Current Biology, Vol. 14, 000–000, May 25, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.cub.2004.xx.xxx

Vision: The World Through Picket Fences

Dispatch

David Burr[1] and John Ross[2]

Our visual system must allow us to see the form of objects in motion. Tracking objects of interest stabilises their images on the retina, but is not sufficient, as untracked images move on the retina. This problem is solved by cells tuned in both space and time, combining information about form with information about motion.

One of the major breakthroughs of the last few decades of vision research has been the discovery of two separate functional streams: a ventral stream for the analysis of form (the 'what' stream), and a dorsal stream for the analysis of position and motion (the 'where' stream) [1]. Interestingly, a similar division of labour has recently been described for the auditory system [2]. The visual system is usually thought to separate its processing of form from its processing of motion, and to subdivide these two main streams further into processing modules, each specialised for different aspects of 'what' and 'where': luminance, colour, texture, depth, complex motion and so on [3]. Indeed so widely accepted is the idea of separate modules for different visual attributes, many neurobiologists believe that there is a 'binding problem' of how to link the different attributes together to recover a coherent holistic percept.

While this neat picture of separate paths of analysis is very appealing, and receives support from various lines of study, there are several lines of evidence suggesting the story is at best incomplete. One clear example of interaction between form and motion is the 'biological motion' first described by Gunnar Johansson [4]: when point light sources are attached to an actor's joints, they are perceived as a meaningless jumble of lights when the actor is stationary, but give an immediate vivid impression of the actor when she or he is walking (see http://www.bml.psy.ruhr-uni-bochum.de/Demos/BMLwalker.html for demonstration). Motion reveals form. This phenomenon has been very well studied and is generally ascribed to the combination of information from the form and motion pathways [5].

Biological motion, however, is not a unique example. In this issue of Current Biology [6], Shin'ya Nishida reminds us that the visual system is capable of extracting complex form information from translating patterns. Furthermore, he shows that by taking advantage of the spatio-temporal information available only to a system tuned to motion, we can pick up information not available from any static view or any collection of views.

Nishida [6] displayed moving targets behind a virtual 'picket fence' that obscured the scene except for the thin slits between the 'pickets', illustrated in Figure 1. He went further by showing that more complex alphanumeric patterns, also perceived under these sampled conditions [12], rely on spatio-temporal interpolation of the motion system. After demonstrating that eye-movements are not essential, he showed that forms created in this way can be selectively masked by a superimposed moving noise source, most effectively when the noise moves in the same direction as the moving letter, and at a similar speed. That masking motion of a certain velocity affects the percept points clearly to its role in form perception under these conditions. His second experiment is still cleverer. Using an adaptation of the reverse-correlation technique, he showed that the spatial frequencies used for the form recognition task are higher than the limit imposed by spatial sampling and therefore only available via temporal information.

What sort of mechanisms may be responsible for extracting this temporal information? Nishida [6] takes up an early suggestion that spatio-temporal interpolation may be an intrinsic property of cells in V1 itself [13]. These cells are not only tuned spatially to particular orientations [14], but many are also directionally tuned, giving them an 'orientation' in space–time [15]. A receptive field appropriately tuned to the up-sloping segment of the 'A' grapheme will also be tuned to motion vectors roughly orthogonal to this direction. The alignment in space–time of receptive field and image creates a strong response. The receptive field does not respond well to other spatial orientations, nor to other directions of drift.

The experiments reported by Nishida [6] provide strong confirmation for such mechanisms in human vision. They also suggest an extension. Without spatio-temporally tuned mechanisms, frequencies higher than the Nyquist sampling rate not only provide no information about form, they also cannot support the perception of motion itself — like the wagon-wheels of an old-time western movie, high spatial frequencies on their own signal motion in the wrong direction. So if these frequencies are used for form perception, as Nishida's results suggest, and a true sense of motion is required to extract the form, then more complex mechanisms are needed than the simple spatio-temporal cells in primary visual cortex [13]. There needs to be a mechanism capable of uniting the true motion information from low spatial-frequency mechanisms with the high spatial frequencies required for fine form analysis. This is of course possible, but perhaps calls for further verification of the role of frequencies above the Nyquist rate in spatio-temporal interpolation. One would want to be

[1]Dipartimento di Psicologia, Università di Firenze, and Istituto di Neuroscienze del CNR, Via Moruzzi 1, Pisa 56100, Italy.
[2]School of Psychology, The University of Western Australia, Nedlands, Western Australia 6009, Australia.

