

## **Global Motion Processing in Human Visual Cortical Areas V2 and V3**

**Michele Furlan and Andrew T Smith**

Department of Psychology, Royal Holloway, University of London, Egham TW20 0EX, UK

Abbreviated title:

Global motion processing in human V2 and V3

Corresponding author: Dr Andy Smith, address as above

Email: [a.t.smith@rhul.ac.uk](mailto:a.t.smith@rhul.ac.uk)

Conflict of interests: none

Acknowledgement: We thank Dr Angelika Lingnau for valuable comments on the manuscript.

## **Abstract**

Global motion perception entails the ability to extract the central direction tendency from an extended area of visual space containing widely disparate local directions. A substantial body of evidence suggests that local motion signals generated in primary visual cortex (V1) are spatially integrated to provide perception of global motion, beginning in the middle temporal area (MT) in macaques and its counterpart hMT in humans. However, V2 and V3 also contain motion-sensitive neurons that have larger receptive fields than those found in V1, giving the potential for spatial integration of motion signals. Despite this, V2 and V3 have been overlooked as sites of global motion processing. To test, free of local-global confounds, whether human V2 and V3 are important for encoding global motion, we developed a visual stimulus that yields a global direction, yet includes all possible local directions and is perfectly balanced at the local motion level. We then attempted to decode global motion direction in such stimuli with multivariate pattern classification of fMRI data. We found strong sensitivity to global motion in hMT, as expected, and also in several higher visual areas known to encode optic flow. Crucially, we found that global motion direction could be decoded in human V2 and, particularly, in V3. The results suggest the surprising conclusion that global motion processing is a key function of cortical visual areas V2 and V3. A possible purpose is to provide global motion signals to V6.

*Keywords:* fMRI; MVPA; visual; V2; V3; motion

## **Significance statement**

Humans can readily detect the overall direction of movement in a flock of birds despite large differences in the directions of individual birds at a given moment. This ability to combine disparate motion signals across space underlies many aspects of visual motion perception and has therefore received considerable research attention. The received wisdom is that spatial integration of motion signals occurs in the motion complex MT+ in both human and non-human primates. We show here that areas V2 and V3 in humans are also able to perform this function. We suggest that different cortical areas integrate motion signals in different ways for different purposes.

## Introduction

Motion within the visual image is initially detected locally, by neurons in primary visual cortex (V1) that have small receptive fields. In order to perceive coherent movement of larger regions within the image, it is necessary to integrate local motion signals over space. The ability to do this is central both to perceiving rigid motion of large objects and to perceiving the overall trend in large non-rigid motions such as the direction of flow of a river or a flock of birds. Spatial integration of local motion signals also provides the basis for sensitivity to optic flow patterns that are used for perception of self-motion. Psychophysical studies show that we are able to integrate motion signals over space and extract the overall direction of motion efficiently even when the local signals are very disparate or noisy (Williams and Sekuler, 1984; Newsome and Paré, 1988).

Many studies suggest that in both humans and other primates, spatial integration of motion signals occurs in area MT, which receives strong, direct input from V1. In macaques, most MT neurons are direction-sensitive (Zeki, 1974), suggesting specialization for motion, and receptive fields are larger than in V1, suggesting spatial integration of V1 inputs. Direct evidence for the involvement of macaque MT in global motion perception comes from the finding that lesions of MT impair global motion perception with noisy stimuli (Pasternak and Merigan, 1994) and that microstimulation of MT can bias perceptual judgements of global motion direction (Salzman et al., 1992). Neuroimaging studies suggest that human MT (hMT) may be broadly homologous with macaque MT, being the first area in the processing hierarchy that is specialized for motion processing. Specifically, it is thought that hMT may be a key site for spatial integration of local motion direction signals, principally because of evidence that the MT complex responds more strongly to coherent global motion than to incoherent motion, whereas V1 responds equally well or better to incoherent motion (Rees et al., 2000; Braddick et al., 2001; Helfrich et al., 2013).

In macaques, MT receives input from V1 not only through a direct connection but also indirectly via V2 and V3, which both project to MT (Felleman et al., 1997; Gattass et al.,

1997). It has long been known that a significant minority of neurons in both V2 and V3 are direction-sensitive (e.g. Van Essen and Zeki, 1978; Zeki, 1978). Indeed, Gegenfurtner et al. (1997) reported that V3 is more direction-sensitive than its afferent input, suggesting that direction selectivity may be generated there *de novo*. Moreover, several physiological properties of V2 and V3 other than direction sensitivity are reminiscent of the properties of neurons in MT and are therefore suggestive of motion processing. V2 neurons respond, on average, to lower spatial frequencies than V1 neurons and are more tuned for temporal frequency (Foster et al., 1985). V3 neurons respond to lower spatial and higher temporal frequencies than even V2 (Gegenfurtner et al., 1997). V3 also has much higher contrast sensitivity than either V1 or V2, in this respect resembling MT (Gegenfurtner et al., 1997). V2 and V3 both project to, among others, V6 (Colby et al., 1988), which is strongly specialized for motion processing (Pitzalis et al., 2013a).

When the properties of macaque V2 and V3 are considered, it becomes likely that the notion of a simple linear processing hierarchy (V1→MT→MST and beyond) for motion may be simplistic. Specifically, with their larger receptive fields relative to V1 and their motion sensitivity, V2 and V3 may have an important role in processing global motion, based on local motion signals from V1. Here, we test this possibility in the human brain by examining the extent to which V2 and V3 are able to signal global motion direction. For comparison, we apply the same methods to V1, hMT and also several higher visual areas (hMST, pVIP, V6 and CSv) that are thought to be involved in processing optic flow during locomotion.

## **Methods**

### *Participants*

Five healthy volunteers took part (three male). Each was scanned on four occasions. All had normal or corrected-to-normal vision. They were screened for MRI contraindications according to standard procedures and written consent was obtained. The procedure was approved by the relevant local research ethics committee.

