Stereopsis from kinetic and flicker edges

K. PRAZDNY
Fairchild Instrument and Camera Corporation
Palo Alto, California

Binocularity seems to appear early in the visual system, before any complex form recognition processes take place. The question of the nature of the primitives used for binocular matching is still an issue despite a large number of studies concerned with this problem (e.g., Julesz, 1971; Julesz & Schumer, 1981; Kaufman, 1974). More than a decade ago, Ramachandran, Madhusudhan Rao, and Vidyasagar (1973) argued that a stereoscopic mechanism sensitive to monocularly discriminable contours, irrespective of the detailed structure of regions bounded by them, is a sufficient—although, as shown by Julesz (1964, 1971), not a necessary—basis for stereopsis. They demonstrated that observers could obtain stereopsis (report the direction of depth of a central patch with respect to its surround) by matching of pure texture edges. Stereopsis was also reported under isoluminant conditions using monocularly recognizable stereograms (Comerford, 1974; De Weert, 1979; Ramachandran et al., 1973) and when the matching contours differed in type. For example, Ramachandran et al. found stereopsis in stereograms in which one image contained only texture or color contours and the other only luminance contours. This latter demonstration argues against theories (Frisby & Mayhew, 1977; Mayhew & Frisby, 1976) that try to explain stereopsis in stereograms containing texture contours by a partial point-for-point correspondence in spatial-frequency-tuned luminance-domain channels without any need to involve more complex mechanisms relying on extended monocular processing.

In general, results scattered in the literature do seem to suggest that stereopsis can be driven by the results of a rather sophisticated monocular processing based on various edge-detecting mechanisms, that is, primitives more complex than simple zero-crossings or local extrema in the output of linear filters based on the Laplacian of Gaussian smoothed images (Marr & Poggio, 1979; Mayhew & Frisby, 1981). This conclusion is strengthened by our observation that stereopsis can be achieved in stereograms in which the relevant information is portrayed only by kinetic or flicker edges.

Kinetic edges are boundaries defined by motion information alone, without any luminance, texture, or color cues. Our displays were random-dot stereo cinematograms. A set of random-dot stereogram frames was constructed so that two stationary vertical bars were correlated across the consecutive frames, while their surround consisted of dynamic noise, that is, was uncorrelated in time and across the two half-images (Figure 1). Viewing one frame in isolation reveals only uncorrelated random-dot fields (black and white have equal probability of occurrence), that is, at any time instant, there is no spatial binocular-disparity information available that specifies even the existence of the two bars. When the sequence of stereograms was displayed in rapid succession as a cinematogram (the stimulus interval was $\frac{1}{30}$ sec, and the interstimulus interval was $\frac{1}{60}$ sec), two stationary bars in each half-image could clearly be seen. The random dots within the bars were also uncorrelated across the eyes. The (static) disparity of the bars (± 16 arc min, i.e., one bar had the nasal and the other the temporal disparity of the same magnitude) was created by manipulating their relative positions in the enclosing random-dot field. All observers could determine the relative depth relationship of two bars (which one was in front) without difficulty (100% success rate).[1]

Lee (1970) performed a similar experiment in which a vertical bar of static noise was slowly moving in the field of static noise (the left and right eyes were presented with uncorrelated random-dot fields). Depending on the disparity, the bar could be seen as moving either in front or behind the window.

Relative depth of two bars can also be perceived when the edges are defined only by flicker. In a field of static or dynamic random-dot noise (uncorrelated across the eyes), we flickered two bars, that is, the black and white dots were continuously changing their polarity. The interior of the bars was uncorrelated across the two images; in fact, they were just an integral part of the random-dot field—each stereo cinematogram frame viewed separately presented only a rivalrous display. Again, our observer could determine the relative depth relationship (nearer/farther) of the two bars (91% success rate) after the flicker was switched on.

