nature neuroscience

Disparity- and velocity-based signals for three-dimensional motion perception in human MT+

Bas Rokers, Lawrence K Cormack & Alexander C Huk

How does the primate visual system encode three-dimensional motion? The macaque middle temporal area (MT) and the human MT complex (MT+) have well-established sensitivity to two-dimensional frontoparallel motion and static disparity. However, evidence for sensitivity to three-dimensional motion has remained elusive. We found that human MT+ encodes two binocular cues to three-dimensional motion: changing disparities over time and interocular comparisons of retinal velocities. By varying important properties of moving dot displays, we distinguished these three-dimensional motion signals from their constituents, instantaneous binocular disparity and monocular retinal motion. An adaptation experiment confirmed direction selectivity for three-dimensional motion. Our results indicate that MT+ carries critical binocular signals for three-dimensional motion processing, revealing an important and previously overlooked role for this well-studied brain area.

Most research on motion processing uses computer-generated patterns that move across two-dimensional displays. Real objects, however, move through a three-dimensional world. To support the binocular perception of three-dimensional motion, the visual system must extract information from different patterns of dynamic stimulation on the retinae. Much is known about how the visual system processes stimuli to encode two-dimensional motion in a fixed depth plane and how static binocular disparities are processed to represent position in depth[1,2].

Much less is known, however, about how the visual system combines two-dimensional motion and binocular disparity to compute three-dimensional motion. Most notable is the dearth of evidence for direct roles of the MT and medial superior temporal (MST) visual areas in three-dimensional motion, which is surprising given their central role in two-dimensional motion processing and well-established disparity sensitivity[3–9].

The visual system could rely on one or both of two binocular cues to compute three-dimensional motion direction: changes in binocular disparity over time (changing disparity) and/or the difference in the interocular image velocities at corresponding points on the two retinae (interocular velocity difference, IOVD). The changing disparity cue is sufficient for the perception of motion through depth[10,11], and recent work has shown that the IOVD cue makes a robust, independent contribution[12–17].

We carried out a series of four functional magnetic resonance imaging (fMRI) experiments to identify distinct three-dimensional motion signals in the human visual system, to elucidate the processing of the disparity-based (changing disparity) and/or velocity-based (IOVD) cues, and to test for three-dimensional motion direction selectivity. Across all of our experiments, MT+ responses were consistently selective for stimuli that specified three-dimensional motion. This selectivity was distinct from responses to instantaneous disparities and monocular motions. We found evidence of three-dimensional motion selectivity in V3A and lateral occipital (LO) complex, but none in primary visual cortex. The pattern of blood oxygen level–dependent (BOLD) responses suggests that the computation of three-dimensional motion occurs subsequent to V1 and that three-dimensional motion signals are evident at the population level in MT+.

RESULTS

Observers viewed the stimuli through a mirror stereoscope while we measured BOLD fMRI responses in a variety of visually responsive regions of interest (ROIs), including V1, V2, V3, V3A/B, hV4, LO1/LO2 and MT+ (combined putative MT/MST). In Experiments 1–3, displays alternated in a blocked design between conditions every 12 s. Experiment 4 used an event-related adaptation protocol. Observers continually performed a challenging detection task to control attention (see Online Methods).

Experiment 1: selectivity for three-dimensional motion

When an object moves directly toward or away from an observer fixating a stationary point, regions on the retinae corresponding to the object's location are stimulated by opposite directions of horizontal motion. These dichoptic horizontal opponent motions must be distinguished from similar patterns of retinal stimulation that do not specify three-dimensional motion, such as locally opposing horizontal motions on the same retina[18,19] or opposing vertical motions, whether across the two retinae (dichoptic) or on the same retina (monocular). Horizontal dichoptically opponent motion contains both changes in binocular disparity over time (the changing disparity cue) and interocular velocity differences (the IOVD cue). To encode either of these

Neurobiology, Psychology, Center for Perceptual Systems, Institute for Neuroscience, and Imaging Research Center, The University of Texas at Austin, Austin, Texas, USA. Correspondence should be addressed to B.R. (rokers@mail.utexas.edu).

Received 4 February; accepted 27 April; published online 5 July 2009; doi:10.1038/nn.2343

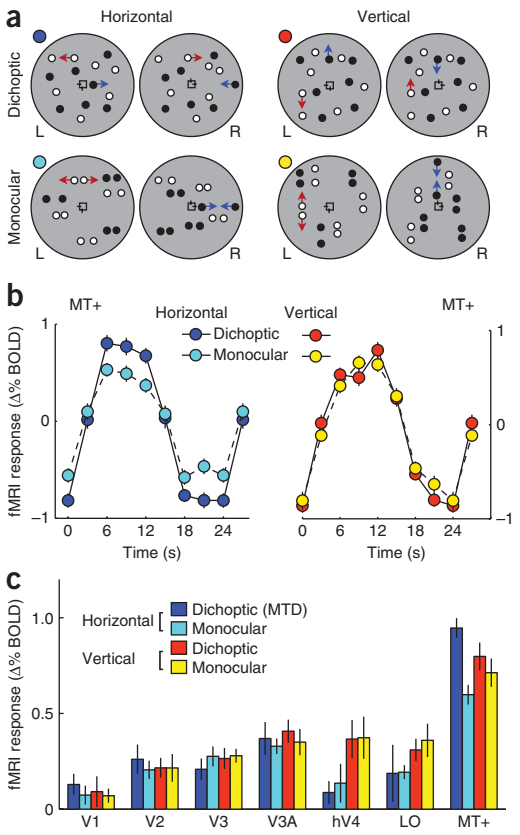

Figure 1 Stimuli and results for Experiment 1. (**a**) Stimulus conditions. Left (L) and right (R) eye half images for each of four conditions are shown. Each image contained 32 dots (14 are shown). Each dot oscillated sinusoidally (arrows illustrate motions) and the two dots in a pair moved in anti-phase. Dots moved along horizontal (left column) or vertical trajectories (right column). In dichoptic conditions (top row), paired dots were split between the left and the right images. In monocular conditions (bottom row), paired dots were presented in the same eye. The horizontal-dichoptic stimulus was perceived as three-dimensional motion; in other conditions, dots were perceived as twinkling, but without any three-dimensional motion. Color codes condition throughout all figures. Left and right half images are fusible stereo-pairs. (**b**) MT+ responses were selective for three-dimensional motion. Time series of fMRI response across moving and stationary epochs (averaged across observers) in MT+ are shown for the four conditions. Left, MT+ responses to horizontal motion are larger for dichoptic than for monocular presentation. Right, MT+ responses to vertical motion are similar for dichoptic and monocular presentation. Error bars represent ± s.e.m. (**c**) MT+ showed the clearest selectivity for three-dimensional motion. fMRI response amplitudes are shown across all visual areas, averaged across subjects. MT+ responses showed a significant interaction between orientation (horizontal/vertical) and presentation (dichoptic/monocular), demonstrating a sensitivity to three-dimensional motion separate from possible sensitivities to horizontal or dichoptic motion alone (see text for statistical information). This selectivity was not evident in other visual areas. Error bars represent ± s.e.m.

cues, the visual system must distinguish the orientation, the direction and the eye of origin of visual motion signals.

