NeuroImage 270 (2023) 119909

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Identifying cortical areas that underlie the transformation from 2D retinal to 3D head-centric motion signals

Puti Wen[a,*], Michael S. Landy[b], Bas Rokers[a,b]

[a] Psychology, New York University Abu Dhabi, United Arab Emirates
[b] Department of Psychology and Center for Neural Science, New York University, United States

ARTICLE INFO

Keywords:
Visual system
Motion perception
Binocular vision
Sensory mechanisms
Stereoscopic displays

ABSTRACT

Accurate motion perception requires that the visual system integrate the 2D retinal motion signals received by the two eyes into a single representation of 3D motion. However, most experimental paradigms present the same stimulus to the two eyes, signaling motion limited to a 2D fronto-parallel plane. Such paradigms are unable to dissociate the representation of 3D head-centric motion signals (i.e., 3D object motion relative to the observer) from the associated 2D retinal motion signals. Here, we used stereoscopic displays to present separate motion signals to the two eyes and examined their representation in visual cortex using fMRI. Specifically, we presented random-dot motion stimuli that specified various 3D head-centric motion directions. We also presented control stimuli, which matched the motion energy of the retinal signals, but were inconsistent with any 3D motion direction. We decoded motion direction from BOLD activity using a probabilistic decoding algorithm. We found that 3D motion direction signals can be reliably decoded in three major clusters in the human visual system. Critically, in early visual cortex (V1-V3), we found no significant difference in decoding performance between stimuli specifying 3D motion directions and the control stimuli, suggesting that these areas represent the 2D retinal motion signals, rather than 3D head-centric motion itself. In voxels in and surrounding hMT and IPS0 however, decoding performance was consistently superior for stimuli that specified 3D motion directions compared to control stimuli. Our results reveal the parts of the visual processing hierarchy that are critical for the transformation of retinal into 3D head-centric motion signals and suggest a role for IPS0 in their representation, in addition to its sensitivity to 3D object structure and static depth.

1. Introduction

From escaping predators to hunting prey, the ability to accurately perceive motion has enabled our survival for as long as we have existed. Although most humans may no longer need to run away from bears, daily activities, such as pouring coffee, driving, and playing sports, are still critically dependent on motion perception. In squash, a player has moments to decide when and where to swing. To support this simple action, the brain carries out challenging perceptual tasks such as motion detection (Borst and Egelhaaf, 1989), the judgment of direction (Britten et al., 1992; Salzman et al., 1990), speed (Liu and Newsome, 2005), and time-to-contact (Tresilian, 1995) in the blink of an eye.

The visual system processes these motion signals along a hierarchical pathway (DeYoe et al., 1994; Maunsell, 1992; Maunsell and van Essen, 1983; Van Essen et al., 1986). Primary visual cortex (V1) and the middle temporal area (MT) are two key sites in this pathway, where the former is thought to primarily process component motion signals and the latter pattern motion signals (Born and Bradley, 2005; Britten et al., 1993; Movshon and Newsome, 1996; Rust et al., 2006). A fundamental challenge of any sensing organism is to recover the distal stimulus, an object's 3D motion relative to the observer in this case, based on the proximal stimulus, the 2D motion signals projected by that object onto the retina. It is not yet clear from previous research how the representation of motion signals contained within the retinal images is transformed into a representation that reflects the motion of objects relative to an observer in our three-dimensional world.

Multiple studies have investigated neural responses to stimuli that either imply or contain motion-in-depth cues and have identified a number of cortical areas that may be involved in the transformation of 2D retinal into 3D object motion signals. Monkey MT neurons are organized in columns by both motion direction (Albright et al., 1984) and

DOI of original article: 10.1016/j.neuroimage.2020.117540
* Corresponding author.
E-mail address: pw1246@nyu.edu (P. Wen).

https://doi.org/10.1016/j.neuroimage.2023.119909.
Received 21 October 2022; Received in revised form 26 January 2023; Accepted 28 January 2023
Available online 17 February 2023.

disparity preference (DeAngelis and Newsome, 1999). About half of neurons in macaque area MT are selective to motion-in-depth direction (Czuba et al., 2014; Sanada and DeAngelis, 2014). Moreover, MT has been linked to the perception of 3D structure from motion (Bradley et al., 1998; Grunewald et al., 2002), and depth from both binocular disparity and binocular motion (Armendariz et al., 2019; Nadler et al., 2008). Human fMRI studies have provided evidence of adaptation and motion aftereffects in hMT+ to 3D motion direction (Czuba et al., 2011; Joo et al., 2016; Rokers et al., 2009). At the same time, while manipulation of neural activity in monkey MT can impact motion percepts (Salzman et al., 1990), other studies found that MT activity at best has a modest relationship with perceptual reports (Britten et al., 1996; Hedges et al., 2011). Transformation of motion signals in regions further downstream in the visual processing hierarchy, such as area FST (Héjja-Brichard et al., 2020), may therefore be required. Similar arguments have been made based on results in humans (Likova and Tyler, 2007). Finally, research on depth perception (Ban et al., 2012) and structure-from-motion (Georgieva et al., 2009; Orban et al., 1999) has suggested that areas such as V3A/B and IPS may play a role in the extraction of 3D motion signals as well.

Here, we used an fMRI-based decoding approach to understand the representation of binocular stimuli that signal 3D motion across the visual hierarchy. We presented various trajectories of 3D motion to observers and asked whether and where we can decode the presented motion direction from the BOLD response. To test that a given cortical area represents the direction of 3D motion rather than the associated direction of the 2D motion signals on the two retinas, we analyzed responses to a control stimulus in which the retinal signals moved vertically, rather than horizontally. This control stimulus contained the same motion energy as the horizontal motion stimulus, but because of the geometry of binocular motion perception, did not produce 3D motion percepts. Specifically, because of the horizontal offset of our two eyes, horizontally opposite retinal motion is consistent with a stimulus approaching or receding from the observer. Vertically opposite retinal motion on the other hand is not consistent with any given 3D motion trajectory. Moreover, since we used relatively sparse dot motion displays, the control stimuli did not induce binocular rivalry, but instead produced percepts of transparent retinal motion (see Supplementary Videos 1 and 2). We hypothesized decoding performance for 3D motion and control stimuli would be equivalent in visual areas that encode 2D retinal motion. However, cortical regions that integrate the retinal motion signals from the two eyes, should exhibit worse decoding performance for the vertical compared to the horizontal motion stimuli.

In summary, we aimed to elaborate our understanding of the classical motion processing pathway using binocular stimuli that signal 3D motion. As motion processing relies on neural mechanisms beyond those required for the processing of fronto-parallel motion, we expected that a critical evaluation of the BOLD responses produced by our binocular motion stimuli would reveal a more complete picture of the visual motion processing hierarchy in the human brain.

2. Materials and methods

2.1. Observers

Nine healthy observers (one male, age 23–46 years) with normal or corrected-to-normal vision participated and provided written informed consent. All observers achieved a score of five or higher (70 s of arc or better) on the Randot Circles Stereotest (Stereo Optical Company, Chicago, IL). Four observers participated at New York University Abu Dhabi, and five observers participated at New York University New York. All observers participated in two scanning sessions for the horizontal condition and two sessions for the vertical condition. Each scanning session lasted 1.5 h. The experiment was approved by the University Committee on Activities Involving Human Subjects at New York University and New York University Abu Dhabi.

2.2. Visual stimulus

Stimuli were generated on a Macintosh computer using MATLAB 9.2 (The MathWorks, Natick, MA, USA) and the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). We presented the binocular stimuli using a ProPixx DLP LED projector (VPixx Technologies Inc., Saint-Bruno-de-Montarville, QC, Canada; screen resolution: 1920 × 1080 pixels, refresh rate: 120 Hz; 60 Hz for each eye) with a rear projection screen (projected screen width: 38.5 cm, viewing distance: 88 cm) positioned at the back of the scanner.

The stimuli were white dots (55.2 cd/m^2) with 0.5 s lifetime drifting against a black (0.2 cd/m^2) background (Fig. 1**A**[1]). All dots were restricted to two black sectors in the left or right hemifield (outer aperture radius: 14 deg, inner aperture radius: 0.7 deg). No dots were presented in two wedges centered on the upper and lower vertical meridians (25 deg of arc wedge width). The wedges allowed us to present stimuli with opposite motion directions on either wedge side, reducing oculo-motor drive (see below). We ensured that dots in the two eyes remained within the sector and remained binocularly visible at all times. The remainder of the display area was filled with $1/f$ noise that was identical in the two eyes and perceptually appeared at the fixation distance. There was a fixation dot at the center of the display.

The total number of white dots displayed in both hemifields at a time for each eye remained at 100 as a new dot would appear at a random location on the screen when a previous dot disappeared. To reduce false binocular matches, any new dot position was set to a minimum of 0.975 deg absolute Euclidean distance from the rest of the dots that were currently on the screen. We restricted the maximum horizontal dot disparity between the two eyes to 0.3 deg. At the onset of the experiment, each dot was initialized with a random disparity chosen from a uniform distribution between the minimum (-0.3 deg) and maximum (0.3 deg) disparity. Depending on the relative velocity of the dots presented to the left and right eye, the observers perceived different motion directions. Dots drifted horizontally in the horizonal condition and vertically in the vertical condition. All other stimulus parameters were identical between the two conditions.

