UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.853 - 76

ANTONIO BORSELLINO COLLEGE ON NEUROPHYSICS

(15 May - 9 June 1995)

"Evidence for second-stage visual mechanisms in humans, selective for radial, circular and translational motion"

David C. Burr
Istituto di Neurofisiologia
Consiglio Nazionale delle Ricerche
56100 Pisa
Italy

These are preliminary lecture notes, intended only for distribution to participants.

Submitted to *Nature* 3 March, 1995: Revised 12 May, 1995

Evidence for second-stage visual mechanisms in humans, selective for radial, circular and translational motion.

M. C. Morrone[1,2], David C. Burr[1,3,5] & Lucia Vaina[4]

[1]Istituto di Neurofisiologia del CNR, Via S. Zeno 51, Pisa, Italy.

[2]Scuola Normale Superiore, Pisa, Italy.

[3]Dipartimento di Psicologia, Università di Roma, Roma, Italy.

[4]Department of Biomedical Engineering, College of Engineering, Boston University, USA.

[5]Corresponding author: Fax +39 50 559725 Tel 559719 E-mail dave@neuro.in.pi.cnr.it

---

As we move through our environment, the flow of the deforming images on our retinae provides rich information about our own motion and about the motion and the three-dimensional structure of the external world. Flow-fields can be decomposed into five independent components including radial, circular and translational motion[1-3]. Here we provide psychophysical evidence for the existence of neural mechanisms in human vision that summate motion signals along these complex trajectories. Radial, circular and translational motions were created on a dynamic random dot visual display, spatially curtailed into symmetrically opposed sectors. Signal-to-noise sensitivity for direction discrimination of all three types of motion increased predictably with the number of exposed sectors, implying the existence of specialised detectors that summate motion signals of different directions across sectors. We present further evidence that complex motion is analysed by two stages of neural mechanisms, an array of first-stage detectors of local translational motion, whose output is then integrated by the specialised second-stage mechanisms mentioned above. Contrast sensitivity for complex motion did not increase greatly with stimulus area, implying an early and local limitation by first-stage mechanisms that impose a contrast threshold for motion discrimination. These findings fit well with recent electrophysiological evidence: in primary visual cortex in monkey, motion-sensitive neurones respond best to local translation, whereas many neurones in the medial superior temporal area of associative cortex have large receptive fields tuned to radial, circular or spiral motion[4-7].

Observers viewed four-frame sequences of random dot displays, in which a proportion of dots were appropriately displaced between frames, while the remainder were replaced at random. The dots fell within a 10° circle, notionally divided into 16 equal sectors. Signal dots could be confined to 1, 2, 4, 8 or all of the sectors, symmetrically arranged, while the remaining sectors were either left blank or filled with incoherently moving "noise" dots of the same density (see figure 1). Sensitivity for motion discrimination was defined as the inverse of the minimum proportion of total dots in the signal sectors at which motion direction could be reliably discriminated (1+N/S). Figure 2 shows how motion sensitivity varied with stimulus area for radial, circular and translational motion. For all types of motion, sensitivity increased with stimulus area, implying that discrimination took advantage of motion signals from all sectors, summed within specialised neural mechanisms. For radial and circular motion, summation could not be based on detectors tuned solely to local direction, particularly when there were few sectors in the display: with only two sectors, the motion in each sector was in opposite directions, and with four, it was orthogonal. Note also that sensitivity was lower when the non-signal sectors were filled with noise than when they were empty. The additional noise from remote sectors reduces sensitivity, probably because it is integrated within a common detector unit. Interestingly, in these conditions the coherent motion did not seem to be confined to the signal sectors, but the whole display appeared to expand, rotate or slide. The broken curves of figure 2 represent stimuli of constant signal-to-noise, assuming linear integration[8,9]. The prediction of an ideal integrator model is for a linear increase in sensitivity when the non-signal sectors were filled with noise dots, and a square-root increase when they were empty. In both conditions, the model provides an adequate fit to the data, indicating that the motion is discriminated by specialised operators of constant efficiency[9].

