

Stereoscopic Vision: Solving the Correspondence Problem

Dispatch

Andreas Nieder

Neurons in early visual areas respond to horizontal disparity in images that do not give rise to stereopsis. False binocular matches, however, are discarded at the apex of the visual pathway: the activity of neurons in the primate inferior temporal cortex correlates directly with conscious depth perception.

In binocular vision, the images that are projected onto the retinae of the left and right eyes are slightly displaced relative to each other. This positional difference is known as ‘horizontal disparity’ (Figure 1A) and gives rise to vivid depth perception, even if the monocular images are unstructured noise and contain no distinctive features, as in random-dot stereograms (Figure 1B) [1]. But how does the visual system know which dot in the left eye belongs to which dot in the right eye to perform appropriate matching of the monocular images? This ‘correspondence problem’ is a central issue that the visual system must solve in order to derive three-dimensional information. A recent study [2] now reports neurons in the inferior temporal cortex that have solved the correspondence problem.

Most neurons at relatively early processing stages of the visual hierarchy — such as areas V1 [3], MT [4] and MST [5] in monkeys, or the visual Wulst in owls [6] — do not have the capacity to discard false binocular matches. They signal horizontal disparity in anti-correlated random-dot stereograms (Figure 1C), which do not lead to depth perception in humans [7] or other animals [2,3,6]. While the observed neural responses to anti-correlated random-dot stereograms were consistent with the idea that such neurons act as locally restricted filters (Figure 1D) [8,9], it was concluded that they cannot be a direct correlate of depth perception [3] — even though an electrical stimulation study [10] showed that stimulation of clusters of MT neurons with a certain preferred disparity changed monkeys’ depth perception in a predictable way. Somewhere in the brain the correspondence problem had to be solved by neurons that respond only to stimuli that give rise to stereopsis (Figure 1E).

A recent study [2] now reports that single neurons at the apex of the ventral visual pathway (Figure 2) in behaving monkeys do not respond to disparity in anti-correlated random-dot stereograms, and thus have solved the correspondence problem. Janssen and co-workers [2] tested disparity-sensitive neurons in the anterior inferior temporal cortex, area TE. As most neurons at this advanced processing stage respond

only to complex object properties, the authors provided as stimuli three-dimensional shapes defined by disparity cues [11]. By introducing smooth disparity gradients in random-dot displays, double curved three-dimensional surfaces were generated that bent towards (convex) or away (concave) from the observer.

Janssen *et al.* [2] found that many TE neurons responded selectively to either concave or convex double curved surfaces in random-dot stereograms, thus encoding the depth ordering of three-dimensional surfaces. In contrast to previous findings in early visual cortex, TE neurons did not convey information about false-matched disparities in the anti-correlated random-dot stereograms. Whether stimulated with anti-correlated patterns or uncorrelated patterns — neither of which supports depth perception — the neurons showed equally low firing rates, but they always responded vigorously to the correlated random-dot stereogram. Some of the investigated neurons were not (only) tuned to three-dimensional shapes, but responded to the plain depth position of a shapeless random-dot stereogram (analogous to cells studied previously in earlier visual areas [3,5,6]). Even these ‘primary-like’ disparity-sensitive neurons were insensitive to anti-correlated random-dot stereograms. It was convincingly demonstrated that the neurons’ insensitivity to anti-correlated random-dot stereograms could not be explained by a lack of attention or by difference in eye movements in the two stimulus conditions.

Together, these results show that disparity-selective TE neurons correlate directly with the monkeys’ three-dimensional shape perception. It would be interesting to see whether neurons in higher visual areas of the second, dorsal pathway (Figure 2) also have the capacity to discard false matches. An ideal candidate for investigation seems to be the caudal part of the intra-parietal sulcus. Neurons in this region are sensitive to the three-dimensional orientation of a surface defined by horizontal disparity, suggesting an important role in the perception of three-dimensional shape [12].

Area TE is the highest-order level of the ventral visual pathway (Figure 2), so the question arises as to where exactly in the hierarchy of the visual system neural responses to disparity stimuli become disambiguated. It may not be within TE: neurons in TE might just reflect the highly processed input they receive from preceding stages. Along the visual pathway, disparity-selective neurons become increasingly more sophisticated. While cells in V1 generally respond only to the absolute disparity of random-dot stereogram surfaces, many cells in area V2 respond to relative disparity between different regions of a random-dot stereogram [13] or detect edges in random-dot stereogram [14]. Area V4, which feeds into TEO and TE, contains even more elaborate neurons which signal the disparity-defined three-dimensional orientation of bars [15].