2.2.1. Horizontally opposite motion stimuli

The main experiment contained three types of motion direction: lateral, in-depth, or oblique (Fig. 1B). Each dot in the lateral-motion stimuli maintained a constant disparity throughout its lifetime, moving strictly leftward or rightward at a rate of 0.85 deg/sec. Dots in the motion-in-depth stimuli changed disparity at 1.7 deg/sec (dot speeds of 0.85 deg/sec in each eye but in opposite directions), with no net leftward or rightward motion. Oblique motion stimuli were a combination of lateral and in-depth motion: dots moved leftward or rightward at a rate of 1.2 deg/sec in one eye and were stationary in the other eye. These dot speeds were used because they allowed for high direction-discrimination sensitivity for 3D motion (Czuba et al., 2010).

Stimuli cycled through motion directions either clockwise or counterclockwise in the horizontal (xz) plane. For example, when cycling counterclockwise, rightward motion was followed by right/away, away, left/away, left, left/toward, toward, right/toward, right again and so on (Fig. 1C). The eight possible lateral/in-depth/oblique motion directions were presented for three seconds each.

[1] See supplementary video 1 for an illustration of the stimuli used in the horizontal motion condition: https://osf.io/8vz4t/files/osfstorage/63c503e4e481030255976783. Stimuli simulated a volume of dots moving in three dimensions. Stimuli cycled through motion directions either clockwise or counterclockwise in the horizontal (xz) plane. For example, when cycling clockwise (as shown in the video) rightward motion was followed by right/toward, toward, left/toward, left, left/away, away, right/away, right again and so on (Fig. 1C). The eight possible lateral/in-depth/oblique motion directions were presented for three seconds each. Please refer to Methods for additional stimulus details.

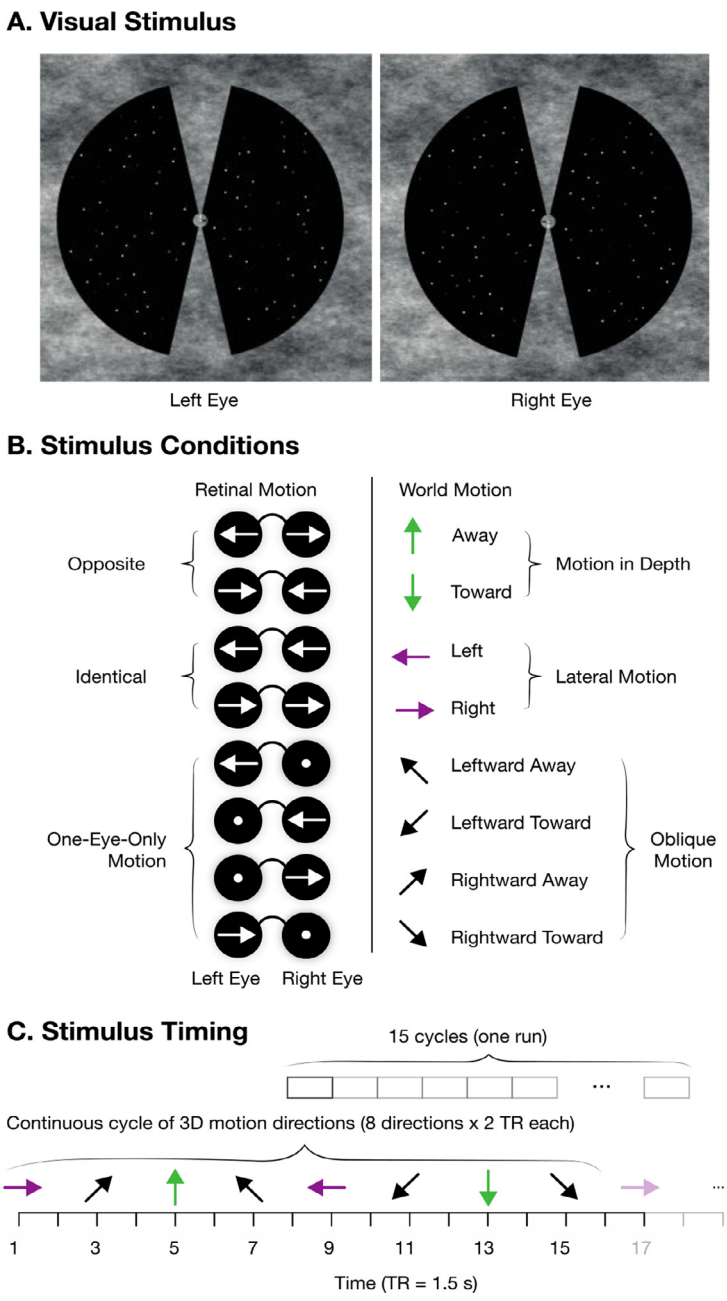

Fig. 1. Stimuli and experimental design for the main experiment (horizontal condition). ***A.* Visual Stimulus.** In each monocular image, white dots drifted horizontally against a black background. Dots were presented within a circular aperture. The region beyond the aperture was filled with pink noise to aid binocular fusion. At all times, opposite motion directions were presented in the two hemifields. ***B*. Stimulus Conditions**. The 2D retinal motion signals produced eight different 3D head-centric motion directions, depending on the relationship of the retinal motion between the two eyes. ***C*. Stimulus Timing.** Timeline for one run of the main experiment. In each scan, stimuli cycled through each of the eight directions 15 times. In the odd runs, the stimuli were presented in a continuous counterclockwise cycle of motion directions. The order was reversed to a clockwise cycle in the even runs.

Each MRI scan run lasted 375 s in total. The first 15 s (5 trials) of each run were removed to minimize the effects of transient magnetic saturation and to allow for the hemodynamic response to reach a steady state. In the remaining 360 s, each cycle of eight directions was repeated 15 times. Direction order was alternated (clockwise vs. anti-clockwise) from scan to scan. At all times, opposite motion directions were presented in the two hemifields. For example, if dots in the right hemifield were moving away from the observer, then dots in the left hemifield would move towards the observer. This was done to discourage eye movements induced by optokinetic drive and encourage consistent fixation on the center of the screen.

2.2.2. *Vertically opposite motion stimuli*

All participants also took part in a vertical-motion experiment. All stimulus properties were identical to that in the main experiment, with the exception that dots drifted vertically rather than horizontally (Fig. 2 [2]). Thus, trial timing as well as stimulus parameters such as speed and luminance were maintained between experiments, with the critical

[2] See supplementary video 2 for an illustration of the stimuli used in the vertical-motion condition: https://osf.io/8vz4t/files/osfstorage/63c503e381150f02651e1eca. The sequence of stimuli shown in the video is identical to the sequence in the horizontal-motion video except that each horizontal retinal-motion trajectory was rotated by 90 deg producing vertical rather than horizontal motion. In the video, downward motion was followed by stationary/down, up/down, up/stationary, up, stationary/up, down/up, down/stationary, down again and so on (Fig. 2). The eight possible motion directions were presented for three seconds each. Please refer to Methods for additional stimulus details.

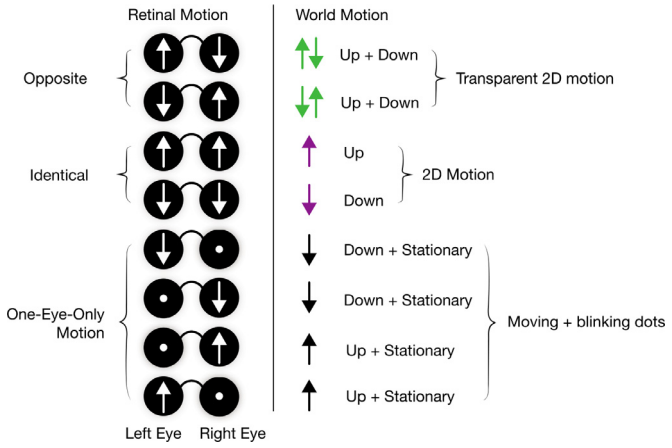

Fig. 2. Stimuli and experimental design for the vertical condition. Each horizontal retinal motion trajectory was rotated by 90 deg and now moved vertically. The control experiment has analogous retinal stimulus properties while no longer producing 3D motion percepts. Opposite retinal motion (upward motion in one eye and downward motion in the other) produced a percept of transparent motion (2D or fronto-parallel) rather than of motion in depth.

difference that the vertically opposite motion stimuli produced transparent motion, rather than 3D motion percepts.

2.3. Behavioral task

To control attention and ensure proper vergence throughout the experiment, subjects performed a depth-detection task at fixation. At random intervals, the binocular disparity of the fixation dot changed slightly (.1 deg of arc or less). Participants indicated the perceived depth change (near/far disparity relative to the screen) by button press. Detection performance did not significantly differ between the horizontal and vertical motion conditions ($F(1,114) = .1539$, $p = .6956$).