We next measured summation of *contrast sensitivity* for these stimuli (only for radial motion). Here the dots in the signal sectors moved coherently, and the contrast of the whole display was adjusted to determine the threshold for discriminating the direction of motion, and for detecting the patterns (figure 3). When the non-signal sectors were filled with noise, there was strong summation, but little or no summation when they were empty. The dashed and dot-dashed lines show the predictions of the integrator model for two experimental conditions (see caption for details). Contrast sensitivity data follow these predictions for only the smallest stimulus areas, particularly for the no-noise condition: for much of the range, sensitivity was far less than that predicted by the integrator, implying that other factors may limit performance. One possibility (consistent with the physiology) is that discrimination of complex motion involves two stages, local-motion detectors followed by integrator mechanisms tuned to complex motion. The dotted line through the open symbols of fig 3 shows how sensitivity should vary with area if it were limited by the contrast threshold of independent detectors responding locally to the direction of motion. The slight dependence on area results from the increased probability that a larger area will excite at least one local detector[10,11]. The stimuli at all points along this line will elicit the same excitation in the integrator unit, so integration can not enhance the limit imposed by the lower stages. Sensitivity for discrimination of complex motion should follow the predictions of the integrator model up to the point where it

meets the limit set by the first-stage motion detectors. The data follow well these predictions, indicating that complex motion analysis involves two stages, local-motion detectors that limit stimulus visibility, and more specialised detectors that analyse expansion, rotation, and translation (and combinations thereof) over extensive regions.

Several lines of psychophysical evidence are consistent with the existence of two stages of motion analysis. Contrast sensitivity for spatially curtailed sinusoidal gratings shows full summation over a relatively limited area (equal to about one cycle of periodicity of the stimulus[12,13]), suggesting that performance for these stimuli is limited by first-stage contrast-dependent detectors, while the area of integration for direction discrimination (of translation) of limited-life random-dot patterns can be much larger, extending up to 9°[14,15] (a value that we have replicated informally in this study for all three types of motion). Morgan[16] has shown that D_{max} (the largest dot-displacement to support coherent motion[17]) varies with dot size and stimulus blur in a way that suggests an early pre-filtering stage followed by a broad band detector, an idea that has received recent support from masking studies[18]. Our results further suggest that the pre-filtering stage is directionally selective. Early adaptation experiments of Regan and Beverley suggested that radial and translational motion may be analysed by different mechanisms[19,20], although the quantitative predictions of the size of the receptive fields for these mechanisms (less than 1.5°) agrees neither with the results reported here nor with the physiology[4-7]: perhaps first-stage mechanisms may have mediated the adaptation to some extent. Despite these differences, however, the adaptation studies also suggested that radial motion may be a second-stage analysis, after a common contrast-dependent first stage[21].

Modern electrophysiological studies have shown that motion is processed at several different cortical levels, including primary visual cortex (V1), middle temporal (MT) and medial superior temporal (MST) cortex. Neurones in V1 prefer translational motion, and have relatively small receptive fields. In MT, receptive fields are larger, but remain selective to translational motion[22]. In the dorsal segment of MST (MSTd), cells have very large receptive fields, and many are selective to radial or circular motion[4-7]. However, selectivity is not restricted to these "cardinal directions" of optic flow, as there also exist neurones selective to combinations of radial and circular motion (spiral motion) and to combinations of spiral and translational motion[5-7]. The experiments reported here do not address the issue of whether selectivity is restricted to the cardinal directions, but do show that their must exist neurones that respond to radial or circular motion trajectories. Other properties of MSTd neurones, such as their insensitivity to density, texture and relative depth of the texture elements[4,5], and particularly to the position of the centre of the radial, circular or spiral motion[6,7] (also supported by psychophysical evidence[23]) make these neurones well suited for optic flow field analysis, and for guiding the heading of ego motion[24,25].

