



Integrating Contours Within and Through Depth

ROBERT F. HESS,*‡ ANTHONY HAYES,† FREDERICK A. A. KINGDOM*

Received 25 April 1996; in revised form 25 July 1996

To better understand the role of disparity in contour integration we compared detection performance of “paths” composed of elements confined either to a single depth plane, or spanning multiple depth planes. In both cases paths defined by alignment of elements were embedded in a noise background-field made up of similar, but randomly positioned, elements covering the same depth range as the path elements. We show that a systematic disparity cue can enhance the detectability of paths which traverse depth, but that this detectability is weak compared to paths made up of elements of the same disparity. These results suggest that the outputs of disparity detectors tuned to different disparities can be linked to define contours. © 1997 Elsevier Science Ltd. All rights reserved.

Stereopsis Contour integration Surface Depth Linking

INTRODUCTION

The visual world is composed of contours at all orientations to our line of sight. While many contours are confined to the fronto-parallel plane, an understanding of how the visual system integrates contours must include contours which traverse depth as well as space.

A recent attempt to understand the rules governing contour integration within a single depth plane (i.e., the fixation plane) was made by Field *et al.* (1993). They showed that the extent to which individual elements are linked to form curved contours depends both on their local alignment and proximity. They defined an “association field” by the orientation alignments which most strongly link oriented elements together, as a function of distance from a particular element. Using a similar paradigm, Kovacs & Julesz (1993) demonstrated that closure is a special case in contour integration, and McIlhagga & Mullen (1996) have shown that rules similar to those producing the “association field” govern the integration of chromatically defined elements.

A previous investigation has shown an improvement with stereoscopic vs monocular presentation in the detection of dotted lines that contain disparity information and traverse depth (Uttal, 1983). To our knowledge, no-one has shown whether contour detection through depth is possible when disparity is the only available cue,

i.e., when the targets are not confined to one depth plane and are undetectable monocularly. Hess & Field (1995) used such cyclopean stimuli to demonstrate that contour integration can be accomplished across two widely separate depth planes, suggesting that the underlying operators receive binocular input. The finding that the separation between two depth planes is not critical, at least out to the fusion limit, leaves open the possibility that while the underlying mechanisms are binocular they may not be disparity tuned. There is, though, ample evidence that disparity is a primary visual attribute (Julesz, 1971) and alone can define complex forms which we can accurately discriminate (e.g. Regan & Hamstra, 1994). On the other hand, there is also evidence that disparity information alone is not sufficient for some types of texture discrimination (Nothdurft, 1985). Furthermore, the role of disparity in contour linking of the Field *et al.* (1993) type has been questioned (Hess & Field, 1995). In the present investigation we approach this issue from a slightly different direction, by asking whether we can integrate contours defined by disparity alone. If the answer is “yes”, the “association field” proposal of Field *et al.* (1993) may be extended to include contours which vary in 3-D as well as 2-D.

METHODS

In all experiments the observers’ task was to discriminate a “path stimulus” (Fig. 1, top stereo-pair) from a “no-path stimulus” (Fig. 1, bottom stereo-pair). A path stimulus consisted of a set of non-oriented elements aligned along a common contour, embedded in a background of similar, but randomly placed non-oriented elements. A no-path stimulus consisted of only randomly placed non-oriented elements. By using such stimuli we

*McGill Vision Research Unit, Department of Ophthalmology, McGill University, Montreal, Quebec, Canada.

†Department of Psychology, University of Melbourne, Parkville, Victoria, Australia.

‡To whom all correspondence should be addressed [Fax +1-517-843-1691; Email Rhess@bradman.vision.mcgill.ca].

