Motion Processing: How Low Can You Go?   Dispatch

Richard J.A. van Wezel and Maarten J. van der Smagt

Neurons at early stages in the visual system can only 'view' small parts of the visual world, impeding their ability to determine correctly an object's motion direction. New studies suggest that this 'aperture problem' is already solved by special neurons at the first stage of motion detection in primary visual cortex.

An important task for the visual system is to determine the direction and speed of moving objects. At the first stages of cortical motion processing, in the primary visual cortex (V1), motion-sensitive neurons each respond to only a very small region of the visual field (Figure 1). This means that, for most of these neurons, the object's contours will extend beyond their small 'receptive fields'. This would seem to present a problem, as these neurons only measure the motion component perpendicular to the moving contour, which is not necessarily the same direction — or speed — as the whole moving object. This is what has become well known as 'the aperture problem' [1,2]. Now Pack et al. [3] have reported that a certain class of primary visual neurons — the so-called 'end-stopped' cells — respond in a direction-selective manner specifically to the end-points of a moving contour, irrespective of its orientation. These neurons should in principle be capable of solving the aperture problem.

Neurons in V1 respond to contours of a specific orientation. Most of these neurons — especially the so-called 'complex' cells — are also direction selective. This means that they are excited when the contour is moving in one direction and inhibited by motion in the opposite direction. Most of those complex cells respond well to short contours or endings of contours, but are suppressed by longer contours. This characteristic of these neurons is referred to as 'end-stopping'. V1 contains neurons with a whole range of degrees of end-stopping [4]. Strongly end-stopped cells were identified in the pioneering days of neurophysiological recordings by Hubel and Wiesel [5], who referred to them as 'hyper-complex cells'.

A number of researchers in computational vision and human psychology [6,7] have suggested that neurons with end-stopping behavior are optimally suited to solving the aperture problem, as these cells respond well to contour-endings. As illustrated in Figure 1, for instance, an end-stopped neuron responds to the contour-ending of the wing of a plane in accordance with the perceived motion direction. Non-end-stopped cells, however, respond to the long contour of the wing but signal the wrong direction. In contrast with these

Helmholtz Institute, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands. E-mail: r.j.a.vanwezel@bio.uu.nl

theoretical considerations, neurophysiological evidence has tended to support the idea that the aperture problem is solved at a later stage of motion processing, where information from V1 neurons is combined. Single unit recordings have shown that a group of strong direction-selective neurons in the middle temporal area (MT, also known as V5) that receive input from V1 have solved the aperture problem [8–10]. The paper by Pack et al. [3] seems to refute this generally accepted hierarchical model.

Pack et al. [3] trained macaque monkeys to maintain visual fixation while a stimulus of white and black bars was presented on a grey background at the receptive field location of a recorded cell. The researchers used stimulus sequences that contained bars moving in different directions in combination with different orientations. End-stopped cells responded best when the bar endpoints were in the receptive field, and poorly when the bar was centred on, and extended beyond, the receptive field. The cell's preferred motion direction was independent of the bar's orientation. It is this independence of stimulus orientation that makes these cells suitable candidates for solving the aperture problem, without the need for further integration in area MT.

