

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

Working Paper 230

April 1982

HIDDEN CUES IN RANDOM LINE STEREOGRAMS

H.K. Nishihara and T. Poggio

Abstract: Successful fusion of random-line stereograms with breaks in the vernier acuity range has been 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.

A.I. Laboratory Working Papers are produced for internal circulation, and may contain information that is, for example, too preliminary or too detailed for formal publication. It is not intended that they should be considered papers to which reference can be made in the literature.

© MASSACHUSETTS INSTITUTE OF TECHNOLOGY 1982

Vernier acuity shows that our visual system can infer 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 (see Barlow 1979, 1981, Westheimer 1976, Westheimer and Hauske 1975, Westheimer and McKee 1977a, 1977b, 1979, Crick et al., 1981, Fahle and Poggio, 1981). A major question still unanswered is whether the process itself is parallel—computing at all times the interpolation over a large part of the visual field—or is selective—operating at isolated locations only when judgments in the hyperacuity range are required.

Julesz and Spivack's (1967) claim that stereo fusion of their random line stereograms is based on vernier clues alone apparently eliminates the second possibility. Their conclusion suggests that monocular interpolation provides inputs to the later matching process over a rather large area of the visual field, including thousands of small vernier breaks (Crick et al., 1981). In this note, however, we show that (a) very low resolution structure exists in such patterns and can be used for stereo matching; and (b) human performance on these patterns is consistent with the use of this large scale structure, but is not easily explained if stereo matching is based on the detection and explicit matching of isolated vernier line breaks. In summary, our computational experiments establish that this stereo matching problem *can* be solved without vernier interpolation, whereas psychophysical data suggest that human perception of the random line stereograms does not rely on vernier acuity clues.

An example of Julesz and Spivack's ingenious random line stereograms is shown in figure 1a. The only obvious monocular cues there are minute breaks occurring at random in the thin vertical (or horizontal) line grids. The stereograms clearly yield stereopsis—the center 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. The key question is whether vernier cues are indeed the only monocular cues present in this type of stereogram. H. Barlow and, independently, W. Richards suggested to us that receptive fields—for instance, center-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. Indeed, a recent theory of human stereo vision (Marr and Poggio 1979, 1980; Grimson, 1981, see Marr 1982) proposes that the two images are filtered through about four independent channels, each with a different center-surround receptive field size; and that stereo-matching is based on zero-crossings—boundaries between regions of positive and negative response—in each of these channels.

When the random line stereogram of figure 1a is filtered in this 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 center-surround operator would give positive (white) and negative (black) responses to the random line pattern shown in 1a. Figure 1c shows the result with a smaller center-surround operator. The details of these operators are explained in the figure legend.

Many matching algorithms can successfully use the monocular features provided by the larger operators to solve the stereo correspondence problem. As an example, figure 1d shows the output obtained with the cooperative algorithm of Marr and Poggio (1976; see also Marr et al., 1977) when applied to the binary output shown in figure 1b. It correctly identifies the different disparities of the center square and the surrounding background. Essentially the same results can be obtained with the other stereo matching algorithms developed in our laboratory when they are applied to the same filtered images. Thus random-line stereograms can be solved without first identifying the vernier breaks in the patterns.

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 alternate 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.^{1,2}

Figure 2 shows a random line stereogram with twice the line spacing as shown in figure 1, but with the same break sizes. They are difficult to fuse when viewed at the distances for which figure 1a can easily be fused. We tested XXX subjects under these conditions and for all of them the stereogram of figure 2a was much more difficult to fuse than figure 1a. This result is consistent with the computer simulations (compare figure 1d with figure 2d). Reducing the break size to half as much with the same line spacing also makes the pattern more difficult to fuse. However, patterns with half the line spacing and half the break size as the original remain easy to fuse.

In conclusion, the question of whether interpolation of the vernier type is computed in parallel and made available to other later processes is still open. The elegant random dot and random line stereograms of Julesz (1971) and Julesz and Spivack if 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.

