Representation of stereoscopic structure in human and monkey cortex

Christopher W. Tyler

Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco CA 4115, USA

Tsao *et al.* have recently used functional magnetic resonance imaging to compare processing for moving stereoscopic forms in macaque and human brains. Most humans exhibited activation in a swath of lateral occipital areas, extending into the intraparietal sulcus, with a limited version of the same pattern in monkeys. However, neither species showed strong activation of the motion area known as MT in monkey or its human homolog.

The present era of magnetic imaging research can perhaps be characterized as moving from issues of primary activation patterns to a focus on secondary properties of complex image processing, such as object structure and 3D depth organization. These are particularly challenging goals because such processing is multifaceted, and every aspect of what we see must, ultimately, be represented somewhere in the brain. For example, in the issue of depth structure from binocular disparity information, there must be representations of absolute disparity, relative disparity, depth edges, surface segmentation and 3D depth form, as a minimum. Although they all derive from the same disparity stimulus, each different aspect requires different cellular processing and a different representation metric. Not until the stimulus design allows each of these aspects of disparity processing to be isolated from the others can we take the full measure of the components making up such a neural processing domain. It is in such 'cartographic' surveys of the complexity of neural circuitry that functional magnetic resonance imaging (fMRI) excels.

fMRI responses to dynamic stereoscopic structure

An advance in this enterprise has just appeared in the form of a paper by Tsao *et al.*, comparing disparity processing for complex dynamic forms in human and macaque monkey [1]. Noting that single-unit studies have shown disparity processing to be widely distributed in monkey cortex, they applied the same methods, stimuli and fMRI analysis techniques to human and, for the first time, to macaque brains, to allow a direct comparison of the disparity activation patterns. In particular, Tsao *et al.* could be the first to avail themselves of the technique of dynamic random-dot stereograms developed in 1960 by Bela Julesz [2], which isolate the purely stereoscopic aspects of dynamic form processing, rendering invisible all monocular aspects of both the spatial form and the temporal changes in form, by the randomization of the monocular images over space and time. Humans exhibited activation by this cyclopean, or purely stereoscopic, stimulus in an extended swath of lateral occipital areas (Figure 1), reaching into the parietal lobe along the intraparietal sulcus (IPS). Monkeys showed a limited version of the same activation pattern, with a surprising degree of variation among individuals. The minimum pattern of activation by disparity structure in both species seemed to be centered in retinotopic area V3A and the caudal (or occipital) extension of the IPS, which is currently considered as non-retinotopic. Retinotopy is the cortical property of one-to-one mapping of the retinal array and is used to define the primary and secondary architecture of the projections to the cortex.

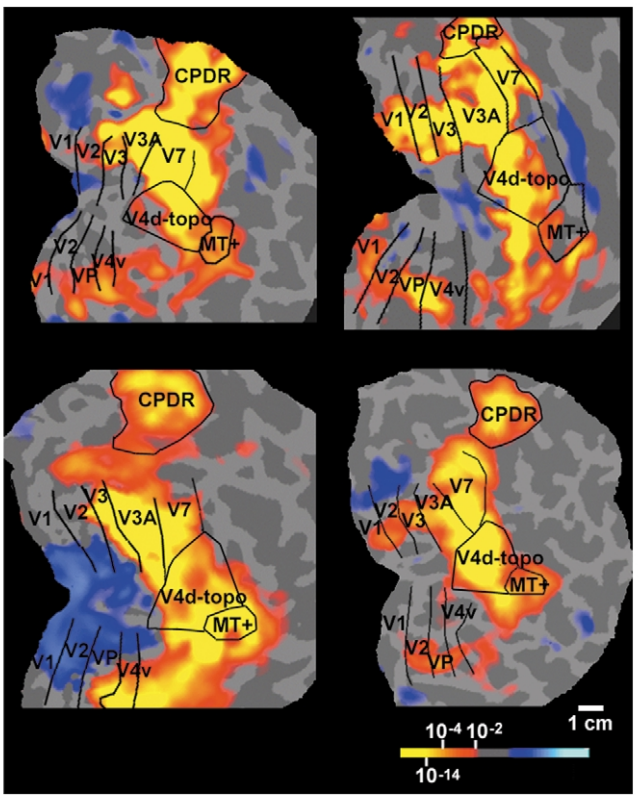

Figure 1. Flattened representations of occipital cortex in four hemispheres (gray), showing the extended activation (red to yellow) and surrounding regions of reduced activation (blue) in response to the stereoscopic stimulus consisting of 3.5° squares of a random range of disparities drifting laterally across the field at 2.2° s^{-1}, presented in a dynamic random-dot display that eliminated monocular cues to the form or the motion. Note that the activation occurs largely in a swath of cortical regions beyond the primary retinotopic projections of V1–V3, extending to the peripheral regions of those retinotopic areas [e.g. to areas V3A, VP, V4v and V7, to the human occipital motion area designated MT+, to the V4d topolog (V4d-topo), and to a caudal parietal disparity region (a non-retinotopic region in the caudal intraparietal sulcus, anterior to V7; CPDR)]. There is substantial variation among the different brains. Reproduced, with permission, from Ref. [1].