### *Stimuli*

Computer generated visual stimuli were projected by a LCD projector onto a rear-projection screen at the end of the scanner bore and were viewed via a mirror mounted on the head coil, giving an image of about 25 x 20 deg visual angle. The stimuli were created using a combination of OpenGL, MATLAB (The Mathwork, Inc), ASF (Schwarzbach, 2011) and Psychtoolbox-3 (Brainard, 1997).

In the main experiment, the stimuli were random-dot kinematograms and were designed to allow global motion direction to be manipulated without altering the balance of local motion directions. In principle this permits the mechanisms of global motion integration to be isolated, but creating such a stimulus poses a significant challenge because global motion direction arises entirely from the population of local directions, so it is hard to change one without changing the other. The most widely used global motion stimulus contains random dots of two types: signal dots that have a common direction (defining the global direction) and noise dots that have random directions (Newsome and Paré, 1988). This stimulus permits parametric manipulation of signal strength and has proved a powerful tool for examining global motion sensitivity in both physiological and psychophysical contexts. However, it is not well suited to the exploration of global motion sensitivity based on multi-voxel pattern analysis (MVPA) of fMRI data. This is because when the direction of global motion (signal direction) is changed, there is a concomitant change in the distribution of local motion directions: a bias in one direction is replaced by a bias in a different direction. If responses to two global motion directions are successfully decoded, it is impossible to know whether decoding was based on the change in global direction *per se* or on the change in the preponderant local direction. Thus, decoding would not necessarily indicate that local directions have been integrated over space. An alternative stimulus (Williams and Sekuler, 1984) has no noise dots but instead employs a rectangular distribution of dot directions centered on the global motion direction. Again, when global direction is changed there is a correspondingly large change to the distribution of local dot directions and this could account for decoding performance. No satisfactory solution to this problem has previously been found. We have developed a novel approach that fully solves the problem, albeit only for stimuli in which global motion is weak.

The stimuli consisted of a circular patch (10 deg diameter) containing 300 high-contrast white dots (each 0.3 deg diameter, approximate luminance 700 cd/m<sup>2</sup>) on a dark background. The dots were initially positioned randomly. Each dot moved along an axis that was chosen randomly but remained constant throughout a 15s stimulus block (see *Design*). The axis was chosen independently for each dot and all possible direction axes were represented, with equal probability (Figure 1). Each dot moved along the selected axis for 500ms, then reversed direction for 500ms, reversed again, and so on, resulting in oscillation back and forth over a short distance. If the temporal phase of oscillation of each dot were assigned randomly, this stimulus would appear as chaotic motion with no global direction, but by synchronizing temporal phase, global motion could be created based on the fact that the brain pools instantaneous local directions over a wide range in order to extract the overall global motion (Williams and Sekuler, 1984).

There were two versions of the stimulus, Class 1 and Class 2. These provided two classes for decoding with multi-voxel pattern analysis (MVPA). Both had the same set of local motion oscillations, with all directions represented. The construction of the stimuli is illustrated in Figure 1. All dots reversed their directions synchronously. Each could take either of two temporal phases, determined by its motion axis, as follows. In a scheme where 0 deg is upward, 90 is rightward, 180 downward and 270 leftward, each dot was assigned a 'forward' motion direction between 0 and 180 deg. The 'reverse' direction was opposite this value i.e between 180 and 360. Each dot was nominally given one of two labels, according to the axis of motion that had been randomly assigned to it. Those dots with forward directions between 0 and 90deg were considered type A for the purpose of assigning their phases and the remainder (90-180) were considered type B. Stimulus classes 1 and 2 differed as follows. In Class 1, each dot moved alternately in its allocated forward and reverse directions for 500ms each, all dots moving forward at the same time. This caused the range of dot directions in the stimulus to alternate between 0-180 (for 500ms) and 180-360 (for 500ms). This appeared as noisy global motion that alternated between rightward (90deg, the mean of the 0-180 range) and leftward (270, the mean of the 180-360 range). Class 2 was derived from Class 1 and comprised exactly the same set of dots, dot positions and motion directions. However, for all type B dots the temporal phase of oscillating motion was reversed with respect to the type A dots (when A is in the forward phase, B is in the reverse phase and vice versa). The effect of

this manipulation is that the range of directions was 90 to 270 in one phase and 270 to 90 (through 0) in the other. Global motion was perceived that alternated between downward in one phase and upward in the other. Thus, the same set of local dot motions could yield either vertical or horizontal global motion. Because of the wide (180deg) range of directions, global motion was very noisy and quite weak, but was discernable. Psychophysical testing showed that naïve observers who judged whether global motion was horizontal or vertical achieved an average of 82% correct (n=5 observers) when presented with 100 examples of such stimuli with no prior practice. Movie 1 illustrates the construction of the stimuli and Movie 2 shows the stimuli.

Global motion here means the integration of local motion signals over space such as to extract an overall trend in a large, spatially noisy stimulus. The purpose of our experiment was to test whether visual areas V2 and V3 are capable of this type of spatial integration. As indicated earlier, the rationale for using the particular stimuli described relates to a key problem of MVPA, that the stimulus classes may all-too-easily differ in more than one respect, in which case it is impossible to know which difference provides the basis for decoding. The experimental paradigm used here relies on temporal integration of BOLD signals over many stimulus cycles (see *Design*), in order to obtain a response that reflects equal contributions from all possible local directions. The stimulus alternation cycle is fast (1Hz) relative to the notoriously sluggish BOLD response (time-to-peak about 6s), so responses to successive 0.5s presentations of the two global directions in each stimulus are integrated. Moreover, because the BOLD response is modelled as the response to many stimulus alternations, temporal integration will occur even if BOLD does follow the alternations to some extent. Thus, the BOLD response averages the 180deg range of local directions presented during one 0.5s phase of the alternation with the remaining 180deg, presented during the other phase, giving integration over 360deg. Global motion signals are also temporally integrated by the BOLD signal but only for the two directions along a single axis. This axis differs between the two stimulus classes whereas the range of local directions contributing to BOLD does not differ (360deg in both cases). Indeed, each individual dot oscillates back and forth along a random axis that is the same in both classes, so the stimuli are matched dot-for-dot at the local level. Thus, at the expense of evoking only weak global motion perception, our stimuli ensure that if the BOLD responses to the two

classes can be decoded (distinguished), it must be on the basis of global motion, not local motion.