2.4. MRI data acquisition

MRI data were acquired on a Siemens Prisma 3T full-body MRI scanner (Siemens Medical Solutions, Erlangen, Germany) using a 64-channel head coil. For each observer, a T1-weighted anatomical scan was acquired (TR: 2400 ms; TE: 2.22; flip angle: 8°; 0.8 mm isotropic voxels). This anatomical volume was used for white/gray matter segmentation and co-registration with the functional scans. T2*-weighted functional scans were acquired using an echo-planar imaging (EPI) sequence (TR: 1500 ms; TE: 38 ms; flip angle: 68°, multiband factor: 4; matrix size: 104 × 104, 2 mm isotropic voxels; 68 slices). For each functional scan, we collected 250 volumes with full brain coverage. Behavioral and neuroimaging data is available at https://openneuro.org/datasets/ds004443/versions/1.0.0.

2.5. Anatomy and ROI identification

We anatomically identified cortical ROIs in the visual system using the Neuropythy atlas (Benson and Winawer, 2018). The following ROIs were selected and used in the analysis: V1v, V1d, V2v, V2d, V3v, V3d, V3A, V3B, hV4, VO1, VO2, PHC1, PHC2, IPS0-5, hMT, MST, LO1, LO2, SPL1, and FEF. We merged V1-3v and V1-3d as V1-3 using *fslmaths*.

2.6. fMRI pre-processing

All scans were organized in the Brain Imaging Data Structure (BIDS) format (Gorgolewski et al., 2017) using dicom2bids and dcm2niix (Li et al., 2016) and then pre-processed using the fMRIPrep pipeline (version 20.2.1) for motion correction, spatial normalization, and co-registration between the functional and anatomical scans (Esteban et al., 2019).

2.7. Decoding motion directions

For functional runs and for each ROI, the data were organized as an $n \times m$ matrix where n is the number of measurements ($n = 240$) for each scan and m is the number of voxels within that ROI. Every column in this matrix was a time series of raw fMRI signals for one voxel. Prior to classification, we applied a high-pass filter using the fast Fourier transform (FFT) to convert the data from the time domain to the frequency domain to remove low-frequency noise (cutoff: 1/72 Hz). The remaining data were then converted back to the time domain with the inverse FFT. We then normalized (z-scored) the time series separately for each voxel. Each scan contained 120 trials – 15 trials for each of the eight motion directions – and each trial lasted 2 TRs (3s). We averaged the trials within a scan that had the same motion direction 6 s after the stimulus onset to account for the hemodynamic delay. This resulted in an eight (number of directions) by m (number of voxels) matrix for each scan. Each row of the final matrix represents the average normalized BOLD response to a given motion direction, and each column represents the data for one given voxel. Pre-processing was performed for each subject, and each scan separately.

We decoded the motion direction using a probabilistic decoding algorithm, TAFKAP (van Bergen and Jehee, 2021). This is an inverted generative model that decodes both a stimulus estimate and its uncertainty. Since we only used eight motion directions, we used an extension of TAFKAP to discrete stimulus sets (van Bergen, pers. comm.). This algorithm returns the direction with the highest probability as the estimated direction and computes the entropy of the decoded distribution as its measure of uncertainty. We applied 32-fold cross-validation and saved the averaged decoding results across validation in an eight-by-eight confusion matrix, where each column represents a presented direction, and each row records the proportion of scans in which the direction was decoded as each of the eight possible directions. Accurate decoding was reflected on the diagonal of this matrix. We obtained a bootstrapped distribution of decoding results for chance performance based on repeatedly randomizing direction labels for the training data. We considered the decoding accuracy for the original data to be statistically significant if the value exceeded the 95th percentile of this chance-performance distribution.

We also applied a searchlight-based analysis to determine decoding performance across the entire cortical surface, including areas beyond the retinotopically defined ROIs. To make the analysis computationally feasible, we used the MATLAB function *classify* with the option 'diaglinear' to decode motion directions in the searchlight analysis, rather than the more demanding probabilistic decoding algorithm. Overall decoding performance was reduced when using the more traditional method, but relative differences between conditions were preserved.

The classifier fit a multivariate normal density to each motion direction with a diagonal covariance matrix using the training data and its label. The diagonal covariance matrix represented an assumption of independent responses across voxels since the data were not sufficient for an accurate estimate of the true covariance. We cross-validated classifier performance using an 80:20 training-to-testing-split method. To compute a chance performance distribution, we resampled data with replacement 5000 times with randomized labels. The p-value was calculated as the percentage of times the resampled chance decoding accuracy was larger than the observed decoding accuracy.

2.8. Motion direction preferences

To evaluate the average motion direction preference of individual vertices across participants, we converted subject surface data into fsaverage6 surface space using FreeSurfer (Reuter et al., 2012). We com-

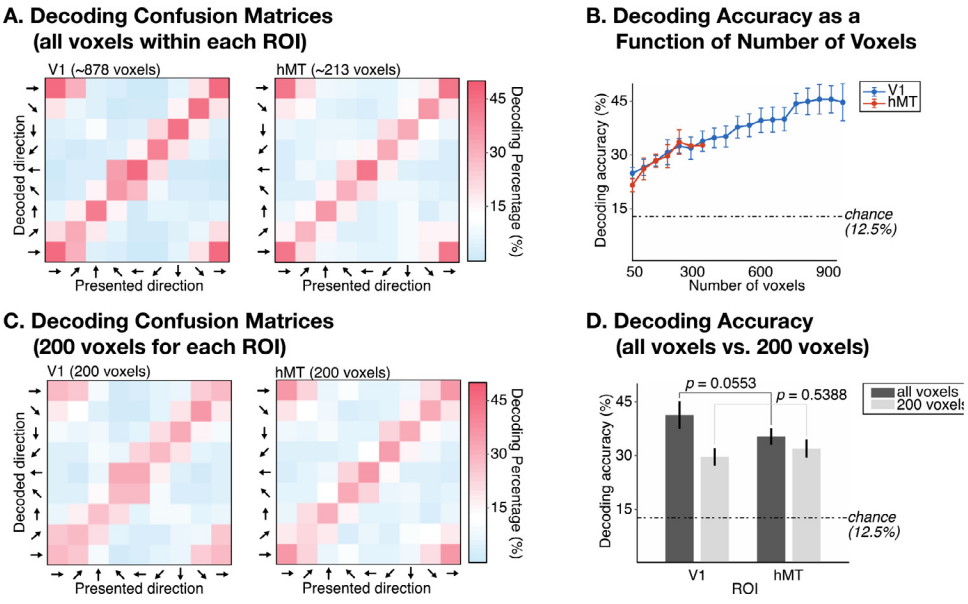

Fig. 3. Reliable decoding of 3D motion signals in both V1 and hMT. ***A***, Decoding confusion matrices using all voxels in V1 and hMT. Matrices are color-coded based on decoding percentage. Theoretical chance performance (12.5%) in white, above-chance performance in red, and below chance in blue. The first row and column of the matrix are repeated at the end of each matrix to illustrate the circular nature of the eight motion directions. ***B***, Decoding accuracy as a function of the number of voxels included in V1 and hMT. Error bars indicate SEM across subjects. Dashed black lines indicate the theoretical chance of decoding one of eight motion directions (12.5%). ***C***, Same analysis as ***A*** except only 200 randomly selected voxels from each ROI were used. ***D***, Decoding accuracy between using all voxels and using only 200 voxels for V1 and hMT. Error bars indicate SEM across subjects. Dashed black lines indicate theoretical chance decoding performance of one of eight motion directions (12.5%). All data reflect results averaged across participants (n = 9).

puted the preferred motion direction for each vertex. We first averaged the BOLD response across trials and scans for each presented motion direction. We then computed the vector sum across motion directions, so that the vector length represents the averaged BOLD response, and the angle represented the preferred motion direction of that vertex. The vector length can also be understood as the strength of the motion direction preference. We computed this vector for each vertex and each subject separately and then calculated the circular mean of this vector across subjects for each vertex. We define the mean vector as the preferred motion direction for that vertex. We plotted the result on a flattened surface for both left and right hemispheres, thresholded by motion preference strength (top 10th percentile).

3. Results

3.1. Reliable decoding of 3D motion signals in V1 and hMT

We probed the representation of 3D motion signals in the canonical motion-sensitive areas, V1 and MT. We presented separate motion stimuli to the two eyes, consistent with a 3D volume of dots that changed their direction every 3 s, cycling through 8 motion directions every 24 s. We extracted the mean response in every voxel to each motion direction in each of 20 scans and trained a classifier to decode the presented motion direction in each ROI. The average decoding results (N=9) are illustrated in Fig. 3A. Correct classification, where the decoded direction matched the presented direction, corresponds to elements on the positive matrix diagonal. Average decoding performance across all directions was 41.19% (+/−3.88% SE) and 35.2% (+/−2.32% SE) in V1 and hMT respectively, reflected in a prominent red positive diagonal across the confusion matrices in both V1 and hMT.