We measured fMRI responses while observers viewed displays comprising 32 moving dot pairs (**Fig. 1a**). Dots in each pair had either opposing horizontal or opposing vertical directions of motion (**Fig. 1a**) and were presented either dichoptically (each dot in a pair to a different eye) or monocularly paired (both dots in a pair in the same eye) (**Fig. 1a**). Dots pairs oscillated back and forth, starting at a random (but opposite) phase on their sinusoidal trajectory to avoid full-field coherent monocular motion and to minimize oculomotor drive. We quantified the amplitude of fMRI response modulation as displays alternated between moving (12 s) and stationary (12 s).

MT+ responded preferentially to horizontally opponent, dichoptic motion: the only stimulus that simulated three-dimensional motion. We examined the average time course in MT+ as the dots alternated between moving and stationary for the four conditions (**Fig. 1b**). MT+ responses were ∼50% larger for dichoptic opponent motion than for monocular opponent motion when the dot paths were horizontal (t test on amplitudes, $t_{16} = 4.77$, $P = 2 \times 10^{-4}$), but were nearly identical for vertical motion ($t_{16} = 0.82$, $P = 0.425$).

We examined the response amplitudes for all of the areas that we studied (**Fig. 1c**). The difference in response to dichoptically separated as compared with monocularly paired stimuli was substantially larger for horizontal than for vertical stimuli in MT+ (repeated measures ANOVA, interaction between presentation (dichoptic/monocular) and direction (horizontal/vertical), $F_{1,35} = 10.22$, $P = 0.006$). This pattern of results was unique to MT+ ($P > 0.6$ for all other ROIs).

The larger MT+ response to dichoptically separated compared with monocularly paired horizontal opponent motion suggests that MT+ carries signals that are specific to three-dimensional motion. The lack of difference for vertical motions demonstrates orientation specificity of the responses consistent with the horizontal offset between the two eyes. This dependence on orientation also confirms that the difference in responses between dichoptically separated and monocularly paired displays was not a result of differential patterns of monocular dot density between the conditions or of differences in overall patterns of dichoptic stimulation.

These results suggest that MT+ processes three-dimensional motion distinctly from similar stimuli that do not specify three-dimensional motion. Of course, horizontal dichoptically opponent motion, as with any real moving object, contains changing disparity and IOVD cues (as well as instantaneous binocular disparities). Although psychophysical evidence supports the notion that both cues are used[12–15], it is unclear how they are computed and combined. We therefore isolated each cue in two additional experiments: one that tested for a contribution of the changing disparity cue (Experiment 2) and one that tested for a contribution of the IOVD cue (Experiment 3). We then confirmed that three-dimensional motion responses were direction selective in an event-related adaptation experiment (Experiment 4).

Experiment 2: disparity-based cue to three-dimensional motion

The changing disparity cue can be isolated in dynamic random dot stimuli that contain systematic changes in disparity over time, but do not contain any coherent monocular motion[20,21]. Such displays only contain disparity-based motion signals in the cyclopean view (that is, after binocular combination); each eye's half image looks like incoherent random dot noise.

We examined this changing disparity–isolating stimulus (**Fig. 2a**). On each stimulus frame, dots were repositioned in random image locations and then assigned one of two disparities depending on the quadrant in which they fell (**Fig. 2a**). Disparities changed smoothly and systematically over time to produce a percept of wedges moving sinusoidally through depth, with adjacent quadrants moving in opposite directions (**Fig. 2a**). Such stimuli do not contain coherent monocular motions and thus no IOVD cue, nor do they contain global coherent motion that might drive eye movements.

We measured fMRI responses as observers viewed a display that alternated between this changing disparity stimulus (changing disparity motion) and a spatiotemporally scrambled version containing the same number of dots and the same distribution of dot disparities

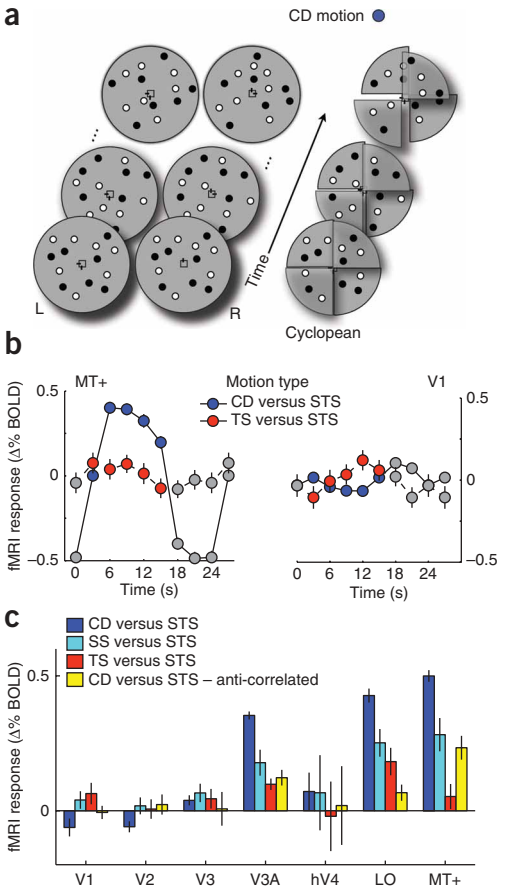

Figure 2 Stimuli and results for Experiment 2. (**a**) Changing disparity (CD) motion stimulus. The left pairs of patches show the left and right eye's half images. The rightmost patches depict a scheme of the cyclopean percept. Binocular disparities between left and right images defined a spatiotemporal pattern of quadrants oscillating sinusoidally in depth. The locations of individual dots on each stimulus frame were randomized, removing any coherent monocular motion. The display alternated between this changing disparity motion stimulus (12 s) and a spatiotemporally scrambled (STS) version (12 s). (**b**) MT+ responses were selective for changing disparity–defined three-dimensional motion. Time series of fMRI response in MT+ (left) and V1 (right) are shown for both changing disparity motion (blue) and temporally-scrambled changing disparity-motion (red) versus STS motion (gray). The MT+ response was large to changing disparity motion, but small to STS motion. V1 responses were generally small, but note the inversion of response to changing disparity and STS motion. (**c**) MT+ showed the clearest selectivity for changing disparity–defined three-dimensional motion. Average fMRI response amplitudes across all visual areas are shown for changing disparity motion, spatially scrambled changing disparity motion (SS), temporally scrambled changing disparity motion (TS) and anti-correlated changing disparity motion (all versus spatiotemporally scrambled changing disparity motion). MT+ responses were large to changing disparity motion, smaller to scrambled changing disparity motion, negligible to temporally scrambled changing disparity motion and greatly reduced when the changing disparity display was anti-correlated. V3A responses were smaller, but followed a similar pattern. LO responses were also selective for changing disparity motion, but the response to temporally scrambled changing disparity motion was relatively large, despite the fact that changing disparity motion was not perceived in this condition.