Corresponding author: Christopher W. Tyler (cwt@ski.org).

The predominance of stereoscopic activation in V3 and V3A confirms the consensus result from previous human neuroimaging studies in [3–11] and from non-stereoscopic studies of depth processing in monkeys [12]. What is notable is that the strongest activation in every human experiment was in the area adjacent to the traditional retinotopic areas. Another property, which is not generally discussed, is that these vivid dynamic stimuli activated no regions of the monkey brain anterior to area MT (at least as far forward as is shown in the flattened cortical maps). A similar pattern is seen in the human brain, even when attention is uncontrolled, implying a sharp break between the processing of complex visual stimuli and some other form of neural processing in the anterior reaches of the occipital lobe. Not even the attentional elaboration of the spatial structure can be located here because such elaboration would be modulated by the presence or absence of the physical stimulus.

Complexity of stereoscopic activation

With careful controls, Tsao *et al.* [1] effectively excluded the role of segmentation edges, attention and eye movements in their complex disparity responses. We thus obtain a picture of disparity processing occurring predominantly in the dorsal pathway projecting out of the primary visual representation. To a large extent, these data have corralled the processing of disparity structure to a selection of late retinotopic areas, progressing to non-retinotopic cortical areas on the lateral posterior surface of the brain. It should be said that the definition of retinotopy in general is based on the use of non-stereoscopic stimuli. It remains an open question whether these newly identified dorsal stereoscopic areas might exhibit retinotopy for local structure in stereoscopically-defined, rather than luminance-defined, retinotopy probes. This issue exemplifies how each new advance in stimulus or paradigm specificity of the cortical responses requires us to rethink the completeness of those that had gone before. For example, the retinotopy of motion area V5 could not be evaluated until an adequate motion-based stimulus was incorporated into the retinotopic paradigm [13]. Areas V3A and V7 respond retinotopically to luminance-based stimuli but regions of the IPS can appear non-retinotopic until adequate stimuli are found to activate them. The strong activation by moving depth structure now suggests how the retinotopy of these cortical regions might be evaluated effectively. The neighboring regions, activated neither by luminance stimuli nor by dynamic depth stimuli, will have to wait until the key stimuli are found to unlock their principles of operation.

Difficulties with the interpretation of results

One difference that emerges between human and macaque is in the response of the well-established motion area, which is in the middle temporal lobe (MT) in monkey but in the homologous occipital region (V5) in human. In macaque, it is noteworthy that neither the moving stereoscopic stimulus nor the planar dynamic noise strongly activated MT. In human, there was some activation of the homologous motion area but no difference between the activation by the moving stereo stimulus and the zero-disparity noise plane. One could explain such results by the concept of coextensive subpopulations below the spatial resolution of fMRI. If a cortical area has the same number of cells responsive to zero disparity and to the other disparities in the range of the present test stimuli, alternating between these two conditions would produce no change in the overall blood-oxygen-level-dependent (BOLD) activation signal, even though the cells would be accurately coding the various disparities. However, the macaque data make it clear that there is little response in MT even to the local motions (as perceived by humans) in the planar dynamic noise stimulus. This result suggests that macaque MT needs coherent regions of luminance motion to show significant activation. Neither incoherent local motion nor coherent disparity-defined motion will do.

Despite the richness of their study, the test paradigms developed by Tsao *et al.* still fall short in certain respects. The principal stimulus chosen for analysis consisted of laterally moving random cyclopean checks versus a static plane. In addition to non-zero disparity, the cues to activation in this stimulus thus included relative disparity differences, 2D segmentation edges, 2D segmentation structure, depth edges, 3D structure, local disparity change over time, and lateral motion – each of which might be processed in a different cortical region. This variety of stimulus properties might go some way to explaining the variation in responsiveness to the stereoscopic stimulus, particularly in the ventral occipital regions. Indeed, half of the group of human observers was excluded from all analyses on the justification that their retinotopic maps were not sufficiently reliable. The quality and distribution of the stereoscopic responses of this excluded group were not discussed, however.

One unusual feature of the stereoscopic results is the fact that all observers showed either no response or an inhibitory, negative response in the foveal projection areas of areas V1–V3, with pronounced activation at greater eccentricities. Tsao *et al.* attribute this switchover to possible differences in disparity tuning with eccentricity, but this interpretation seems to be invalidated by the control experiment of a disparity plane presented at the same range of disparities as the original checkerboard stimulus. In that case, the early retinotopic activation pattern inverts to predominate in the foveal representations of areas V1–V3, with no significant activation at greater eccentricities. Such data vitiate any analysis based on summing voxel responses over the whole extent of a particular retinotopic area because the responses vary from negative to positive with increasing eccentricity, and would therefore tend to sum to zero overall. As demonstrated in Figure 2 of Ref. [1], the fovea is readily able to provide a percept of disparity structure over the $\pm 22^{\circ}$ range of the stimuli used. Further studies will be needed to resolve the implications of this activation reversal for the two stimulus types containing the same local disparity information.