REFERENCES

1. Helmholtz, H. (1858) *Crelles J.* **55** 25.
2. Koenderink,J.J. (1986) *Vision Res.* **26** 161-168.
3. Verri, A., Girosi, F & Torre, V (1990) *J. Opt. Soc. Am.* **A7** 912-922.
4. Tanaka, K. & Saito,H. (1989) *J. Neurophysiol.* **62** 626-641.
5. Duffy, C.J. & Wurtz, R.H. (1991) *J. Neurophysiol.* **65** 1329-1345.
6. Orban G.A., Lagae L., Verri A., Raiguel S., Xiao D., Maes H., Torre V. (1992) *Proc. Natl. Acad. Sci.* **89** 2595-2599.
7. Graziano M.S.A., Andersen R.A., Snowden R.J. (1994) *J. Neuroscience* **14** 54-67.
8. Green, D.A. & Swets, J.A. (1966) *Signal detection theory and psychophysics.* John Wiley & sons, New York.
9. Barlow, H.B. (1978) *Vision Res.* **18** 637-655.
10. Graham, N. (1977) *Vision Res.* **17** 637-652.
11. Pelli, D.G. (1985) *J. Opt. Soc. Am.* **A2** 1508-1532.
12. Anderson, S.J. & Burr, D.C. (1987) *Vision Res.* **27** 621-635.
13. Anderson, S.J & Burr, D.C. (1991) *J. Opt. Soc. Am.* **A 8** 1330-1339.
14. Watamaniuk, S.N.J. & Sekuler, R. (1992) *Vision. Res.* **32** 2341-2347.
15. Fredericksen, R.E., Verstraten, A.J. & van de Grind, W.A. (1994) *Vision Res.* **34** 3171-3188.
16. Morgan, M.J. (1992) *Nature* **355** 344-346.
17. Braddick, O. (1974) A short-range process in apparent motion *Vision Res.* **14** 519-527.
18. Yang, Y. & Blake, R. (1994) *Nature* **371** 793-796.
19. Regan, D. & Beverly, K.I. (1978) *Vision Res.* **18** 415-421.
20. Beverley, K.I. & Regan, D. (1979) *Vision Research* **19** 1093-1104.
21. Petersik, J.T., Beverley, K.I. & Regan, D. (1981) *Vision Research* **21** 829-832.
22. Saito, H.-A., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986) *J. Neurosci.* 7 177-191.
23. Snowden, R.J. & Milne, A.B. (1995) *Vision Res.* under review.
24. Hannon, W.H. & Hannon, D.J. (1988) *Nature* **336** 162-163.
25. Warren, W.H. & Hannon, D.J. (1990) *J. Opt. Soc. Am. A* **7** 160-169.
26. Watson, A.B. & Pelli, D.G. (1983) *Percept. Psychophys.* **33** 113-120.
27. Nelder, J.A. & Mead, R. (1964) *Computer Journal* **7** 308-313.

Acknowledgements: We thank Denis Pelli for helpful comments on the manuscript. LMV was supported by Office of Naval Research contract and by NIH grant.

