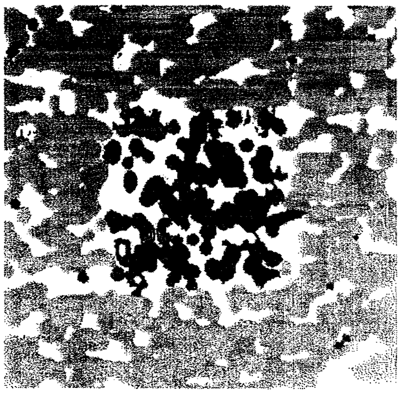
(c). The result of applying the stereomatching algorithm to the array of fig. 3b after 5 iterations. No match is obtained (white regions) over a larger portion of the array compared with fig. 1c. The algorithm used here, as in fig. 1c, is the cooperative algorithm of Marr and Poggio[16,17] operating on the sign of the stereo images after convolution with a DOG mask. Parameters are the same as described in ref. 16. The network corresponding to the algorithm is loaded by an "and" operation on the "binarized" convolved images. In this way the cooperative algorithm originally described for random dot stereograms can be successfully used on stereo pairs of natural images—at a number of different resolutions (set by the size of the DOG mask). In fig. 1c and 3c we have used 7 disparity layers covering a total disparity range of ±7 pixels. We have run the algorithm with different disparity ranges and also with somewhat different parameters obtaining similar results.



(a)



(b)



(c)

REFERENCES AND NOTES

[1] Barlow, H.B. *Nature* **279**, 189-190 (1979)

[2] Barlow, H.B. *Proc. R. Soc.Lond. B.*, **212**, 1-34 (1981).

[3] Westheimer, G. *Am. J. Optom. physiol. Opt.* **53**, No. 7, 362-364 (1976)

[4] Westheimer, G., Hauske, G. *Vis. Res.* **15**, 1137-1141 (1975)

[5] Westheimer, G., and McKee, S.P. *Vis. Res.* **17a**, 89-93 (1977a)

[6] Westheimer, G., McKee, S.P. *Vis. Res.* **17b**, 941-947 (1977b)

[7] Westheimer, G., McKee, S.P. In: *Investigative Ophthalmology and Visual Science*. St. Louis **118**, 614-621 (1979)

[8] Crick, F.H.C., Marr, D., Poggio, T. In: "*The Cortex*", ed. F.O.Schmitt. M.I.T. Press (1981). Also available as M.I.T.A.I. Memo 557 (1980)

[9] Fahle, M. and Poggio, T. *Proc. Roy. Soc. Lond.B.*, **213**, 451-477 (1981)

[10] Julesz,B., Spivack, G.J. Science, **157**, 563-565 (1967)

[11] Marr, D., Poggio, T. *Proc. R. Soc. Lond. B* **204**, 301-328 (1979). Also available as M.I.T. A.I. memo 451 (1977).

[12] Marr, D.C., Poggio, T. *A.I. Memo* No. 558, (July 1980)

[13] Grimson, W.E.L.: *From images to surfaces* M.I.T. Press (1981)

[14] Marr, D.: *Vision*. W.H. Freeman, San Francisco (1982)

[15] Mayhew, J.E.W., Frisby, J.P. *Artificial Intelligence*, **17**, 349-385 (1981)

[16] Marr, D., Poggio, T. *Science* **194**, 283-287 (1976)

[17] Marr, D., Poggio, T., Palm, G. *Biol. Cybernetics*, **28**, 223-239 (1978)

[18] Julesz,B. *Foundations of cyclopean perception* Chicago:University of Chicago Press (1971)

[19] Nishihara, H.K., Larson, N.G. *Proc. DARPA image Understanding Workshop* 114-120 Science Application Inc., Arlington, VA (1981)

[20] Wilson, H.R., Bergen,J.R. *Vis. Res.*, **19**, 19-32 (1979)

[21] Marr, D., Hildreth, E. *Proc. Roy. Soc. Lond. B* **207**, 187-217 (1980)