
Spatial-gradient limit on perception of multiple motion

Yoseph Hermush, Yehezkel Yeshurun

Department of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel

Received 10 February 1994, in revised form 14 February 1995

Abstract. Motion is perceived whenever a subject is presented with an appropriate spatio-temporal visual pattern. Like many other visual tasks, motion perception involves both local and global processing, and thus might be subject to the well-known paradox that arises from the fact that local features and observations form the basis for global perception, but sometimes this global percept can not be easily derived from any single local observation, as is best exemplified by the aperture problem. Globally, dual (transparent) motion can be readily perceived. Spatial limits on the local ability to perceive multiple motion are sought. By using the framework of apparent motion, it is found that dual, orthogonally oriented motion can be perceived only when the dots that constitute the two motions are separated by some spatial limit. For short-range apparent motion, the limit is found to be comparable to D_{max} , and the visual system cannot perceive more than a single coherent motion in a local 'patch' of radius D_{max} . It was also found that this spatial limit on local-motion perception is not constant, but depends linearly on the spatial organisation of the stimuli, and vanishes for stimuli having reverse contrast. The lower bound on the ability to perceive multiple motion is compared with some well-known bounds in stereopsis, and a cortical columnar architecture that might account for it is proposed.

1 Introduction

Apparent motion is perceived when a series of stimuli separated by a spatial and temporal distance is displayed. Motion is perceived although the stimulation is discrete and not continuously moving. It has long been assumed (Braddick 1974, 1980; Anstis 1980) that apparent-motion phenomena are divided into two classes: long-range apparent motion, in which the spatial and temporal intervals are relatively long and the perceived motion is believed to be a result of high-level cognitive processes, and the second class, short-range apparent motion, in which the spatial and temporal intervals are limited. The perceived motion is believed to be a result of the output of the low-level-motion detectors in the visual system.

Although some doubts have recently been raised regarding the existence of this dichotomy (Cavanagh 1991; Petersik 1991), there are numerous experimental findings supporting it, at least for simple cases involving random-dot stimuli (Petersik 1989). To mention a few, it was found that short-range apparent motion is not dichoptic (but see Carney and Shadlen 1993), is followed by motion aftereffects, and does not respond to colour, whereas long-range apparent motion is dichoptic, is not followed by aftereffects, and responds to colour.

There exist quite a few models for visual-motion perception. Most of the models could be classified as either correlation models, where spatial patterns are matched, or various motion-energy/filters models, where the spatiotemporal pattern of the visual field is responded to [see Nakayama (1985) and Hildreth and Koch (1987) for reviews, and Watson and Turanu (1995) for a discussion of correspondence and motion-energy models]. Most of the correlation-based models involve a local estimate of similarity between visual patterns across a short spatial and temporal distance. In principle, it is not necessarily the similarity that is the right metric, but perhaps the covariance (Werkhoven et al 1990) and, for the motion-energy models, there is no direct comparison of any two patterns, but rather an estimation of gradients or spatiotemporal pattern.

However, any motion model must involve some spatial-scale parameter that defines the spatial range over which it operates (multiple scales notwithstanding).

The spatial range of motion detectors is related to one of the most fundamental problems for models of visual perception: local vs global phenomena. The visual system extracts local features that are used to estimate a global percept which, in turn, might redefine the operation of the local-feature extractors. A small gap in a contour can be either noise or a significant part of the image. The global percept based on the output of local receptive fields depends heavily on this local information, but the global percept defines whether this gap is signal or noise. The prominent example of the local-global problem in motion perception is the aperture problem: local data are ambiguous, but a global percept can definitely be computed.

The visual system can readily perceive multiple motions, as is demonstrated by transparencies (eg Koenderink et al 1985). Recently, it has been demonstrated (Qian et al 1994) that if motion is locally balanced, by stimuli moving in opponent directions, then global transparent motion is no longer perceived. However, experimental stimuli of motion transparency are usually global, and it is not clear if local and isolated transparency can be detected, in the sense of perceiving two coherent distinct motions in a small window. This is, again, analogous to the aperture problem: global multiple motion can be easily perceived, but if the image is viewed through a small window, it might not be possible to determine uniquely the two motion vectors.