(see Online Methods). Perceptually, the displays alternated between crisp, wedge-shaped surfaces moving through depth and a dynamic three-dimensional cloud of dots with no coherent structure or motion. These stimuli were indistinguishable when viewed monocularly.

When we plotted the average time course of the MT+ response as the display alternated between changing disparity motion and spatiotemporally scrambled motion, the response modulation was readily apparent (**Fig. 2b**). This contrasts with the V1 response, which is not only much weaker, but also inverted, probably as a result of the abundance of unstructured disparities continuously present in the spatiotemporally scrambled blocks (**Fig. 2b**).

To verify that the MT+ modulation was a result of a changing disparity signal *per se* (rather than to spatial or temporal disparity structure present in the changing disparity blocks), we ran three additional conditions: spatially scrambled (scrambling the spatial positions of the dots while leaving the temporal structure intact), temporally scrambled (scrambling the order in which the frames were presented while leaving the spatial structure intact; **Fig. 2b**) and anti-correlated (pairing each white dot in one eye with a black dot in the other eye and vice versa). This last manipulation greatly degrades the disparity signal on which the changing disparity cue relies[13–15,21,22]. Each condition alternated against the spatiotemporally scrambled stimulus.

In all three conditions, the response modulation was substantially reduced relative to the primary changing disparity motion condition, indicating that MT+ responses were largest for spatiotemporal patterns of disparity that specify three-dimensional motion. We plotted the response amplitudes for all of the areas studied in the main changing disparity condition and the three other conditions (**Fig. 2c**). Comparison across areas showed a strong contrast between the earliest areas (for example, V1, V2 and V3) and later areas (V3A, LO and MT+; **Fig. 2c**). In MT+, responses varied significantly between the changing disparity condition and the other three conditions (ANOVA main effect of motion type, $F_{3,71} = 16.90$, $P = 0.003$; changing disparity versus: spatially scrambled, $t_{22} = 2.45$, $P = 0.023$; temporally scrambled, $t_{22} = 8.04$, $P = 5 \times 10^{-8}$; anti-correlated, $t_{22} = 3.42$, $P = 3 \times 10^{-3}$).

Even though the spatially and temporally scrambled conditions contained the same overall distribution of disparities, the spatiotemporal structure uniquely present in the changing disparity condition yielded considerably larger MT+ responses. Similar modulations of response were also observed in LO ($F_{3,71} = 16.80$, $P = 0.003$) and V3A ($F_{3,71} = 18.30$, $P = 0.002$), but not in earlier areas ($P > 0.06$).

The relatively weak responses to the temporally scrambled stimulus suggest that MT+ is selective for motion from the changing disparity cue, distinct from disparity-defined edges or other depth-defined structure generated by the presence of wedges in our displays. The temporally scrambled frames had the same number and average magnitude of disparity-defined edges, but lacked the temporal structure consistent with smooth three-dimensional motion. We further verified this by repeating the main changing disparity condition with displays that contained different numbers of wedges (**Supplementary Fig. 1**). Note that anti-correlated displays containing only the changing disparity cue yielded a ~50% response reduction. In Experiment 3, when IOVDs were present, the identical manipulation had a very different effect on MT+ responses.

Experiment 3: velocity-based cue to three-dimensional motion

Having established a changing disparity response in MT+, we tested for the presence of an IOVD signal. Although it was possible in Experiment 2 to remove the IOVD cue from the stimulus and to isolate the changing disparity cue, it is not geometrically possible to do the converse. One can, however, selectively degrade the changing disparity signal using binocularly anti-correlated stimuli without affecting the monocular velocities on which the IOVD cue relies[13,14].

We measured fMRI responses as subjects viewed displays that alternated between moving (12 s) and stationary (12 s). We presented either

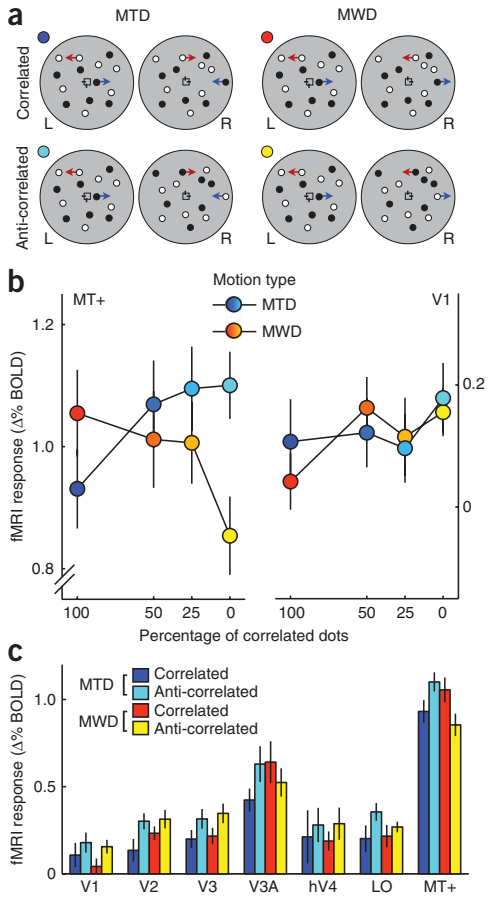

Figure 3 Stimuli and results for Experiment 3. (**a**) Stimulus conditions. Each pair of patches represents a left and right eye half image. In MTD conditions, dots in each dichoptic pair moved in opposite horizontal directions in the two eyes. In MWD conditions, dots in each pair moved in the same horizontal directions (in phase). The range of disparities was the same in MWD and MTD conditions. In correlated conditions, dot pairs were of the same contrast polarity. In anti-correlated conditions, dot pairs were of opposite contrast polarity. In anti-correlated displays, MWD stimuli appeared flat, but MTD stimuli still appeared to contain motion through depth. (**b**) MT+ responses to motion containing IOVDs were robust to anti-correlation. Average fMRI response amplitudes are shown for MTD and MWD as a function of the percentage of correlated dots (100%, perfect correlation; 0%, complete anti-correlation). MT+ responses (left) to MWD decreased as the percentage of correlation decreased (95% confidence interval, on the slope = [–0.33, –0.03], units of percentage BOLD change per 100% correlation), but MT+ responses to MTD increased with anti-correlation. V1 responses (right) showed no such dependence. (**c**) MT+ showed selectivity for IOVD-biased three-dimensional motion. Average fMRI response amplitudes across all areas are shown. MT+ responses to MWD decreased as the percentage of correlated dots decreased, whereas responses to MTD increased. A similar, but weaker, pattern of response was evident in V3A. Trends in other areas were non-significant (see text for statistical information).

motion through depth (MTD) displays or motion within depth (MWD) displays that contained dots moving in frontoparallel planes occupying the same range of disparities traversed by the MTD stimulus (**Fig. 3a**). Dot pairs oscillated sinusoidally with random starting phase. The only difference between the MTD and MWD stimuli was the relationship of a dot pair's motion across the two eyes, which was either in phase (MWD) or anti-phase (MTD) (**Fig. 3a**).

