



Bistable illusory rebound motion: Event-related functional magnetic resonance imaging of perceptual states and switches

P.-J. Hsieh, G.P. Caplovitz, and P.U. Tse*

Department of Psychological and Brain Sciences, Moore Hall, H.B. 6207, Dartmouth College, Hanover, NH 03755, USA

Received 12 October 2005; revised 21 March 2006; accepted 27 March 2006

The neural correlates of a recently discovered visual illusion that we call ‘illusory rebound motion’ (IRM) are described. This illusion is remarkable because motion is perceived in the absence of any net motion energy in the stimulus. When viewing bars alternating between white and black on a gray background, the percept alternates between one of flashing bars (veridical) and the IRM illusion, where the bars appear to shoot back and forth rather like the opening and closing of a zipper. The event-related functional magnetic resonance imaging (fMRI) data reported here reveal that (1) the blood-oxygen-level-dependent (BOLD) signal in the human analog of macaque motion processing area MT (hMT+) increases when there is a perceptual change from “no-IRM” to “see-IRM” and decreases when there is a perceptual change from “see-IRM” to “no-IRM,” although the stimulus remains constant; and (2) the BOLD signal in early retinotopic areas (V1, V2, and V3d) shows switch-related activation whenever there is a perceptual change, regardless whether from IRM to no-IRM or vice versa. We conclude that hMT+ is a neural correlate of this novel illusory motion percept because BOLD signal in hMT+ modulates with the perception of IRM.

© 2006 Elsevier Inc. All rights reserved.

Introduction

We report here the neural correlates of a recently discovered motion illusion called “illusory rebound motion” (IRM; Hsieh et al., 2005). Low-level stimulus characteristics of the effect have been characterized elsewhere, as has the role of attention in this effect (Hsieh and Tse, 2006; Hsieh et al., 2005), but the neural mechanisms underlying this effect remain unknown. The goal of this study is to cast light on the neural mechanisms underlying the perception of IRM.

IRM appears to be a variant of transformational apparent motion (TAM; Tse, in press; Tse and Logothetis, 2002) and illusory line motion (ILM; Hikosaka et al., 1993a,b) but has certain new properties that would not have been predicted by past theories that attempted to account for TAM or ILM. ILM occurs when a bar is instantaneously presented shortly after a cue. The bar will appear to move continuously away from the cue (Fig. 1A). When a bar of a different color instantaneously replaces a bar over which ILM has just occurred, IRM occurs when the second bar seems to shoot back in the opposite direction relative to the previous direction of ILM (Fig. 1B). Additionally, if bars of different colors are presented one after another at a constant stimulus onset asynchrony (SOA) following ILM, IRM can be perceived to occur over every bar with alternating direction, much like the opening and closing of a zipper, even though there is no actual net motion energy in any direction in the stimulus (Fig. 1C). This illusion is called “illusory rebound motion” because illusory motion is perceived to rebound repeatedly. Of the many subjects we have tested, none has ever reported seeing repeated motion in the same direction; Rather, motion is always observed to rebound back and forth.

It has also been shown recently that IRM can be perceived spontaneously without being preceded by ILM or a transient cue and can revert spontaneously to the (veridical) percept of flashing bars (Hsieh and Tse, 2006). For example, when viewing bars flashing between white and black on a gray background (Fig. 2A), the percept is bistable because it alternates between IRM and simply flashing (Fig. 2B). The effect has to be experienced to be fully understood. The reader is encouraged to visit <http://www.freewebs.com/hsieh/> to experience IRM first hand.

Traditional ‘translational’ apparent motion, where an object appears to jump back and forth even though it only appears instantaneously in two different locations, requires two stimuli that are appropriately separated in both temporal and spatial domains (Korte, 1915). However, IRM shows that two stimuli having an identical spatial shape, extent, and position, but differing only in the temporal domain, can still generate a powerful impression of motion even in the absence of net motion energy in any direction. This property of IRM provides a good opportunity to test an inherent confound existing in most previous fMRI studies on

* Corresponding author.

E-mail addresses: Po-Jang.Hsieh@dartmouth.edu (P.-J. Hsieh), Gideon.Caplovitz@dartmouth.edu (G.P. Caplovitz), peter.u.tse@dartmouth.edu (P.U. Tse).

Available online on ScienceDirect (www.sciencedirect.com).

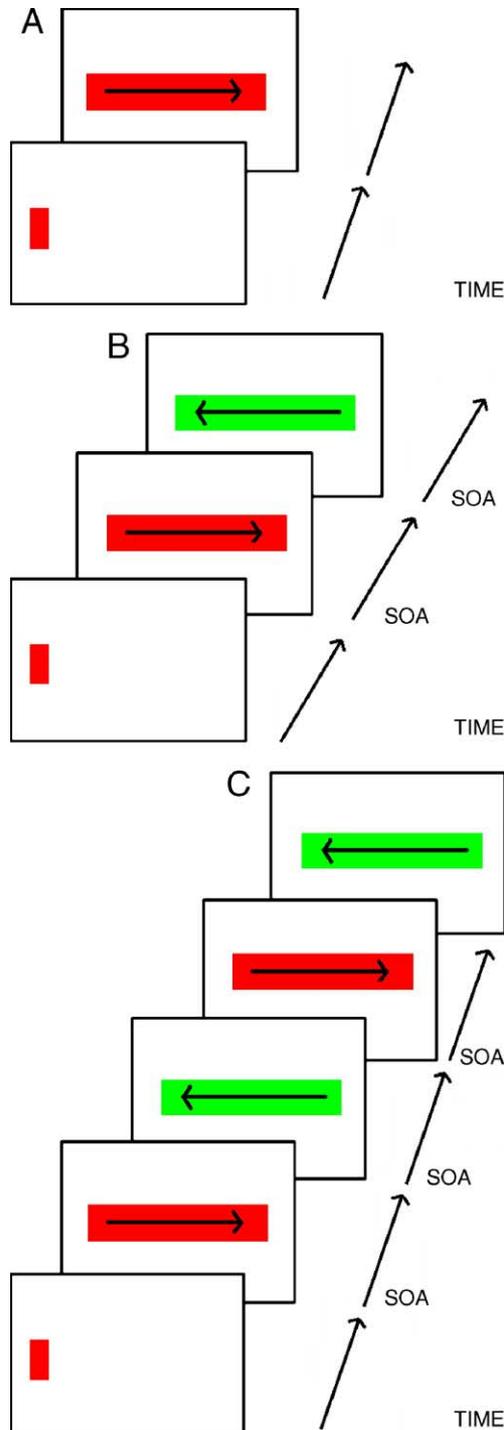


Fig. 1. (A) Illusory line motion: when a horizontal bar is presented shortly after an initial stimulus, the bar is perceived to shoot smoothly away from the initial stimulus. (B) Illusory rebound motion: when a second bar of a different color instantaneously replaces a bar over which ILM has just occurred, observers report that the bar appears to shoot smoothly in the opposite direction. (C) Repeated IRM: if bars of alternating colors are repeatedly presented after an ILM (one after another with a constant SOA), IRM can be perceived to occur over every bar with alternating direction. The arrows on the bars (not present in the actual stimulus) indicate the perceived motion direction. All bars are in fact presented all at once. Any perceived motion is illusory. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

apparent motion. Previous fMRI studies on apparent motion typically relied on comparing an “apparent motion” condition to a “flickering” condition (Goebel et al., 1998; Liu et al., 2004). In the apparent motion condition, two stimuli that are temporally and spatially separated are presented. In the flickering condition, two stimuli that are still spatially separated are presented simultaneously to get rid of the apparent motion. Any difference in BOLD response between these two conditions would presumably be due to the perceptual difference (seeing apparent motion versus not). However, one possible confound that is inherent to this apparent motion paradigm is that a difference in the spatiotemporal domain exists between the stimuli that generate perceived motion and perceived flashing. Therefore, it is difficult to distinguish whether differences in BOLD response during the two perceptual states are due to perceptual differences per se or spatiotemporal differences in the stimuli. By using IRM as a stimulus, this potential confound is eliminated because the spatial–temporal characteristics of the stimulus are the same whether IRM is perceived or flashing is perceived.