It is tempting to conclude that human stereoscopic vision, like human motion perception (Anstis, 1980; Braddick, 1974), consists of at least two separate mechanisms. There are, in principle, two different strategies that could be used by a visual system in either motion or stereo. The first strategy can be based on some form of correlational, that is, point by point, analysis of the two images without any extensive monocular preprocessing (Anstis, 1980; Julesz, 1971; Marr & Poggio, 1979; Mayhew & Frisby, 1981). In this scheme, the extraction of binocular information would precede the extraction of edge-based shape information. A conventional random-dot stereogram (Julesz, 1971) is the best example: a shape can be discerned only after the two images are fused. The second strategy would use the output of monocular edge extraction to establish correspondence between images containing contours but no point-by-point correlation (Ramachan-

LEFT IMAGE   RIGHT IMAGE
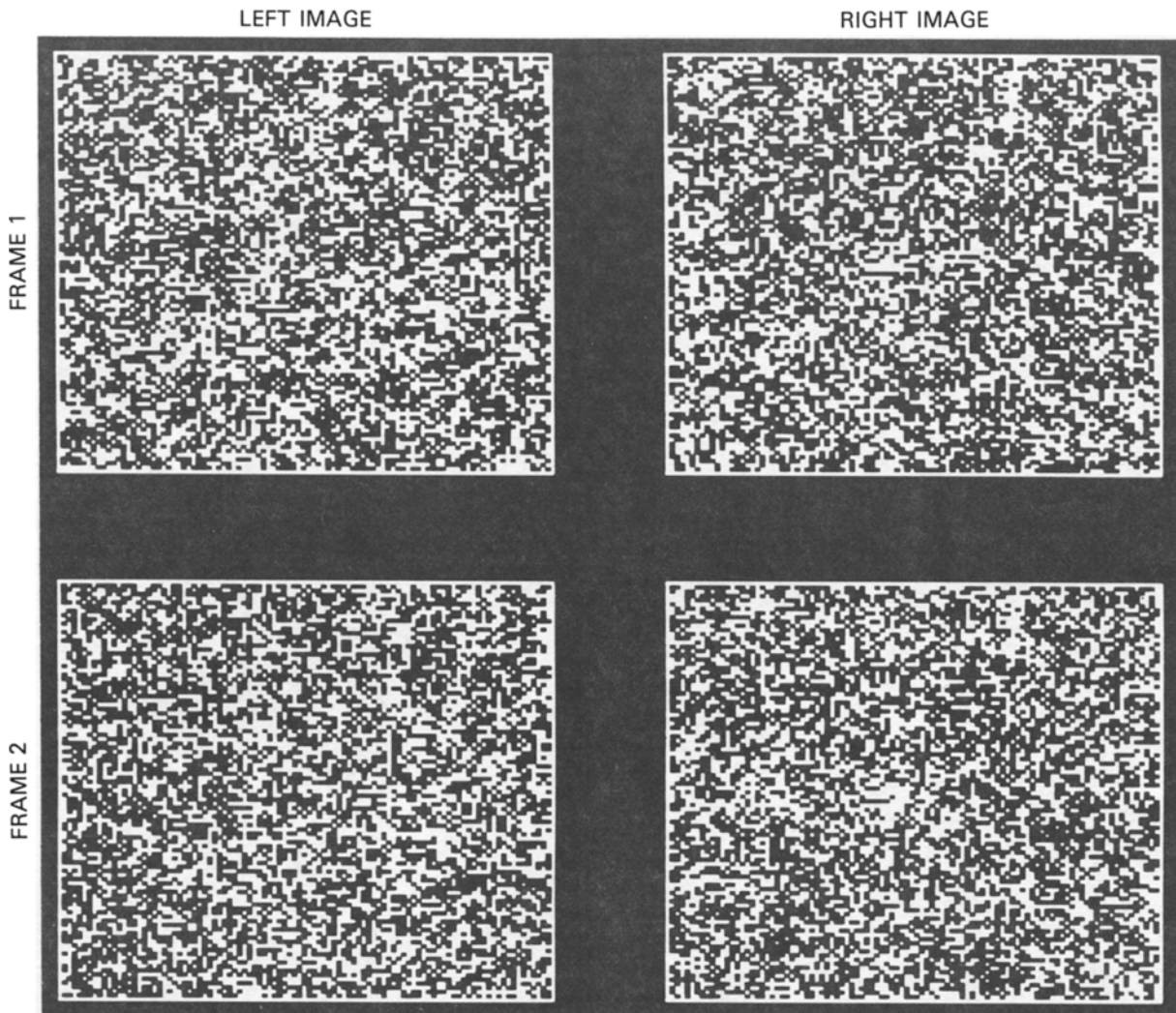
FRAME 1

FRAME 2

Figure 1. Two consecutive frames of a stereo cinematogram. Kinetic contours portraying two vertical bars were created by shifting a number of columns (15) by an integral number of rows (1) across two consecutive frames. Each individual frame (stereogram) in isolation contains no information about the bars: only a rivalrous random-dot field can be perceived. The information for the vertical bars is contained in correlation across time (i.e., across two consecutive frames) and can be observed by turning the figure by 90° and fusing the two time frames as a stereogram. The spatial disparity is carried by the difference in positions of the kinetic contours delineating the bars in the two half images.

dran et al., 1973). In this scheme, various contour-extracting mechanisms feed into a common monocular processing center, the output of which may then be available for binocular combination.

Such parallels between stereopsis and the motion-detecting system are suggestive indeed. On the other hand, the stereopsis based on matching contours in the absence of point-by-point correlation has been reported to be rivalrous, even to the point that the impression of depth is only "qualitative" (Ramachandran et al., 1973). This is unlike apparent motion where the subjective motion impressions are not at all vague. It is thus possible that the disparities associated with such contours cannot be measured accurately by the binocular system (perhaps only their signs can be monitored). The impression of depth