Displays contained fully correlated dots, fully anti-correlated dots (**Fig. 3a**), or 50/50% or 25/75% correlated/anti-correlated dots (data not shown). For the correlated displays, the dots either appeared to be moving toward and away through depth (MTD) or side-to-side at various depths in the same three-dimensional volume (MWD). For the anti-correlated displays, the dots either appeared to be moving toward and away through depth (MTD, just as for the correlated displays) or side-to-side with little or no relative depth (MWD, as anti-correlation degraded the disparity signal). This loss of percepts of three-dimensional position in light of preserved percepts of three-dimensional motion qualitatively replicates our previous psychophysical isolation of the IOVD cue using anti-correlation[14].

MT+ responses to MWD decreased as binocular correlation decreased (**Fig. 3b**). In contrast, MT+ responses to MTD did not (and may have increased slightly; **Fig. 3b**). These slopes were both different from zero and from one another (95% confidence interval on the difference between the slopes = [0.16, 0.54]; see **Fig. 3** for additional confidence intervals). If MT+ responses were only driven by net monocular motion, we should have observed identical responses across all conditions. Indeed, the same analysis for V1 showed that neither the MTD nor the MWD slope was significantly different from zero and, notably, they were not different from one another (95% confidence interval on the difference between the slopes = [–0.15, 0.25]; **Fig. 3b**).

The smaller MWD response to anti-correlated dots than to correlated dots is consistent with observed MT+ responses in previous fMRI studies of disparity processing[23]. This reduction probably reflects the reduced disparity-based component of the MT+ response. Given that the only difference between the MWD and MTD stimuli is whether the dots were in phase across the eyes, we interpret the lack of a similar MTD reduction as revealing a distinct contribution of the (remaining) IOVD signal.

We analyzed the responses across visual areas (omitting the intermediate correlations; **Fig. 3c**). The significant interaction between correlation (correlated/anti-correlated) and direction of motion (MTD/MWD) in MT+ (ANOVA, $F_{1,35} = 9.36$, $P = 8 \times 10^{-3}$) supports our parametric analysis of slopes (**Fig. 3c**). V3A showed a similar pattern ($F_{1,35} = 6.92$, $P = 0.018$), but no other areas did (all $P > 0.1$).

This pattern indicates that MT+ carries a signal associated with MTD even when strong disparity-based (changing disparity) cues are weak or absent. It is therefore likely that, in addition to responding to the changing disparity cue (Experiment 2), MT+ can also be driven by the IOVD cue to three-dimensional motion.

Experiment 4: selective adaptation to three-dimensional motion

If, as the preceding experiments indicate, MT+ is important in three-dimensional motion perception, we should expect to find direction selectivity for three-dimensional motion. We performed an fMRI-adaptation experiment to test whether adaptation to unidirectional three-dimensional motion (containing both changing disparity and IOVD cues) differentially affected fMRI responses to a subsequent probe stimulus that moved in either the same or opposite direction as the adaptor. In each scanning session an adapting random dot field moved either toward or away from the observer. After 100 s of adaptation, observers viewed a series of 7.5-s trials, each containing a 4-s top-up adaptor followed by a 1-s probe (with same or opposite directions; **Fig. 4a**).

We observed strong direction-selective adaptation; MT+ responses were smaller when the probe stimulus moved in the same direction as the adaptor compared with when the adaptor moved in the opposite direction ($F_{1,1476} = 85.72$, $P \approx 0$; **Fig. 4b**). Pronounced effects were evident in every observer in every scanning session (**Supplementary Fig. 2**) and were consistent with results from ongoing psychophysical studies (**Supplementary Fig. 3**). We observed similar, but smaller, effects in V3A ($F_{1,1476} = 14.94$, $P = 0.0001$) and LO ($F_{1,1476} = 13.86$,

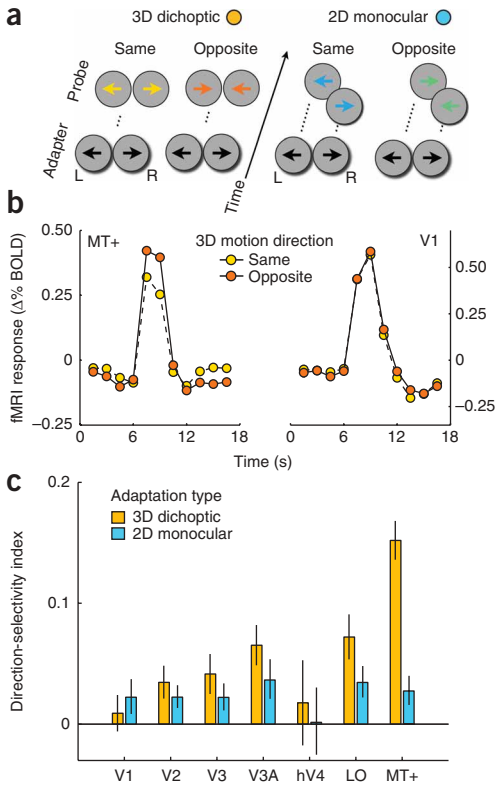

Figure 4 Stimuli and results for Experiment 4. (**a**) Stimulus conditions. Each pair of circular patches represents the left and right eye half images. In the three-dimensional (3D) dichoptic condition (left), a top-up three-dimensional motion adaptor containing random dots appeared to drift either away (shown) or toward the subject. The subject subsequently viewed a probe stimulus that moved in either the same or opposite direction. In the two-dimensional (2D) monocular conditions (right), the two eyes' probe stimuli were presented in immediate temporal succession, rather than simultaneously; the two conditions were otherwise identical. (**b**) MT+ responses reflect direction-selective adaptation to MTD. Time series of fMRI response in MT+ (left) and V1 (right) are shown for probe stimuli during the three-dimensional motion adaptation condition. In MT+, but not V1, responses to opposite-direction probes (orange) were larger than to same-direction probes (yellow). Error bars are smaller than plotting symbols. (**c**) MT+ showed the strongest three-dimensional motion direction selectivity. Average direction-selectivity indices across all areas for three-dimensional motion (orange bars) and monocular two-dimensional motion (cyan bars) adaptation. Three-dimensional motion direction selectivity was strongest in MT+, intermediate in areas V3A and LO, and not evident in V1. Direction selectivity to two-dimensional monocular motion was weak, but present across visual areas. Error bars indicate bootstrapped 68% confidence intervals (equivalent to ± 1 s.e.m.).