To our knowledge, only one study conducted using traditional translational apparent motion does not have the abovementioned confound (Muckli et al., 2002). In this study, two spatially separated squares flashed on and off alternately at a specific rate that generated bistable perceptual switching between apparent motion and flashing. Our stimuli differ in that IRM does not require two spatially separated stimuli and has no net motion energy as the stimulus consists of nothing more than rectangles changing color at a fixed SOA.

It has been hypothesized that high-level motion processing units (hMT+) pool responses from low-level units (V1) through spatial–temporal summation (e.g., Grossberg and Rudd, 1992). Thus, existing models of motion processing require a spatiotemporal offset in luminance or some other feature profile in order for units to indicate the presence of motion in the world. IRM is an important phenomenon because motion is perceived in the absence of a spatial change in the profile of luminance or any other stimulus feature. By using IRM as a stimulus, we can test whether motion processing units can be activated in the absence of any net motion energy. Because motion is perceived in IRM, it is reasonable to predict that such units, presumably in hMT+, can code for motion even in the absence of net motion energy in any stimulus dimension.

In summary, we conducted the present experiment using event-related fMRI in order (1) to better understand the neural mechanisms underlying the perception of IRM by determining brain areas involved in the processing of IRM; (2) to remove the confound from the temporal domain that exists in most (e.g., Goebel et al., 1998; Liu et al., 2004) but not all (Muckli et al., 2002) previous fMRI studies on apparent motion; (3) to test whether temporal alternation alone, in the absence of any spatial offset along any featural dimension, is sufficient to activate higher-level units (hMT+); and (4) to better understand the mechanisms underlying perceptual bistability more generally.

Materials and methods

Participants

Seventeen healthy right-handed volunteers (of both genders between the ages of 18 and 40) were run in the IRM/no-IRM

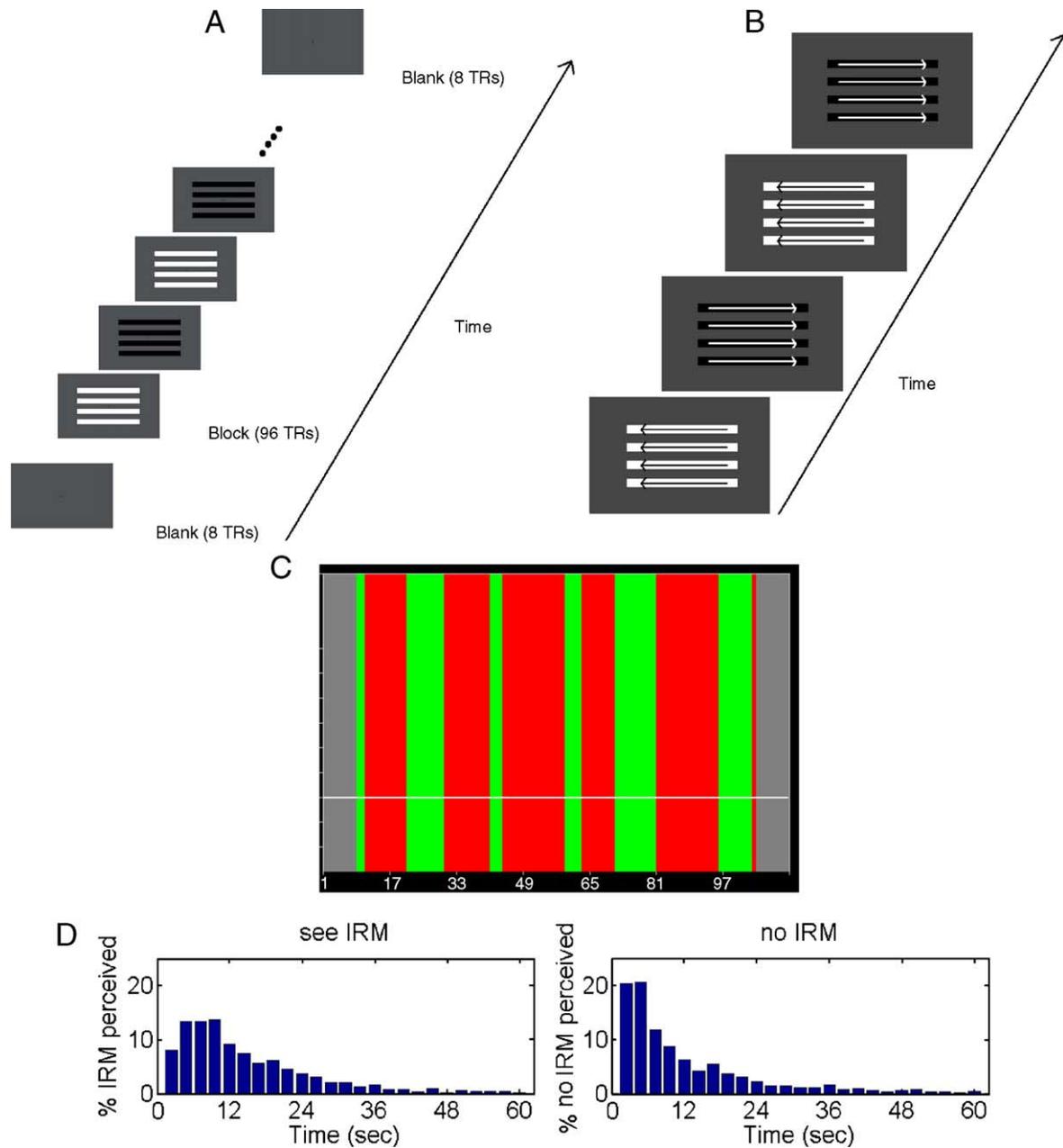


Fig. 2. Stimuli and psychophysics results. (A) An example run of the stimuli. In each run, there was one stimulation block containing 96 volumes ($TR = 2.5$ s) situated between two blank periods of 8 volumes each. In each stimulation block, four horizontal bars, all either black or white, were flashed on an intermediate gray background. Subjects were asked to report their current perceptual state by pressing a button with their right hand when they saw IRM and release the button when they saw flashing. (B) An illustration of the IRM percept. Arrows on the bars indicate the perceived motion direction. In this case, the white bars are perceived to be moving leftward and the black bars are perceived to be moving rightward. (C) An example result of subjects' perceptual states. Red phases correspond to the “see-IRM” states. Green phases correspond to the “no-IRM” states. Gray phases correspond to the blank conditions. Time in TR (1 TR = 2.5 s) is shown on the x axis. (D) Histograms of perceptual durations (s) during fMRI experiments accumulated over 17 subjects are shown for the “see-IRM” and “no-IRM” percepts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

experiment. All gave informed consent within a protocol passed by the Dartmouth committee for the protection of human subjects and its internal review board. All had normal depth perception and normal or corrected-to-normal visual acuity. Subjects were paid twenty dollars per session. All of our stimuli were presented binocularly.

Stimuli and task

In the current study, we used luminance-defined stimuli to induce IRM. When viewing bars flashing between white and black on a gray background (Fig. 2A), the percept of the bars alternates between “flashing bars” and “IRM” (Fig. 2B). During the

“flashing bars” or “no-IRM” percept, the bars are perceived to be simply flashing between black and white. During the “IRM” percept, the bars appear to shoot back and forth with each color change, even though they are in reality just flashing (Fig. 2B).