3.1.1. Accounting for the number of voxels

A common finding in the decoding of motion direction from BOLD responses is that decoding accuracy is greater in V1 than MT (Kamitani and Tong, 2006; Serences and Boynton, 2007; Wang et al., 2014), even when equating for the difference in size between these areas. For example, at our 2 mm isotropic imaging resolution, V1 contained ~878 voxels, and hMT only contained ~213 voxels on average (n = 9). To equate for area size, we re-ran the classifier in V1 and hMT, each time increasing by 50 the number of voxels used for decoding. When equating the number of voxels by repeatedly randomly sampling a subset of voxels, decoding accuracy was near-identical in V1 and hMT (Fig. 3B). Thus, in contrast to prior work on the decoding of 2D retinal motion, we found no evidence for a difference in decoding accuracy between V1 and MT.

Prior work has suggested that superior performance in the decoding of motion (and orientation) in early cortical visual areas, including V1, may be due to vignetting, biases in the BOLD response to visual stimuli near the edges of stimulus apertures (Carlson, 2014; Wang et al., 2014). Our results suggest that these biases may be less pronounced for stimuli that are not constrained to the image plane, a question we will return to at the end of the Results section.

To directly compare decoding performance across ROIs, we set 200 voxels as the upper limit for every ROI for all subsequent analyses. If an ROI had more than 200 voxels, we randomly selected 200 voxels for each cross-validation. Note that some ROIs contained less than 200 voxels on average and the results from especially small ROIs should be considered with caution: V3B (170 voxels), LO2 (139), MST (32), IPS4 (42), IPS5 (11), VO1 (133), VO2 (164), SPL1 (91), PHC1 (104), PHC2 (106), and FEF (65). Limiting the number of voxels did impact classifier performance, reducing decoding performance in both V1 ($t_{V1}(8) = 4.74$, p = .002) and hMT ($t_{hMT}(8) = 3.49$, p = .008), although the reduction in the percentage of accurately decoded stimuli in V1 (from 41.2% to 29.6%) was significantly larger than in hMT (from 35.2% to 31.8%; $t_{V1\text{-}hMT}(8) = 3.41$, p = .009). However, the overall pattern of results remained relatively unaffected (compare Fig. 3A and C), where a red positive diagonal line is still apparent in both regions, although now less prominent in V1. In contrast to all previous work (Kamitani and Tong, 2006; Serences and Boynton, 2007; Wang et al., 2014), decoding in hMT is no worse than in V1 once equating for number of voxels ($t_{V1\text{-}hMT(200)}(8) = -.64$, p = .539; Fig. 3D). Having put overall decoding performance in V1 and MT on equal footing, we next turned to understanding differences in the representation of motion signals in each region, based on two additional analyses.

3.1.2. Decoding misclassification

Reliable and comparable decoding of 3D motion signals in V1 and MT does not necessarily imply classification based on similar information. Inspection of the confusion matrices indicates that decoding errors were not randomly distributed. In the confusion matrices, the off-diagonal squares represent the percentage of trials the presented motion direction is misclassified based on the pattern of BOLD responses across

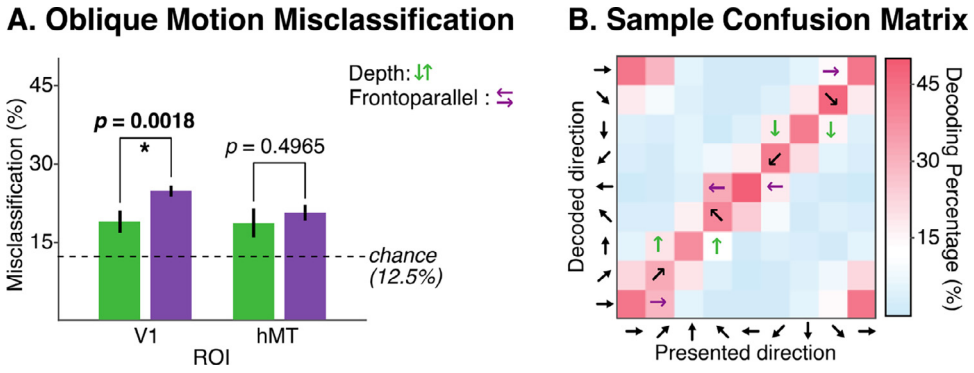

Fig. 4. Decoding misclassification: Differences in decoding bias reveal differences in representation in V1 and hMT. ***A***, Oblique motion misclassification. Results are taken directly from the confusion matrices in Fig 3C. Green and purple bars each represent the percentage of the oblique trials that were misclassified as either the oblique motion's depth or frontoparallel component. Dashed line: chance performance (12.5%). Error bars: SEM across subjects (n=9). ***B***, Sample confusion matrix. The results from the matrices that were included are indicated with purple (frontoparallel) and green (depth) arrows.

voxels in an ROI. The decoder often misclassifies the presented motion direction as an immediately neighboring one. For example, the decoder will often misclassify rightward-away oblique motion as either rightward or away rather than any other motion direction. Conversely, the decoder rarely misclassifies the presented motion direction with a direction 180 degrees away (e.g., misclassifying leftward as rightward or approaching as receding motion). As a result, many red diagonal lines appear to be three squares wide and are bordered by blue regions with below-chance decoding (Fig. 3C).

The oblique motion direction stimuli in our experimental design only contain non-zero motion signals in one eye. For example, the rightward-away motion stimulus presents rightward motion to the right eye, but a stationary field of dots to the left eye (Fig. 1B). If a voxel in an ROI predominantly encodes 2D (monocular) retinal motion rather than 3D motion, we expect that ROI to show a bias towards the presented monocular retinal motion. As monocular visual inputs integrate and transform into a representation of 3D motion, we expect a decrease in that bias, and a more uniform misclassification to either the neighboring fronto-parallel or exclusively in-depth motion stimulus (Fig. 4A). To test for this bias, we computed the percentage of times oblique motion trials were misclassified as either the neighboring fronto-parallel or in-depth motion stimulus (Fig. 4B) based on the confusion matrices from the previous analysis (Fig. 3C). We fit a linear mixed-effects model including the fixed effects of ROI (V1/hMT), motion direction (fronto-parallel/motion-in-depth), and their interaction. We also included by-subject random effects of intercept and slope. Misclassification percentage was the dependent variable. The model revealed a significant interaction between ROI (V1/hMT) and motion direction (fronto-parallel/motion-in-depth) on misclassification (F(1,1148) = 6.6192, p = .0102). A test of the simple effects showed a greater tendency in V1 to misclassify an oblique direction into its nearest fronto-parallel direction ($M_{V1\_fronto\text{-}parallel}$ = 24.61%, $SE_{V1\_fronto\text{-}parallel}$ = .84%) instead of its nearest motion-in-depth direction ($M_{V1\_motion\text{-}in\text{-}depth}$ = 18.36%, $SE_{V1\_motion\text{-}in\text{-}depth}$= .87%; F(1,574) = 9.8682, p = .0018). Misclassification in hMT on the other hand was not significantly biased toward either direction ($M_{hMT\_fronto\text{-}parallel}$ = 20.36%, $SE_{hMT\_fronto\text{-}parallel}$ = .84%; $M_{hMT\_motion\text{-}in\text{-}depth}$ = 18.36%, $SE_{hMT\_motion\text{-}in\text{-}depth}$= .92%; F(1,574) = .4629, p = .4965; Fig. 4A). Our fixed stimulus sequence (clock- or counter-clockwise sequence, see Methods) did not confound the analysis. We found no significant decoding bias that misclassified a given stimulus as either the previously presented stimulus or the stimulus presented subsequently (t_{V1} (8) = -.4291, p = .6792; t_{hMT} (8) = −1.4064, p = .1972). The misclassification result suggests that V1 predominantly represents retinal motion signals, as would be expected from previous literature (Hubel and Wiesel, 1974; Snowden et al., 1991; Wuerger et al., 1996). For hMT, the pattern of BOLD responses to oblique motion appears to be less determined by its fronto-parallel component, as would be expected for regions that contain more binocular neurons. The 3D trajectory is represented beyond simply encoding 3D motion based on the motion's weaker fronto-parallel component. To further test this hypothesis, we conducted additional experiments that compared decoding performance to horizontal and vertical motion stimuli in V1 and hMT.

3.1.3. Decoding horizontal versus vertical motion

Since our two eyes are horizontally offset, opposite retinal motion (leftward motion in one eye and rightward in the other) results from objects moving toward/away from the participant. In comparison, vertically opposite 2D retinal motion (upward motion in one eye and downward in the other) is inconsistent with the 3D motion of a single stimulus. To dissociate the representation of 2D retinal and 3D head-centric motion in cortical areas we therefore presented vertical rather than horizontal retinal motion to all participants in a control experiment. The vertical motion stimuli contained identical retinal motion energy, but produced percepts of transparent 2D rather than 3D motion. In ROIs that represent 2D retinal motion, we expected decoding performance to be identical in the horizontal and vertical motion conditions. Conversely, in ROIs that represent 3D head-centric rather than 2D retinal motion, we expected significantly better decoding performance in the horizontal compared to the vertical conditions.