In this paper we approach this problem by analysing the lower spatial limits on perception of multiple motion. We construct a relatively small stimulus (less than 1 deg), consisting of vertically and horizontally moving dots, and vary the distance between the two apparent motions. The experiments we have conducted are aimed at answering the following question: is there a lower spatial limit on the ability to perceive simultaneously more than a single coherent motion?

2 Method

2.1 Stimuli

Our main experimental paradigm is based on presenting subjects with dual-motion displays. These displays consist of two points of size 3 min arc that apparently 'move', one horizontally and one vertically. The point moving horizontally is located in the middle of the display. The point moving vertically is located at a distance R from the center of the horizontal motion path, either to its left or to its right. The motion magnitude, d , of the two points is equal. The points were coloured black and the background white. In preliminary trials we observed that when the dots were far apart, two concurrent and coherent motions could be perceived. However, when the two dots were too close, the percept of two well-defined motions was replaced by a percept of an 'unstructured' and fuzzy motion (eg four uncorrelated dots moving randomly). We therefore conducted a series of experiments to investigate the ability to perceive a multiple coherent motion as a function of the relation between the dots. In different displays, the horizontal distance R , the motion magnitude d , and the contrast of the points were varied. Figure 1 illustrates a common dual-motion display.

2.2 Procedure

The presentation was done by means of a Silicon Graphics computer and terminal, on a 12.5 deg by 12.5 deg (500 pixels \times 500 pixels) window. Subjects were seated 0.6 m from the display screen. In all experiments, the subjects were shown a series of dual-motion panels, each series with different parameters. For each display panel, subjects were asked to press a key according to whether they did or did not see two concurrent and coherent motions (a two-alternative forced-choice design). Each display was presented for 1 s, during which subjects made the decision. If no response was

made during that period, the display was erased until the subject made a choice. At that time, a new display was presented. The subject started a series by pressing a key. For each series, the probability for seeing dual motion was then computed as a function of the horizontal distance and the motion magnitude. The probability was set as the ratio between the number of times the subject identified dual motion and the total number of times he saw the stimulus with the specific parameters. The total set of series was divided into subsets of at most three series. After each subset, subjects were given a resting period. The sets were partitioned into two groups, each done on a different day.

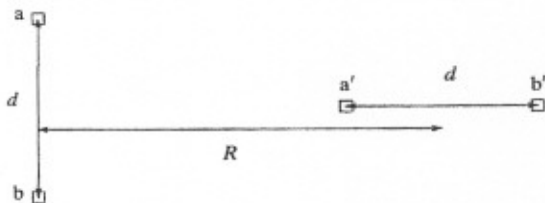


Figure 1. A dual-motion display. Points *a* and *a'* were displayed for 80 ms, were erased (ISI 10 ms), and were replaced by points *b* and *b'*; and vice versa. The vertical movement could also be symmetrically located to the right of the horizontally moving point.

2.3 Subjects

Four male subjects, two of the authors (YH, YY) and two other subjects (YW, EE), all with normal or corrected-to-normal vision, participated in the experiments. Subject YW had some prior experience in psychophysical experiments but, as subject EE, was naive regarding the purpose of this study. For each dual-motion panel, subjects were asked to indicate whether they did or did not perceive a concurrent and coherent motion of the two dots and not a motion per se. Prior to the experiment, each subject was given a demonstration of two dual-motion displays, both with a motion magnitude $d = 12$ min arc. In one of the displays the two moving dots were distant (separated by horizontal distance bigger than 1 deg) and in the other the two dots were overlapping ($R < d/2$ or, in this case, $R < 6$ min arc). The first was intended to demonstrate multiple coherent motion and the other an ambiguous, fuzzy, and unstructured display.

3 Experiment 1

In the first experiment we intended to question the existence of a spatial limit on the ability to perceive multiple motion. Dual-motion displays with black dots moving against a white background, motion magnitude of 12 min arc and varying horizontal distances were used. The horizontal distance was varied by steps of 3 min arc in the range of 0 to 48 min arc. Each horizontal distance was repeated four times in random order, to make a total of $4 \times 17 (= 68)$ display panels.