Evolutionary factors in response diversity

The paper by Tsao *et al.* concludes with a discussion of the notable differences in complex disparity processing

in human relative to macaque, emphasizing that the two species are separated by 30 million years of evolution. (They might have added the even more significant difference in brain size of about a factor of ten.) Such differences cast a critical light on attempts at interspecies brain mapping [14]. Most likely, the mapping will overlap well in early visual areas but progressively diverge as the overlay of complex processing in humans overtakes the relatively more simple capabilities of the macaque brain. One might have hoped that something as simple as the processing of 3D structure would fall on the side of the early matching processes, and indeed that the genetic similarity and homogeneous laboratory environment of the monkeys would result in a uniformity of cortical responses. The diversity in patterns of response among the four macaques suggests that, as in humans, the various depth cues (e.g. disparity, vergence, motion parallax, shape-from-shading, perspective and texture gradient) are weighted with pronounced individual variation, although presumably with the same veridical outcome of accurate perception of distance and graspable 3D structure. Perhaps there are so many cues available that there is no need for the developmental process to mandate a particular weighting across individuals. In view of this diversity, the most useful approach might be to perform the fMRI mapping on the particular monkey to be investigated in single-unit studies.

Acknowledgements
Supported by NIH/NEI grant EY 7890.

References

1 Tsao, D.Y. *et al.* (2003) Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39, 555–568
2 Julesz, B. (1971) *Foundations of Cyclopean Perception*, University of Chicago Press
3 Ptito, M. *et al.* (1993) Localization and lateralization of stereoscopic processing in the human brain. *NeuroReport* 4, 1155–1158
4 Gulyas, B. and Roland, P.E. (1994) Binocular disparity discrimination in human cerebral cortex: functional anatomy by positron emission tomography. *Proc. Natl. Acad. Sci. U. S. A.* 91, 1239–1243
5 Mendola, J. *et al.* (1999) The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J. Neurosci.* 19, 8560–8572
6 Kwee, I.L. *et al.* (1999) Perceptual processing of stereopsis in humans: high-field (3.0-tesla) functional MRI study. *Neurology* 53, 1599–1601
7 Kourtzi, Z. and Kanwisher, N. (2001) Representation of perceived object shape in the human lateral occipital complex. *Science* 293, 1506–1509
8 Backus, B. *et al.* (2001) Human cortical activity correlates with stereoscopic depth perception. *J. Neurophysiol.* 86, 2054–2068
9 Merboldt, K.D. *et al.* (2002) Functional MRI of self-controlled stereoscopic depth perception. *NeuroReport* 13, 1721–1725
10 Negawa, T. *et al.* (2002) M pathway and areas 44 and 45 are involved in stereoscopic recognition based on binocular disparity. *Japan. J. Physiol.* 52, 191–198
11 Gillaie-Dotan, S. *et al.* (2002) Shape-selective stereo processing in human object-related visual areas. *Hum. Brain Map.* 15, 67–79
12 Sereno, M.E. *et al.* (2002) Three-dimensional shape representation in monkey cortex. *Neuron* 33, 635–652
13 Huk, A.C. *et al.* (2002) Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205
14 Van Essen, D.C. *et al.* (2001) An integrated software system for surface-based analysis of cerebral cortex. *J. Amer. Med. Inf. Assoc.* 8, 443–459

0166-2236/$ - see front matter
doi:10.1016/j.tins.2003.12.009

Research Focus Response

Response to Tyler: Representation of stereoscopic structure in human and monkey cortex

Doris Tsao[1,2] and Roger Tootell[2,3]

[1]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA
[2]Massachusetts General Hospital, NMR Center, Athinoula A. Martinos Center, Charlestown, MA 02129, USA
[3]Department of Radiology, Harvard Medical School, Boston, MA 02115, USA

Stereoscopic depth perception presumably depends on disparity-tuned cells in the visual system, but what is the network of cortical areas that ultimately brings about 3D perception? Lesion studies have distinguished between dorsal stream areas, which process where an object is, and ventral stream areas, which process what an object is. Because the primary importance of 3D structure is to define spatial layout (after all, the identity of almost any object can be shown with a perspective-less cartoon), one might expect 3D perception to activate mainly dorsal stream areas. Surprisingly, physiologists have found disparity-tuned cells in almost every region of the visual cortex [1], including the anterior temporal lobe [2], raising the possibility that a much larger network of cortical areas is crucially involved.

To investigate this, we performed functional magnetic resonance imaging (fMRI) in alert macaque monkeys (and in humans, for comparison) [3]. We found that the strongest disparity-driven activation in the macaque was restricted to a small cluster of dorsal areas at the junction between the lunate and intraparietal sulci, namely areas V3, V3A and the caudal intraparietal sulcus (CIPS). In contrast to single-unit results, our fMRI data suggest that the cortical processing of stereopsis is not equally distributed throughout the entire brain.

Moreover, we found this selective activation of V3, V3A and CIPS in all eight hemispheres (four monkeys) tested, in response to a wide variety of stereoscopic stimuli (disparity-defined checkerboard, disparity annulus and

Corresponding author: Doris Tsao (doris@nmr.mgh.harvard.edu).