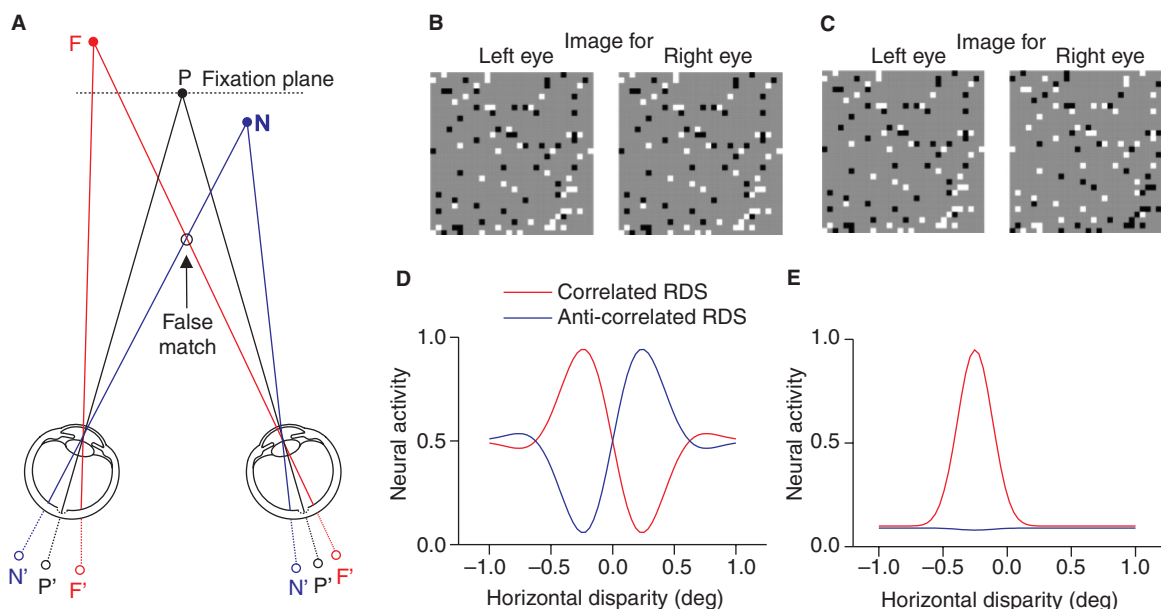


Figure 1. Geometric relationships in binocular vision and disparity detectors.

(A) An observer fixates point P, the images of which (P') are projected onto corresponding retinal locations (the fovea). Because of the lateral displacement of the eyes, points beyond the fixation point are projected on disparate retinal locations. For example, the angular differences between the images of point F (F) relative to corresponding retinal locations in the right and left eye defines this point's horizontal disparity. Negative and positive disparities occur for objects nearer or farther away than the fixation plane, respectively. The correspondence problem is indicated by the arrow: in principal, the image of N in the left eye could be matched falsely with the image of F in the right eye. But the visual system discards such false matches, and no object is perceived at that depth. (B) In correlated random-dot stereograms, black and white dots appear at identical locations of the image shown to the left and right eye. Presented with horizontal disparity, such random-dot stereograms allow depth perception. (C) In anti-correlated random-dot stereograms, the left eye sees a white dot at the same location where the right eye sees a black dot, and vice versa. Anti-correlated random-dot stereograms cannot be fused and do not support stereopsis. (D) Schematic disparity response profile of a disparity-sensitive neuron in early visual cortex. The neural activity as a function of disparity in correlated random-dot stereograms (red) is completely inverted during stimulation with anti-correlated random-dot stereograms (blue). For anti-correlated random-dot stereograms, such a neuron would indicate an image farther away than the fixation plane — by increased activity to positive disparity — which is not perceived. (E) A neuron that has solved the correspondence problem should only be tuned to disparity in correlated random-dot stereograms, but respond with baseline activity to anti-correlated random-dot stereograms. The neurons reported by Janssen *et al.* [2] belong to this class of global disparity detectors.

But even in early visual forebrain areas, some disparity-sensitive neurons already discard false matches in anti-correlated random-dot stereograms, leading to an average false-match activity that is only about half the size that a pure local filtering model would predict [3,6,9]. Interestingly, neurons in the behaving owl's visual Wulst that exhibited longer response latencies were less sensitive to anti-correlated random-dot stereograms, suggesting a functional hierarchy of disparity processing leading from spatial filters to more global disparity detectors [6]. It is thus very likely that the correspondence problem is not solved at one defined area of the visual system, but rather gradually from one stage to the other.

What mechanisms might eliminate false binocular matches? Coding ambiguities in sensory systems are generally caused by the narrow filter characteristics of peripheral sensory neurons. In spatial hearing [16] and electrolocation [17], for example, ambiguities are often eliminated at higher computational levels by integration of multiple processing streams. But such across-channel integration alone is insufficient to

generate global disparity detectors, as it cannot explain the suppression of responses to disparity in anti-correlated random-dot stereograms. One simple, yet very effective mechanism to eliminate responses to false matches is implementation of higher discharge thresholds for higher-order neurons that get input from local detectors. This would lead to a decrease of baseline activity in high-order detectors and would enable the visual system to 'clip' response dips in a cell's response profile before it could become inverted in anti-correlated random-dot stereograms (Figure 1D,E).