3.1 Results

The probability of perceiving dual motion is plotted as a function of the horizontal distance in figure 2. It is clear that there is a threshold distance, only above which one can detect dual motion. The midpoint between the last distance for which the probability is 0 and the first for which the probability is 1 was found. The threshold was set as the midvalue minus $d/2$ (ie relative to the edge of the horizontal movement and not to its centre). The threshold values obtained were 15, 10.5, and 7.5 min arc for YH, EE, and YW, respectively. It should be emphasized again that the threshold

distance found is a limit on the perception of a concurrent and coherent motion of two dots (horizontal and vertical motion, in this experiment) and not of motion per se. Thus, when the dots are closer than the limit distance, the percept of two well-defined motions is replaced by a percept of an unstructured and fuzzy motion.

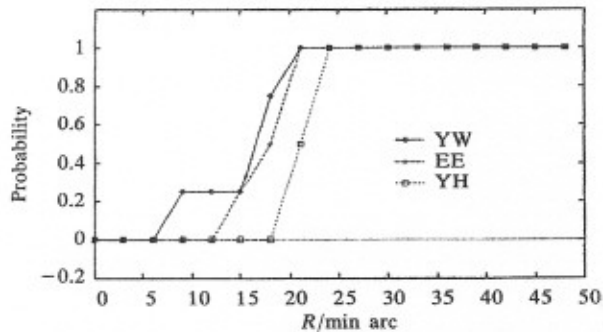


Figure 2. Results of experiment 1: the probability of perceiving dual motion as a function of horizontal distance R . The motion magnitude, d , was 12 min arc. A threshold distance, only above which the subjects could observe two motions, is clearly visible.

4 Experiment 2

In the second experiment we checked the nature of the spatial limit found previously. This was done by measuring the dependence of the dual-motion threshold on the motion magnitude d .

The procedures in the experiment were the same as in experiment 1. The subjects were presented with a series of dual-motion panels. Each panel had a different motion magnitude and a different horizontal distance. The motion magnitude was varied in the range of 9 to 21 min arc (again in a step of 3 min arc) for short-range experiments, and in the range of 45 to 60 min arc (steps of 5 min arc) for the long-range experiments. Each panel was repeated four times in random order, to make a total of $17 \times 5 \times 4 = 340$ panels. The series of panels was divided into blocks of 85 panels.

4.1 Results

The dual-motion thresholds as a function of the motion magnitude are plotted in figure 3. It can be seen that the dual-motion threshold is linearly dependent upon the motion magnitude for all subjects. It can be seen from the results that the threshold gradient, rather than the threshold itself, is the limiting factor.

The linear dependence found for the short-range apparent motion resembles the known linear dependence of D_{\max} upon stimulus size. We were therefore interested in comparing the threshold values found above with dual-motion displays with D_{\max} values achieved with random-dot-kinematogram displays. An experiment similar to that of Morgan (1992) was conducted (on subject YH). By means of a display with varying dot size, D_{\max} was found to be constant up to a stimulus size of 9 min arc and then increased with a gradient similar to the gradient limit achieved in our experiment 2 (see figure 4). The similarity implies that the spatial extent of the two phenomena is the same.

4.1.1 Eye-fixation effects. Control experiments, similar to experiments 1 and 2, in which the subject was asked to fixate on a cross sign positioned in the centre of the display, 30 min arc above the horizontally moving dot (see figure 5), were conducted on one of the subjects (YH). The results obtained were identical to the ones without the cross sign, thus showing that eye fixation does not affect the results.

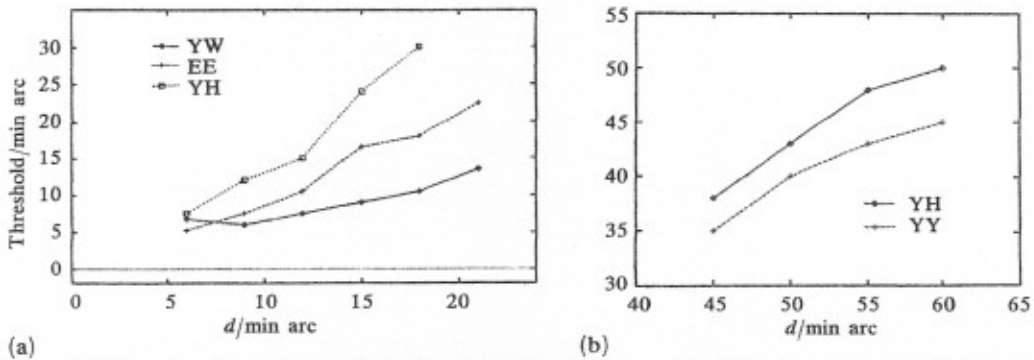


Figure 3. Results of experiment 2: dual-motion thresholds as a function of motion magnitude, d , for (a) short-range and (b) long-range motion. The dual-motion threshold depends linearly on the motion magnitude, ie their ratio is constant.