Stimuli were projected from a digital data projector (refresh rate 60 Hz) onto a Plexiglas screen outside the bore of the magnet and viewed via a tangent mirror inside the magnet that permitted a maximum of $22^\circ \times 16^\circ$ visible area. The projected image was smaller than this and subtended approximately $17^\circ \times 12^\circ$. The fixation spot was a square subtending 0.2° of visual angle, which changed color every 3.2 s on average. The background was always gray (52.4 cd/m^2), even during the blank period. The stimuli consisted of four bars, separated vertically by 0.67° . Each of the four bars subtended 0.67° in height and 8° in width and was centered 0.67° above, 0.67° below, 2.01° above, and 2.01° below the fixation point respectively (spaced 0.67° apart vertically). The retinotopic mapping scans were performed using the same projection system covering the entire $17^\circ \times 12^\circ$ projected image. As a result, all the stimuli were restricted within the retinotopic mapping stimulus areas. The four bars alternated between white (102.8 cd/m^2) and black (1.7 cd/m^2) every 500 ms (Fig. 2A). Each subject carried out an average of 9.10 ± 0.25 (range from 7 to 10) runs in the scanner. In each run, there was one stimulation block containing 96 volumes (TR = 2.5 s) situated between two blank periods of 8 volumes each (Fig. 2A). In each stimulation block, four horizontal bars, all either black or white, were flashed on an intermediate gray background. Subjects were asked to report their current perceptual state by pressing a button with their right hand when they saw IRM (Fig. 2C, red periods) and releasing the button when they saw flashing (Fig. 2C, green periods).

Fixation task

Eye movements, wakefulness, and attention to the fovea were controlled for by requiring subjects to perform a reaction time task in which the subject had to respond, within 2500 ms (via button press), to a randomly occurring change in fixation point color using a button press. The fixation point changed color from blue/yellow (yellow inner square surrounded by a blue border) to red/green (green inner square surrounded by a red border) on average every 3.2 s. This color change occurred an equal number of times during each block. Subjects were instructed to pay attention to both the fixation task and the IRM judgment task simultaneously. The fixation task (pressing a second button with the left hand whenever the fixation point changed color) was impossible to perform in the absence of foveation of the fixation spot.

fMRI acquisition

Continuous whole-brain BOLD signal was acquired at the Dartmouth Brain Imaging Center on a GE 1.5 T Signa scanner using a standard head coil. Standard T2*-weighted echoplanar functional images were collected using 25 slices (4.5 mm thickness and 3.75-by-3.75 mm in-plane voxel resolution, inter-slice distance 1 mm, TR = 2500 ms, flip angle = 90° , field-of-view = $240 \times 240 \times 256 \text{ mm}$, descending interleaved slice acquisition, matrix size = 64×64) oriented approximately along the anterior-commissure posterior-commissure plane. These slices were sufficient to encompass the entire brain of each subject. Cushions were used to minimize head motion. T1-weighted

anatomical images were acquired using a high-resolution 3D spoiled gradient recovery sequence (SPGR; 124 sagittal slices, TE = 6 ms, TR = 25 ms, flip angle = 25° , $1 \times 1 \times 1.2 \text{ mm}$ voxels) as well as a T1-weighted coplanar anatomical image with the same slice orientation as the EPI data which was used for coregistration of the functional data to the high resolution scan.

fMRI data analysis

Data were analyzed offline using BRAIN VOYAGER (BV) 4.9.6 and MATLAB software developed in house. Effects of small head movements were removed using BV's motion correction algorithms. Functional data were not smoothed in the space domain. Low frequency oscillations in the time course with periods greater than or equal to 37 TRs (3 cycles per run) were removed.

Regions of interest and time course data analysis

Regions of interest (ROIs) in the current study include individually specified retinotopic areas V1v, V1d, V2v, V2d, V3v, V3d, V3A/B, and V4v (Fig. 3A). In all, four sets of ROIs were created for the retinotopic regions (Fig. 3B). One set was localized to correspond to the 4.6 visual degrees surrounding fixation where the flashing bars were presented (<4.6 visual degrees), another set was localized to the periphery where no stimulus was present (>4.6 visual degrees), and a third set used the entire retinotopic region (i.e. the union of the previous two). As a control, to guard against the possibility of contamination from voxels in the central region, a fourth peripheral ROI was defined for each retinotopic area as the region that responded maximally to retinotopic mapping stimuli beyond 7.8° visual angle from the foveal representation. Comparisons between these sets of ROIs can discriminate local-stimulus-driven activity from more global non-stimulus-driven activity.

In addition to these areas, localizer scans were performed to isolate the lateral occipital complex (LOC) and hMT+. hMT+ is an area of the brain known to process motion (Castelo-Branco et al., 2002; Goebel et al., 1998; Liu et al., 2004; Muckli et al., 2002, 2005). LOC on the other hand is classically thought to process object shape. hMT+ and LOC are of particular interest in this study because (1) the perceived direction of IRM is influenced by multiple stimulus factors, including contour relationships, texture, and motion energy (Hsieh and Tse, 2006), and (2) IRM may be related to TAM, in which the perceived moving direction relies essentially on figural parsing (a comparison of contour and surface relationships among successive scenes) (Tse, in press; Tse and Logothetis, 2002; Tse et al., 1998).

Event-related average time courses were computed within ROIs. Time course segments representing the same perceptual states were averaged across runs. In each subject, the time course data for each voxel within an ROI was averaged together to create a single mean waveform per ROI. The mean BOLD signal in each ROI was averaged by condition using the subjects' button presses as a trigger, indicated by the '0' point on the horizontal axis of Figs. 4, 5, and 6. Note that numbers in corresponding visual areas of the two hemispheres were averaged within subject and then averaged across subjects. Standard errors of the mean were computed across subjects.

Variable durations of perceptual states tend to blur the later part of the time course of individual responses. To avoid this mixed-state contamination, we only show the signals corresponding to the

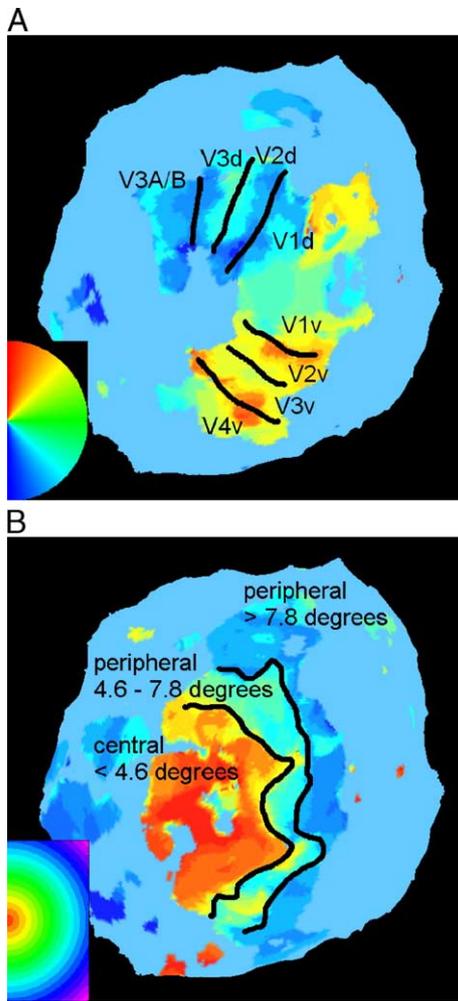


Fig. 3. Retinotopy. (A) A typical retinotopic map of the flattened left hemisphere occipital pole for one subject is shown with the approximate borders between the retinotopic areas specified in black. Retinotopic area masks were individually specified for each hemisphere of each subject. Blue here represents the lower vertical meridian, cyan/green the horizontal meridian, and red the vertical meridian. (B) A typical retinotopic map of the flattened left hemisphere occipital pole for one subject is shown with the approximate borders specified in black between the central (<4.6 visual degrees), middle (4.6–7.8 visual degrees), and peripheral (>7.8 visual degrees) areas. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

first three TRs before and after a perceptual state change, which was based on the distributions of the durations for each perceptual state shown in the Results section.

Two-tailed paired *t* tests comparing the BOLD signal at TR = -1 to that at TRs 1, 2, and 3, respectively, were carried out to test whether there is a BOLD signal rise upon a perceptual switch. We chose to use this statistic because TR = -1 represents the last TR before a perceptual switch, and TRs = 1, 2 and 3 represent the first three TRs after a perceptual switch. A comparison between them can presumably reveal if there is a change in BOLD signal before and after a perceptual shift.