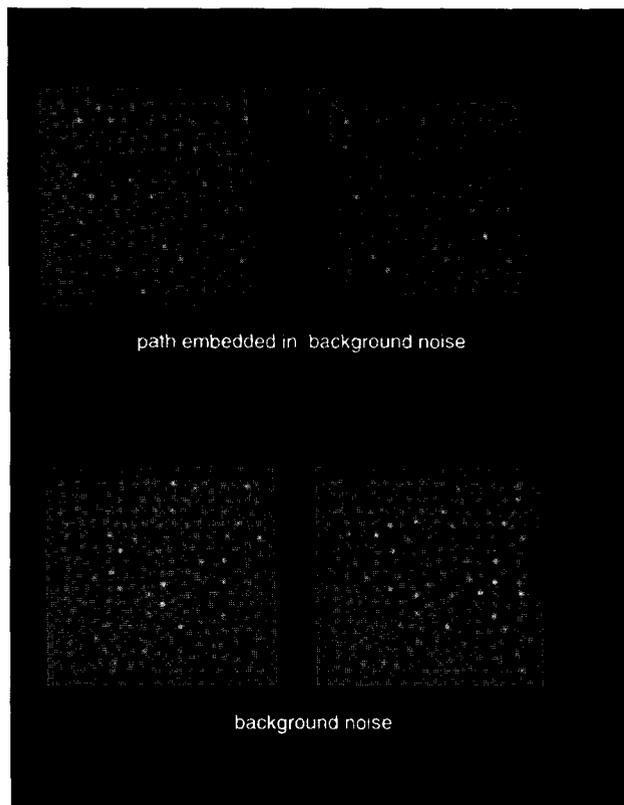


FIGURE 1. A path embedded in a cube of randomly positioned background elements is displayed in the top stereo-pair, background elements alone are displayed in the bottom stereo-pair. The path in the top figure consists of a near colinear and near regularly spaced string of dots traversing the pattern vertically in the middle. On each trial subjects were presented with, in random order, either the path embedded in background elements or the background elements alone.

hope to infer rules which govern the outputs of visual neurones used in the extraction of the path from the background elements.

Stimuli

Non-oriented spatial-frequency bandpass elements were used in this study. The non-oriented elements were defined by the equation:

$$g(x, y) = c \cdot \cos\left(2\pi \frac{R}{P}\right) \exp\left(-\left(\frac{R^2}{2\sigma^2}\right)\right) \quad (1)$$

where c is the contrast and was equal to 35%, $R = (\sqrt{x^2 + y^2})$, P is the spatial period of the sinusoid and was equal to 3.7 c/deg and σ is the space constant and was equal to 4 min.

A no-path stimulus was constructed in the following way. A 4.16 deg wide square was divided into a 16×16 grid of equally sized cells. An element was placed in each display cell, with the restriction that each cell contain the centre of only one element. This eliminated the clumping of elements due to random placement. The elements were also placed to avoid overlap as much as possible. Once an empty cell was chosen, its neighbours were examined to see if they contained an element. If this was not the case, a background element was laid down at some random point within the cell. If one or more of the neighbours

contained an element, then the new element was laid down at a position within the cell to avoid overlap. If this was impossible it was laid down in a position within the cell to minimize overlap.

A path stimulus consisted of two parts; the path itself and the background (Fig. 1, top stereo-pair). Details of the construction of the path are given elsewhere (Field *et al.*, 1993; Hess & Field, 1995). The path was constructed from a backbone of 12 invisible line segments; each line segment was of length 0.31 deg and the line segments joined at an angle uniformly distributed from $\alpha - 10$ to $\alpha + 10$ deg. α is called the path angle, and was throughout the experiments set to 0 deg (i.e., 0 ± 10 deg). Elements were placed at the middle of each line segment (\pm a step jitter of 0.1 deg). To avoid path detection due to random path closure, which can apparently have significant effects on path detection (Kovacs & Julesz, 1993), the path was checked to ensure that it neither intersected itself, nor looped back on itself. The entire path was pasted into the display at a random location, ensuring the centres of the elements occupied different cells. Empty cells were filled with elements, as described for the no-path stimulus above. The average length of each backbone line segment (0.31 deg) was the same as the average distance between neighbouring elements in the background. Previous studies (Field *et al.*, 1993; McIlhagga & Mullen, 1996) have shown that path detection varies inversely with the length of the backbone line segments, but in a smooth manner, so the choice of segment length is not critical.

Neither the local nor the global element density served as a cue to discrimination of path (Fig. 1, top) from no-path stimuli (Fig. 1, bottom). The average distance from an element to its neighbour was no different for path and no-path stimuli. If element density is not a cue, then path detectability should be solely due to the alignment of elements on a path, since nothing else distinguishes path from no-path stimuli. The path elements only differed from the background elements in that they were aligned along an invisible backbone and, therefore, had an orientation correlation. The use of non-oriented elements should therefore ensure, assuming there are no local or global density cues (see Methods), that curved 2-D paths so constructed are undetectable. McIlhagga & Mullen (1996) and Hess & Field (1995) confirmed this finding in control experiments, where either the orientation of the path elements was randomized or non-oriented elements were used for both path and background; they found that neither straight nor curved paths could be detected.