The same nine subjects completed the same number of runs of the vertical condition and we decoded the vertical data using the same analysis pipeline. We fit a linear mixed-effects model including the fixed effects of ROI (V1/hMT), motion condition (horizontal/vertical), and their interaction. We also included by-subject random effects of intercept and slope. Decoding accuracy was the dependent variable. The model revealed a significant interaction between ROI (V1/hMT) and motion condition (horizontal/vertical) on decoding accuracy (F(1,1148) = 15.689, p < .0001). In V1, testing for simple effects showed decoding accuracy was not significantly different between the horizontal ($M_{V1\_horizontal}$ = 29.64%, $SE_{V1\_horizontal}$ = .75%) and vertical motion conditions ($M_{V1\_vertical}$ = 26.97%, $SE_{V1\_vertical}$ = .76%; F(1,574) = 1.0071, p = .3011). In hMT on the other hand, we found significantly greater decoding accuracy in the horizontal motion condition ($M_{hMT\_horizontal}$ = 31.84%, $SE_{hMT\_horizontal}$ = .76%) compared to the vertical motion condition ($M_{hMT\_vertical}$ = 24.09%, $SE_{hMT\_vertical}$ = .65%; F(1,574) = 10.343, p = .0014; Fig. 5).

Next, we asked if the pattern in hMT was driven by a specific motion type. We divided the eight motion directions into three groups: fronto-parallel, oblique, motion-in-depth. We tested if the difference between the horizontal and vertical conditions in hMT was driven by one of these motion types. While decoding performance varied somewhat between motion types (fronto-parallel, oblique and in-depth) as well as condition (horizontal and vertical), a two-way ANOVA showed a significant main effect of horizontal vs. vertical condition in hMT (F(1,48) = 7.47, p = .0087) but not in V1 (F(1,48) = .26, p = .6104). Further, when considering the simple effects within the fronto-parallel motion type, we found no significant difference in decoding performance for either V1 (t(8) = -1.1315, p = .2906) or hMT (t(8) = 1.2475, p = .2475). This result suggests that the difference between conditions in hMT was not driven solely by fronto-parallel stimuli where both eyes received the

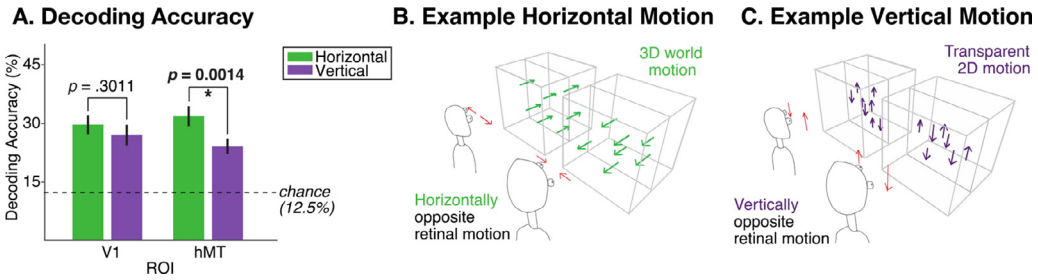

Fig. 5. ***A***, Decoding accuracy in the horizontal and vertical conditions. Dashed line: chance performance (12.5%). Error bars: SEM across subjects (n=9). ***B-C***, Example horizontal and vertical motion directions compared in this analysis.

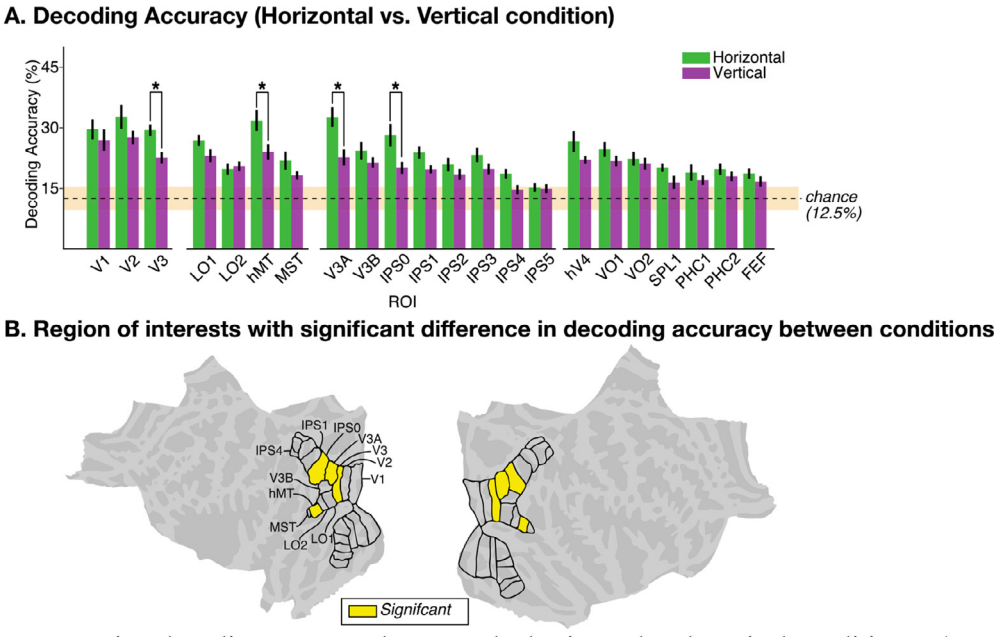

Fig. 6. Motion decoding accuracy between the horizontal and vertical conditions. ***A***, Decoding accuracy averaged across subjects. Significant paired-sample t-tests between the horizontal and vertical conditions are indicated at the top of the bars (*: significant with a FDR of 10%; df = 8). Dashed black line: chance performance (12.5%). Light orange horizontal patch: 95% confidence interval of chance (9.55-15.25%). Error bars indicate SEM across subjects (n=9). ***B***, Regions of interest with significant differences in decoding accuracy between conditions colored in yellow in both hemispheres. The black outlines indicate each of the 22 ROIs.

same motion direction signals. The results are consistent with the misclassification biases reported in the previous section. Taken together, these findings suggest that the signals in V1 predominantly represent 2D retinal motion while the signals in hMT at least in part represent 3D head-centric motion.

3.2. Motion decoding performance beyond V1 and hMT

V1 and hMT are two critical regions for the processing of visual motion. However, the complete motion pathway involves additional areas. The classical V1→MT motion-processing pathway is based on studies of retinal motion in both humans and non-human primates. Our current work aims to advance our understanding of the motion pathway that underlies 3D motion perception. Thus, in the next sections, we look beyond V1 and hMT. We first investigate the role of ROIs in retinotopically characterized visual cortex. Subsequently, we search for evidence for the representation of 3D motion signals beyond those ROIs in a searchlight analysis across the entire cortical surface.

3.2.1. Representation of motion signals in retinotopically-characterized cortex

We decoded motion signals in 20 additional ROIs using our analysis pipeline (Fig. 6A). Note that the V1 and hMT results in Fig. 6A are identical to those in Fig. 5. Many regions showed above-chance decoding accuracy in both the horizontal and vertical conditions, which was not expected for some regions, especially those located in the ventral pathway. Decoding accuracy also varied across the cortex. Unsurprisingly, a one-way ANOVA revealed a significant difference in decoding accuracy across ROIs for both conditions ($F_{horizontal}(21,176)$= 6.77, p < .0001; $F_{vertical}(21,176)$= 5.59, p < .0001). Early visual areas as well as dorsal regions tend to have relatively greater decoding accuracy compared to the ventral regions in the horizontal condition.

To identify regions that show greater representation of 3D head-centric motion over 2D retinal motion, we then ran a paired sample t-test of the horizontal decoding accuracy against the vertical conditions. In addition to hMT, three ROIs showed a significant difference between the horizontal and vertical conditions. We found significant differences between conditions in V3, hMT, V3A, and IPS0 ($t_{V3}(8)$ = 3.02, p = .0166; $t_{hMT}(8)$ = 3.03, p= .0163; $t_{V3A}(8)$ = 3.53, p = .0077; $t_{IPS0}(8)$ = 3.19, p = .0127), corrected for FDR using the Benjamini-Hochberg procedure (q = 0.1). At q = .1 we accept that on average 10% of our significant results are false positives. These regions are colored in yellow in Fig. 6B and they are clearly not randomly distributed across visual cortex.