in such displays may be due predominantly to the vergence eye-movement control system, which is known to be sensitive to monocular contours (Kidd, Frisby, & Mayhew, 1979). We have found, however, that it is possible to make the bars (defined only by kinetic edges) move in depth when their disparity is varied across the frames. It remains to be seen if such perceptions can be accounted for by an appeal to vergence eye movements.

REFERENCES

ANSTIS, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society* (London), Series B, **290**, 153-168.
BRADDICK, O. J. (1974). A short range process in apparent motion. *Vision Research*, **14**, 519-528.

COMERFORD, J. P. (1974). Stereopsis with chromatic contours. *Vision Research*, **14**, 975-982.

DE WEERT, C. M. (1979). Colour contours and stereopsis. *Vision Research*, **19**, 555-564.

FRISBY J. P., & MAYHEW, J. E. W. (1977). Global processes in stereopsis. *Perception*, **6**, 195-206.

JULESZ, B. (1964). Binocular depth perception without familiarity cues. *Science*, **145**, 356-362.

JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago: Chicago University Press.

JULESZ, B., & SCHUMER, R. A. (1981). Early visual perception. *Annual Review of Psychology*, **32**, 575-627.

KAUFMAN, L. (1974). *Sight and mind: An introduction to visual perception*. Oxford: Oxford University Press.

KIDD, A. L., FRISBY, J. P., & MAYHEW, J. E. W. (1979). Texture contours can facilitate stereopsis by initiating appropriate vergence eye movements. *Nature*, **280**, 829-832.

LEE, D. N. (1970). Binocular stereopsis without spatial disparity. *Perception & Psychophysics*, **9**, 216-218.

MARR, D., & POGGIO, T. (1979). A theory of human stereopsis. *Proceedings of the Royal Society*, (London), Series B, **204**, 301-328.

MAYHEW, J. E. W., & FRISBY, J. P. (1976). Rivalrous texture stereograms. *Nature*, **264**, 53-56.

MAYHEW, J. E. W., & FRISBY, J. P. (1981). Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence*, **16**, 349-384.

RAMACHANDRAN, V. S., MADHUSUDHAN RAO, V., & VIDYASAGAR, T. R. (1973). The role of contours in stereopsis. *Nature*, **242**, 412-414.

NOTE

1. Five experienced observers with normal stereo vision viewed the displays. The stimuli were presented on a high-resolution color CRT controlled by a Symbolics 3600 Lisp machine and viewed from a distance of approximately 60 cm. From this distance, a half image (100 $\times$ 100 pixels) subtended about 5.3 arc deg. The width of both bars was 15 pixels (47 arc min).