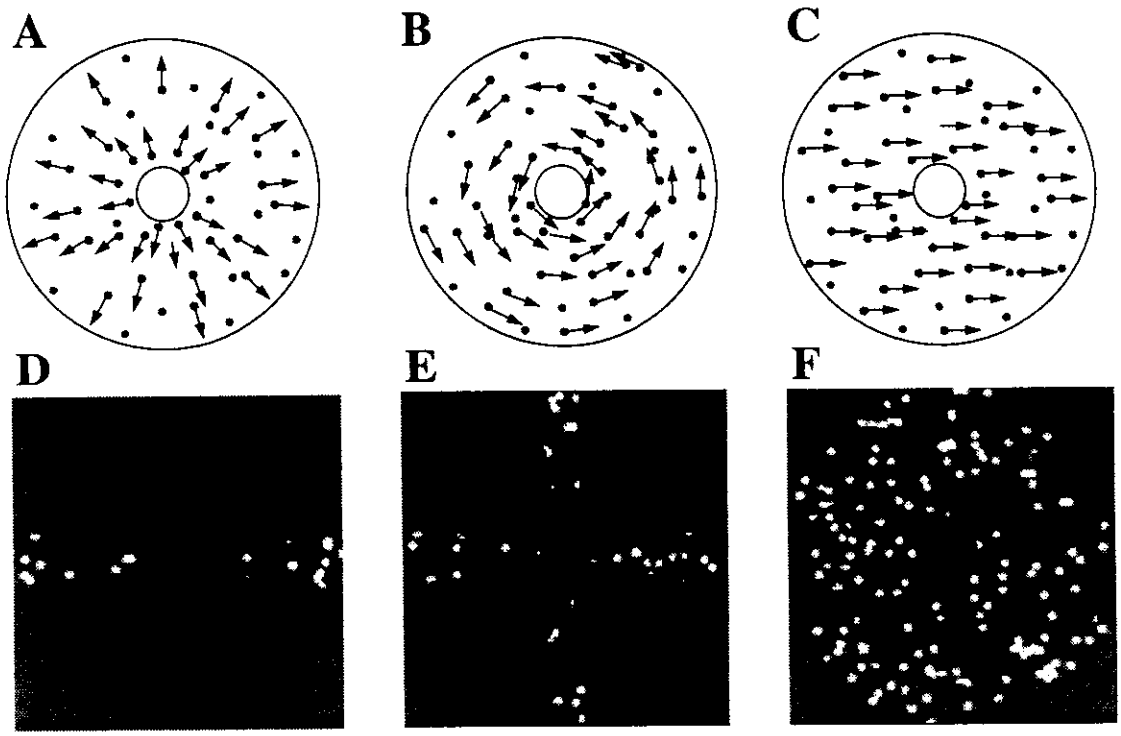

FIGURE 1

Examples of stimuli. Figs A-C are schematic illustrations of the three types of motion, and figs D-F photographs of single frames of the actual stimuli. The complete pattern (fig. F) subtended 10° at the eye, and was filled with 360 small black and white Gaussian patches of space constant 0.5' (excluding the central 1.5°). Four frames were presented, each for 80 ms, within a temporal Gaussian window of σ = 50 ms. On each frame, a proportion of the dots was displaced by 2' to produce radial, rotational or translational motion (illustrated by the arrows in figs A-C), and the remainder of the dots replaced at random (dots without arrows). The pattern was progressively curtailed to 1, 2, 4 or 8 sectors, each subtending 22.5° and maximally separated from the others (e.g. figs D & E). For the "noise" conditions, the empty sectors were filled with incoherent dots of the same density, so that the spatial pattern on each frame resembled fig. F, but coherent motion was confined to specific sectors.

METHODS Three performance measures were obtained: (1) Signal-to-noise sensitivity for motion direction, defined as the inverse of the minimal proportion of coherent dots in the signal sectors (1+N/S) at which the direction of motion could be discriminated; (2) contrast sensitivity (the inverse of Michaelson contrast at threshold) for motion discrimination; and (3) contrast sensitivity for pattern detection. Within a given session, subjects were required to discriminate motion direction (expansion from contraction, clockwise from counter-clockwise rotation, leftward from rightward or upward from downward translation); or to discriminate the pattern from one in which the signal sectors were set to mean-luminance (detection thresholds). From trial to trial, dot proportion or stimulus contrast was determined by the adaptive QUEST[26] procedure. Thresholds were calculated by fitting a cumulative Gaussian curve to the psychometric data, about 200 points for each condition. The stimuli were prepared in advance on a Silicon Graphics workstation. Given the random nature of the stimuli and the problem of "wrap-round" (when the motion took dots outside their sector they were replaced at random), the nominal signal-to-noise ratio was not always precise. We therefore measured it by auto-correlation along the motion trajectory (divided by overall signal power), and reclassified the stimuli accordingly. The stimuli were displayed on a Barco calibrator monitor by VSG framestore under PC control, at a resolution of 128 X 128 pixels, 190 Hz framerate and 25 cd/m^2 mean luminance.