As we expected early visual cortex to encode 2D retinal motion, we were not surprised to find that decoding accuracy for vertical and horizontal motion did not differ significantly in V1 and V2 ($t_{V1}(8)$ = 0.95, p = .3718; $t_{V2}(8)$ = 1.94, p= .0879). We also found no significant dif-

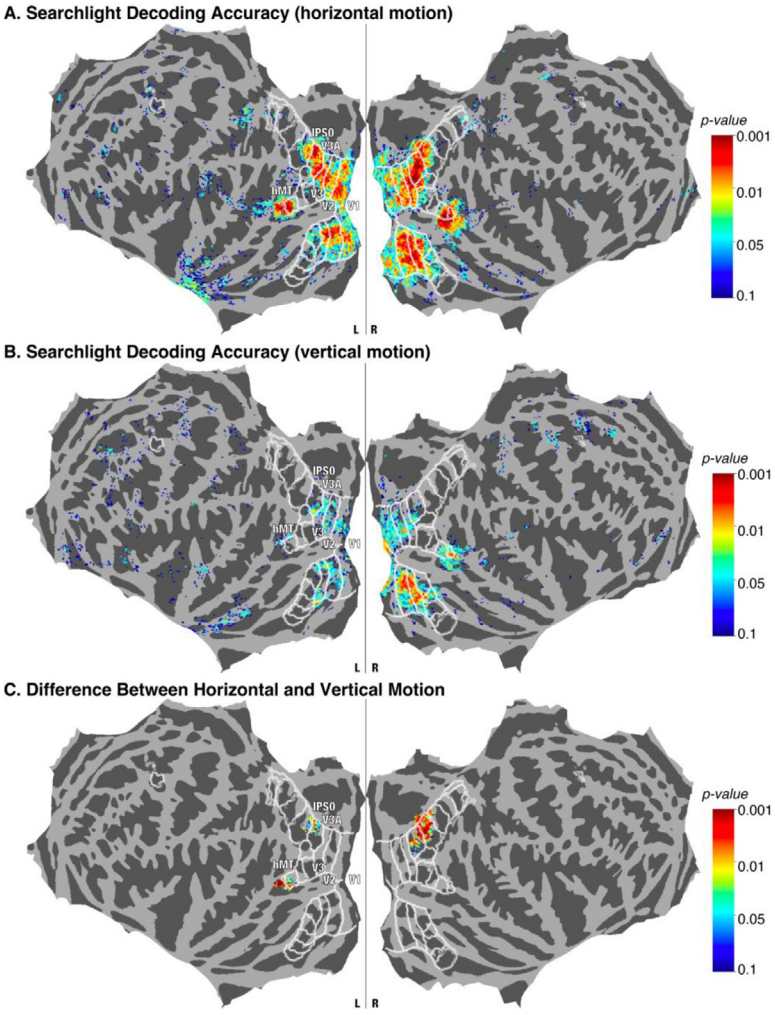

Fig. 7. Searchlight decoding results on the cortical surface. ***A***, Classification accuracy for the horizontal motion condition. ***B***, Classification accuracy for the vertical motion condition. ***C***, The difference in decoding accuracy between the horizontal and vertical motion conditions thresholded by cluster size. The color bar indicates *p*-values in log scale and ranges from 0.001 to 0.01 to 0.05 and to 0.1 for all three panels. The white outlines indicate each of the 22 ROIs.

ferences in ventral regions such as PHC1-2, which are not commonly thought to be motion-responsive. The differences between conditions start to arise in later dorsal regions. V3, V3A, and IPS0 are adjacent to one another and could process similar cues in our stimulus.

In most visual regions decoding performance was slightly better in the horizontal compared to the vertical condition. However, this global effect was substantially smaller than the specific ROI-based effects. The areas in visual cortex are interconnected, and the smaller global effects across areas might be inherited from the larger effects we identified. Perceived dot density and motion opponency might be another factor: Perceptually, the dot pattern in the vertical condition will appear denser because stimulus elements in the two eyes are not fused. Moreover, in many of the conditions the dots will appear in pairs that appear to move in opposite directions, which has been shown to reduce motion discrimination and to reduce neural activity in extra-striate cortex (Qian et al., 1994). While we cannot unequivocally identify a cause for the overall trend, the ROIs that show effects above and beyond that trend are well-established members of the motion processing hierarchy, while most areas that show the smaller trend effect are typically not considered motion selective.

The preceding analyses were restricted to ROIs within the retinotopically mapped visual processing hierarchy. However, hMT appears relatively isolated, and one might have expected a role for MST as well. However, MST contains significantly fewer than 200 voxels (32 voxels on average) and this might have obscured results. Additionally, the cortical areas neighboring hMT are not typically identified using retinotopic mapping, but may nonetheless show evidence for the representation of motion signals. We therefore carried out a searchlight analysis to identify additional regions in which 2D retinal and/or 3D head-centric motion signals may be reliably decoded.

3.2.2. *Searchlight classification*

Our analyses so far have focused on visual ROIs that are part of the well-established parieto-occipital and temporo-occipital hierarchy and have previously been shown to contain clear retinotopic organization. To investigate the representation of binocular-motion signals in additional visual, or indeed any cortical areas, we conducted a searchlight analysis across the entire cortical surface. We used a MATLAB classifier (see Materials and Methods) in the searchlight analysis instead of TAFKAP due to the latter's high computational demand. This approach also verifies that our results do not depend on a particular choice of multivariate pattern analysis algorithm. Cross-validated classifier decoding results averaged across subjects are shown in Fig. 7. Every vertex served as the center of a searchlight sample consisting of the nearest 100 vertices. We compared the decoding accuracy of each sample to the bootstrapped chance distribution to calculate the *p*-value, which was then represented at each vertex. We obtained the chance distribution from the searchlight analysis on bootstrapped data with shuffled labels. Despite variability across subjects, most subjects displayed the same pattern: reliable decoding performance around the canonical motion processing pathway (V1-hMT) as well as IPS0. The whole pathway is evident in the averaged map for the horizontal condition (Fig. 7A). Large

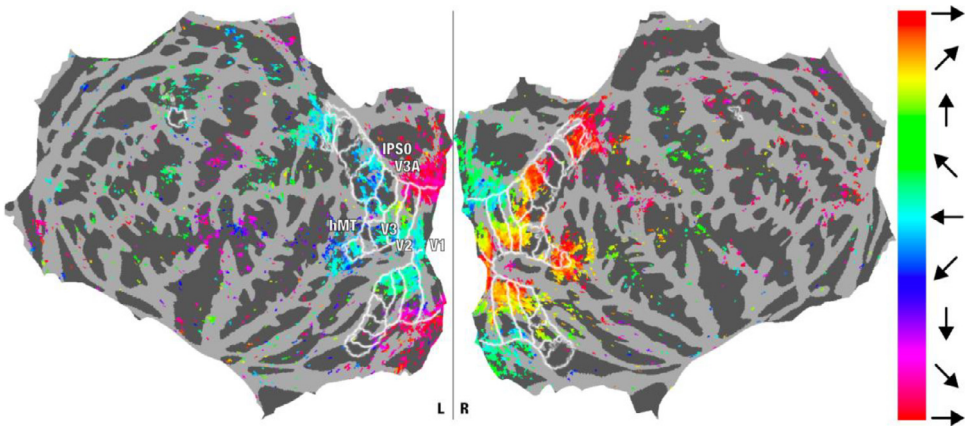

Fig. 8. Motion direction preferences averaged across nine subjects on a flattened map for both hemispheres. The color of each dot represents the preferred motion direction for that specific vertex after averaging across subjects. The color bar ranges from red (rightward motion) to green (away motion) to teal (left motion) and to purple (toward motion) in a counterclockwise manner on the HSV colormap. White contours indicate the outline of ROIs. The results are thresholded by only showing vertices in the top tenth percentile of averaged and *z*-scored BOLD responses.

patches cover the entirety of the early visual areas, the dorsal pathway, and regions around the hMT complex.

To highlight regions that represent binocular-motion signals beyond their retinal constituents, we contrasted decoding performance between the horizontal and vertical conditions. We bootstrapped the data 5000 times. The *p*-value was the proportion of times the randomized condition difference was larger than the observed difference. For every random sample, we found the largest significant cluster across both hemispheres and recorded the size of that cluster into our reference distribution of significant cluster sizes. After 5000 resamples, we found that the largest cluster size resulting from random sampling was 66 vertices. We thresholded our data using this criterion, leading to the map of *p*-value results in Fig. 7C. Note that subtracting the results in Fig. 7A and 7B does not directly produce Fig. 7C. Fig. 7A and B depict *p*-values of the averaged data across all subjects and reflect the ability to decode in the horizontal and vertical conditions, respectively. Fig. 7C depicts the result of a paired *t*-test between horizontal and vertical accuracies across the 9 participants. This analysis revealed three major clusters. We identified a region overlapping with and extending slightly anterior beyond hMT in the left hemisphere. Two additional clusters were located around V3A and IPS0 bilaterally, with the regions in the left hemisphere centering on V3A and the right hemisphere centering on both V3A and IPS0. The cluster around V3A/IPS0 is the largest cluster identified by the searchlight result. Interestingly, no early visual area (V1-3) was evident on this map, consistent with our other results that horizontal and vertical motion are decoded equally well in these regions in early visual cortex.