LEGENDS to the FIGURES

Figure 1

- a. Random-line stereograms of thin vertical line segments with small breaks, portraying a center square in front of the surround. When viewed at XXX cm., the vernier break is 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 part (a) filtered through a center-surround operator consisting of the difference of two gaussians. The width of the operator's center 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' center would compute. This corresponds to the medium sized channel revealed by psychophysical experiments (Wilson and Bergen, 1979; Marr and Poggio, 1979; Marr and Hildreth, 1980). Positive values are shown white and negative black; white (black) values would then represent the activity of the corresponding on-(off) center-surround ganglion cells.
- c. A center-surround operator (as in (b)) with a center width of four times the break size in (a). For a 15" break size this would correspond to a 1' channel. When sampled at 30" intervals—the cone spacing—little more than the vertical line organization of the pattern in (a) can be detected without interpolation.
- d. The result of applying the stereo matching algorithm of Marr and Poggio³ to the (binary) array shown in (b) sampled at 30" intervals. The grey levels here indicate the disparity of the matches obtained by the algorithm after 5 iterations, white corresponding to +6 pixels disparity (+3'), black to -6 pixels (-3') (black is also assigned to non-matched pixels). The algorithm successfully extracts the correct disparity information without need of vernier clues. The same algorithm fails to provide correct matches on the input shown in (c) because of the absence of coarse structure.

Figure 2

- a. A random line stereogram as in figure 1(a) with the same break size but twice the line spacing. The disparity of the central square is the same as in figure 1.
- b. The sign of the convolution of figure 2a with a center-surround operator as in figure 1b. The coarse structure is much weaker.
- c. Same as figure 1c for the stereogram of figure 2a.
- d. The result of applying the stereomatching algorithm to the array of figure 2b after 5 iterations. Matching is more difficult. It is impossible with this size of receptive field for an even sparser stereogram.