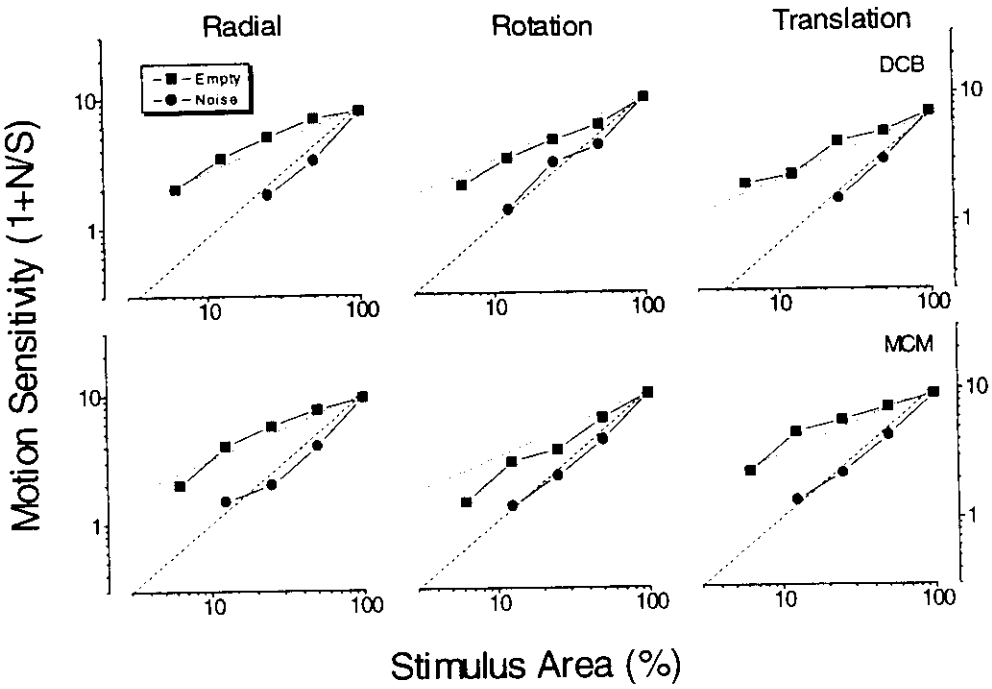

FIGURE 2

Motion sensitivity for correctly discriminating the direction of radial, circular and translational motion (vertical for DCB, horizontal for MCM), as a function of stimulus area. Stimulus contrast was 0.5 (Michelson contrast). The squares refer to the condition where the non-signal sectors were set to average mean-luminance, and the circles to the condition where these sectors were filled with non-coherent dots of the same average density. For all three types of motion, sensitivity increased approximately linearly with stimulus area in the mask condition (dashed lines), and with the square-root of stimulus area (dotted lines) in the no-mask condition. The broken lines represent constant signal-to-noise ratios of an ideal integrator, that summates motion signals over the whole display, assuming the noise to be Poisson-distributed (so its variance is proportional to the number of noise dots). The results follow closely these predictions, implying the existence of specialised mechanisms that sum motion energy between sectors. Similar results were replicated in two other observers, one of whom was naive.