Retinotopy stimulus

Retinotopy was carried out on all subjects ($n = 17$) run in the experiment using standard phase-encoding techniques (Serenio et

al., 1995; 4.5 mm thickness and 3.75-by-3.75 mm in-plane voxel resolution, inter-slice distance 1 mm, TR = 1600 ms, flip angle = 90°, field-of-view = 240 × 240 × 256 mm, interleaved slice acquisition, matrix size = 64 × 64; 16 slices oriented along the calcarine sulcus) with the modification that two wedges of an 8-Hz flicker black and white polar checkerboard grating were bilaterally opposite (like a bowtie) to enhance signal to noise (Slotnick and Yantis, 2003). Wedges occupied a given location for 2 TRs (3.2 s) before moving to the adjacent location in a clockwise fashion. Each wedge subtended 18° of 360°. In order to bring spins to baseline, 9.6 s (6 TRs of dummy scans) was discarded before each run. 168 volumes were collected on each run. A minimum of 7 wedge runs were collected for each subject and then averaged to minimize noise before retinotopic data analysis in BV 4.9.6. At least three runs were collected per subject using expanding 8-Hz flickering concentric rings that each spanned approximately 0.6° of visual angle in ring width. Each ring was updated after one TR (1.6 s) after which it was replaced by its outward neighbor, except that the outermost ring was replaced by the innermost ring, whereupon the cycle was repeated. Retinotopic areas (V1d, V1v, V2d, V2v, V3d, V3v, V4v, and V3A/B) were defined as masks on the basis of standard criteria (Serenio et al., 1995), assuming a contralateral quadrant representation for V1d, V1v, V2d, V2v, V3d, and V3v and a contralateral hemifield representation for V4v/VO and V3A/B (Tootell et al., 1997). V4v and the hemifield representation just anterior to it, called VO (Brewer et al., 2004), were combined into a common mask because the border between these regions was not distinct in all subjects, as was true for the combination of V3A and V3B into a common V3A/B mask. A typical retinotopic map of the flattened left hemisphere occipital pole for one subject is shown in Fig. 3A.

Subregion ROIs within the retinotopic areas were further created. Because the stimulus was centrally located and subtended 4.7° in height and 8° in width, we identified the central subregion of each retinotopic area that was activated by the ring stimulus within a radius of 4.6 visual degrees from fixation. We also identified the peripheral subregion of each retinotopic area that was activated by the ring stimulus within a radius greater than 4.6 visual degrees from fixation by subtracting voxels comprising the above-defined subregion ROI from the ROI corresponding to each complete retinotopic area (Fig. 3B). Lastly, we also defined a conservative ROI corresponding to the peripheral subregion of each retinotopic area that was activated by the ring stimulus beyond a radius of 7.8 visual degrees from fixation.

Individual hMT+ mask localization

The human analog of macaque motion processing area MT has been called V5 or human hMT+. Left and right hMT+ were localized in 11 of the 17 subjects using a localizer scan comprised of three to six runs of 3 min each. The hMT+ localizer stimuli consisted of a grid of 3 × 3 subgrids of white solid squares on a black background. The length and height of each square were approximately 1° by 1°. This was constructed by eliminating the zeroth, ±fourth, and ±eighth rows and columns from a regular grid of squares. Square centers were separated by approximately 3°. In baseline blocks, the grid remained stationary for a 20-s epoch followed by an epoch where the grid rotated clockwise around its center at a speed of 270° per second. Each run contained nine epochs of alternating motion and non-motion stimulation. As in the main experiment, subjects carried out a simple fixation task, pressing a button in the right hand any time the fixation point

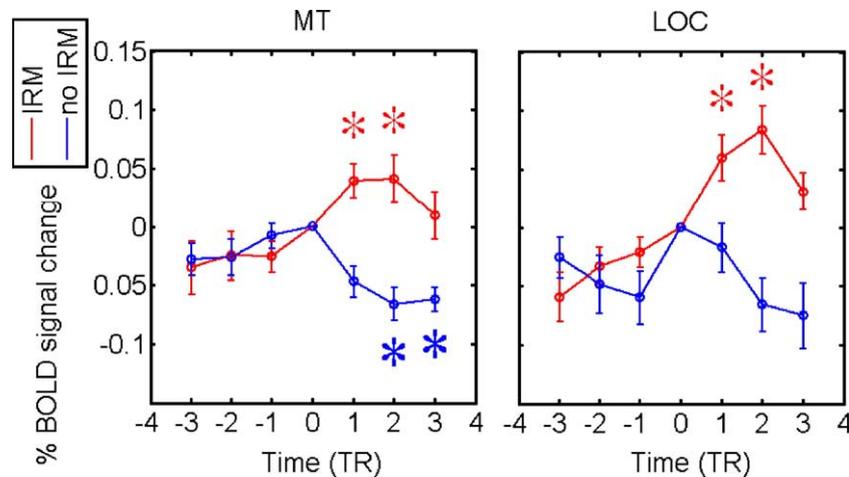


Fig. 4. BOLD signal time courses within areas hMT+ and LOC ($n = 11$). The BOLD signal change was averaged across voxels within subjects' ROIs and across hemispheres. In hMT+, the BOLD signal increases significantly when there is a perceptual change from "no-IRM" to "see-IRM" and decreases significantly when there is a perceptual change from "see-IRM" to "no-IRM." In LOC, the BOLD signal increases significantly when there is a perceptual change from "no-IRM" to "see-IRM" but fails to reach significance when the opposite perceptual state switch occurs, relative to the BOLD signal at TR = -1. Note, in Figs. 4, 5, and 6, the BOLD signal change was averaged across voxels within subjects' ROIs and across hemispheres and compared to the TR = -1 position, which corresponds to the last volume before a subject reported a perceptual switch. In the two-tailed paired t test (comparing the BOLD signal at TR = -1 to that at TR = 1, that at TR = 2, and that at TR = 3), those areas that reach significance are marked as "*" ($P < 0.05$).

changed color. hMT+ was localized as activity in the motion > non-motion GLM contrast that survived at least the threshold $P < 0.0001$ corrected (fixed effects). In addition, activation had to occupy the inferior occipital gyrus or inferior temporal sulcus in order to be localized as hMT+.

Individual LOC mask localization

An individual LOC mask was also determined individually following standard procedures (Kourtzi and Kanwisher, 2000a) for the same 11 subjects for whom individual hMT+ masks were made. Object images ($7^\circ \times 7^\circ$) were placed on a white background and were embedded within a black grid. Their centroid position was updated randomly every TR within a 1° radius of the fixation point in order to prevent perceptual fading. Control images were comprised of the same images scrambled within the same grid. The left and right hemisphere LOC masks were created from the fixed effects GLM analysis contrast of unscrambled objects > scrambled objects for each of the 11 subjects, at a P level of 0.001 uncorrected. In many subjects, there is an anterior portion of the LOC located in the middle fusiform gyrus and a posterior portion located just inferior to hMT+ that is activated by this contrast. The present LOC masks were selected as the posterior region since the two subregions were not abutting in any subject and could well comprise areas with different functionalities.

The separation of hMT+ and LOC

To distinguish LOC and hMT+, any overlap region was eliminated from the LOC mask and the hMT+ mask in all subjects. An "MT-LOC" mask was created for each hemisphere of each subject by removing any voxels shared by the LOC and hMT+ mask from the hMT+ mask. These 'subtraction' masks were used because, depending on the threshold at which hMT+ and LOC masks are specified, there can be overlap voxels shared between these masks. In order to control for the possibility that any BOLD signal differences between perceptual states (i.e. see-IRM vs. see

flashing) seen here is driven solely by this overlap region, the overlap region was removed from the LOC and hMT+ masks for the eleven subjects for whom individual LOC and hMT+ masks had been determined. The mean (\pm SE) Talairach coordinates of MT-LOC mask in the left hemisphere were $x = -41.2 \pm 1.2$, $y = -68.5 \pm 1.6$, and $z = 3.6 \pm 1.4$, and in the right hemisphere: $x = 45.9 \pm 1.0$, $y = -65.6 \pm 1.0$, and $z = 2.4 \pm 1.4$.