REFERENCES and NOTES

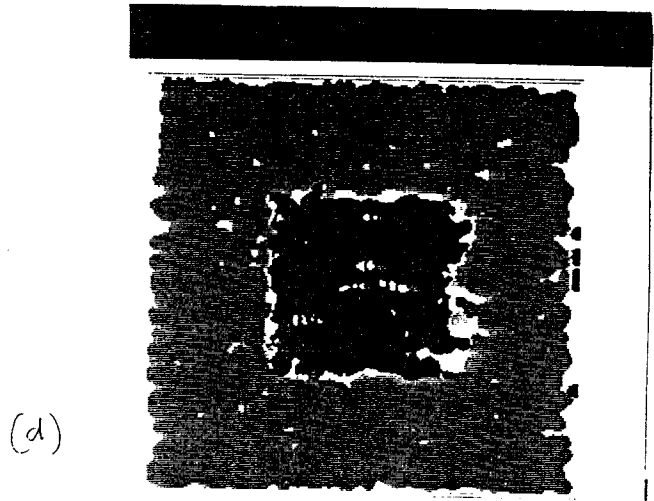
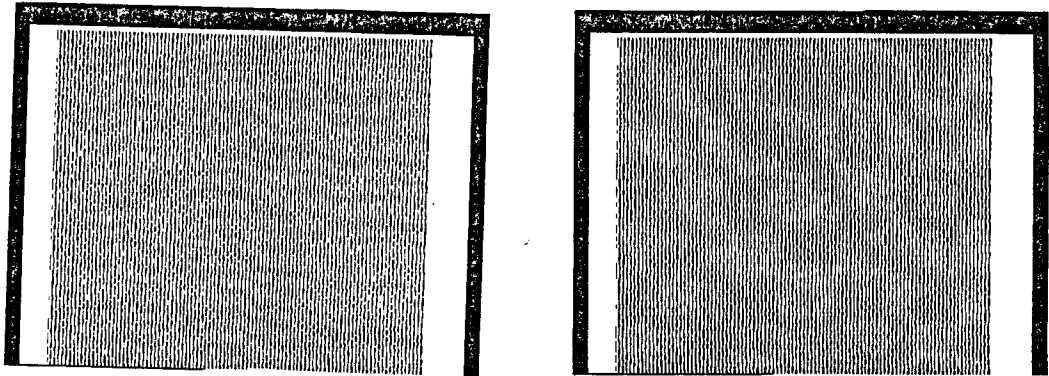
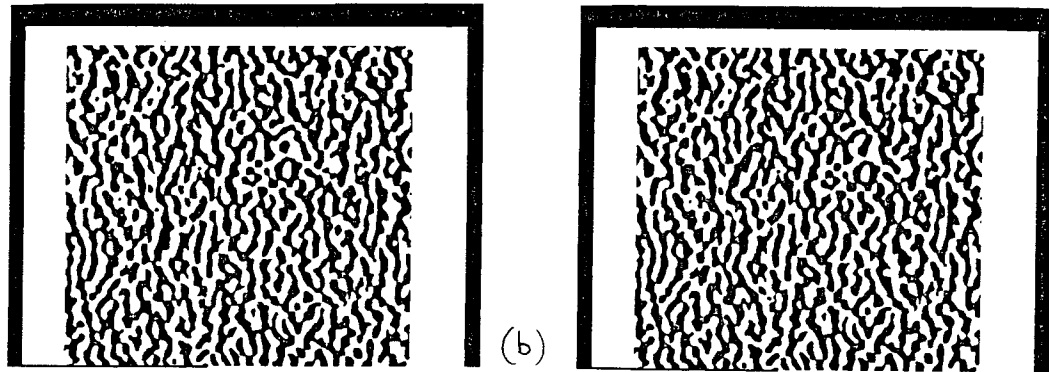
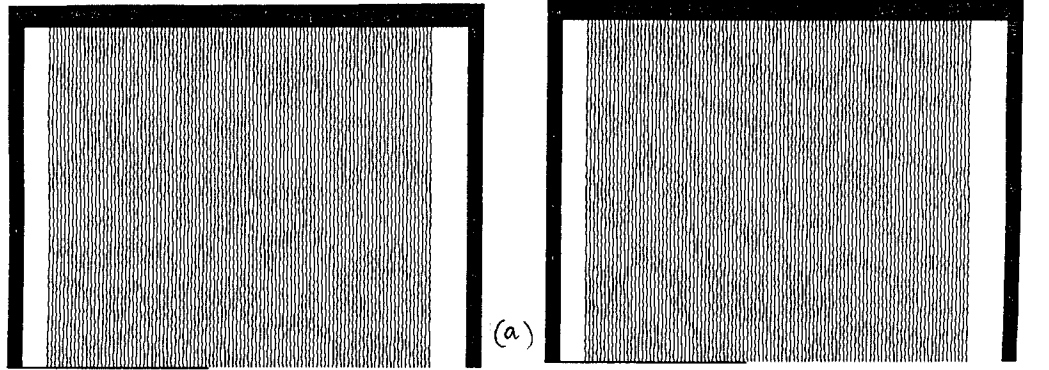
- Barlow, H.B.: Reconstructing the visual image in space and time. *Nature* 279, 189-190 (1979)
- Barlow, H.B.: Critical limiting factors in the design of the eye and visual cortex. *Proc. R. Soc. Lond. B.*, 212, 1-34 (1981).
- Crick, F.H.C., Marr, D., Poggio, T.: An Information-Processing Approach to Understanding the Visual Cortex. In: "The Cortex", ed. F.O.Schmitt. M.I.T. Press (1981). Also available as M.I.T.A.I. Memo 557 (1980)
- Fahle, M. and Poggio, T.: Visual Hyperacuity: Spatiotemporal interpolation in human stereo vision. *Proc. Roy. Soc. Lond.B.*, 213, 451-477 (1981)
- Grimson, W.E.L.: From images to surfaces. M.I.T. Press (1981)
- Julesz, B. Foundations of cyclopean perception. Chicago: University of Chicago-Press (1971)
- Julesz, B., Spivack, G.J.: Stereopsis based on vernier acuity cues alone. *Science*, 157, 563-565 (1967)
- Marr, D.: Vision. W.H. Freeman, San Francisco (1982)
- Marr, D., Hildreth, E.: Theory of Edge Detection. *Proc. Roy. Soc. Lond. B* 207, 187-217 (1980)
- Marr, D.C., Poggio, T.: From Understanding Computation to Understanding Neural Circuitry. In: Neuronal Mechanisms in Visual Perception. *Neurosciences Res. Prog. Bull.* 15, No. 3, 470-488. Eds. E. Poppel, R. Held, J.E. Dowling (1977)
- Marr, D., Poggio, T.: Cooperative Computation of Stereo Disparity. *Science* 194, 283-287 (1976)
- Marr, D., Poggio, T.: A Computational Theory of Human Stereo Vision. *Proc. R. Soc. Lond. B* 204, 301-328 (1979)
- Marr, D.C., Poggio, T.: Some Comments on a Recent Theory of Stereopsis. A.I. Memo No. 558, (July 1980)
- Nishihara, H.K., Larson, N.G.: Toward a real time implementation of the Marr-Poggio stereo matcher. Proc. DARPA image Understanding Workshop 114-120 Science Application Inc., Arlington, VA (1981)
- Marr, D., Poggio, T., Hildreth, E.: Smallest Channel in Early Human Vision. *J. Opt. Soc. Am.* 70, No. 7, 868-870 (1980).
- Westheimer, G.: Diffraction theory and visual hyperacuity. *Am. J. Optometry & Physiological Optics* 53, No. 7, 362-364 (1976)
- Westheimer, G., Hauske, G.: Temporal and spatial interference with vernier acuity. *Vision Res.* 15, 1137-1141 (1975)
- Westheimer, G., and McKee, S.P.: Integration regions for visual hyperacuity. *Vision Research* 17a, 89-93 (1977a)
- Westheimer, G., McKee, S.P.: Spatial configurations for visual hyperacuity. *Vision Research* 17b, 941-947 (1977b)
- Westheimer, G., McKee, S.P.: What prior unocular processing is necessary for stereopsis? In: *Investigative Ophthalmology and Visual Science*. St. Louis 118, 614-621 (1979)