Stereo-images

Stereo-image pairs, each of 4.16×4.16 deg angular subtense, were generated and displayed on the monitor. These images were binocularly combined with a mirror (four sets of two mirrors set at 90 deg to one another) haploscope. In principle, any element, whether it be a path element or a background element, could be displayed with an arbitrary disparity within the limits imposed by the size of the images and the overall display

TABLE 1. Stimulus parameters (Appendix)

<i>Element parameters</i>	
Gabor wavelength	10.4 min (carrier spatial wavelength)
Envelope space constant	2.6 min (Gaussian standard deviation)
Carrier phase	0 (sine phase)
Element contrast	35%
Element size	12.3 min
<i>Display parameters</i>	
Cell size	15.6 min
No. of cell rows	16
No. of cell columns	16
Stereo-image pair displacement	325 min (from left top of left image to left top of right image)
Maximum disparity	0–7.8 min
	0–15.6 min
<i>Path parameters</i>	
Path length	12 (number of elements in path)
Path step size	18.8 min (length of each path step)
Step-size jitter	6.5 min (variability of path step)
Path-angle jitter	10 deg (variability of 2-D angle of path backbone)
Depth sequence	+3.9 min and +7.8 min

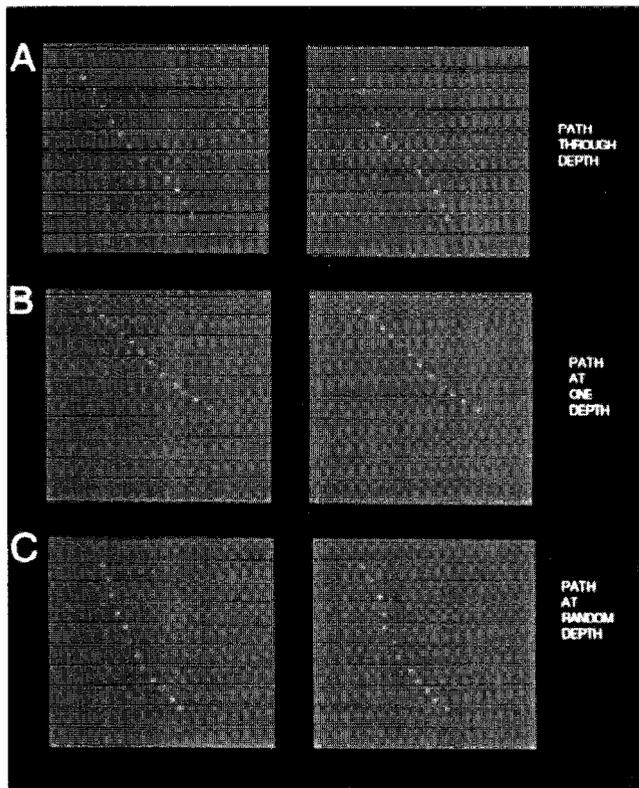


FIGURE 2. Paths devoid of random background elements. The paths either traverse depth (A), occupy one depth plane (B), or are randomly distributed in depth (C).

size. Disparities were produced by adding equal and opposite horizontal shifts to the elements in each stereo-pair. In practice, depth planes were limited to 12, six on either side of the fixation plane. Each stereo image-pair contained an abrupt luminance border, which facilitated fusion within the plane of the screen. The perception of elements in depth was strong and once initiated maintained itself for each of the brief presentations

within each block of trials. Each path element had its own disparity determined by the experimental condition. Background elements were randomly assigned disparity values between the extreme values represented by path elements. Table 1 in the Appendix details the stimulus parameters used. Figure 2 shows stereo-pairs for the different path stimuli used. These are shown, devoid of background elements, for demonstration purposes.

Apparatus and experimental procedures

All stimuli were displayed on a Sony Trinitron monitor driven by a Sun Sparc station 2 computer, which generated stimuli on-line and controlled the display and data collection. The mean luminance of the display was 35 cd/m². The monitor was driven by an 8 bit D/A converter and an 8 bit frame buffer, and was software gamma corrected. The gamma-corrected monitor behaved linearly when displaying high spatial frequencies (12 c/deg square wave) up to 50% contrast. The monitor was viewed in an otherwise dark room through a four-mirror haploscope, which had an equivalent optical path of 130 cm. Each experimental run consisted of a block of 50 trials in which 25 different path and 25 different no-path stimuli were presented in random order. Each presentation was of 1 sec duration, cued to a "beep", and the subject's task was to decide if a path or no-path stimulus had been presented. Each block was repeated six times to obtain 300 trials per condition. *d'* values were derived from the results.