$P = 0.0002$), smaller, but still significant, effects in V2 and V3 ($P < 0.05$), and no reliable effects in V1 ($P = 0.58$; **Fig. 4b**). This confirms the importance of MT+ in the processing of three-dimensional motion (as well as the secondary roles of V3A and LO), as identified in the first three experiments.

We confirmed that these effects were not inherited from the adaptation of earlier monocular two-dimensional direction mechanisms by presenting the left and right eye's probe stimuli in immediate succession, rather than simultaneously (**Fig. 4a**). These staggered probes yielded small adaptation effects of constant magnitude across areas (marginal in MT+, $F_{1,1477} = 3.63$, $P = 0.057$). The pattern of response is consistent with downstream areas inheriting V1 adaptation, but the magnitude of monocular two-dimensional adaptation cannot account for the large three-dimensional motion adaptation effects that we observed in MT+.

To quantify selectivity for three-dimensional motion, we computed a direction-selectivity index reflecting the difference in response to the opposite- versus same-direction probes normalized by their sum (Online Methods). The direction-selectivity index for three-dimensional motion was largest in MT+ (**Fig. 4c**). Other areas showed lesser degrees of selectivity, which appeared to grow in magnitude across the hierarchy (that is, V1 < V2 < V3 < V3A, LO < MT+). We plotted the selectivity indices for the staggered probes (**Fig. 4c**). Notably, the three-dimensional dichoptic index was markedly different from the staggered (monocular) index only in MT+ ($P < 10^{-4}$ bootstrapped, $P > 0.05$ in all other areas).

The distinction between the patterns of three-dimensional dichoptic and two-dimensional monocular direction-selective adaptation enhance the conclusions drawn from the prior blocked design experiments: MT+ carries strong three-dimensional direction signals. Although V3A and LO may also process three-dimensional signals, their selectivity could not be distinguished from sensitivity to monocular motions. The results in two subregions in MT+ (putative human MT and MST) were generally similar across experiments (**Supplementary Fig. 4**). The greater sensitivity of the adaptation protocol did reveal a somewhat larger three-dimensional motion direction-selectivity index in MST (0.15 ± 0.02) than in MT (0.09 ± 0.02). The conservative method for identification of MT and MST left many MT+ voxels unassigned (Online Methods), so the overall MT+ selectivity is not the simple average of MT and MST selectivity.

DISCUSSION

Our results demonstrate that the two binocular signals produced by three-dimensional motion, changing disparity and IOVD, are processed in human MT+. Despite extensive investigation of the role of human MT+ (as well as macaque MT and MST) in the perception of both frontoparallel (two dimensional) motion and static disparity[3,7,24,25], electrophysiological evidence for three-dimensional motion sensitivity has been middling. Some physiological work in monkeys and cats has suggested the existence of cells tuned to trajectories of three-dimensional motion[26–31], and a study on motion parallax has pointed to a role for motion-selective MT neurons in three-dimensional perception[32], but recent work has not observed responses that are unambiguously selective for three-dimensional motion in macaque MT[5,7]. It seems incongruous, however, that a brain area might have evolved simply to process two-dimensional motion and static depth.

Our results differ from some prior electrophysiological studies, which optimized stimuli to suit the two-dimensional frontoparallel selectivity of individual neurons[5]. In contrast, we optimized our stimuli to yield robust three-dimensional motion percepts to human observers. A previous fMRI study did find responses to structure from two-dimensional motion in human MT+[33]. Another fMRI study emphasized responses to cyclopean stereomotion anterior to canonical human MT+[20]. Our results are difficult to compare directly with those of that study because of differences in stimuli (we only used smooth changing-disparity motion alternating with a variety of scrambled controls and separately investigated the contribution of coherent monocular motions) and differences in data analysis (we first identified and subdivided MT+ and then made additional measurements in these ROIs to quantify the responses to motion though depth). Although our findings in MT+ do not preclude the involvement of areas outside MT+, they do demonstrate clear three-dimensional motion selectivity in canonical MT+ and motivate future work to assess the contributions of subregions and adjacent areas[20].

Our results also suggest a role for V3A and lateral occipital areas (LO1/LO2), which are less well understood in both human and monkey models. Our findings suggest that they deserve further consideration in the processing of disparity and velocity signals. In contrast, we observed very little evidence for three-dimensional motion selectivity in V1.

Although our measurements could have lacked the sensitivity or resolution to identify three-dimensional motion selectivity in V1, we suggest that V1 most likely serves to extract the building blocks[34–36] for three-dimensional motion computations that are ultimately performed in extrastriate areas. Subcortical pathways may also contribute to these computations, possibly bypassing V1 altogether.

There are numerous reasons why the pattern of results across all four experiments cannot be explained by sensitivity to disparity alone[5]. First, the first three experiments contrasted, in different ways, responses to stimuli with the same distributions of disparities. Second, we observed patterns of responses to three-dimensional motion across areas that differed from previous fMRI characterizations of sensitivity to static disparities[37]. Third, the results of our adaptation experiment demonstrated direction selectivity by comparing responses to stimuli that differed solely in their direction of three-dimensional motion, while containing identical patterns of disparities.

Furthermore, we observed a response reduction to anti-correlated stimuli in MT+, but only for stimuli that did not contain IOVDs. Thus, the effects of anti-correlation were different for displays that did and did not contain IOVDs, revealing a neural computation for three-dimensional motion that cannot be explained in terms of sensitivity to disparity alone.

In summary, our results reveal an important and previously overlooked role for human MT+ in three-dimensional motion processing. Our results provide evidence that MT+ responds to both changing disparities and IOVDs that specify movement through three-dimensional space. This three-dimensional motion sensitivity derived from both disparity- and velocity-based cues motivates reconsideration of the well-studied sensitivities to frontoparallel motion and static disparity in MT and MST. Likewise, our observation of an IOVD signal in MT+ may explain why several computations in MT appear to be monocular[38,39]. Such eye-specific processing may not be an inherited artifact of two-dimensional motion processing, but rather a reflection of a computational strategy for computing interocular velocity differences. Canonical neurophysiological models of MT motion processing[40–42] will need extension to incorporate the comparison of monocular velocity signals, as well as sensitivity to changing disparities over time.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureneuroscience/.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank D. Ress and T. Czuba for assistance with magnetic resonance imaging and comments on the manuscript and P. Neri for commenting on an earlier version of the manuscript. This work was supported by a National Science Foundation CAREER Award (BCS-0748413), a Mind Science Foundation Research Grant, and a pilot scanning grant from the University of Texas at Austin Imaging Research Center to A.C.H., and a Netherlands Organisation for Scientific Research grant to B.R. (2006/11353/ALW).