Evidence for such a threshold operation comes from a study in owls [6] where the most unequivocally responding cells had significantly lower discharge rates to non-preferred disparities. This mechanism is also supported by a neural network model [18]. A threshold operation alone, however, is not sufficient to explain the lack of activity to anti-correlated random-dot stereograms. Inhibitory influences may also contribute. Cells that suppressed responses to anti-correlated random-dot stereograms showed significantly more inhibition

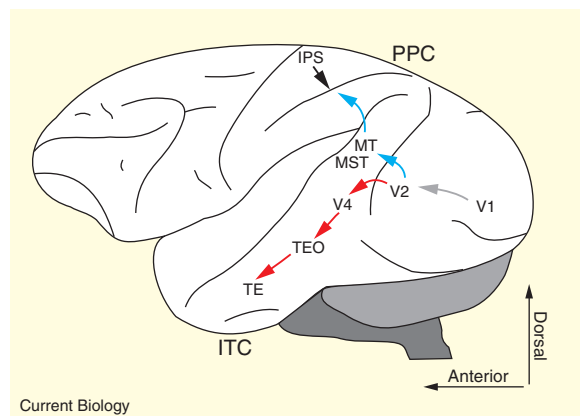


Figure 2. Simplified scheme of the ventral (red) and dorsal (blue) visual pathways in the monkey brain.

IPS, intra-parietal sulcus; ITC, inferior temporal cortex; MT, middle temporal area; MST, medial superior temporal area; PPC, posterior parietal cortex.

compared to neurons that signaled disparity in opposite-contrast patterns [6].

In the past decade, the binocular disparity energy model [8,9] has been an extremely fruitful mathematical description of primary disparity detectors, making predictions that can be tested physiologically. Neurons that solve the correspondence problem, as described by Janssen *et al.* [2], however, are a clear deviation from the local filtering model [9]. Future physiological and computational work will have to elaborate modifications of the disparity energy model to account for discrepancies that become progressively more evident in real neurons [6,9,19].

References

1. Julesz, B. (1971) Foundations of Cyclopean Perception. (University of Chicago Press).
2. Janssen, P., Vogels, R., Liu, Y. and Orban, G.A. (2003). At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron* 37, 693–701.
3. Cumming, B.G. and Parker, A.J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389, 280–283.
4. Krug, K., Cumming, B.G. and Parker, A.J. (1999). Responses of single MT neurons to anti-correlated stereograms in the awake macaque. *Soc. Neurosci. Abstr.* 25, 275.
5. Takemura, A., Inoue, Y., Kawano, K., Quaia, C. and Miles, F.A. (2001). Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding. *J. Neurophysiol.* 85, 2245–2266.
6. Nieder, A. and Wagner, H. (2001). Hierarchical processing of horizontal disparity information in the visual forebrain of behaving owls. *J. Neurosci.* 21, 4514–4522.
7. Cumming, B.G., Shapiro, S.E. and Parker, A.J. (1998). Disparity detection in anticorrelated stereograms. *Perception* 27, 1367–1377.
8. Ohzawa, I., DeAngelis, G.C. and Freeman, R.D. (1990). Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* 249, 1037–1041.
9. Ohzawa, I. (1998). Mechanisms of stereoscopic vision: the disparity energy model. *Curr. Opin. Neurobiol.* 8, 509–515.
10. DeAngelis, G.C., Cumming, B.G. and Newsome, W.T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature* 394, 677–680.
11. Janssen, P., Vogels, R. and Orban, G.A. (1999). Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc. Natl. Acad. Sci. USA* 96, 8217–8222.
12. Shikata, E., Tanaka, Y., Nakamura, H., Taira, M. and Sakata, H. (1996). Selectivity of the parietal visual neurones in 3D orientation of surface of stereoscopic stimuli. *Neuroreport* 7, 2389–2394.

13. Thomas, O.M., Cumming, B.G. and Parker, A.J. (2002). A specialization for relative disparity in V2. *Nat. Neurosci.* 5, 472–478.
14. von der Heydt, R., Zhou, H. and Friedman, H.S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Res.* 40, 1955–1967.
15. Hinkle, D.A. and Connor, C.E. (2002). Three-dimensional orientation tuning in macaque area V4. *Nat. Neurosci.* 5, 665–670.
16. Mazer, J.A. (1998). How the owl resolves auditory coding ambiguity. *Proc. Natl. Acad. Sci. U.S.A.* 95, 10932–10937.
17. Heiligenberg, W. (1991). *Neural Nets in Electric Fish*. (MIT Press, Cambridge, MA).
18. Lippert, J. and Wagner, H. (2001). A threshold explains modulation of neural responses to opposite-contrast stereograms. *Neuroreport* 12, 3205–3208.
19. Cumming, B.G. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature* 418, 633–636.