There is one other stimulus quality that could in principle provide a basis for decoding. This is the difference between the temporal direction-reversal phases of Type A dots and Type B dots. It is extremely difficult to imagine that the brain might detect and encode this difference. Since dots are allocated to Types A and B arbitrarily, they are spatially intermingled. Each set contains all directions, so A and B cannot be identified in terms of directions present. To decode the stimuli based on temporal phase it would be necessary to group together the set of dots that move on axes between the 0-180 axis and the 90-270 axis and encode their phase, group the set of dots moving on axes between 90-180 and 270-360 and encode their phase, then detect that the relationship between these phases differs between Class 1 and Class 2 (the stimuli to be decoded). Some neurons would need to preferentially respond to one of these arbitrary phase combinations and other neurons to the other. This seems highly improbable.

It may be useful to reiterate why this rather elaborate stimulus scheme is necessary. If conventional motion coherence stimuli were used, a set of 'signal' dots with a common direction would be presented among 'noise' dots having random directions. The two classes would have different signal directions, vertical and horizontal. In each case, the signal dots create a bias in the distribution of local motion directions and this bias covaries with global motion direction. It would not be possible to know whether decoding was truly based on global motion direction or instead reflected the predominant local dot direction. Whereas in MVPA with standard coherent dot stimuli, classification performance might reflect the differences in either global or local motion, in our experiment it is impossible to inadvertently decode local dot direction.

### *Design*

Throughout each scan run, lasting approx. 5 mins, participants fixated centrally. To divert attention from the motion stimuli and maintain a constant attentional state, a demanding letter identification task was carried out at fixation. A continuous random letter stream appeared, the letter changing at 2Hz. Moving dots were masked off



immediately around the letter to avoid overlap. The participant searched the stream for the occurrence of the letter E. On seeing E, they incremented a mental count and they reported the final count verbally at the end of the scan run.

Each oscillating motion stimulus lasted for 15s. Stimulus blocks were separated by 7.5s blocks in which there was no stimulus other than the fixation task. Each scan run contained 12 stimulus blocks that alternated between Class A (horizontal global motion) and Class B (vertical global motion). At the beginning of each block, a new set of random dot positions and motion axes was generated and used throughout the block. Each run commenced and ended with a 15s period with no stimulus except fixation. Each participant completed 8 runs, giving 48 blocks of each Class in total.

#### *Data acquisition*

Data were acquired with a 3T Siemens TIM Trio MR scanner with a 32-channel array head coil. Functional images were acquired with T2\* weighted gradient-recalled echo planar imaging (EPI) sequence (31 slices, TR 2500 ms, TE 31 ms, flip angle 85°, voxel size 2.5 mm isotropic). Parallel imaging (GRAPPA, acceleration factor 2) was used. In each scan session (see below), structural data were acquired using a T1-weighted 3D anatomical scan (MPRAGE, Siemens; TR 1830ms, TE 5.56ms, flip angle 11°, resolution 1x1x1 mm).

#### *Eye movement recording*

Eye position was continuously monitored with infrared video photography. The image from a camera positioned close to the left eye (NordicNeuroLab, Norway) was fed to pupil-detection software (Arrington Inc, USA) and x/y position was sampled at 60Hz. A short calibration run, in which eye movements of known size were made, was conducted at the beginning of the experiment and again at the end.

#### *Functional localizers*

Functional data were analysed in terms of mean activity across all the voxels within each of a number of visual areas defined on the basis of separate localizer scans. Each participant was scanned on four occasions, once for the main experiment and three

times for visual localisers. On one occasion, hMT and hMST were defined, based on a standard method (Dukelow et al., 2001; Huk et al., 2002) A circular patch of dots (8° diameter) was presented with its centre placed 10° to the left or right of fixation. Blocks of 15s in which the dots were static were alternated with blocks of 15s in which the dots moved alternately inward and outward along the radial axes, creating alternating contraction and expansion. Sixteen blocks (eight static and eight moving) were presented in each scan run; one run was completed with the stimulus on the left and another with it on the right. With this procedure, two regions that have been called hMT and hMST can be differentiated in terms of the absence or presence, respectively, of ipsilateral drive. It is likely that “hMST” comprises two or more regions that respond to motion and have large receptive fields, but further refinement requires demanding high-resolution mapping techniques (Kolster et al., 2010) that are beyond the scope of this study. In the same session, a 3D anatomical scan with high contrast between grey and white matter (MDEFT, (Deichmann et al., 2004) was acquired for use in retinotopic mapping.

On another occasion, occipital areas V1, V2, V3 and V3A were identified with a standard retinotopic mapping procedure employing an 8Hz counterphasing checkerboard wedge stimulus (a 24° sector) of radius 12°. Check size was scaled by eccentricity in approximate accordance with the cortical magnification factor. The wedge rotated clockwise at a rate of 64 s/cycle and 8 cycles were presented. This stimulus was presented twice to each participant, and the data from the two scan runs were averaged.

A third localizer was used to identify areas hV6, pVIP and CSv (Wall and Smith, 2008; Cardin and Smith, 2010) This consisted of two time-varying optic flow stimuli (light dots on a dark background). The first was egomotion-compatible optic flow that cycled through spiral space to simulate back-and-forth spiral motion of the observer. The second was an egomotion-incompatible 3x3 array of similar spiral motions. Each stimulus was presented for 3 s in an event-related design, with inter-trial intervals (ITIs) in which the screen was blank apart from a central fixation spot. The ITIs varied between 2 and 10 s, following a Poisson probability distribution. Each scan run had 32 trials (16 per condition) presented in a pseudorandom order, and lasted approx. 5 mins. Six such scan runs were conducted. Participants were continuously engaged in a colour

counting task at fixation. Contrasting the activity elicited by these two stimuli isolates regions (hV6, pVIP and CSv) that favour egomotion-compatible flow from those that respond well to any flow stimulus. CSv is as originally defined in the human brain (Wall and Smith, 2008) and it is unknown whether a counterpart exists in macaque brain. Area hV6 corresponds closely to human V6 of Pitzalis (2006) and it seems likely that it has functions and connections that are similar to macaque V6 (Pitzalis et al., 2013a). The status of pVIP (putative VIP) is less certain. It appears to be the same region as human VIP of Bremmer et al (2001), who suggested that it may be homologous with macaque VIP. It is possible that pVIP corresponds to IPS4 of Swisher et al (2007) or IPS5 of Konen et al (2008).