AUTHOR CONTRIBUTIONS

The authors jointly conceived the project, conducted the experiments and wrote the manuscript. B.R. programmed the visual displays and conducted the data analyses.

Published online at http://www.nature.com/natureneuroscience/.
Reprints and permissions information is available online at http://www.nature.com/reprintsandpermissions/.

1. Born, R.T. & Bradley, D.C. Structure and function of visual area MT. *Annu. Rev. Neurosci.* **28**, 157–189 (2005).
2. Cumming, B.G. & DeAngelis, G.C. The physiology of stereopsis. *Annu. Rev. Neurosci.* **24**, 203–238 (2001).
3. DeAngelis, G.C. & Newsome, W.T. Organization of disparity-selective neurons in macaque area MT. *J. Neurosci.* **19**, 1398–1415 (1999).
4. Huk, A.C., Dougherty, R.F. & Heeger, D.J. Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* **22**, 7195–7205 (2002).
5. Maunsell, J.H. & Van Essen, D.C. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.* **49**, 1148–1167 (1983).
6. Albright, T.D. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* **52**, 1106–1130 (1984).
7. DeAngelis, G.C. & Newsome, W.T. Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biol.* **2**, e77 (2004).
8. Nguyenkim, J.D. & DeAngelis, G.C. Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *J. Neurosci.* **23**, 7117–7128 (2003).
9. Smith, A.T. & Wall, M.B. Sensitivity of human visual cortical areas to the stereoscopic depth of a moving stimulus. *J. Vis.* **8**, 1–12 (2008).
10. Julesz, B. *Foundations of Cyclopean Perception* (The University of Chicago Press, Chicago, 1971).
11. Cumming, B.G. & Parker, A.J. Binocular mechanisms for detecting motion-in-depth. *Vision Res.* **34**, 483–495 (1994).
12. Brooks, K.R. Interocular velocity difference contributes to stereomotion speed perception. *J. Vis.* **2**, 218–231 (2002).
13. Harris, J.M. & Rushton, S.K. Poor visibility of motion in depth is due to early motion averaging. *Vision Res.* **43**, 385–392 (2003).
14. Rokers, B., Cormack, L.K. & Huk, A.C. Strong percepts of motion through depth without strong percepts of position in depth. *J. Vis.* **8**, 1–10 (2008).
15. Fernandez, J.M. & Farell, B. Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Res.* **46**, 1307–1317 (2006).
16. Beverley, K.I. & Regan, D. Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *J. Physiol. (Lond.)* **235**, 17–29 (1973).
17. Shioiri, S., Saisho, H., & Yaguchi, H. Motion in depth based on inter-ocular velocity differences. *Vision Res.* **40**, 2565–2572 (2000).
18. Heeger, D.J., Boynton, G.M., Demb, J.B., Seidemann, E. & Newsome, W.T. Motion opponency in visual cortex. *J. Neurosci.* **19**, 7162–7174 (1999).
19. Qian, N., Andersen, R.A. & Adelson, E.H. Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J. Neurosci.* **14**, 7357–7366 (1994).
20. Likova, L.T. & Tyler, C.W. Stereomotion processing in the human occipital cortex. *Neuroimage* **38**, 293–305 (2007).
21. Norcia, A.M. & Tyler, C.W. Temporal frequency limits for stereoscopic apparent motion processes. *Vision Res.* **24**, 395–401 (1984).
22. Cumming, B.G. & Parker, A.J. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* **389**, 280–283 (1997).
23. Bridge, H. & Parker, A.J. Topographical representation of binocular depth in the human visual cortex using fMRI. *J Vis* **7**, 1–14 (2007).
24. Bradley, D.C., Qian, N. & Andersen, R.A. Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* **373**, 609–611 (1995).
25. Dodd, J.V., Krug, K., Cumming, B.G. & Parker, A.J. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J. Neurosci.* **21**, 4809–4821 (2001).
26. Akase, E., Inokawa, H. & Toyama, K. Neuronal responsiveness to three-dimensional motion in cat posteromedial lateral suprasylvian cortex. *Exp. Brain Res.* **122**, 214–226 (1998).
27. Toyama, K., Komatsu, Y., Kasai, H., Fujii, K. & Umetani, K. Responsiveness of Clare-Bishop neurons to visual cues associated with motion of a visual stimulus in three-dimensional space. *Vision Res.* **25**, 407–414 (1985).
28. Cynader, M. & Regan, D. Neurons in cat visual cortex tuned to the direction of motion in depth: effect of positional disparity. *Vision Res.* **22**, 967–982 (1982).
29. Zeki, S.M. Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *J. Physiol. (Lond.)* **242**, 827–841 (1974).
30. Poggio, G.F. & Talbot, W.H. Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *J. Physiol. (Lond.)* **315**, 469–492 (1981).
31. Regan, D. & Cynader, M. Neurons in cat visual cortex tuned to the direction of motion in depth: effect of stimulus speed. *Invest. Ophthalmol. Vis. Sci.* **22**, 535–550 (1982).
32. Nadler, J.W., Angelaki, D.E. & DeAngelis, G.C. A neural representation of depth from motion parallax in macaque visual cortex. *Nature* **452**, 642–645 (2008).
33. Orban, G.A., Sunaert, S., Todd, J.T., Van Hecke, P. & Marchal, G. Human cortical regions involved in extracting depth from motion. *Neuron* **24**, 929–940 (1999).
34. Ponce, C.R., Lomber, S.G. & Born, R.T. Integrating motion and depth via parallel pathways. *Nat. Neurosci.* **11**, 216–223 (2008).
35. Barlow, H.B., Blakemore, C. & Pettigrew, J.D. The neural mechanism of binocular depth discrimination. *J. Physiol. (Lond.)* **193**, 327–342 (1967).
36. Hubel, D.H. & Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)* **195**, 215–243 (1968).
37. Tsao, D.Y. *et al.* Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* **39**, 555–568 (2003).
38. Tailby, C., Majaj, N. & Movshon, T. Binocular integration of pattern motion signals by MT neurons and by human observers [Abstract]. *J. Vis.* **7**, 95a (2007).
39. Majaj, N.J., Tailby, C. & Movshon, J.A. (2007). Motion opponency in area MT of the macaque is mostly monocular [Abstract]. *J. Vis.* **7**, 96a (2007).
40. Rust, N.C., Mante, V., Simoncelli, E.P. & Movshon, J.A. How MT cells analyze the motion of visual patterns. *Nat. Neurosci.* **9**, 1421–1431 (2006).
41. Perrone, J.A. & Thiele, A. A model of speed tuning in MT neurons. *Vision Res.* **42**, 1035–1051 (2002).
42. Simoncelli, E.P. & Heeger, D.J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).