An "LOC-MT" mask was also created for each hemisphere of each subject by removing any voxels shared by the LOC and hMT+ mask from the LOC mask. The mean left LOC mask location in Talairach coordinates was $x = -42.7 \pm 1.2$, $y = -73.2 \pm 1.6$, and $z = -5.8 \pm 2.1$, and the mean right hemisphere LOC mask location was $x = 45.2 \pm 1.5$, $y = -70.6 \pm 1.4$, and $z = -9.6 \pm 1.5$.¹

Note that we could not identify retinotopic subregions within hMT+ and LOC, probably because these areas are only marginally retinotopic. Therefore, unlike the retinotopic areas, the time courses were averaged by using the whole hMT+ and LOC masks.

Results

Fixation task

The average button press reaction time, for runs that were included in the analysis, occurred within 1 TR after the change in the color of the fixation point. Button presses followed $84.795\% \pm 2.34\%$ of fixation point changes within 1 TR; and the average reaction time was 768.03 ± 32.77 ms. Based on these results, we conclude that the subjects were maintaining fixation on a consistent basis throughout the experiment. No motor areas were found to be activated differentially between see-IRM and no-IRM

¹ By way of comparison, the mean location reported by Kourtzi et al. (2003) (left hemisphere Talairach coordinates: $-41.9, -64.8, -2.7$; right hemisphere: $39.1, -65.6, -12.0$), found using an LOC localizer in each of ten subjects, fits well within the bilateral activations found in the present study.

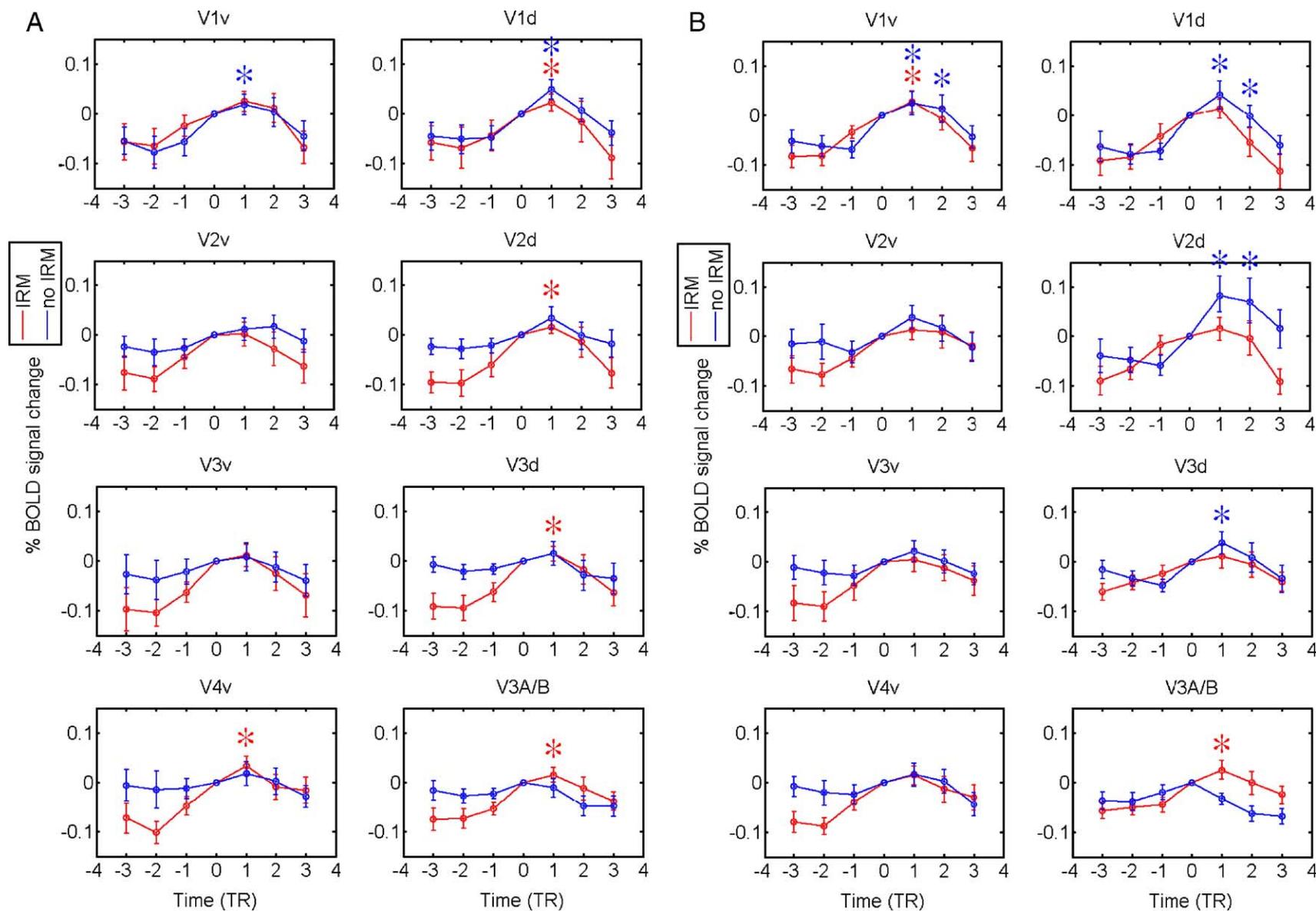


Fig. 5. BOLD signal time courses within central sub-ROIs (corresponding to the location of the stimulus) and peripheral sub-ROIs across all subjects who had been retinotopically mapped ($n = 17$). (A) In retinotopic cortex within a radius of 4.6 visual degrees from the foveal representation, the BOLD signal tends to rise when there is a transition into either perceptual state. Though the t test results only reach significance in area V1d when there is a transition into either perceptual state, the basic pattern of the BOLD signal in other areas is similar to that in V1d. (B) In the peripheral retinotopic areas (>4.6 visual degrees) that do not receive direct bottom-up input from the location of the stimulus, the basic pattern of BOLD modulation is similar to that in the central retinotopic areas. The BOLD signal tends to rise when there is a transition into either perceptual state. In the two-tailed paired t test (comparing the BOLD signal at TR = -1 to that at TR = 1, that at TR = 2, and that at TR = 3), those areas that reach significance are marked as “*” ($P < 0.05$).

states, corroborating that the motor task was equivalent across both conditions. When analyzing color changes of the fixation point as events, the results show that the fixation point color switches do not modulate the BOLD signal in any retinotopic areas (data not shown). When analyzing button presses to color changes of the fixation point as events, the results show that motor responses do not modulate the BOLD signal in retinotopic areas (data not shown).

Behavioral results

Histograms of perceptual durations measured in the scanner accumulated over all 17 subjects are shown for each of the two different perceptual states (Fig. 2D). The mean perceptual duration for the “no-IRM” state, in which subjects did not see IRM, was

14.23 s (median = 7.5 s; mode = 5 s). The mean perceptual duration for the “see-IRM” state, in which subjects saw IRM, was 15.53 s (median = 12.5 s; mode = 12.5 s). The distribution of the perceptual durations was skewed toward the right and can be approximated by a gamma distribution, as has been observed in other examples of perceptual bistability (Leopold and Logothetis, 1999; Meng and Tong, 2004).