### *Data analysis*

Data were analysed using Brain Voyager QX 2.3 (BrainInnovation, The Netherlands), MATLAB (The Mathwork Inc, USA) and R (R Foundation for Statistical Computing). The first 2 volumes of each run were discarded. 3D motion correction and slice time correction were performed. The data were temporally high-pass filtered at 3 cycle/run (approx. 0.01Hz). The pre-processed EPI scans were then co-registered with the anatomy. Finally, both functional and anatomical data were aligned into AC-PC space. The pre-processed data were analysed within the General Linear Model (GLM), separately for each participant.

For the localizers, each motion condition was modelled separately, with a regressor formed by convolving the stimulus time-course with a canonical hemodynamic impulse response function (HRF) and then scaling to unity. Head motion regressors were also included. For the retinotopic mapping data, the temporal phase of the response to the rotating wedge at each voxel was obtained by fitting a model to the time-series. Phases were superimposed as colours on a segmented and flattened representation of the grey matter derived from the MDEFT scan. Phase was taken as an indicator of visual field position in terms of polar angle and the boundaries of V1, V2, V3 and V3A were drawn by eye using conventional criteria. The hMST localizer data were analysed by fitting a model and the results were superimposed on the flattened grey matter representation as a *t*-map. hMST was defined as a cluster of voxels at the expected location that

responded significantly to ipsilateral motion (Smith et al., 2006). For the third localizer, to localize CSv, hV6 and pVIP, separate regressors were fitted to the blocks of the two types (single motion patch versus 9 patches), according to a standard procedure developed in our lab (Wall and Smith, 2008; Cardin and Smith, 2010) in which a cluster of voxels that was significantly more strongly activated by egomotion-compatible than egomotion-incompatible motion was identified in each of three expected locations.

The multivariate analysis was based on exemplars that consisted of beta values (effect sizes) from GLM analysis conducted as above. The pre-processed time-courses at each voxel were averaged across runs. Multivoxel pattern analysis (MVPA) was then performed on the beta values derived from these averaged time-courses. Inclusion of voxels as features in the MVPA was based on the ROIs and a separate analysis was performed for each ROI. A limitation of this approach is that small visual areas such as pVIP and hV6 may contain as few as 15-20 functional voxels whereas MVPA requires a larger number of features (voxels) to be efficient. To ameliorate this problem, data were combined across participants prior to MVPA analysis (Brouwer and Heeger, 2009; Furlan et al., 2014). For each visual area, decoding performance was assessed as follows. An estimate of the response at every voxel was obtained by fitting a General Linear Model (GLM) that included a regressor to model the trial response, obtained by convolving a box-car function representing the stimulus timing with the HRF. Separate regressors modelled horizontal and vertical global motion. The resulting beta values were normalised to remove any overall difference between the two classes and then used as response values (exemplars) for the two stimuli. Decoding performance was examined for each ROI as a function of the number of features included, by progressively including more voxels, selected randomly, and repeating the analysis. For each sample size, the analysis was repeated 100 times with a different random selection of voxels and the resulting decoding performances were averaged. In the main analysis (Figures 3-8), voxels were selected from the overall pool without regard for participant-of-origin, to ensure that the largest possible number of voxels could be included. In a second analysis (Figure 9) they were selected such that the same number was taken from each participant, as a check that all participants contributed equally to the results.

For each MVPA analysis, a subset of observations was used to train the classifier, which was a Support Vector Machine (SVM) with a linear kernel. The SVM was trained to identify the optimal separating boundary (hyperplane) between the two Classes. A 'leave-one-out' method was used. Of the 8 scans, 7 were used for training and the 8<sup>th</sup> was used for testing. This was repeated 8 times, leaving out each run in turn, and the 8 performances were averaged. Finally, for each ROI, the hypothesis that the classification accuracy was different from chance level was tested by comparing it against the test accuracy on the same dataset after having randomly permuted (shuffled) the labels, which should produce chance-level accuracies with a similar variance to the main analysis. 1000 such analyses were performed with different random permutations, employing the same leave-one-out method, giving 1000 performance estimates per permutation. The 95<sup>th</sup> percentile of the distribution of permuted performance results (typically in the range 60-65% correct) was taken as a critical value for regarding unpermuted performance values as significantly above chance. GLM analysis was performed with BrainVoyager and all analyses beyond GLM (merging the ROIs, voxel selection, SVM classification) were performed with MATLAB (The Mathwork, USA) using the LIBSVM library for support vector machines (Chang and Lin, 2011).

## Results

### *Localizers*

All cortical regions of interest were successfully defined in both hemispheres of each participant. Their locations are very similar to previous studies that used the same localizer methodology and are shown in Table 1. The locations of V1, V2 and V3, the regions of central interest in this study, are illustrated as colored overlays on the medial surface of the occipital cortex in Figure 2.

	Location (left)	Location (right)
V1d	-8 -92 -2	9 -94 -2
V1v	-7 -85 -7	8 -84 -8
V2d	-13 -99 2	12 -97 4

V2v	-12 -84 -10	13 -82 -10
V3d	-19 -97 2	19 -94 7
V3v	-20 -80 -10	18 -79 -9
hMT	-42 -73 1	42 -73 1
hMST	-41 -66 2	43 -67 2
V3A	-17 -90 13	25 -88 17
CSv	-10 -32 37	10 -28 37
pVIP	-26 -56 50	28 -55 50
V6	-14 -83 23	15 -80 25

**Table 1.** Mean location (Talairach co-ordinates of centroid) of each cortical region studied, shown separately for the left and right hemispheres. The centroid co-ordinates were extracted with the algorithm provided by *BrainVoyager*.