ONLINE METHODS

Subjects. fMRI data were collected in three subjects (the three authors, males aged 33–44), all with normal or corrected-to-normal vision. All were experienced psychophysical observers in both motion and depth experiments. Experiments were undertaken with the written consent of each subject and all procedures were approved by the University of Texas at Austin Institutional Review Board and the University of Texas at Austin Imaging Research Center safety guidelines for magnetic resonance research. Each observer participated in a total of 19 scanning sessions.

Magnetic resonance imaging. Magnetic resonance imaging was performed at the University of Texas at Austin Imaging Research Center on a GE Signa HD 3T scanner, using a GE 8-channel phased array head coil.

Anatomical imaging. A whole-brain anatomical volume at 1-mm^3 resolution was acquired for each subject. Brain tissue was segmented into gray matter, white matter and cerebrospinal fluid by an automated algorithm followed by manual refinement[43]. The inplane anatomical volume from each fMRI session was then co-registered with this whole brain anatomical volume with an accuracy of ~1 mm (ref. 44).

fMRI. For Experiments 1–3, we used a two-shot spiral BOLD fMRI sequence[45] ((2.2-mm)^3 voxels, 3-s volume acquisition duration, repetition time = 1.5 s, echo time = 30 ms, 73 degree flip angle), with 24 pseudo-coronal slices starting at the occipital pole and continuing anterior to provide occipital and parietal coverage. Data from the first 24 s of each fMRI scan were discarded to minimize transient effects of both the scanner's magnetic saturation and observer's hemodynamics.

For Experiment 4, we used a two-shot spiral sequence with coarser spatial resolution and finer temporal resolution ((3.2-mm)^3 voxels, 1.5-s volume acquisition duration, repetition time = 750 ms, echo time = 30 ms, 56 degree flip angle), with 14 quasi-axial slices covering the posterior visual cortices, oriented roughly parallel with the calcarine sulcus. Data from the first and last two trials (15 s) of each fMRI scan were discarded to ensure that magnetic and hemodynamic steady-state had been reached and so that BOLD responses for each trial were similarly convolved with the responses from preceding and following trials.

In each session, T1-weighted inplane anatomical images were acquired for co-registration. Each fMRI time series of each voxel was highpass filtered (0.015 Hz cutoff frequency) to compensate for the slow signal drift typical in fMRI signals. Each voxel's time series was also divided by its mean intensity to convert the data from arbitrary image intensity units to percent signal modulation (% BOLD signal change) and to compensate for variations in mean image intensity across space. The resulting time series were then averaged across voxels in a given cortical area or ROI (see below) restricted to the portion of the gray matter that was responsive to the stimulus after reference scan restriction.

Defining the visual areas. The fMRI data were analyzed in each of the visual cortical areas separately for each subject. Mapping of visual areas V1, V2, V3, V3A, hV4, LO, MT+, MT and MST was performed in separate experimental sessions for each subject using standard techniques (see **Supplementary Fig. 5**)[4,46,47].

Statistical data analysis. For Experiments 1–3, we computed the average cycle time series responses shown in **Figures 1b** and **2b** by averaging the fMRI time series in each visual area across stimulus cycles (six per scan), repeats (three per condition) and subjects (three) at each time point. S.e.m. were computed for each time point, and were generally only slightly larger than the size of the plotting symbols. We estimated the fMRI response amplitudes (**Figs. 1c**, **2c** and **3b**,**c**) using standard Fourier-based block-design methods implemented by the Stanford VISTA group software package, mrVISTA/mrLoadRet[18] (see **Supplementary Fig. 5**).

For the analysis of response as a function of binocular correlation (**Fig. 3b**), we fit the response amplitudes to each condition (MTD and MWD) with a line using a least-squares algorithm. We then tested whether the slopes (or differences between slopes) were significantly different from zero using bootstrapping; we resampled the data with replacement 10^4 times, fit each resampled dataset with a line and calculated the 95% confidence intervals on the resulting distributions of bootstrapped slopes. Experiment 4 used an event-related adaptation protocol and analysis scheme that was identical to that described previously[48].

General visual apparatus, stimuli and task. Subjects viewed three-dimensional motion displays using a custom-built magnetic resonance–compatible mirror stereoscope. Stimuli were presented on a 36.5- × 27.4-cm rear projection screen (19.6 × 14.6 degrees of visual angle) illuminated by a 60 Hz LCD projector (linearized with 109.4-cd m^{-2} mean luminance). The viewing distance was 109 cm. Dots in the display were anti-aliased to obtain subpixel position accuracy. We used a pair of adjustable (pan/tilt) mirrors (one above each eye) to project the appropriate part of the stimulus display onto each eye. A mirror that is inclined with respect to the visual axis will produce a rotation of the projected image when slanted about its vertical axis. Observers visually corrected this rotation by counter-rotating the two halves of the visual display through the display software, implemented using the Psychophysics Toolbox[49]. A pair of adjustable septums (one immediately between the mirrors and one further along the optical path) blocked each eye's (undesired) view of the contralateral half image. This apparatus has been used and discussed in additional detail in our prior psychophysical work dissociating the changing disparity and IOVD cues[14].

In all of the experiments that involved moving dot pairs (Experiments 1, 2 and 4), the two dots in each pair had the same center of motion, regardless of whether they were presented in the same monocular half image or across the two half images. In Experiments 1–3, 32 dots pairs were presented in a 7.6-degree circular gray background aperture. In each monocular half image, half of the dots were white and half of the dots were black. Each dot in one half image was paired with a dot in the other half image; there were no unpaired dots in any displays. Dot size was 0.15 degree and binocular disparity ranged between ±22 arcmin. The average dot density was 0.83 dots per degree^2. Dots were spaced in the circular apertures to minimize spatial interactions across nearby dot pairs by requiring that the center of motion of each dot (corresponding to its 0 arcmin disparity point) was at least 0.88 degrees away from the same point of any dot belonging to another pair. A fixation point with nonius lines and four reference points outside of the aperture were present at all times to aid with proper vergence, fixation and half-image rotation. In Experiment 4, a number of features were modified to maximize the psychophysical adaptation effects. Details are reported below and in **Supplementary Figure 6**.

Behavioral task. To control attention across experimental conditions, the observers performed a challenging visual dot-change task throughout all fMRI runs[50]. Observers were required to press a button within 1 s of detecting a change in contrast polarity (Experiments 1 and 3) or the color (Experiments 2 and 4) of a randomly selected dot pair. Time of dot pair change was drawn from a 4-s average exponential distribution with a 1-s minimum and 12-s maximum (**Supplementary Fig. 6**).