RESULTS AND DISCUSSION

In principle, perfectly aligned paths composed of perfectly regularly spaced elements can be detected owing to their consistent alignment and regularity. To ensure that this was not the case for our stimuli we added a ± 10 deg jitter to the path angle and a ± 0.1 deg jitter to the path step and reduced the detectability of our paths to

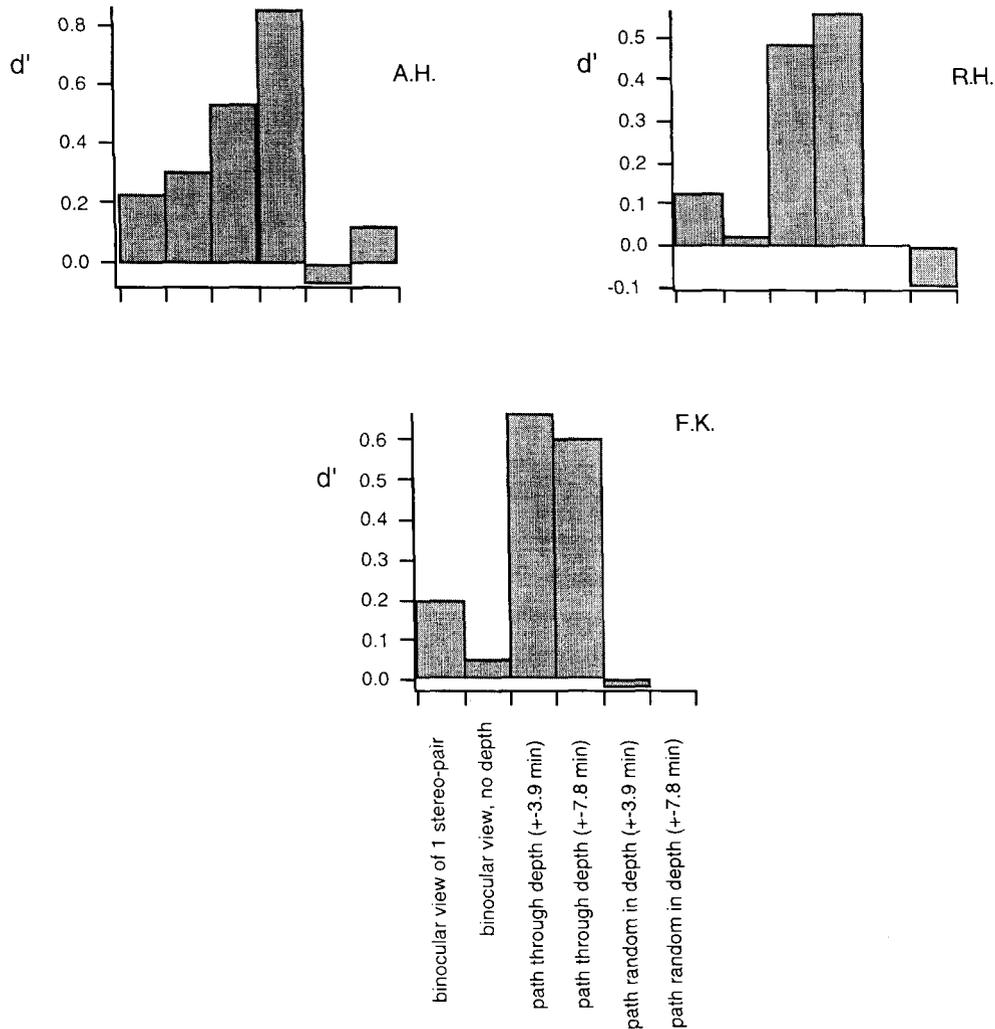


FIGURE 3. Detectability of monocular, binocular and stereoscopic contours.

chance performance. This manipulation was crucial because we wanted to determine whether the visual system can take advantage of linking across depth and to achieve this end such paths needed to be invisible monocularly. In other words we wanted to use cyclopean stimuli. To first verify that our stimulus paths were unable to be detected monocularly, or binocularly when no depth variation was present, we measured path detectability (i.e., d') for (a) one of the stereo-pairs viewed binocularly; and (b) a stereo-pair in which no depth cue was present viewed binocularly. The results shown in Fig. 3 for three subjects (two leftmost histogram bars) established that the stimuli were almost undetectable (d' close to zero).