### *Decoding global motion direction*

The ability of the SVM to decode (predict) which class a given stimulus belonged to is shown, separately for each visual area examined, as a function of the number of voxels included in the analysis, in Figures 3- 5. Also shown for each area are chance performance (50%) and a  $p < 0.05$  significance level based on permutation testing in the relevant visual area.

Figure 3 shows (left panel) the result for hMT, which provides a comparison for all other areas studied. If a cortical region is sensitive to global motion direction, it is expected that decoding performance will build rapidly as the number of exemplars (voxels) is increased and then stabilize at a high level. This pattern is clearly evident in hMT. Decoding performance reaches about 90% accuracy for higher voxel numbers, indicating strong sensitivity to global motion direction. Also shown in Figure 3 are the results for hMST, which is presumed, by analogy with macaques, to receive strong inputs from hMT, and V3A, which is also thought to be strongly motion-sensitive (Huk and Heeger, 2000; Fischer et al., 2012) and has been reported to respond more strongly to coherent than noisy motion (Braddick et al., 2001). In these regions, performance

reaches about 80% (MST) or 85% (V3A) for large voxel samples, again suggesting significant sensitivity to global motion direction. Figure 4 shows results for three regions that are strongly associated with optic flow related to self-motion, namely CSv, pVIP and hV6. As might be expected, all three show good sensitivity to direction of global motion. Decoding performance reaches 90% in CSv and about 80% in V6 and pVIP, well above the criterion for statistical significance.

The key results of the study, decoding performance for V1, V2 and V3, are shown in Figure 5. In the case of V1, performance is at chance level, as expected. However, in V2 a clear trend emerges. Performance reaches about 70%, significantly above chance at  $p < 0.05$  for the largest voxel samples. In V3, surprisingly, global motion direction could be decoded very robustly. The results were computed separately for the dorsal (V3d) and ventral (V3v) portions of V3 (in view of a long-standing debate about whether the two halves of macaque V3 have the same properties; see Discussion), and are shown superimposed. Performance in V3 reaches about 85% overall and is similar in the two sub-regions, although slightly greater in V3d than in V3v.

Strictly speaking it is the axis, not the direction, of global motion that is decoded in our analysis. However, in general, neurons in primate visual cortex that are motion-sensitive respond much more strongly to one direction of motion along the preferred axis than the other. In some cases, motion in the non-preferred direction even results in suppression. However, neurons also exist that respond well to either direction along one axis but not to motion on other axes. If the human brain is similar to non-human primates in this regard, it is likely that such bidirectional axis-specific responses contribute to decoding performance, but probable that decoding more strongly reflects direction-specific responses. Moreover, when global motion is considered from a functional perspective, it is easy to see the value of combining direction-specific outputs to yield global motion direction but harder to see the utility of computing global axis of motion with no direction label. We therefore speak of decoding global direction, even though decoding axis of motion would more accurately describe our analysis, because in our view it reflects better the likely underlying physiology.

Figure 5 suggests strongly increasing global motion sensitivity from V1 to V2 and again from V2 to V3. However, decoding performance is known to depend strongly on the magnitude of the signals being decoded (Smith et al., 2011). This makes it difficult to make reliable quantitative comparisons across different brain regions. To facilitate interpretation, we extracted the voxel-wise response magnitude (beta weights from the GLM), averaged across the voxels used for decoding in each visual area. Figure 6 shows these magnitudes, separately for each region, based on the maximum number of voxels used (250). Areas V1, V2 and V3 elicited responses of similar magnitude, so the comparison among them is probably fair and the increase in sensitivity to global motion from V1 to V2 and from V2 to V3 is probably real. As has been noted in some previous fMRI studies of visual motion, responses were substantially smaller in hMT and hMST than in V1-V3. Decoding performance (for a given degree of neural specificity) scales with response amplitude (Smith et al., 2011). Therefore the numerically superior decoding performance in hMT and hMST relative to V1-V3 occurs despite much smaller signals and so it may reflect a much greater difference in neural specificity than comparison of Figures 3 and 5 suggests. In the same vein, it should be noted that some of the visual areas where decoding performance was best also showed the weakest responses in terms of amplitude, suggesting a very high level of neural sensitivity to global motion in these areas compared even to hMT and hMST. Most strikingly, in CSv the mean amplitude of the response across all voxels was low, yet decoding was highly robust. This is quite remarkable: CSv shows the highest decoding performance of any area (jointly with hMT) and this is achieved despite being based on the smallest signal. Areas pVIP and V6 also have high decoding performance in relation to response amplitude, suggesting a high degree of global motion specificity.

Another way to evaluate the results is to take only the largest voxel sample (end point of plots in Figures 3-5) and consider the distribution of probability values obtained when decoding performance for samples of this size is compared by *t*-test with performance in permutation tests. Strong decoding performance should be reflected in the *p* values clustering at low (significant) values. The results of such an analysis are shown in Figure 7. For each visual area, 250 voxels were randomly selected. Decoding performance was evaluated for this sample with correct and randomly permuted labels. This was repeated 1000 times with a different random sample of 250 voxels and a different permutation



each time, yielding a total of 1000  $p$  values. Figure 7 shows that clustering at low  $p$  values is evident in all areas except V1. It is strongest in CSv and hMT, followed by V3A. Crucially, it clearly occurs in both V2 and V3. The clustering is quantified in Figure 8, which shows the mean  $p$  value for each area along with the proportion of  $p$  values that are less than 0.05. The results considered in this way are in line with the raw decoding performances shown in Figures 3-5. Again, however, quantitative comparison across areas is not strictly appropriate and the ordering of visual areas in this figure is only indicative.