Another recent study, by Tinsley et al. [11], has shown that a specific class of V1 neurons in the marmoset monkey responds to the pattern direction of a moving plaid. A plaid is composed of two sinusoidal gratings of different orientations added together. Even though the stimulus consists of two independently moving component gratings, it is perceived as a plaid moving in a single direction. For component gratings, the aperture problem holds — only the direction perpendicular to the orientation is perceived. For plaids, the aperture problem is solved — there is only one unique possible plaid direction.
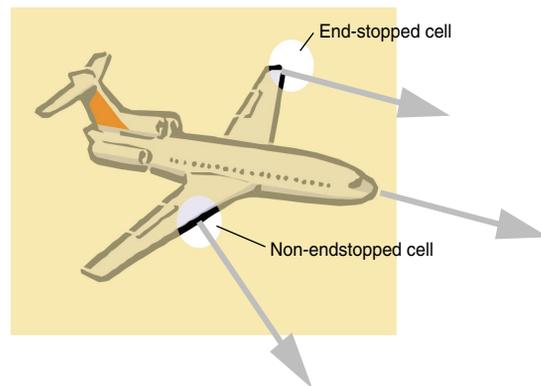
Figure 1. The aperture problem of motion vision.
A non-end-stopped V1 neuron can signal an incorrect motion direction of a contour of a moving object. Through an aperture the contour seems to move perpendicular to the orientation of the contour, and this does not always coincide with the direction of the whole object. An end-stopped neuron signals the correct motion direction for a contour end.
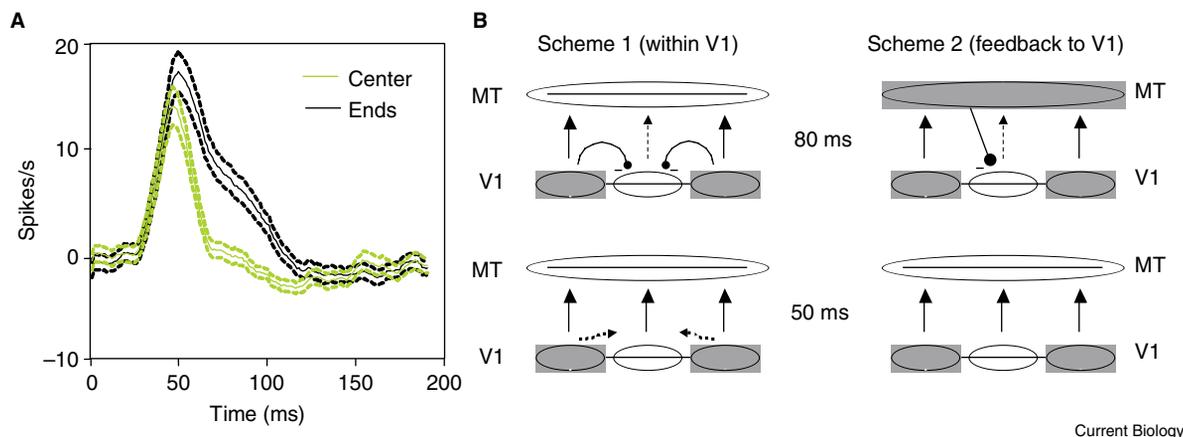
Figure 2. Response dynamics of end-stopped cells.
(A) Time course of the response in V1 when a long bar is presented in the center of the receptive field (green line), or when the end-point of the bar appeared in the receptive field (black line); data averaged over 29 end-stopped V1 neurons, and taken from [3]. The cells initially respond well to a bar placed anywhere in the receptive field, but after 20–30 milliseconds, they respond only when the endpoints are in the receptive field. (B) Schematic representation of how the dynamic response profile in (A) could emerge, either by connections within V1 or by feedback signals from higher-order cortical areas like MT.

Again, traditional views hold that gratings are processed in V1, and that those signals are combined in a subsequent stage of visual motion processing. Evidence for this integration stage has been found in area MT, which contains neurons responding specifically to moving plaids, the so-called 'pattern neurons' [8,9,12]. Tinsley *et al.* [11] show for the first time that neurons in V1, which are broadly tuned for direction and have short, wide receptive fields prefer a direction that coincides with the direction of the moving plaid. Even though Tinsley *et al.* [11] do not report the end-stoppedness of the cells that they recorded from, it is very well possible that the group of neurons is similar to the end-stopped neurons described by Pack *et al*. [3].

The conclusion drawn by both Pack *et al.* [3] and Tinsley *et al.* [11] is that the aperture problem can be solved as early as in V1, and that integration at higher cortical stages is not necessary to solve the aperture problem and explain phenomena such as plaid motion. Yet the hierarchical (two-stage) model of motion processing has long been the consensus view in the field. Indeed, several studies have demonstrated that motion integration takes place in MT, for instance in the pattern neurons described above [8,9,12]. Furthermore, Pack *et al.* [3] show that their results on end-stopped neurons are very similar to what they find in MT. Do these pattern-selective MT neurons perhaps receive their input from end-stopped V1 cells? Currently, there is no direct evidence for this, but it might be obtained from simultaneous recording of cells in V1 and MT. MT pattern cells respond differently to moving plaid patterns that are changed slightly to mimic transparency [13], in parallel with the change in perceptual performance [14,15]. What about those end-stopped V1 cells?