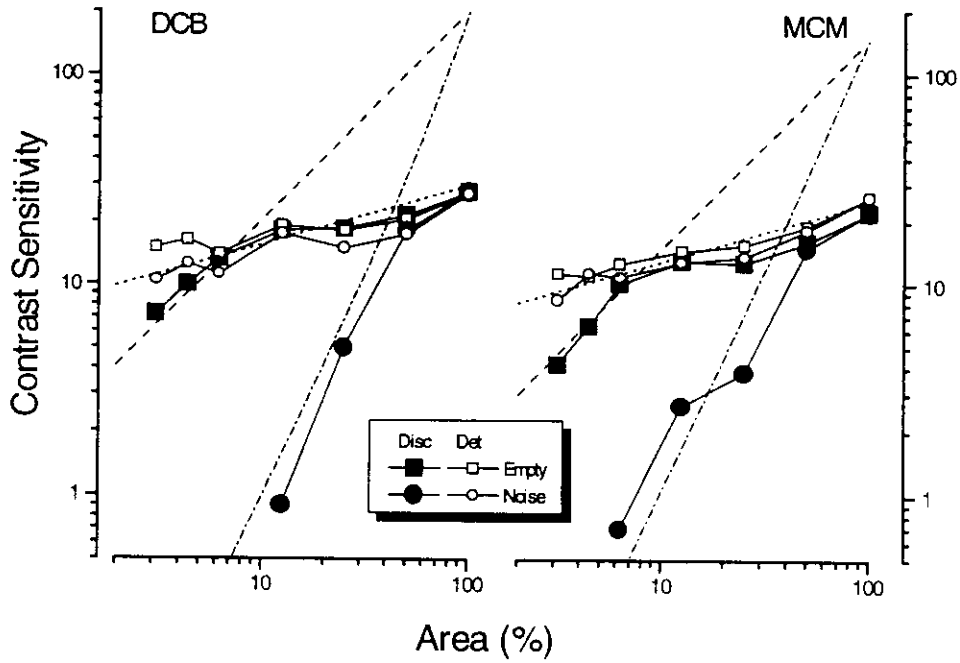

FIGURE 3

Contrast sensitivity (Michelson contrast) for discriminating the direction of radial motion (filled symbols) and for detection of these patterns (open symbols). In this experiment, the three smallest stimuli had only one signal sector, varying in size from 10 to 22.5°. For all other stimulus areas, the sectors subtended 22° and were symmetrically opposed, as before. Squares refer to the condition when the non-signal sectors were set to average mean-luminance, and circles to when they were filled with noise dots of the same density and contrast. In the signal sectors S/(S+N)=63%. The three curves show theoretical predictions for variation in sensitivity. The dotted curve shows the prediction if detection and discrimination were limited by the action of local independent first-stage motion detectors. The slight increase in sensitivity with area results from "probability summation" between the detectors[10], given by:

$$S_c = kA^{1/\beta}$$

Where S_c is contrast sensitivity, k an arbitrary constant adjusted to achieve the best fit of the detection data, A stimulus area and β the slope of the psychometric function, chosen as 3.5, on the basis of the present data and previous studies[11]. The dashed and dot-dashed curves show the prediction of threshold sensitivity for a perfect integrator limited only by its internal noise N_I, given for constant d'_e =1 (by definition at threshold)[8], given by:

$$d'_e = 1 = \frac{S_E}{N_E} = \frac{\alpha c D_s}{\sqrt{N_I^2 + \alpha c D_n}}$$

Where S_E and N_E are the signal and noise of the detector, D_S and D_n the number of signal and noise dots, c contrast and α a constant governing contrast gain (of pre-integrator units). The last equality relies on the assumptions that the integration is quasi-linear at threshold, and that the noise in the stimulus is Poisson-distributed. Hence, the integrated signal will be proportional to the product of contrast and dot number, while the total noise will be given by the squareroot of the sum of the internal and external noise variance. α and N_I were determined to give the best fit of the discrimination data, where the predicted sensitivity was less than that predicted from probability summation: otherwise the upper equation was used for the fit. α was determined by the Simplex minimisation procedure[27] to give the best fit of discrimination sensitivity in the noise condition (where N_I has negligible effect). N_I was then adjusted to fit simultaneously the data of both noise and no-noise conditions. The estimates of α and N_I (respectively) were 0.65, 0.3 for DCB and 0.7, 0.7 for MCM. See text for implications of the data.