3.2.3. Motion direction preferences

Previous research has suggested that high decoding accuracy of neural signals in early visual areas might arise from confounds such as global preference maps for both orientation and motion (Beckett et al., 2012; Wang et al., 2014). Thus, we used additional analyses to scrutinize the representation of motion signals in V1 and hMT that could underlie the decoding results. The motion-preference map (Fig. 8) revealed a horizontal fovea-inward bias pattern along the majority of the visual pathway except in peripheral early visual areas (V1-3), where a horizontal fovea-outward bias was observed instead. Horizontal fovea-inward biases are indicated by a pattern where vertices consistently indicate a largest response to motion towards the fovea. We observe that the left hemisphere (i.e., the right visual field) prefers leftward motion (motion towards the fovea), and the right hemisphere (i.e., the left visual field) prefers rightward motion in foveal parts of V1, V2, and V3, as well as V3AB, LO1-2, hMT, MST, and IPS0. The peripheral regions of V1, V2, and V3 show a fovea-outward bias where each vertex prefers motion away from the fovea. In addition to the predefined ROIs, an area anterior to IPS5 and an area anterior and inferior to hMT and MST also showed horizontal fovea-inward bias. Previous work showed that retinotopically biased preference contributes to the decoding of stimulus orientation (Freeman et al., 2011) and motion direction (Beckett et al., 2012; Wang et al., 2014). We see that the stimulus location only affected bias early in the visual hierarchy, which explains the high decoding accuracy in V1 for the horizontal motion condition.

4. Discussion

We presented independent motion signals to the two eyes to identify the cortical regions that underlie the transformation from 2D retinal to 3D head-centric motion signals. We identified three major clusters. In early visual cortex (V1-V3), we found no significant difference in decoding performance between stimuli that contained horizontal motion signals, which produced 3D motion percepts, and matched control stimuli that contained vertical motion signals and produced transparent 2D motion percepts, suggesting that these areas contain representations of retinal motion. In voxels in and surrounding hMT however, decoding performance was consistently superior for horizontally opposite compared to vertically opposite motion stimuli. We identified a third cluster in V3A/IPS0 that showed a similar pattern. Our results reveal the parts of the visual processing hierarchy that are critical for the transformation of 2D retinal into 3D head-centric motion signals and suggest a role for V3A/IPS0 in the representation of 3D motion signals in addition to its sensitivity to 3D object structure and static depth.

4.1. Representation of retinal motion signals in early visual cortex

We found that early visual areas V1-3, and especially V1, yield a strong representation of retinal motion and very little sign of the representation of 3D motion. The stimuli in the vertical and horizontal condition contained the same retinal motion energy. Although vertical motion does not produce a coherent 3D-motion percept, in these regions we were able to decode 2D vertical transparent motion just as well as 3D horizontal motion.

Additionally, we found a strong bias towards decoding retinal motion in V1. When oblique motion directions are misclassified, they were more likely to be classified according to the constituent retinal motion signals. For example, the response in V1 to rightward-away motion more closely resembles rightward than away motion. This makes sense if a region is at a stage of the visual processing hierarchy where the motion signals from the two eyes have yet to merge into a unified percept. Thus, our results suggest that early visual areas encode retinal motion signals, consistent with prior work on component motion signals (Born and Bradley, 2005; Britten et al., 1993; Movshon and Newsome, 1996; Rust et al., 2006).

4.2. Representation of motion signals in later regions

Our hMT results suggest a mixture of 2D retinal and 3D head-centric motion representations. The eight-way decoding result shows that hMT can decode horizontally opposite motion well, but performance dropped significantly when decoding vertically opposite retinal motion in the two eyes. hMT also did not show a bias towards retinal motion in the misclassification results, similar to IPS0.

It is important to consider what information is represented in hMT. Motion processing begins in V1 with retinal direction selectivity as well as tuning for other properties (size/spatial frequency, disparity, etc.). Downstream, higher-level motion cues must be derived, such as the interocular velocity difference (IOVD) and changing disparity over time (CD). These cues are then integrated to allow the brain to reconstruct the world around us. Motion direction can be reliably decoded throughout many of these motion-processing stages. It is important to distinguish exactly what information is exploited by the decoder. hMT showed a significantly different pattern of results compared to V1 in our study. We made a clear distinction that V1 encodes retinal motion signals and hMT does not. However, we cannot be sure that hMT fully represents 3D motion from these results alone. The same argument holds for the other regions (IPS0, V3A, and LO1) that are highlighted in our study. These regions showed a different pattern of results compared to early visual areas, but additional data are needed to elucidate the nature of the representation of motion signals in these areas.

Our findings of significant decoding of 3D motion signals using stereoscopic visual cues in IPS0 were somewhat surprising, especially in the absence of evidence for the representation of retinal motion signals. Scant previous research has suggested that IPS0 is part of the motion-processing pathway. Instead, IPS0 is a highly multisensory area and frequently appears in studies of working memory (Bray et al., 2015; Brigadoi et al., 2017), motor planning (Buneo and Andersen, 2006), spatial representation (Makin et al., 2007), and multimodal integration (Ladda et al., 2020; Regenbogen et al., 2018). Future work is needed to determine if the area's involvement is purely visual, or contains a representation of world space across multiple different modalities, such as auditory cues that signal 3D motion.

The remaining question circles back to potential differences in the representations in IPS0 and the classical motion region hMT. Representation of binocular cues in hMT is critical for many visual tasks other than motion perception, such as structure from motion and depth perception (Backus et al., 2001; Bradley et al., 1998; Neri et al., 2004; Orban et al., 1999; Preston et al., 2008). These cues become meaningful after the integration across the two retinas and can be used by multiple other regions such as MST, LO1, V3A, and IPS0.

Could the decoder have exploited alternative sources of information? One possibility to consider is binocular disparity, particularly because our horizontal-motion stimuli contain (horizontal) binocular disparities, whereas our vertical motion stimuli do not. It is therefore a reasonable concern that the observed differences in decoding performance between horizontal and vertical motion condition were due to binocular disparity rather than motion cues, especially in regions like V3A and hMT, that are known to be sensitive to static binocular disparity signals. However, these binocular disparity signals by themselves are insufficient to discriminate between the presented motion directions. For example, rightward- and leftward-motion stimuli contained identical disparity cues, but were nonetheless discriminable by our decoder. Regions that only encode static depth and not 3D motion would not produce superior decoding accuracy in the horizontal condition.

A second potential contributor to differences in decoding accuracy between conditions is the arrangement of monocular and binocular neurons in different visual areas. Regions containing predominantly monocular neurons will have less conflicting signal within each neuron. In the most extreme case, half of the neurons in this region receive input from the left eye and the other half receive input from the right eye, then MVPA could exploit this eye-specific information to make classification of the 3D motion direction even if this region does not contain representations of 3D motion direction. A difference in the proportion of monocular neurons between regions could result in the difference in decoding performance. However, this contributor is most impactful in single-neuron studies. With fMRI, the voxel is the smallest unit of signals, and each voxel contains hundreds of thousands of neurons. It would be much more difficult for the MVPA classifier to exploit the arrangement of monocular neurons at the voxel level. On the other hand, if our classifier was indeed exploiting eye-specific motion signals, then we would not see a difference in decoding performance between the horizontal and vertical motion conditions as the motion energy contained within each eye was matched.

In conclusion, multivariate pattern analysis methods have been highly successful in decoding sensory input based on neural signals. However, it has not always been clear in previous work exactly what signals any particular decoder exploits. In the domain of orientation and motion specifically, previous work has shown that artifacts (vignetting) may drive decoding performance. Here we provide one way to dig deeper. By carefully manipulating visual stimuli and evaluating the impact on decoder performance across conditions, we can start to identify the representational transformations in cortex that must underlie behavior. We hope to extend this method to other stimulus properties and modalities and move beyond the reconstruction of sensory input from neural signals and towards identifying the representational transformations that underlie perceptual inference.

Data availability

Behavioral and neuroimaging data is available at https://openneuro.org/datasets/ds004443/versions/1.0.0.

Credit authorship contribution statement

Puti Wen: Conceptualization, Data curation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Michael S. Landy:** Supervision, Formal analysis, Writing – review & editing. **Bas Rokers:** Supervision, Conceptualization, Formal analysis, Writing – review & editing.

Acknowledgements

Support: Global PhD Student Fellowship to PW. NIH EY08266 to MSL. We thank Ari Rosenberg and Lowell Thompson for valuable comments on the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.119909.

References

Albright, T.D., Desimone, R., Gross, C.G., 1984. Columnar organization of directionally selective cells in visual area MT of the macaque. J. Neurophysiol. 51, 16–31.
Armendariz, M., Ban, H., Welchman, A.E., Vanduffel, W., 2019. Areal differences in depth cue integration between monkey and human. PLoS Biol. 17, e2006405.
Backus, B.T., Fleet, D.J., Parker, A.J., Heeger, D.J., 2001. Human cortical activity correlates with stereoscopic depth perception. J. Neurophysiol. 86, 2054–2068.
Ban, H., Preston, T.J., Meeson, A., Welchman, A.E., 2012. The integration of motion and disparity cues to depth in dorsal visual cortex. Nat. Neurosci. 15, 636–643.
Beckett, A., Peirce, J.W., Sanchez-Panchuelo, R.-M., Francis, S., Schluppeck, D., 2012. Contribution of large scale biases in decoding of direction-of-motion from high-resolution fMRI data in human early visual cortex. Neuroimage 63, 1623–1632.
Benson, N.C., Winawer, J., 2018. Bayesian analysis of retinotopic maps. Elife 7. doi:10.7554/eLife.40224.
Born, R.T., Bradley, D.C., 2005. Structure and function of visual area MT. Annu. Rev. Neurosci. 28, 157–189.
Borst, A., Egelhaaf, M., 1989. Principles of visual motion detection. Trends Neurosci. 12, 297–306.