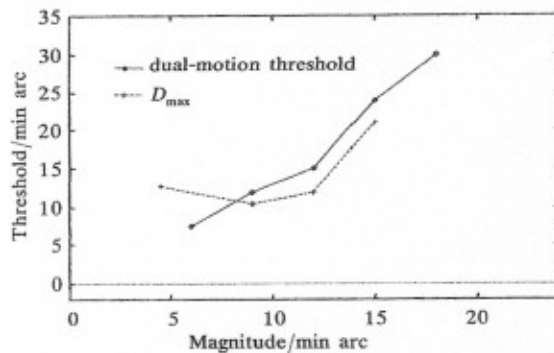


Figure 4. Results of experiment 2 (thresholds as a function of motion magnitude) compared with D_{max} values (as a function of stimulus magnitude) obtained from viewing random-dot-kinematogram displays with varying dot size. D_{max} relates to the dot size in the same manner as the dual-motion threshold relates to the motion magnitude. The experiment was similar to that in Morgan (1992).

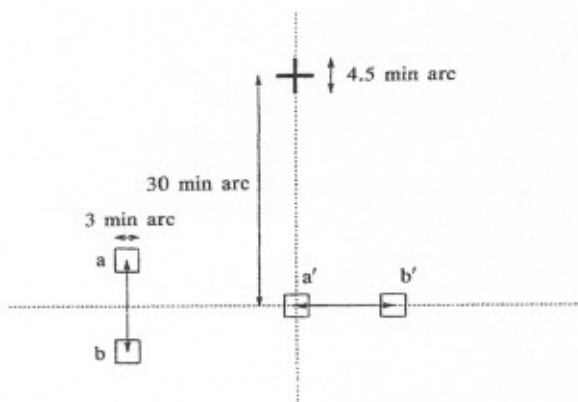


Figure 5. Checking the effect of eye fixation. Experiments 1 and 2 were repeated with a dual-motion display with a fixation cross positioned in the centre of the display, 30 min arc above the horizontally moving dot. The results were identical to those obtained without the cross, thus eliminating a possible effect of eye fixation upon the results.

5 Experiment 3

In this experiment we checked the influence of contrast reversal between the moving points on the value of the dual-motion threshold. In these experiments, points a and b were of the same contrast (white on gray) and points a' and b' were of a reversed contrast (black on gray). All the other parameters were the same as those in experiment 1 (see figure 6).

The same paradigm as in experiment 1 was used. The series of panels contained different horizontal distances. The total number of panels was $4 \times 17 (= 68)$.

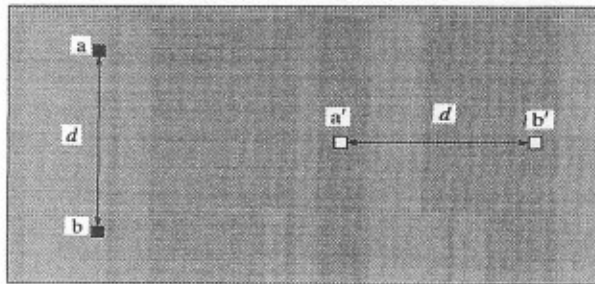


Figure 6. A dual-motion display of dots with reversed contrast. Points a and b are of the same contrast (white on gray) and points a' and b' are of a reversed contrast (black on gray). All the other parameters are the same as those in experiment 1.

5.1 Results

The results are summarised in figure 7 and table 1. The results indicate that contrast reversal significantly decreases the dual-motion threshold (eg from 10.5 min arc to 1.5 min arc for subject EE). These results are expected, since black dots and white dots

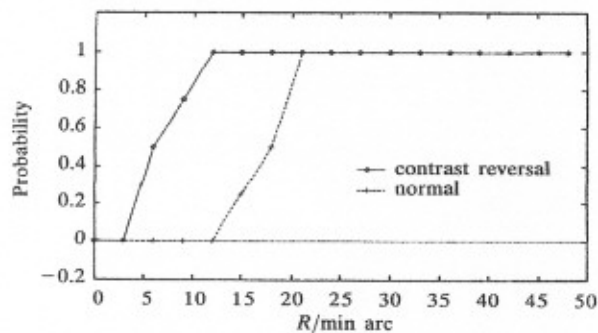


Figure 7. Checking the dependence of the dual-motion threshold upon contrast reversal. The probability of perceiving dual motion is plotted as a function of horizontal distance R for subject EE. The results for the normal curve are from experiment 1 and for contrast reversal from experiment 3.