Numbers in parentheses indicate the percentage that a subject was in the ‘see-IRM’ state at a given TR around a perceptual switch from the ‘no-IRM’ state to the ‘see-IRM’ state: TR = -4 (51%); TR = -3 (40%); TR = -2 (20%); TR = -1 (0%); TR = 1 (100%); TR = 2 (92%); TR = 3 (79%); TR = 4 (66%). Similarly, the percentage that a subject was in the ‘no-IRM’ state at a given TR around a perceptual switch from the ‘see-IRM’ state to the ‘no-IRM’ state is: TR = -4 (34%); TR = -3 (21%); TR = -2 (8%);

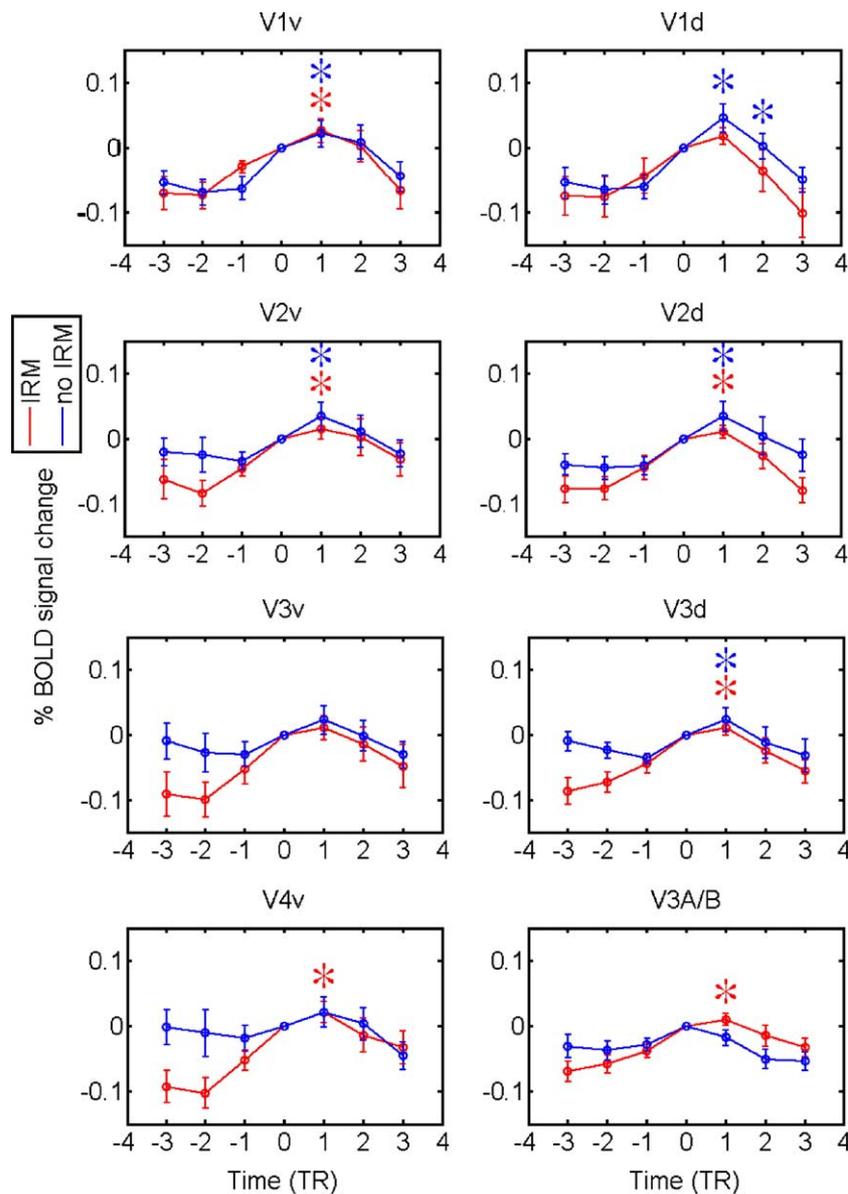


Fig. 6. BOLD signal time courses averaged within the whole retinotopic areas across all subjects who had been retinotopically mapped ($n = 17$). In V1v, V2v, V2d, and V3d, the BOLD signal rises when there is a transition into either perceptual state. This basic pattern of the BOLD signal exists in all retinotopic areas, even those that did not reach significance. In the two-tailed paired t test (comparing the BOLD signal at TR = -1 to that at TR = 1, that at TR = 2, and that at TR = 3), those areas that reach significance are marked as “***” ($P < 0.05$).

TR = 1 (0%); TR = 1 (100%); TR = 2 (80%); TR = 3 (60%); TR = 4 (49%). Because the percentage of perceptual state contributing to the BOLD signal at TR = -4 and TR = 4 was below 50% for one of the conditions, time course data were shown between TR = -3 and TR = 3.

Event-related time course data

In Figs. 4, 5, and 6, the time courses were plotted in terms of percent signal change relative to the baseline defined by the level of the BOLD signal at the time of the transition as indicated by the button press, set to the value zero here. In the two-tailed paired *t* test (comparing BOLD signal at TR = -1 to that at TRs = 1, 2, and 3), those areas that reach significance are marked as “*” ($P < 0.05$).

Fig. 4 shows the averaged BOLD signals in hMT+ and LOC across the eleven subjects who carried out the hMT+ and LOC mappings. In hMT+, the BOLD signal increases significantly after a perceptual change from “no-IRM” to “see-IRM” and decreases significantly after a perceptual change from “see-IRM” to “no-IRM.” In LOC, the BOLD signal increases significantly after a perceptual change from “no-IRM” to “see-IRM”.

Fig. 5A shows the averaged BOLD signals in the central retinotopic areas (<4.6 visual degrees, corresponding to the topographic location of the stimulus in cortex) across all subjects who had been retinotopically mapped ($n = 17$; note that, for the eleven subjects who had hMT+ and LOC ROIs, the pattern of BOLD modulation in retinotopic areas was similar to that of all seventeen subjects; we therefore show the data from all seventeen subjects). In area V1d (Fig. 5A), signal intensity rises after perceptual transitions not only from “no-IRM” to “see-IRM” (blue) but also from “see-IRM” to “no-IRM” (red). A similar, but weaker pattern of BOLD signal response could be observed in other retinotopic areas.

This switch-related modulation of the BOLD signal occurs in the same manner regardless of which perceptual state the observer is entering, suggesting that this is a response to the transition between states. If the BOLD modulation in retinotopic areas is a response to the transition between perceptual states but not to the stimulus directly, then all the voxels in retinotopic areas should have a similar response. Averaging across voxels outside those that received bottom-up stimulus activation would be expected to show the same pattern as seen in Fig. 5A if this hypothesis is correct. However, if the response is primarily due to the stimulus, no BOLD signal modulation should be found upon a perceptual switch in cortex that does not receive bottom-up stimulus input. Fig. 5B shows the averaged BOLD signals in the peripheral retinotopic (>4.6 visual degrees) areas across the seventeen subjects. In general, a similar switch-related modulation as that apparent in Fig. 5A can be observed in every region.

In order to guard against the possibility that there were stimulus-responsive voxels remaining in the peripherally defined ROIs, a further analysis was carried out in which peripheral ROIs were defined as regions that responded maximally to retinotopic mapping stimuli beyond 7.8°. Results (data not shown) show switch-related BOLD signal activity even when peripheral masks were defined extra conservatively to mitigate against any possible contamination from stimulus-driven voxels. From this, we can conclude that BOLD signal modulation in retinotopic areas upon a perceptual state change is due to perceptual switching per se, and not to the stimulus.

Because BOLD responses within central and peripheral retinotopic areas were essentially the same, we combined data from these areas. By averaging across complete retinotopic ROIs, the response to perceptual state switches in retinotopic areas should become more apparent as the signal to noise ratio improves. The result in Fig. 6 shows that, when averaging across all voxels within respective retinotopic areas, more areas (V1v, V2v, V2d, and V3d) reach significance, again implying that the BOLD signal modulation in retinotopic cortex is a response to the transition between states rather than a direct response to anything about the stimulus. Though the effect is weaker in V3v and V4v, the basic pattern of the BOLD signal is similar to that found in areas that reach significance.