The voxel samples used for MVPA were selected randomly from the available pool without regard for the participant of origin. Because the size of a given ROI inevitably varies among participants, this means that the participants did not contribute equal numbers of voxels to the analyses. It is therefore important to check that the results shown in Figures 3-5 were not driven by one or two atypical participants. To do this, we repeated the entire analysis with the additional constraint that all participants contributed equally. Voxels were again selected randomly, independently for each iteration. This constraint meant that the maximum number of voxels that could be used was that available from the participant with the smallest ROI. For large ROIs, such as V1, this constraint was easily accommodated but for small areas it was more severe, most notably in pVIP. The results of this refined MVPA analysis are shown in Figure 9 for all nine visual regions examined. In every case, the results are in line with the full analysis shown in Figures 3-5, demonstrating that the results do not reflect a small number of rogue participants. As a further check for consistency of performance across participants, we conducted analyses based on data for each individual participant for the crucial areas (V2 and V3), together with V1 for comparison. In these relatively large areas there were sufficient voxels to support such analyses. The results are shown in Figure 10. Decoding performance is similar to that obtained with pooled data (Figure 5) and the variance across participants is satisfactory.

Finally, it is essential to establish that the decoding performance we obtained was not based on differences in eye movements between conditions. If participants tracked the global motion, the eyes would move primarily horizontally in one condition and vertically in the other. Any brain region that is influenced by eye movements might then

appear to decode global motion direction when actually decoding eye movement direction. To test this possibility, we computed and compared the variances of horizontal and vertical eye position, after editing the eye traces to remove blinks and occasional saccades away from fixation. Good traces could not always be obtained and measurements were based only on periods when the quality was high, so that small following eye movements would not be obscured by instability of the pupil measurement. The variance of eye position was low, indicating that fixation was generally very good. Table 2 shows the breakdown for vertical and horizontal eye position during vertical and horizontal global motion. Also shown for comparison is the variance during inter-trial intervals when no motion was present. The four values obtained during motion presentation are similar to each other and to the values obtained with no motion. There is a reversal of the direction of the difference between the two motion conditions but the differences are very small and variance is greater for eye movement orthogonal to global motion than parallel to it, the opposite of the expected pattern due to following eye movements. We conclude that there were no detectable following movements and that decoding must have been based on the stimulus itself.

	Horizontal eye stability (SD)	Vertical eye stability (SD)
Horizontal global motion	0.28	0.32
Vertical global motion	0.36	0.27
No motion	0.29	0.32

**Table 2.** Stability (standard deviation in degrees) of the x and y components of eye position during horizontal and vertical global motion trials and during inter-trial intervals.

## Discussion

In this study, sensitivity to direction of global motion was examined by perfectly balancing local motion across stimuli that had different global directions, then conducting MVPA on the BOLD responses they elicited. The principal purpose of the

study was to test the existence of a previously overlooked role of V2 and V3 in encoding global motion.

Global motion direction could not be decoded from responses in V1. Neurons in the LGN are not direction-sensitive in macaques and are not thought to be so in humans. Therefore V1 is the first visual region with direction-sensitive neurons and is the site of initial local motion detection. Global motion perception is thought to arise by subsequent integration of local motion signals arising in V1. Many psychophysical studies have characterized this process (e.g. Hiris and Blake, 1995; Burr et al., 1998; Edwards and Badcock, 1998; Smith et al., 1999) and physiological studies have demonstrated and characterized integration of local motion signals in MT neurons (e.g. Newsome et al., 1989; Britten et al., 1993). In a two-stage scheme of this kind, global motion sensitivity is not expected in V1, and was not found. In contrast, strong evidence of global motion sensitivity was found in hMT, in line with expectations. The striking feature of our results is that global motion direction could be decoded in V3 and even, albeit more weakly, V2 (Figure 5). Superficially, this is surprising, since even local motion sensitivity is not particularly strongly associated with V2 or V3. In macaques, direction-sensitive neurons are found in both regions but most reports suggest that they exist only in similar proportions to V1, on average about 10-15%, depending on the study and measurement method (e.g. Van Essen and Zeki, 1978; Zeki, 1978; Baizer, 1982), giving no reason to think that either area is specialised for motion processing. Some studies have suggested higher proportions of direction-selectivity in V2 and V3. For example Foster et al. (1985) reported that 38% of cells are direction-selective in macaque V2 compared to 20% in V1, while Felleman et al. (1987) and Gegenfurtner et al. (1997) both reported that around 40% of cells are direction sensitive in V3. In view of these physiological studies, we might expect to be able to decode local motion direction in V2 and V3 (as has been demonstrated in humans: Kamitani and Tong, 2006; Hogendoorn and Verstraten, 2013). However, V2 and V3 seem unlikely locations for the extraction of global direction from directionally noisy local motion signals. Despite multiple reports of sensitivity in macaque V2/V3 to the direction of rigidly moving stimuli, we know of no evidence of the kind that exists in MT (e.g. Newsome et al., 1989) for spatial integration of disparate directions. Gegenfurtner et al. (1997) has shown that some V3 cells show “pattern” responses to plaids (which have spatially overlapping

components) but it appears that integration of direction signals across space has not been examined physiologically in either V2 or V3 of macaque.

The fact that we were able to decode the direction of global motion from responses in human V2 and V3 allows us to speculate about how afferent signals are combined in V2 and V3. In macaques, V2 and V3 receptive fields are built primarily from V1 afferents. In V2, receptive fields are larger than in V1 and in V3 they are larger still, as evidenced both in single-unit recording in macaques (Gattass et al., 1981; Baizer, 1982; Gattass et al., 1988) and also in population RF estimates from fMRI in humans (Smith et al., 2001; Harvey and Dumoulin, 2011). These larger receptive fields presumably arise by combining inputs from V1 neurons with slightly different receptive field centers. Spatial integration is therefore a natural consequence. However, global motion sensitivity does not automatically fall out of having large receptive fields but requires specific neural wiring. To preserve direction selectivity, it is necessary for spatial pooling to be constrained to specific subsets of neurons. Most fundamentally, the V1 afferents to be combined must be drawn from the minority that are direction-sensitive. To create narrow direction tuning, they must be drawn from subsets with quite similar direction preferences. To create global motion signals that are robust to large local variations, they must be drawn from broader subsets covering a range of local directions (up to 180deg). Our results suggest that the latter process occurs in V2 and V3, as well as hMT. An alternative possibility is that global motion sensitivity in V2 and V3 might reflect feedback rather than feed-forward connections.