Table 1. Dual-motion thresholds (in min arc) for normal stimuli (experiment 1) and with contrast reversal (experiment 3).

Subject	Normal	Contrast reversal
YW	7.5	3
YH	15	10.5
EE	10.5	1.5

on a gray background are differentially responded to by 'on-centre' and 'off-centre' cells, possibly processed by different visual pathways (Schiller 1983), and usually not grouped psychophysically (Glas and Switkes 1976). Thus, it is plausible that two motion signals with a weak interaction are generated, resulting in the percept of two coherent motions even though the two signals are nearly spatially overlapping.

6 General discussion

We have shown that there exists a limit on the ability to perceive coherent dual motion in a local 'window' of the visual field. It can be seen from experiment 1 that when two pairs of dots are spatially distinct, their apparent motion can be simultaneously perceived. When the distance between the pairs decreases, there is an abrupt decrease in our ability to perceive it as a concurrent dual motion, and instead a 'fuzzy-motion' percept appears. In experiment 2 we have shown that this distance depends linearly on the magnitude of the motion, and in experiment 3 we demonstrated that for short-range apparent motion the phenomenon does not transfer to reverse-contrast stimuli.

6.1 Relevance to motion-perception models

Our results show that, in the same local area, interaction between two motions caused by the apparent-motion scheme can not be adequately resolved. This finding could be interpreted in the correlation-model framework as well as in the motion-energy model.

For the correlation-based models, motion is estimated by interaction between two spatially distinct cells, and our results mean that within a small area, if more than a single 'correlation-cell pair' is active, a confused percept arises. For the motion-energy/filters models, motion is estimated by the response of cells to a spatiotemporal pattern. If the dual motion is perceived by two cells that are separately tuned to different spatiotemporal regions, then the situation is similar to that of the correlation-based models. On the other hand, if a single cell is tuned to respond to both motion vectors, it means that the visual system does not resolve the ambiguous response of this cell. Thus, our findings do not directly support, nor depend on, the specific motion model.

6.2 Geometric interpretation of the results

A simple model that could account for the results of our experiments can suggest that the dual-motion threshold is imposed by the Gestalt principle of proximity: apparent motion will be perceived between the most adjacent dots. According to this model, when point a' is closer to point a than to point b' , coherent dual motion will cease to appear (see figure 8). By defining $x = R - d/2$, the above can be translated into a

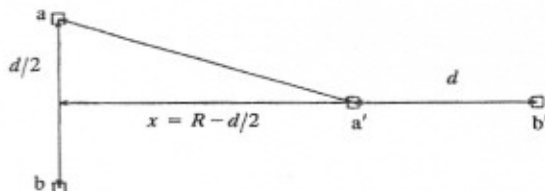


Figure 8. A possible model accounting for our results can suggest that the dual motion threshold is defined by the relative distance between the points consisting the stimuli. When point a' is closer to point a than to point b' , dual motion will cease to appear. Therefore, the condition determining the expected threshold is: $(d/2)^2 + x^2 < d^2$ or $x < \sqrt{3}d/2$.

condition determining the expected threshold. Dual motion will be seen until

$$\left(\frac{d}{2}\right)^2 + x^2 < d^2, \quad (1)$$

$$x < \frac{\sqrt{3}}{2} d. \quad (2)$$

The threshold suggested by this principle is not too far from the threshold obtained for the long-range apparent motion (experiment 2, figure 3b). The results for the short-range apparent motion in experiment 2 show that the subject ceases to observe dual motion for much larger horizontal distances, thus rendering this explanation implausible for short-range stimuli.

6.3 Motion perception and stereopsis

Visual acuity is in the range of 1 min arc. In this paper, we deal with spatial limits that are an order of magnitude coarser than normal acuity. It might be interesting, in this regard, to point out some similarities between the limits discussed here and well-known spatial limits in stereopsis.