The two middle bars of the histograms in Fig. 3 show results for the condition where the path (angle 0 ± 10 deg) traversed depth from +3.9 to -3.9 min (left bar) or +7.8 to -7.8 min (right bar) within a field of background elements whose depth was randomly varied between these same limits. In each of these cases, the paths are considerably more detectable. The two rightmost bars of the histograms show results for the condition where we repeat the experiment but now randomize the

depth values within the path over the same limits as those governing the background elements (± 3.9 pixels in the left bar and ± 7.8 pixels in the right bar). This manipulation reduces performance to chance. The results of Fig. 3, when taken together, suggest that the visual system can indeed take advantage of correlations across disparity. We next ask whether the visual system is more sensitive to continuity within a disparity plane or over a number of disparity planes.

In Fig. 4, we compare results for a path that traverses depth (leftmost bar in Fig. 4) with one that is confined to a single fixed depth plane (right-of-middle bar in Fig. 4). In both cases the background elements span a range of ± 3.9 min about the fixation plane. The path at one depth plane is much more detectable than the same path that traverses depth planes. If, however, the single depth plane at which the path is located is not fixed from trial to trial, but randomly varied within the depth range spanned by the background elements, then the path is again less detectable (left-of-middle bar in Fig. 4).

Finally, if the path which traverses depth travels through a depth range that is significantly less than that covered by the background then its detectability is