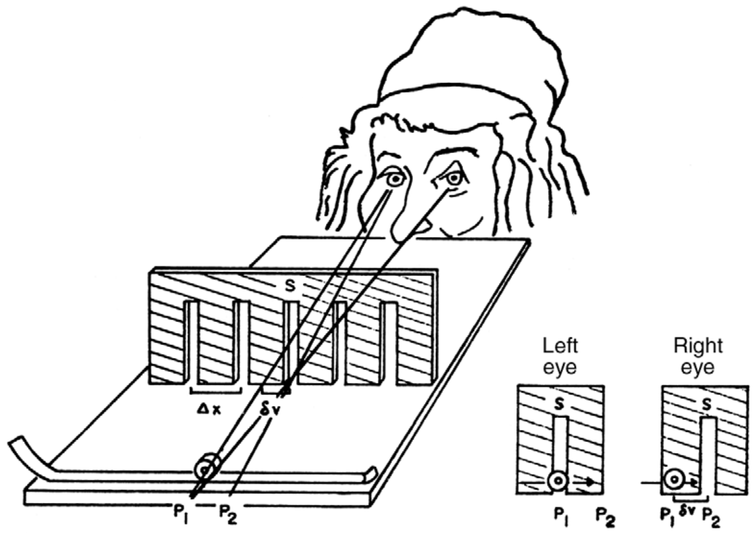

Figure 1. A simple scene viewed through pickets.

The whole of the object moving behind the slits is never visible at any instant, and if it is in depth it may be visible to one eye but not to the other. Yet observers can see what it is and its depth. This is the phenomenon of spatio-temporal interpolation, first observed for simple tasks such as vernier or stereo alignment [7–11]. These early studies showed that the phenomenon is not due to artefacts such as pursuit eye movements — which actually 'paint' the scene onto the retina — and requires that motion be perceived.

certain, for example, that eye-movements — clearly not sufficient on their own — did not supplement the pattern recognition.

Why have complex spatio-temporal visual mechanisms evolved? To allow us to view objects moving behind picket fences? Unlikely. As well as handling sampled images, they will increase the efficiency with which vision handles any image in motion. In order to increase the signal to noise level, the visual system integrates over time, for about 100 milliseconds in daylight. During this interval, motion across a receptive field could cause considerable smear. If the receptive field itself is tuned to the image motion, however, the integration can occur without necessarily producing smear [16].

Interestingly, recent evidence has suggested that motion smear, usually detrimental, may be useful in disambiguating the direction of motion [17]. This is a further example of the interaction between form and motion. The effects of motion smear are mimicked by Glass patterns generated from random dot patterns — by superposition of a rotated or other transformed pattern with the original — that give a strong sense of global form. When a sequence of these patterns, uncorrelated with each other, is displayed over time, there is a strong impression of global motion along the path of global form, even though there is no motion energy in this direction [18]. Indeed, cells in areas MT/MST of the macaque monkey visual cortex respond to these sequences in the same way as they do to real motion [19]. Furthermore the individual patterns in a sequence can alter our perception of direction when shown as a background for motion [20] and change the direction preferences of cells in the monkey brain [19].

We do not suggest that an understanding of the spatio-temporal mechanisms that vision possesses will make the binding problem disappear. But it may well be time to place less emphasis on the separate feats of analysis the visual system can perform and more on how its different components work in cooperation to allow us to see the dynamic world we live in.

References

1. Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. Trends Neurosci. *6*, 414-417.
2. Rauschecker, J.P., and Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U.S.A. *97*, 11800-11806.
3. Zeki ,S. (1993). A vision of the brain. Oxford: Blackwell Scientific.
4. Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Percept. Psychophy. *14*, 201-211.
5. Giese, M.A., and Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. Nat. Rev. Neurosci. *4*, 179-192.
6. Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. Curr. Biol., this issue.
7. Morgan, M.J., and Thompson, P. (1975). Apparent motion and the Pulfrich effect. Perception *4*, 3-18.
8. Burr, D.C., and Ross, J. (1979). How does binocular delay give information about depth? Vision Res. *19*, 523-532.
9. Burr, D.C. (1979). Acuity for apparent vernier offset. Vision Res. *19*, 835-837.
10. Fahle, M., and Poggio, T. (1981). Visual hyperacuity: Spatio-temporal interpolation in human vision. Proc. Roy. Soc. Lond. B *213*, 451-477.
11. Fahle, M., Biester, A., and Morrone, C. (2001). Spatiotemporal interpolation and quality of apparent motion. J. Opt. Soc. Am. A Opt. Image Sci. Vis. *18*, 2668-2678.
12. Ross, J. (1977). A new type of display relying on vision's sensitivity to motion. J. Physiol. (Lond) *271*, 3P.
13. Burr, D.C., and Ross, J. (1986). Visual processing of motion. Trends Neurosci. *9*, 304-306.
14. Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. (London) *195*, 215-243.
15. Emerson, R.C., Bergen, J.R., and Adelson, E.H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. Vision Res. *32*, 203-218.
16. Burr, D.C. (1980). Motion smear. Nature *284*, 164-165.
17. Geisler, W.S. (1999). Motion streaks provide a spatial code for motion direction. Nature *400*, 65-69.
18. Ross, J., Badcock, D.R., and Hayes, A. (2000). Coherent global motion in the absence of coherent velocity signals. Curr. Biol. 2000, *10*, 679-682.
19. Krekelberg, B., Dannenberg, S., Hoffmann, K.P., Bremme, F., and Ross, J. (2003). Neural correlates of implied motion. Nature *424*, 674-677.
20. Burr, D.C., and Ross, J. (2002). Direct evidence that "speedlines"

influence motion mechanisms. J. Neurosci. *22*, 8661-8664.