A closer inspection of the temporal dynamics of the cell responses can also give more insight in the flow of information. One very interesting aspect of Pack *et al.*'s [3] study is that they used a receptive-field mapping technique that allows for the disclosure of temporal

dynamics of neural responses in detail. This so-called 'reverse correlation' technique is based on a very rapid presentation of a noisy stimulus sequence (in the 10–20 millisecond range). By reverse correlating the neural response with this stimulus sequence, it is possible to obtain the temporal aspects of receptive field characteristics very efficiently. This technique, developed to study auditory cortical neurons [16] is nowadays increasingly in vogue in vision research to study dynamical processes in the visual cortex [17,18]. Pack *et al.* [3] observed that end-stopping grows over time. At first, the cell responds well to a bar placed anywhere in the receptive field, and only 20–30 milliseconds later the end-stopped characteristic is clearly visible (Figure 2A). This means that end-stopping takes time to evolve, just like many surround effects documented in V1 [19,20]. It is interesting to note that such a time course for end-stopped neurons has also been previously suggested from the results of human psychophysical experiments [6].

In conclusion, the new studies [3,11] put forward the possibility that the motion direction (and speed) is already adequately coded in primary visual cortex, yet important questions remain. The most important of these questions is how end-stopping emerges in V1. Is this accomplished at the level of V1, for instance by horizontal connections, or is it a consequence of feedback from higher cortical areas like area MT (see Figure 2B)? In the latter case the aperture problem would not be 'solved' in primary visual cortex, even though we now know it can at least be observed this low in the cortical hierarchy.

References
1. Stumpf, P. (1911). Uber die abhangigkeit der visuellen bewegungsrichtung und negativen nachbildes von den reizvorgangen auf der netzhaut. Zeitschr. Psychol. *59*, 321-330.
2. Wallach, H. (1935). Uber visuell wahrgenommene bewegungsrichtung. Psychol. Forsch. 20, 325-380.
3. Pack, C.C., Livingstone, M.S., Duffy, K.R., and Born, R.T. (2003). End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. Neuron *39*, 671-680.

4. DeAngelis, G.C., Freeman, R.D., and Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. J. Neuro-physiol. *71*, 347-374.

5. Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. *195*, 215-243.

6. Lorenceau, J., Shiffrar, M., Wells, N., and Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. Vis. Res. *33*, 1207-1217.

7. van den Berg, A.V., and Noest, A.J. (1993). Motion transparency and coherence in plaids: the role of end-stopped cells. Exp. Brain Res. *96*, 519-533.

8. Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1986). The analysis of moving visual patterns. Exp. Brain Res. 11, 117-152.

9. Rodman, H.R., and Albright, T.D. (1989). Single-unit analysis of pattern motion selective properties in the middle temporal visual area (MT). Exp. Brain Res. *75*, 53-64.

10. Pack, C.C., and Born, R.T. (2001a). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature *409*, 1040-1042.

11. Tinsley, C.J., Webb, B.S., Barraclough, N.E., Vincent, C.J., Parker, A. and Derrington, A.M. (2003). The nature of V1 neural responses to 2D moving patterns depends on receptive field structure in the marmoset monkey. J. Neurophysiol. *90*, 930-937.

12. Pack, C.C. and Born, R.T. (2001b). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. Nature *414*, 905-908.

13. Stoner, G.R., and Albright, T.D. (1992). Neural correlates of percep-tual motion coherence. Nature *358*, 412-414.

14. Stoner, G.R., and Albright, T.D., and Ramachandran, V.S. (1990). Transparency and coherence in human motion perception. Nature *344*, 153-155.

15. Noest, A.J. and van den Berg, A.V. (1993). The role of early mecha-nisms in motion transparency and coherence. Spat. Vis. *7*, 125-147.

16. de Boer, R. and Kuyper, P. Triggered correlation (1968). IEEE Trans Biomed. Eng. *15*, 169-179.

17. Livingstone, M.S., Pack, C.C. and Born, R.T. (2001). Two-dimen-sional sub-structure of MT receptive fields. Neuron *30*, 781-793.

18. Borghuis, B.G., Perge, J.A., Vajda, I., van Wezel, R.J.A., van de Grind, W.A., and Lankheet, M.J.M. (2003). The motion reverse cor-relation (MRC) method: a linear approach in the motion domain. J. Neurosc. Methods *15*, 153-166.

19. Bair, W., Cavanaugh, J.R., and Movshon, J.A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. J. Neurosci. *23*, 7690-7701.

20. Lamme, V.A.F., Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. *11*, 571-579.