### *Purpose of global motion processing in V2 and V3*

It is likely that motion signals are refined in different ways in different cortical areas. The most studied area in macaques is MT, where motion energy signals are combined in specific ways for specific purposes, including solving the aperture problem to give specificity to “pattern direction” as well as combining signals over space. In MSTd, the emphasis is on extracting components of optic flow. Most MSTd neurons respond best to specific combinations of expansion and rotation and many do so in a position-invariant manner (Graziano et al., 1994; Lagae et al., 1994). This clearly requires highly selective

combination of inputs. Macaque VIP is in many ways similar to MSTd in terms of visual response properties, having many flow-sensitive neurons (Bremmer et al., 2002). Macaque V6 has also been associated with optic flow arising from self-motion. However, V6 emphasises near-space and a recent review (Pitzalis et al., 2013a) suggests that it may extract information about objects in the presence of flow, rather than signaling flow *per se*. Thus, these macaque cortical regions all respond to motion on a global scale but probably encode it for different purposes. The same applies to the human cortical regions we have studied. We have previously shown that hV6, pVIP and CSv are not only responsive to optic flow but selectively responsive to egomotion-compatible visual motion (Wall and Smith, 2008; Cardin and Smith, 2010). Two of them, hV6 and pVIP, may have similar properties to their putative macaque counterparts. They respond well in fMRI studies to sustained simulated self-motion in a constant direction (e.g. forward motion). In contrast, CSv responds only weakly to such stimuli but appears to be specifically concerned with changes in heading (Furlan et al., 2014). As in macaques, different motion-sensitive regions may extract different types of global motion information.

What, then, is the purpose of global motion processing in V2 and V3? There are few clues in the literature, which is limited compared to that on V1 and focuses on a search for responses properties not evident in V1. Properties proposed are typically spatial rather than temporal. For example, Hegdé and Van Essen (2000) report sensitivity to complex shapes in V2 while Merigan et al. (1993) showed that macaque V2 lesions affect complex spatial tasks such as detecting the orientation of a row of dots. Although V2 and V3 project to MT in macaques, there is little reason to think that they supply MT with global motion information. Indeed, reversibly inactivating V2/V3 affects MT neurons primarily in terms of their sensitivity to depth rather than motion (Ponce et al., 2008; Ponce et al., 2011; Smolyanskaya et al., 2015). Motion-sensitive neurons in V2 and V3 may therefore primarily project elsewhere.

A possible destination for the global motion signals that we have shown to be present in V2/V3 is area V6. V6 is highly specialised for motion and appears to carry signals relating to self-motion. Although there are anatomical connections between MT and V6 in macaques (Galletti et al., 2001), it is not thought that V6 derives its motion sensitivity

from MT/MST. In humans, hMT+ and V6 have similar response latencies, suggesting that they receive separate, parallel inputs from V1 and that V6 generates self-motion sensitivity *de novo* rather than inheriting it from MT+ (Pitzalis et al., 2013b). In macaques, a direct connection between V1 and V6 has been demonstrated (Galletti et al., 2001), adding plausibility to this suggestion. We suggest, however, that the global motion sensitivity that is so characteristic of V6 in both macaque and human may not be generated primarily from V1 afferents but may instead be based in large part on afferents from V2 and V3. Galletti et al. (2001) showed that V6 connects strongly with both V2 and V3, as well as V1, and indeed Shipp et al. (1998) claimed that V6 connects to V2 and V3 but not V1. Earlier tracer studies (Colby et al., 1988; Gattass et al., 1997), performed prior to the delineation of V6, demonstrated connections between V2/V3 and area PO, which corresponds loosely to V6. Indeed, Colby et al. claimed that V2 is the strongest source of visual input to area PO. In view of the specialization of V6 for motion, it seems likely that its V2/V3 afferents carry motion signals, and plausible that supplying V6 with global motion information may be the purpose of the global motion sensitivity that we have demonstrated in V2 and V3.

## References

- Baizer J (1982) Receptive field properties of V3 neurons in monkey. *Investigative Ophthalmology and Visual Science* 23:78-95.
- Braddick OJ, O'Brien JMD, Wattam-Bell J, Atkinson J, Hartley T, Turner R (2001) Brain areas sensitive to coherent visual motion. *Perception* 30:61-72.
- Brainard DH (1997) The psychophysics toolbox. *Spatial Vision* 10:433-436.
- Bremmer F, Duhamel JR, Hamed SB, Graf W (2002) Heading encoding in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience* 16:1554-1568.
- Bremmer F, Schlack A, Shah N, Zafiris O, Kubischik M, Hoffmann K-P, Zillies K, Fink G (2001) Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29:287-296.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1993) Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience* 10:1157-1169.
- Brouwer GJ, Heeger DJ (2009) Decoding and Reconstructing Color from Responses in Human Visual Cortex. *Journal of Neuroscience* 29:13992-14003.
- Burr DC, Morrone MC, Vaina LM (1998) Large receptive fields for optic flow detection in humans. *Vision Research* 38:1731-1743.
- Cardin V, Smith AT (2010) Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. *Cerebral Cortex* 20:1964-1973.
- Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines *ACM Transactions on Intelligent Systems and Technology* 2:21-27.
- Colby C, Gattas R, Olson R, Gross C (1988) Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study. *Journal of Comparative Neurology* 269:392-413.
- Deichmann R, Schwarzbauer C, Turner R (2004) Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. *NeuroImage* 21:757-767.
- Dukelow SP, DeSouza JFX, Culham JC, van den Berg AV, Menon RS, Vilis T (2001) Distinguishing Subregions of the Human MT+ Complex Using Visual Fields and Pursuit Eye Movements. *Journal of Neurophysiology* 86:1991-2000.