Discussion

Several studies examining the neural basis of apparent motion (Goebel et al., 1998; Liu et al., 2004) have found increased activation in hMT+ when apparent motion was seen relative to a no-motion control. In these studies, the control stimulus was not identical in temporal profile with the motion stimulus. It is therefore possible that differences in activation in these studies were not due to the motion percept per se but to stimulus differences. In the present study, the stimulus was identical regardless of perceptual state, eliminating any concerns of low-level or stimulus-driven confounds. Thus, any area where the BOLD signal tracks perceptual state most likely reflects processing related to perceptual state rather than low-level stimulus differences or contributions of motion energy per se. Our findings are consistent with past fMRI findings that used translational apparent motion as a probe (Goebel et al., 1998; Liu et al., 2004), showing a greater activation for apparent motion than flicker in hMT+. Data show that the BOLD signal in hMT+ increases significantly after a perceptual change from “no-IRM” to “see-IRM” and decreases significantly after a perceptual change from “see-IRM” to “no-IRM”. We can therefore conclude that hMT+ embodies neural correlates of perceptual state in IRM. Moreover, in contrast with Muckli et al. (2002) whose stimulus was also identical regardless of perceptual state, our IRM stimulus differs in that there is no spatial change and no net motion energy. One of our key findings is therefore that temporal alternation alone, in the absence of any spatiotemporal offset (i.e. motion energy) along any featural dimension, is sufficient to induce apparent motion and activate higher-level units (hMT+) when illusory motion is perceived.

Another notable difference is that, unlike the translational apparent motion stimulus used previously (Goebel et al., 1998; Liu et al., 2004; Muckli et al., 2002), the IRM stimulus used here invokes a BOLD modulation in the LOC similar to that in hMT+, in that there is an increase in BOLD signal when there is a perceptual switch to the ‘see-IRM’ state from the opposite state. The LOC has been implicated in the processing of form (Malach et al., 1995; Kanwisher et al., 1996; Grill-Spector et al., 2001; Haxby et al., 2001). Evidence is emerging that LOC processes global 3D object shape rather than local 2D shape features (Grill-Spector et al., 1998, 1999; Malach et al., 1998; Kourtzi and Kanwisher, 2000b, 2001; Avidan et al., 2002; Mendola et al., 1999; Gilaie-Dotan et al., 2002; Moore and Engel, 2001; Kourtzi et al., 2003) and may even mediate aspects of object recognition (Grill-Spector et al., 2000). The present fMRI data suggest that the LOC plays a greater role in the IRM perceptual state than the no-

IRM perceptual state. Why an area that is presumably involved in the processing of form should exhibit neural activity that is correlated with the perception of IRM is unclear. The most likely possibility is that IRM is a variety of TAM, which it phenomenally resembles. Tse (in press) has recently shown that the LOC is more active during the presentation of TAM than during a control stimulus. The LOC may play an important role in TAM because form cues are used to segment figures, and figure to figure matching is thought to underlie TAM. Another possibility is that the activation of LOC might be necessary in IRM and TAM for maintaining an object representation so that a continuous object transformation can be perceived. In support of these possibilities, a number of recent papers have come to the conclusion that there is potential anatomical and functional overlap between hMT+ and the LOC (Bar et al., 2001; Kourtzi et al., 2003; Ferber et al., 2003; Murray et al., 2003; Zhuo et al., 2003; Liu et al., 2004; Moutoussis et al., 2005), which is supported by behavioral data as well (Liu and Cooper, 2003; Stone, 1999). Future work will have to determine whether form, figural, or global analyses are involved in generating the IRM percept to a degree that does not occur during the percept of flashing.

In contrast to hMT+ and LOC, our results show that the BOLD signal in early retinotopic areas increases when there is a transition into either perceptual state. Similar switch-related activity was observed previously by Muckli et al. (2002) in individual subjects when using bistable translational apparent motion. However, no such activity was found in their group GLM analysis (Muckli et al., 2002; Sterzer et al., 2002, 2003), which is probably due to morphological differences in the retinotopic areas between subjects that will result in signal cancellation in the group GLM analysis. In the current study, our results from subject-specific retinotopic areas show that all retinotopic areas appear to modulate upon a perceptual switch, not just the sub-ROIs that receive bottom-up activation by the stimulus. We show that early retinotopic areas respond to switches in perceptual state to or from perceived motion; they just do not distinguish between perceptual states. These data suggest that early retinotopic areas are not correlated with the conscious perceptual state during IRM.

This conclusion contrasts with a recent finding that the line motion illusion (Jancke et al., 2004) induces neuronal activity in area 18 of the anesthetized cat that is similar to that seen in response to real motion. However, even if species differences do not pose a problem, it cannot be concluded that the neural correlates of conscious perceptual state of line motion occur in early retinotopic areas, just because the global pattern of blood signal observed using optical imaging is similar to that seen when using real motion as a probe.

More puzzling, in light of the present findings, are recent reports by Muckli et al. (2005) and Tse (in press). In Muckli et al. (2005), increased activity was found in retinotopic regions of V1 along the perceived path of translational apparent motion. In Tse (in press), transformational apparent motion induces greater BOLD activity in all retinotopic areas than a flashing control. One possible explanation is that the translational apparent motion stimulus used by Muckli et al. and the TAM stimulus used by Tse both contained motion energy, whereas the current stimulus contained none. Another possibility is that IRM, while phenomenologically similar to translational apparent motion and TAM, might be realized through different neuronal mechanisms.

It is interesting to note that, upon a perceptual switch from “see-IRM” to “no-IRM”, the BOLD signal in hMT+, LOC, and

V3A displays a transient peak in amplitude at $TR = 0$. This transient peak is phase-shifted relative to all the early retinotopic areas (V1v, V1d, V2v, V2d, and V3d), which demonstrate a transient peak at $TR = 1$. This difference suggests that there might be a temporal difference in signal processing or some other mechanism that sets V3A/B, hMT+, and LOC apart from the early visual areas. More importantly, because the same pattern of BOLD response occurred in peripheral early retinotopic cortex, that did not receive bottom-up input from the stimulus, as in central early retinotopic cortex that did, we hypothesize that the BOLD response to perceptual switches is driven by some factor other than the stimulus itself. While only speculation at this point, it is possible that the switch-related modulations in early retinotopic areas that we observe are due to a feedback mechanism that is associated with the switching process from one perceptual state to another. For example, this mechanism may be associated with higher level processing correlated with perceptual switching. Such a feedback signal might arise because of attentional enhancement following a perceptual change (Liu et al., 2005). Another possibility is that BOLD signal modulation seen in V1 and other retinotopic areas arises from top-down signals that are not attentional in nature. For example, feedback to V1 has been hypothesized to be related to feature integration and conjunction binding (Treisman and Gelade, 1980) and the generation of visual awareness (Super et al., 2001; Bullier, 2001; Tong, 2003; Pollen, 2003). The large receptive fields of higher cortical areas may also feed back to V1 and other early retinotopic areas to gain the spatial precision that neurons in higher areas lack (Ahissar and Hochstein, 2004; Hochstein and Ahissar, 2002). Another possibility is that BOLD signal changes that we observe are driven by microsaccades (or eyeblinks) that trigger or are triggered by perceptual switches or states, which would make BOLD signal modulations observed in early retinotopic areas in essence artifactual. Unfortunately, to date, no group, to our knowledge, has succeeded in determining the BOLD signal correlates of microsaccades. Until this is accomplished, microsaccades remain a potential confound in all those experiments where microsaccades could conceivably vary with the conditions of interest. Future experiments are therefore required to understand the mechanisms that underlie the response to perceptual state switches, but not perceptual state, in early visual areas.

Conclusion

The present data show that hMT+ contains neural correlates of the perception of illusory rebound motion because BOLD signal in this area correlates with perceptual state (IRM or no-IRM) even when the stimulus remains unchanged. Neural activity in hMT+ increases when IRM is perceived, and decreases when IRM is not perceived, even though the stimulus does not change in any way. Our data also show that activity in V1 and other retinotopic areas is correlated with switching from one perceptual state to another, independent of the percepts themselves. The present data place useful constraints upon future theories of the neural mechanisms underlying illusory rebound motion, and related effects, such as translational and transformational apparent motion. More generally, our data show that there are percepts whose neural correlates do not lie in early retinotopic visual areas, but rather which lie in at least hMT+ and perhaps also in LOC or elsewhere. Since there is no spatial change and no net motion energy in the IRM stimulus, our key finding is that temporal alternation alone, in the absence of

any spatiotemporal offset along any featural dimension, is sufficient to induce apparent motion and activate higher-level units (hMT+) when illusory motion is perceived. That is, neural activity in hMT+ covaries with perceptual state in the absence of any net motion energy in the stimulus. Because hMT+ tracks perceptual state, the neural correlates of perceived motion, at least for IRM, appear to lie in at least hMT+.