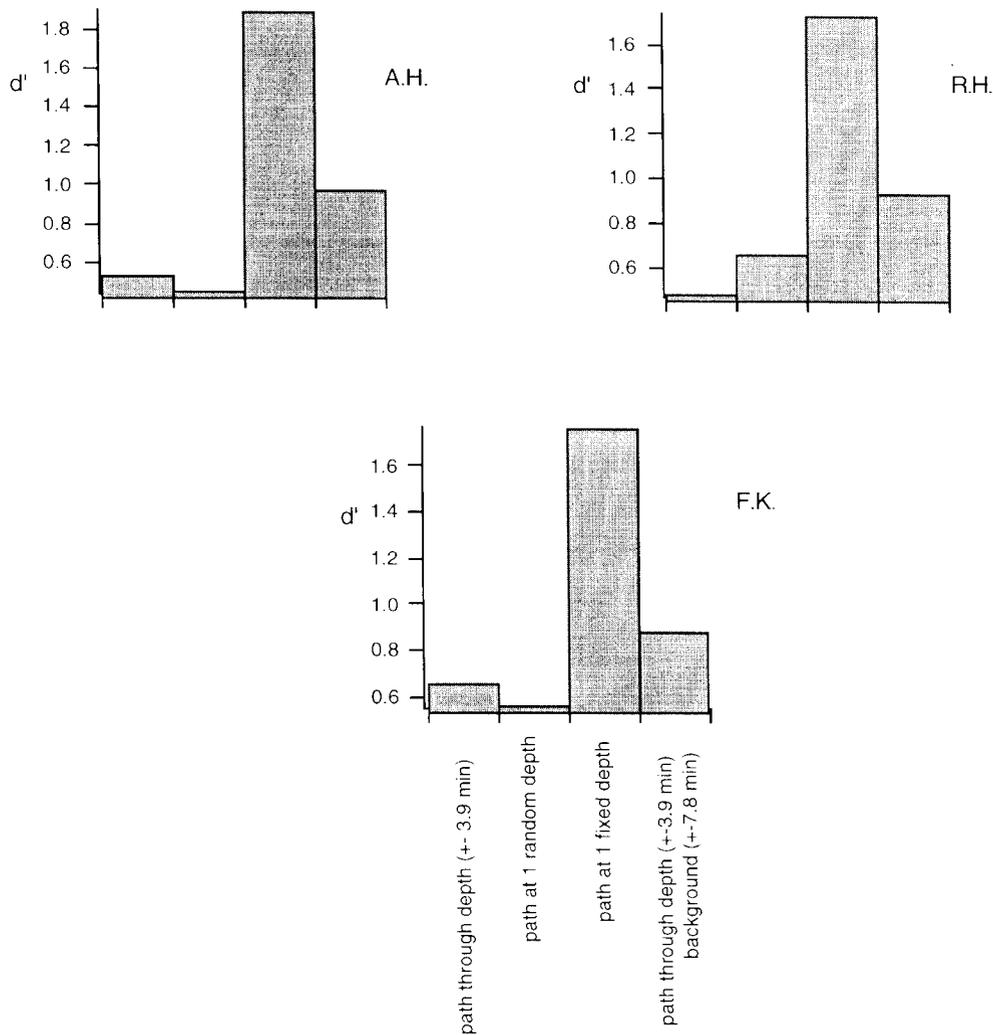


FIGURE 4. Detectability of paths within and through depth planes.

increased compared to when path and background elements cover the same depth range (rightmost bar in Fig. 4). These results suggest that paths confined to one known depth plane are much more detectable than paths that traverse depth if the background elements are distributed over the same depth range. This advantage can be partially eroded by two different manipulations: first, if the depth plane to which the path is confined varies from presentation to presentation; second, if the path traverses a depth range, even one that is smaller than that of the background elements. These findings suggest that the relevant variable is the (path) signal to (background) noise ratio per depth plane. When the path signal elements are confined to a smaller range (either confined to one plane or traversing a smaller depth range) than the background noise elements, then the s/n ratio in any given plane is higher and so is detectability.

Uttal (1983) has also considered the extent to which the detection of contours can be improved by stereopsis. He measured the detectability of straight, regularly spaced dotted lines traversing depth in a cube of 3-D dynamically generated noise dots. The dots in each target-line were presented in temporal sequence at various inter-dot

intervals, and the background noise dots were of various densities. Performance in the stereo conditions was compared to the same stimuli viewed monocularly. Uttal found that at all noise levels and inter-dot intervals, stereo presentation resulted in performance about 10% superior to monocular presentation. Our study primarily differed from Uttal's in two ways. First, Uttal's targets were potentially detectable by an apparent motion cue provided by the temporal sequencing of the target dots. Second, monocular performance in Uttal's task was in most cases above chance level, due primarily to differences between target and noise dot density. Our stimuli on the other hand were deliberately designed to eliminate all other cues to contour detection other than disparity: the targets could not be detected monocularly. Nevertheless, our finding of an approximately 10% improvement as a result of stereo presentation is consistent with Uttal's findings obtained under quite different stimulus conditions. The improvement in contour detection provided by stereopsis can be considered to be a robust result.

Our results support the notion that the visual system links the outputs of disparity-tuned neurones in a similar

way to that previously demonstrated in the 2-D case for orientationally tuned neurones (Field *et al.*, 1993). For our stimulus, stronger associations were made between elements that have similar disparity compared to elements whose disparity is continually changing. This finding almost certainly reflects the very different disparity-based signal/noise ratios in these two conditions. What appear to be different is the relative strengths of linking operations involving orientation in the 2-D case (Field *et al.*, 1993) and disparity in the 3-D cases; linking in the latter case being considerably weaker. This difference is perhaps not surprising because there are very few examples in the natural environment where contours are defined only by depth. In the more usual case, well defined fronto-parallel contours change in depth along their course. What we have demonstrated here is that the visual system can take advantage of continuity through depth as a sole cue to enhance the detectability of contours.

REFERENCES

- Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision Research*, *33*, 173–193.
- Hess, R. F. & Field, D. J. (1995). Contour integration across depth. *Vision Research*, *35*, 1699–1711.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- Kovacs, I. & Julesz, B. (1993). A closed curve is much more than an incomplete one: effect of closure in completion of segmented contours. *Proceedings of the National Academy of Sciences USA*, *90*, 7495–7497.
- McIlhagga, W. & Mullen, K. T. (1996). The detection of colour and luminance contours. *Vision Research*, submitted.
- Nothdurft, H. C. (1985). Texture discrimination does not occur at the cyclopean retina. *Perception*, *14*, 527–537.
- Regan, D. & Hamstra, S. J. (1994). Shape discrimination for rectangles defined by disparity alone, by disparity plus luminance and by disparity plus motion. *Vision Research*, *34*, 2277–2292.
- Uttal, W. R. (1983). *Visual form detection in 3-dimensional space*. Hillsdale, New Jersey: Lawrence Erlbaum.