Binocular stereopsis is one of the many cues which humans use for inferring the structure of a three-dimensional scene. This cue is based on utilisation of the slightly different views of a scene, as projected onto the right and left retinas. It is known that a stereoscopically presented object that is shown for a short time will appear fused and single if its binocular disparity falls within Panum's fusional area. When the disparity exceeds that limit, objects will appear double. Panum's fusional area is foveally about 7–10 min arc (Ogle 1950). This limit can drastically change under other experimental conditions (see discussion of the disparity gradient below), but for simple stimuli and for short viewing time (to avoid the pooling effect), this limit is quite robust.

Short-range apparent motion is limited by D_{\max} , which is believed to be around 15 min arc at the fovea [the lower limit on D_{\max} is estimated to be 10 min arc (Morgan 1992)]. Both Panum's area and D_{\max} scale with eccentricity (Ogle 1950; Tyler 1974; Baker and Braddick 1985; Cleary and Braddick 1990).

For both Panum's area and the D_{\max} limit, it should be emphasised that while the two limits probably reflect the experimental conditions under which they were measured (Chang and Julesz 1983; Morgan 1992), their common attribute is that they represent a spatial limit on an all-or-none phenomenon.

There are other basic findings pertaining to stereoscopic vision that might be related to our findings—the capacity limit and the disparity-gradient limit. In a series of experiments (Tyler 1975) it was shown that using only pure disparity information the visual system cannot locally perceive multiple depth values within an area that is comparable in size to Panum's area. To quote Tyler (1973, page 277), "Stimuli [vertical lines containing sinusoidal disparity variations] with a finer grain than about 3 cycles per degree did not elicit depth perception, even though the sinusoidal curvature was clearly visible monocularly" (in the body of the article Tyler restricts the limit to 3–5 cycles deg^{-1}). Thus within a region of smaller size than 6–10 min arc, depth differences were not apparent, ie multiple depth values could not be perceived. This can be interpreted as a capacity limit on the stereoscopic system.

In another experiment (Burt and Julesz 1980) it was established that Panum's area is modified by the stimuli used, which were two pairs of dots possessing two disparity values. It was found that the fusion limit varied linearly with the distance between the pair of dots, and was thus not a constant, but rather a gradient limit.

Our finding can thus be described as the motion-related counterpart of the two stereopsis-related phenomena: (i) the capacity limit—within a small area, only a single value (disparity and motion direction, respectively) could be coherently observed; (ii) these limits are not constant, but rather are gradient limits. Our results point in the same direction as those of Qian et al (1994). While their results show that two opponent motion vectors cancel one another, our results show that even orthogonal motion vectors produce similar effects. Our results also show that the spatial limit of this phenomenon applies only to stimuli of the same contrast, and practically vanish for reverse-contrast displays.

7 Conclusion

How relevant are the analogies we have described between motion perception and stereopsis? For the experiments on long-range apparent motion, it is possible that the geometric considerations of proximity, described in similar terms to the scaling-invariance principle of Burt and Sperling (1981), can account for the results. Thus, the analogy to stereopsis might be only phenomenological. The short-range limits, on the other hand, seems to scale differently from the long-range ones, and do not agree too well with the proximity model.

Thus, if the analogy (at least for perception of short-range apparent motion) stems from a deeper common root, the natural candidates for being the common substrate of the spatial limits on stereopsis and motion perception are cortical hypercolumns, which represent foveally about 10 arc min (Yeshurun and Schwartz 1989). According to this interpretation, the limits on disparity estimation and the limits on perception of short-range apparent motion, as well as the capacity limit we have described, could be explained by simply assuming that, within a single hypercolumn, only a single value of disparity or motion vector could be made, when only random-dot-stereogram or apparent-motion information is available. Various forms of this hypothesis, namely relating perceptual phenomena to the size of cortical hypercolumns, have appeared before (Schwartz 1980; Chang and Julesz 1984; Levi et al 1985; Yeshurun and Schwartz 1989; Virsu et al 1994). Validation of this hypothesis requires a clear indication of the relation between the scaling of the relevant phenomena and the cortical-magnification factor. For human subjects, this task might not be trivial at this stage. However, if the spatial limits reported here reflect a fundamental principle of cortical anatomy, it might shed more light on the role of the columnar architecture in the design of the visual cortex.