References

- Ahissar, M., Hochstein, S., 2004. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464.
- Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., Malach, R., 2002. Contrast sensitivity in human visual areas and its relationship to object recognition. *J. Neurophysiol.* 87, 3102–3116.
- Bar, M., Tootell, R.B., Schacter, D.L., Greve, D.N., Fischl, B., Mendola, J.D., Rosen, B.R., Dale, A.M., 2001. Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29, 529–535.
- Brewer, A.A., Liu, J., Wade, A.R., Wandell, B.A., 2004. Human ventral occipitotemporal cortex contains several visual field maps with differential stimulus selectivity. *Soc. Neurosci. Meet. (abstract 300.23)*.
- Bullier, J., 2001. Integrated model of visual processing. *Brain Res. Rev.* 36, 96–107.
- Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., Goebel, E., 2002. Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc. Natl. Acad. Sci.* 99, 13914–13919.
- Ferber, S., Humphrey, G.K., Vilis, T., 2003. The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cereb. Cortex* 13, 716–721.
- Gilaie-Dotan, S., Ullman, S., Kushnir, T., Malach, R., 2002. Shape-selective stereo processing in human object-related visual areas. *Hum. Brain Mapp.* 15, 67–79.
- Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., Singer, W., 1998. The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* 10, 1563–1573.
- Grill-Spector, K., Kushnir, T., Edelman, S., Itzhak, Y., Malach, R., 1998. Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21, 191–202.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203.
- Grill-Spector, K., Kushnir, T., Hendler, T., Malach, R., 2000. The dynamics of object-selective activation correlate with recognition performance in humans. *Nat. Neurosci.* 3, 837–843.
- Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. *Vision Res.* 41, 1409–1422.
- Grossberg, S., Rudd, M.E., 1992. Cortical dynamics of visual motion perception complex and its role in object recognition. *Vision Res.* 41, 1409–1422.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Hikosaka, O., Miyauchi, S., Shimojo, S., 1993. Focal visual attention produces illusory temporal order and motion sensation. *Vision Res.* 33, 1219–1240.
- Hikosaka, O., Miyauchi, S., Shimojo, S., 1993. Voluntary and stimulus-induced attention detected as motion sensation. *Perception* 22, 517–526.
- Hochstein, S., Ahissar, M., 2002. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.
- Hsieh, P.-J., Tse, P.U., 2006. Stimulus factors affecting illusory rebound motion. *Vision Res* 46, 1924–1933.
- Hsieh, P.-J., Caplovitz, G.P., Tse, P.U., 2005. Illusory rebound motion and the motion continuity heuristic. *Vision Res.* 45, 2972–2985.
- Jancke, D., Chavane, F., Naaman, S., Grinvald, A., 2004. Imaging cortical correlates of illusion in early visual cortex. *Nature* 428, 423–426.
- Kanwisher, N., Chun, M.M., McDermott, J., Ledden, P.J., 1996. Functional imaging of human visual recognition. *Brain Res. Cogn. Brain Res.* 5, 55–67.
- Korte, A., 1915. Kinematoskopische untersuchungen. *Z. Psychol.* 72, 193–206.
- Kourtzi, Z., Kanwisher, N., 2000. Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.* 12, 48–55.
- Kourtzi, Z., Kanwisher, N., 2000. Cortical regions involved in perceiving object shape. *J. Neurosci.* 20, 3310–3318.
- Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* 24, 1506–1509.
- Kourtzi, Z., Erb, M., Grodd, W., Bulthoff, H.H., 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb. Cortex* 13, 911–920.
- Leopold, D.A., Logothetis, N.K., 1999. Multistable phenomena: changing views in perception. *Trends Cogn. Sci.* 3, 254–264.
- Liu, T., Cooper, L.A., 2003. Explicit and implicit memory for rotating objects. *J. Exp. Psychol., Learn. Mem. Cogn.* 29, 554–562.
- Liu, T., Slotnick, S.D., Tantis, S., 2004. Human MT+ mediates perceptual filling-in during apparent motion. *NeuroImage* 21, 1772–1780.
- Liu, T., Pestilli, F., Carrasco, M., 2005. Transient attention enhances perceptual performance and fMRI response in human visual cortex. *Neuron* 3, 469–477.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci.* 92, 8135–8139.
- Malach, R., Grill-Spector, K., Kushnir, T., Edelman, S., Itzhak, Y., 1998. Rapid shape adaptation reveals position and size invariance. *NeuroImage* 7, S43.
- Mendola, J.D., Dale, A.M., Fischl, B., Liu, A.K., Tootell, R.B., 1999. The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J. Neurosci.* 19, 8560–8572.
- Meng, M., Tong, F., 2004. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* 4, 539–551.
- Moore, C., Engel, S.A., 2001. Neural response to perception of volume in the lateral occipital complex. *Neuron* 29, 277–286.
- Moutoussis, K., Keliris, G., Kourtzi, Z., Logothetis, N.K., 2005. A binocular rivalry study of motion perception in the human brain. *Vision Res.* 45, 2231–2243.
- Muckli, L., Kriegeskorte, N., Lanfermann, H., Zanella, F.E., Singer, W., Goebel, R., 2002. Apparent motion: event-related functional magnetic resonance imaging of perceptual switches and states. *J. Neurosci.* 22 (RC219).
- Muckli, L., Kohler, A., Kriegeskorte, N., Singer, W., 2005. Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biol.* 3, 1–10.
- Murray, S.O., Olshausen, B.A., Woods, D.L., 2003. Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cereb. Cortex* 13, 508–516.
- Pollen, D.A., 2003. Explicit neural representations, recursive neural networks and conscious visual perception. *Cereb. Cortex* 13, 807–814.
- Sereno, M.I., Dale, A., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893.
- Slotnick, S.D., Yantis, S., 2003. Efficient acquisition of human retinotopic maps. *Hum. Brain Mapp.* 18, 22–29.
- Sterzer, P., Russ, M.O., Preibisch, C., Kleinschmidt, A., 2002. Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *NeuroImage* 15, 908–916.

- Sterzer, P., Eger, E., Kleinschmidt, A., 2003. Responses of extrastriate cortex to switching perception of ambiguous visual motion stimuli. *NeuroReport* 14, 2337–2341.
- Stone, J.V., 1999. Object recognition: view-specificity and motion-specificity. *Vision Res.* 39, 4032–4044.
- Super, H., Spekreijse, H., Lamme, V.A., 2001. Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat. Neurosci.* 4, 304–310.
- Tong, F., 2003. Primary visual cortex and visual awareness. *Nat. Rev., Neurosci.* 4, 219–229.
- Tootell, R.B., Mendola, J.D., Hadjikhani, N., Ledden, P.J., Liu, A.K., Reppas, J.B., Sereno, M.I., Dale, A., 1997. Functional analysis of V3A and related areas in human visual cortex. *J. Neurosci.* 17, 7060–7078.
- Treisman, A., Gelade, G., 1980. Feature analysis in early vision: evidence from search asymmetries. *Psychol. Rev.* 95, 15–48.
- Tse, P.U., in press. Neural correlates of transformational apparent motion. *NeuroImage*.
- Tse, P.U., Logothetis, N.K., 2002. The duration of 3-D form analysis in transformational apparent motion. *Percept. Psychophys.* 64, 244–265.
- Tse, P.U., Cavanagh, J.R., Nakayama, K., 1998. The role of parsing in high-level motion processing. In: Watanabe, T. (Ed.), *High-level Motion Processing: Computational, Neurobiological, and Psychophysical Perspectives*. MIT Press, Cambridge, MA, pp. 249–266.
- Zhuo, Y., Zhou, T.G., Rao, H.Y., Wang, J.J., Meng, M., Chen, M., Zhou, C., Chen, L., 2003. Contributions of the visual ventral pathway to long-range apparent motion. *Science* 299, 417–420.