Wilson, H.R., Bergen, J.R. : A four mechanism model for threshold spatial vision. *Vis. Res.*, 19, 19-32 (1979)

(1) Larger spacings should actually improve vernier acuity (Westheimer, 1977b).

(2) The large channel hypothesis would also predict that (a) random line stereograms could be fused for very small break sizes if line spacings were made proportionally smaller; 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 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 (1967) report both phenomena—fusion continuing weakly with break sizes beyond the threshold for vernier acuity, and larger disparity range for horizontal patterns. Notice that scaling of the line spacings would be limited by the loss of contrast as the receptive field extends across more than several lines.

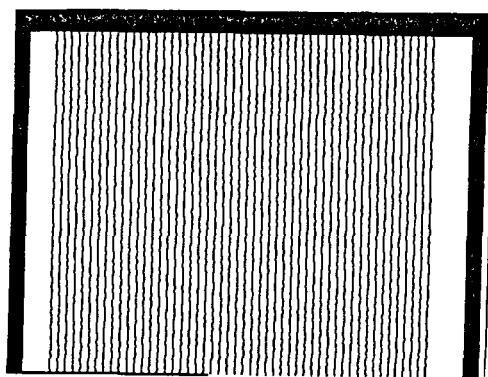
(3) This is the cooperative algorithm of Marr and Poggio (1976, see also Marr et al., 1977) operating on the sign of the stereo images after convolution with a ΔG mask. Parameters are the same as described in Marr and Poggio (1976). The network corresponding to the algorithm is loaded by an "and" operation 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 ΔG mask). In figure 1c and 2c 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.