References

- Anstis S M, 1980 "The perception of apparent movement" *Philosophical Transactions of the Royal Society of London, Series B* **290** 153–168
- Baker C L, Braddick O J, 1985 "Eccentricity-dependent scaling of the limits for short-range apparent motion perception" *Vision Research* **25** 803–812
- Braddick O J, 1974 "A short-range process in apparent motion" *Vision Research* **14** 519–527
- Braddick O J, 1980 "Low-level and high-level processes in apparent motion" *Philosophical Transactions of the Royal Society of London, Series B* **290** 137–151
- Burt P, Julesz B, 1980 "A disparity gradient limit for binocular fusion" *Science* **208** 615–617
- Burt P, Sperling G, 1981 "Time, distance, and feature trade-offs in visual apparent motion" *Psychological Review* **88** 171–195
- Carney T, Shadlen M, 1993 "Dichoptic activation of the early motion system" *Vision Research* **33** 1977–1995
- Cavanagh P, 1991 "Short-range vs long-range motion: Not a valid distinction" *Spatial Vision* **5** 303–309
- Chang J J, Julesz B, 1983 "Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion" *Vision Research* **23** 1379–1385

- Chang J J, Julesz B, 1984 "Cooperative phenomena in apparent movement perception of random-dot cinematograms" *Vision Research* **24** 1781-1788
- Cleary R, Braddick O J, 1990 "Masking of low frequency information in short range apparent motion" *Vision Research* **30** 317-327
- Glas L, Switkes E, 1976 "Pattern recognition in humans: correlations which cannot be perceived" *Perception* **5** 67-72
- Hildreth E C, Koch C, 1987 "The analysis of visual motion: From computational theory to neuronal mechanisms" *Annual Review of Neuroscience* **10** 477-533
- Koenderink J, Doorn A van, Grind W van de, 1985 "Spatial and temporal parameters of motion detection in the peripheral visual field" *Journal of the Optical Society of America A* **2** 252-259
- Levi D M, Klein S A, Aitsebaomo A P, 1985 "Vernier acuity, crowding and cortical magnification" *Vision Research* **25** 963-977
- Morgan M J, 1992 "Spatial filtering precedes motion detection" *Nature (London)* **355** 344-346
- Nakayama K, 1985 "Biological image motion processing: A review" *Vision Research* **25** 625-660
- Ogle K N, 1950 *Researches in Binocular Vision* (Philadelphia, PA: W B Saunders Company)
- Petersik T J, 1989 "The two-process distinction in apparent motion" *Psychological Bulletin* **106** 107-127
- Petersik T J, 1991 "Comments on Cavanagh and Mather (1989): Coming up short (and long)" *Spatial Vision* **5** 291-301
- Qian N, Anderson R, Adelson E, 1994 "Transparent motion perception as detection of unbalanced motion signals" *Journal of Neuroscience* **14** 7357-7366
- Schiller P H, 1983 "Separate on/off pathways in retino-geniculate projection" *Nature (London)* **297** 580-582
- Schwartz E L, 1980 "Computational anatomy and functional architecture of striate cortex. A spatial mapping approach to perceptual coding" *Vision Research* **20** 645-669
- Tyler C W, 1973 "Stereoscopic vision: Cortical limitations and a disparity scaling effect" *Science* **181** 276-278
- Tyler C W, 1974 "Depth perception in disparity gratings" *Nature (London)* **251** 140-142
- Tyler C W, 1975 "Spatial organization of binocular disparity sensitivity" *Vision Research* **15** 583-590
- Virsu V, Nasanen R, Osmoviita K, 1987 "Cortical magnification and peripheral vision" *Journal of the Optical Society of America A* **4** 1568-1578
- Virsu V, Rovamo J, Larinen P, 1994 "Illusory perception of gratings stimulating a small number of neurons" *Vision Research* **34** 3253-3263
- Watson A, Turanu K, 1995 "The optimal motion stimulus" *Vision Research* **35** 325-336
- Werkhoven P, Snippe H, Koenderink J, 1990 "Metrics for the strength of low level motion perception" *Journal of Visual Communication and Image Representation* **1** 176-188
- Yeshurun Y, Schwartz E L, 1989 "Cepstral filtering on a columnar image architecture: A fast algorithm for binocular stereo segmentation" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 759-767