Stimulus details. The stimuli in Experiment 1 (**Fig. 1a**) served to contrast signals generated by stimuli that yield three-dimensional motion percepts (dichoptically separated opposite-direction horizontal motion) with signals generated by stimuli that do not produce three-dimensional motion percepts (monocularly paired or vertical opposite-direction motion).

Moving dots were presented in circular patches in the left and right eye half images in four different experimental conditions. Dots were presented in pairs, and dots in each pair oscillated in opposite directions on a sinusoidal trajectory producing a maximum separation of ±22 arcmin. In horizontal conditions, the dots moved along horizontal trajectories. In vertical conditions, the dots moved along vertical trajectories. In dichoptic conditions, the oppositely moving dots in each pair were split between the left and the right half images. In monocular conditions, each oppositely moving dot pair was randomly positioned in either the left or the right eye. The horizontal-dichoptic stimulus consisted of changing horizontal binocular disparities and horizontal IOVDs and produced a percept of three-dimensional motion through depth (MTD). The other conditions contained similar motion properties (that is, the same orientation, but with monocular dot pairings, or the same dichoptic dot pairings, but with different orientation) and were generally perceived as twinkling moving dots (that is, motion opponent displays[18,19]).

Experiment 2. The stimuli in Experiment 2 (**Fig. 2a**) served to contrast disparity signals that produce strong three-dimensional motion percepts (changing disparity over time) with disparity signals that do not produce strong three-dimensional motion percepts (spatially scrambled, temporally scrambled or anti-correlated versions of the stimuli).

Dots were presented in the circular patches in the left and right eye half images in four different experimental conditions. In each condition, dot pair disparity varied between ±22 arcmin and the 12-s presentation of the experimental stimulus alternated with a 12-s spatiotemporally scrambled version in a blocked design. In the main changing disparity motion condition, the binocular disparities of the dots were systematically manipulated over time to create a spatiotemporal pattern of quadrants oscillating at 1.6 Hz toward and away from the observer. This was accomplished by assigning the same disparity to all dots in a quadrant and sinusoidally modulating that disparity over time. Adjacent quadrants moved in opposite directions (anti-phase). Critically, the image-plane locations of individual dots on each stimulus frame (60-Hz refresh rate) were randomized. This removed any coherent monocular motion from the display, but preserved the systematic changes in each quadrant's disparity over time[20]. Likewise, the presence of opposite directions of motion through depth across the quadrants removed full-field cyclopean motion and any corresponding drive to change vergence. We compared BOLD responses to spatially scrambled, temporally scrambled and anti-correlated versions of the same stimulus.

Experiment 3. The stimuli in Experiment 3 (**Fig. 3a**) served to contrast retinal motion signals (IOVDs) that produce strong three-dimensional motion percepts (MTD stimuli) with retinal motion signals that do not produce strong three-dimensional motion percepts (MWD stimuli). We used binocular anti-correlation to investigate the contribution of IOVDs, when the contribution of changing disparity was strongly reduced.

We presented 32 dots pairs in the circular patches of the left and right eye half images in eight different experimental conditions. In MTD conditions, the dots in each dichoptic pair moved horizontally in opposite directions (anti-phase; note that this stimulus is identical to the dichoptic-horizontal stimulus described in Experiment 1). In MWD conditions, the dots in each dichoptic pair moved in the same horizontal direction (same phase), at a randomly selected disparity. This manipulation is equivalent to changing in-phase dot pair sinusoidal motion (MWD) to anti-phase motion (MTD). The monocular motions and range of disparities were identical across MTD and MWD conditions. In correlated conditions, the dots in each dichoptic pair were of the same contrast polarity. In anti-correlated conditions, the dots in each dichoptic pair were of opposite contrast polarity.

Note that anti-correlated MWD displays appeared perceptually flat, but anti-correlated MTD displays were still perceived as containing three-dimensional motion. This selective robustness of MTD to anti-correlation reveals the contribution of the IOVD cue. We have used the same stimulus manipulation to dissociate the IOVD cue from the changing disparity cue in quantitative psychophysical experiments[14].

Experiment 4. This experiment employed a direction-selective adaptation protocol to investigate the selectivity of cortical areas to specific directions of three-dimensional motion. Each three-dimensional motion adaptation scanning session started with 100 s of initial adaptation to unidirectional three-dimensional motion (either toward or away). The adaptation direction was constant throughout each scanning session, which contained 12 adaptation runs. Each scanning run contained 36 7.5-s trials in random order. A trial contained 4 s of top-up adaptation, a 1.25-s interstimulus interval (mean gray screen), a 1-s probe stimulus and a final 1.25-s intertrial interval. In 14 of the 36 trials, the probe stimuli moved in the same direction as the adaptor, in another 14 it moved in the opposite direction, and in the remaining eight trials the adaptor was followed by a blank (no probe) stimulus. These adaptor-blank trials were used to estimate the baseline response to the adaptor (see **Supplementary Fig. 5**). Each subject participated in two scanning sessions, one with three-dimensional motion adaptation toward and one with adaptation away.

The monocular two-dimensional adaptation control experiment presented the same stimuli with slightly different timing. The 4-s top-up adaptation was followed by a 0.75-s interstimulus interval and a 2-s probe stimulus, during which the two monocular images were shown for 1 s each in random order (followed by a 0.75-s intertrial interval). The three-dimensional adaptation and two-dimensional monocular (staggered probe) control experiments were thus nearly identical, except for the temporally staggered (and thus monocular) presentation of the probe stimulus in this control experiment. Each subject participated in two scanning sessions, one for each direction of three-dimensional motion adaptation (toward and away).

43. Wandell, B.A., Chial, S. & Backus, B.T. Visualization and measurement of the cortical surface. *J. Cogn. Neurosci.* **12**, 739–752 (2000).
44. Nestares, O. & Heeger, D.J. Robust multiresolution alignment of MRI brain volumes. *Magn. Reson. Med.* **43**, 705–715 (2000).
45. Glover, G.H. & Lai, S. Self-navigated spiral fMRI: interleaved versus single-shot. *Magn. Reson. Med.* **39**, 361–368 (1998).
46. Engel, S.A. *et al.* fMRI of human visual cortex. *Nature* **369**, 525 (1994).
47. Sereno, M.I. *et al.* Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* **268**, 889–893 (1995).
48. Larsson, J., Landy, M.S. & Heeger, D.J. Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *J. Neurophysiol.* **95**, 862–881 (2006).
49. Brainard, D.H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
50. Huk, A.C., Ress, D. & Heeger, D.J. Neuronal basis of the motion aftereffect reconsidered. *Neuron* **32**, 161–172 (2001).

doi:10.1038/nn.2343