Bradley, D.C., Chang, G.C., Andersen, R.A., 1998. Encoding of three-dimensional structure-from-motion by primate area MT neurons. Nature 392, 714–717.
Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.
Bray, S., Almas, R., Arnold, A.E.G.F., Iaria, G., MacQueen, G., 2015. Intraparietal sulcus activity and functional connectivity supporting spatial working memory manipulation. Cereb. Cortex 25, 1252–1264.
Brigadoi, S., Cutini, S., Meconi, F., Castellaro, M., Sessa, P., Marangon, M., Bertoldo, A., Jolicœur, P., Dell'Acqua, R., 2017. On the role of the inferior intraparietal sulcus in visual working memory for lateralized single-feature objects. J. Cogn. Neurosci. 29, 337–351.
Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., Movshon, J.A., 1996. A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis. Neurosci. 13, 87–100.
Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1993. Responses of neurons in macaque MT to stochastic motion signals. Vis. Neurosci. 10, 1157–1169.
Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1992. The analysis of visual motion: a comparison of neuronal and psychophysical performance. J. Neurosci. 12, 4745–4765.
Buneo, C.A., Andersen, R.A., 2006. The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. Neuropsychologia 44, 2594–2606.
Carlson, T.A., 2014. Orientation decoding in human visual cortex: new insights from an unbiased perspective. J. Neurosci. 34, 8373–8383.
Czuba, T.B., Huk, A.C., Cormack, L.K., Kohn, A., 2014. Area MT encodes three-dimensional motion. J. Neurosci. 34, 15522–15533.
Czuba, T.B., Rokers, B., Guillet, K., Huk, A.C., Cormack, L.K., 2011. Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. J. Vis. 11(10):18.
Czuba, T.B., Rokers, B., Huk, A.C., Cormack, L.K., 2010. Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. J. Neurophysiol. 104, 2886–2899.
DeAngelis, G.C., Newsome, W.T., 1999. Organization of disparity-selective neurons in macaque area MT. J. Neurosci. 19, 1398–1415.
DeYoe, E.A., Felleman, D.J., Van Essen, D.C., McClendon, E., 1994. Multiple processing streams in occipitotemporal visual cortex. Nature 371, 151–154.
Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack, R.A., Gorgolewski, K.J., 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods 16, 111–116.
Freeman, J., Brouwer, G.J., Heeger, D.J., Merriam, E.P., 2011. Orientation decoding depends on maps, not columns. J. Neurosci. 31, 4792–4804.
Georgieva, S., Peeters, R., Kolster, H., Todd, J.T., Orban, G.A., 2009. The processing of three-dimensional shape from disparity in the human brain. J. Neurosci. 29, 727–742.
Gorgolewski, K.J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M.M., Churchill, N.W., Cohen, A.L., Craddock, R.C., Devenyi, G.A., Eklund, A., Esteban, O., Flandin, G., Ghosh, S.S., Guntupalli, J.S., Jenkinson, M., Keshavan, A., Kiar, G., Liem, F., Raamana, P.R., Raffelt, D., Steele, C.J., Quirion, P.O., Smith, R.E., Strother, S.C., Varoquaux, G., Wang, Y., Yarkoni, T., Poldrack, R.A., 2017. BIDS apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS Comput. Biol. 13, e1005209.
Grunewald, A., Bradley, D.C., Andersen, R.A., 2002. Neural correlates of structure-from–motion perception in macaque V1 and MT. J. Neurosci. 22, 6195–6207.
Hedges, J.H., Gartshteyn, Y., Kohn, A., Rust, N.C., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 2011. Dissociation of neuronal and psychophysical responses to local and global motion. Curr. Biol. 21, 2023–2028.
Héjja-Brichard, Y., Rima, S., Rapha, E., Durand, J.-B., Cottereau, B.R., 2020. Stereomotion processing in the nonhuman primate brain. Cereb. Cortex 30, 4528–4543.
Hubel, D.H., Wiesel, T.N., 1974. Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. J. Comp. Neurol. 158, 295–305.
Joo, S.J., Czuba, T.B., Cormack, L.K., Huk, A.C., 2016. Separate perceptual and neural processing of velocity- and disparity-based 3D motion signals. J. Neurosci. 36, 10791–10802.
Kamitani, Y., Tong, F., 2006. Decoding seen and attended motion directions from activity in the human visual cortex. Curr. Biol. 16, 1096–1102.
Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., 2007. What's new in psychtoolbox-3. Perception 36, 1–16.
Ladda, A.M., Wallwork, S.B., Lotze, M., 2020. Multimodal sensory-spatial integration and retrieval of trained motor patterns for body coordination in musicians and dancers. Front. Psychol. 11, 576120.
Li, X., Morgan, P.S., Ashburner, J., Smith, J., Rorden, C., 2016. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. J. Neurosci. Methods 264, 47–56.
Likova, L.T., Tyler, C.W., 2007. Stereomotion processing in the human occipital cortex. Neuroimage 38, 293–305.
Liu, J., Newsome, W.T., 2005. Correlation between speed perception and neural activity in the middle temporal visual area. J. Neurosci. 25, 711–722.
Makin, T.R., Holmes, N.P., Zohary, E., 2007. Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. J. Neurosci. 27, 731–740.
Maunsell, J.H., 1992. Functional visual streams. Curr. Opin. Neurobiol. 2, 506–510.
Maunsell, J.H., van Essen, D.C., 1983. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. J. Neurosci. 3, 2563–2586.
Movshon, J.A., Newsome, W.T., 1996. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J. Neurosci. 16, 7733–7741.
Nadler, J.W., Angelaki, D.E., DeAngelis, G.C., 2008. A neural representation of depth from motion parallax in macaque visual cortex. Nature 452, 642–645.
Neri, P., Bridge, H., Heeger, D.J., 2004. Stereoscopic processing of absolute and relative disparity in human visual cortex. J. Neurophysiol. 92, 1880–1891.
Orban, G.A., Sunaert, S., Todd, J.T., Van Hecke, P., Marchal, G., 1999. Human cortical regions involved in extracting depth from motion. Neuron 24, 929–940.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–442.
Preston, T.J., Li, S., Kourtzi, Z., Welchman, A.E., 2008. Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. J. Neurosci. 28, 11315–11327.
Qian, N., Andersen, R.A., Adelson, E.H., 1994. Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. J. Neurosci. 14, 7357–7366.
Regenbogen, C., Seubert, J., Johansson, E., Finkelmeyer, A., Andersson, P., Lundström, J.N., 2018. The intraparietal sulcus governs multisensory integration of audiovisual information based on task difficulty. Hum. Brain Mapp. 39, 1313–1326.
Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. Neuroimage 61, 1402–1418.
Rokers, B., Cormack, L.K., Huk, A.C., 2009. Disparity- and velocity-based signals for three--dimensional motion perception in human MT+. Nat. Neurosci. 12, 1050–1055.
Rust, N.C., Mante, V., Simoncelli, E.P., Movshon, J.A., 2006. How MT cells analyze the motion of visual patterns. Nat. Neurosci. 9, 1421–1431.
Salzman, C.D., Britten, K.H., Newsome, W.T., 1990. Cortical microstimulation influences perceptual judgements of motion direction. Nature 346, 174–177.
Sanada, T.M., DeAngelis, G.C., 2014. Neural representation of motion-in-depth in area MT. J. Neurosci. 34, 15508–15521.
Serences, J.T., Boynton, G.M., 2007. The representation of behavioral choice for motion in human visual cortex. J. Neurosci. 27, 12893–12899.
Snowden, R.J., Treue, S., Erickson, R.G., Andersen, R.A., 1991. The response of area MT and V1 neurons to transparent motion. J. Neurosci. 11, 2768–2785.
Tresilian, J.R., 1995. Perceptual and cognitive processes in time-to-contact estimation: analysis of prediction-motion and relative judgment tasks. Percept. Psychophys. 57, 231–245.
van Bergen, R.S., Jehee, J.F.M., 2021. TAFKAP: An improved method for probabilistic decoding of cortical activity. bioRxiv doi:10.1101/2021.03.04.433946.
Van Essen, D.C., Newsome, W.T., Maunsell, J.H., Bixby, J.L., 1986. The projections from striate cortex (V1) to areas V2 and V3 in the macaque monkey: asymmetries, areal boundaries, and patchy connections. J. Comp. Neurol. 244, 451–480.
Wang, H.X., Merriam, E.P., Freeman, J., Heeger, D.J., 2014. Motion direction biases and decoding in human visual cortex. J. Neurosci. 34, 12601–12615.
Wuerger, S., Shapley, R., Rubin, N., 1996. on the visually perceived direction of motion" by Hans Wallach: 60 years later. Perception 25, 1317–1367.