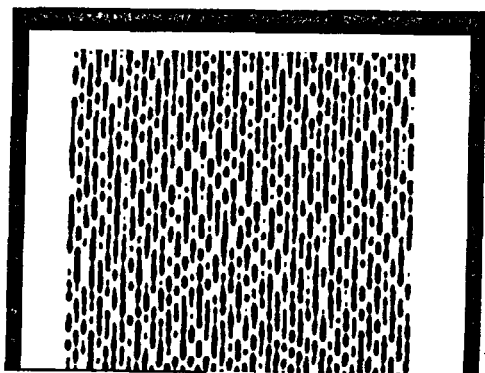
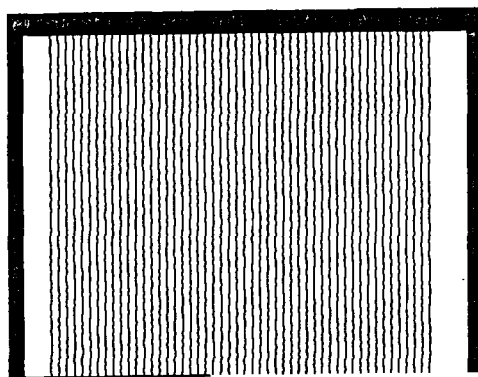
DE2 res:1 min:0 max:255

Fig. 1

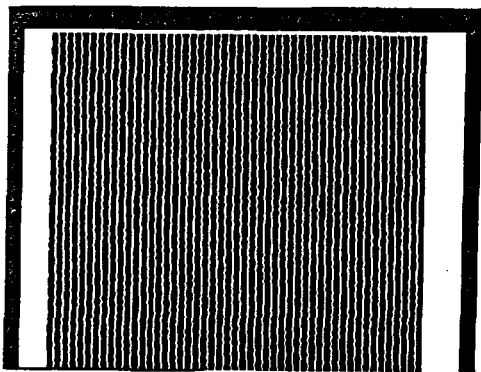
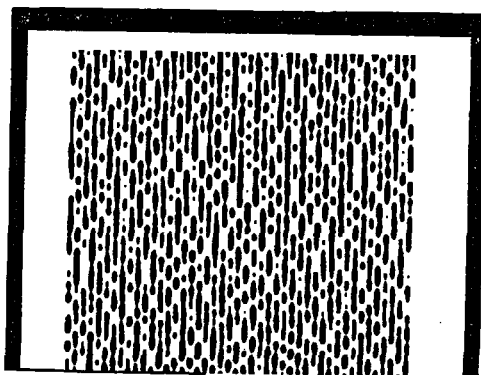
9



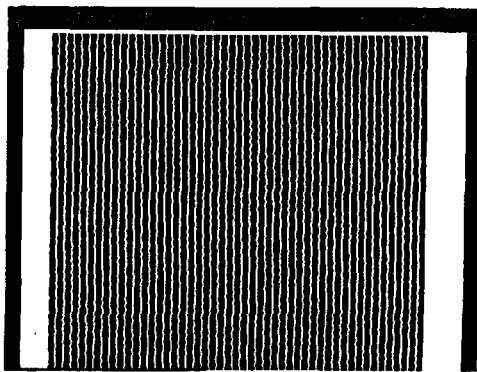
(a)



(b)



(c)



(d)

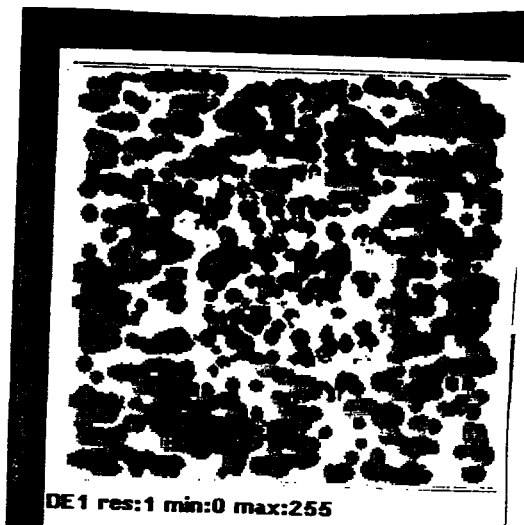


Fig. 2