- Edwards M, Badcock DR (1998) Discrimination of global-motion signal strength. *Vision Research* 38:3051-3056.
- Felleman DJ, van Essen DC (1987) Receptive field properties of neurons in area V3 of macaque extrastriate cortex. *Journal of Neurophysiology* 57:889-920.
- Felleman DJ, Burkhalter A, Van Essen CC (1997) Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex. *Journal of Comparative Neurology* 379:21-47.
- Fischer E, Bühlhoff HH, Logothetis NK, Bartels A (2012) Human Areas V3A and V6 Compensate for Self-Induced Planar Visual Motion. *Neuron* 73:1415-1240.
- Foster KH, Gaska JP, Nagler M, Pollen DA (1985) Spatial and temporal frequency selectivity of neurons in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology (London)* 365:331-363.
- Furlan M, Wann JP, Smith AT (2014) A representation of changing heading direction in human cortical areas pVIP and CSv. *Cerebral Cortex* 24:2848-2858.
- Galletti C, Gamberini M, Kutz DF, Fattori P, Luppino G, Matelli M (2001) The cortical connections of area V6: an occipito-parietal network processing visual information. *European Journal of Neuroscience* 13:1572-1588.
- Gattass R, Gross CG, Sandell JH (1981) Visual topography of V2 in the macaque. *Journal of Comparative Neurology* 201:519-539.
- Gattass R, Sousa A, Gross C (1988) Visuotopic organization and extent of V3 and V4 of macaque. *Journal of Neuroscience* 8:1831-1845.
- Gattass R, Sousa APB, Mishkin M, Ungerleider LG (1997) Cortical projections of area V2 in the macaque. *Cerebral Cortex* 7:110-129.
- Gegenfurtner KR, Kiper DC, Levitt JB (1997) Functional properties of neurons in macaque area V3. *Journal of Neurophysiology* 77:1906-1923.
- Graziano M, Andersen R, Snowden R (1994) Tuning of MST neurons to spiral motions. *Journal of Neuroscience* 14:54-67.
- Harvey BM, Dumoulin SO (2011) The Relationship between Cortical Magnification Factor and Population Receptive Field Size in Human Visual Cortex: Constancies in Cortical Architecture. *Journal of Neuroscience* 31:13604-13612.
- Hegd  J, Essen DCV (2000) Selectivity for Complex Shapes in Primate Visual Area V2. *Journal of Neuroscience* 20:RC61: 61-66.

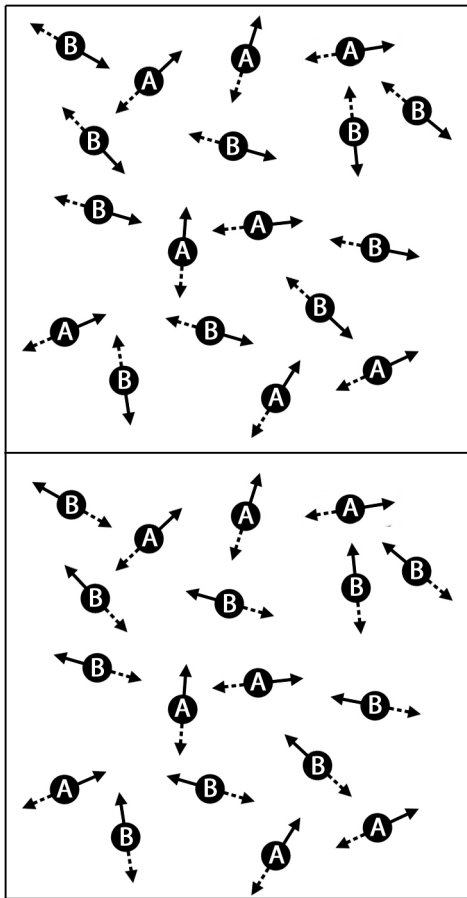


- Helfrich R, Becker H, Haarmeier T (2013) Processing of coherent visual motion in topographically organized visual areas in human cerebral cortex. *Brain Topography* 26:247-263.
- Hiris E, Blake R (1995) Discrimination of coherent motion when local motion varies in speed and direction. *Journal of Experimental Psychology; Human Perception and Performance* 21:308-317.
- Hogendoorn H, Verstraten FAJ (2013) Decoding the motion aftereffect in human visual cortex. *NeuroImage* 82:426-432.
- Huk AC, Heeger DJ (2000) Task-Related Modulation of Visual Cortex. *Journal of Neurophysiology* 83:3525-3536.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and Functional Subdivision of Human Areas MT and MST. *Journal of Neuroscience* 22:7195-7205.
- Kamitani Y, Tong F (2006) Decoding Seen and Attended Motion Directions from Activity in the Human Visual Cortex. *Current Biology* 16:1096-1102.
- Kolster H, Peeters R, Orban GA (2010) The Retinotopic Organization of the Human Middle Temporal Area MT/V5 and Its Cortical Neighbors. *Journal of Neuroscience* 30:9801-9820.
- Konen CS, Kastner S (2008) Representation of Eye Movements and Stimulus Motion in Topographically Organized Areas of Human Posterior Parietal Cortex. *Journal of Neuroscience* 28:8361-8375.
- Lagae L, Maes H, Raiguel S, Xiao D-K, Orban GA (1994) Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *Journal of Neurophysiology* 71:1597-1626.
- Merigan WH, Nealey TA, Maunsell JHR (1993) Visual effects of lesions of cortical area V2 in macaques. *Journal of Neuroscience* 13:3180-3191.
- Newsome WT, Paré EB (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience* 8:2201-2211.
- Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52-54.
- Pasternak T, Merigan WH (1994) Motion perception following lesions of the superior temporal sulcus in the monkey. *Cerebral Cortex* 4:247-259.

- Pitzalis S, Fattori P, Galletti C (2013a) The functional role of the medial motion area V6. *Frontiers in Behavioral Neuroscience* 6 (Article 91):1-13.
- Pitzalis S, Bozzacchi C, Bultrini A, Fattori P, Galletti C, Russo FD (2013b) Parallel motion signals to the medial and lateral motion areas V6 and MT + *NeuroImage* 67:89-100.
- Pitzalis S, Galletti C, Huang R-S, Patria F, Committeri G, Galati G, Fattori P, Sereno MI (2006) Wide-Field Retinotopy Defines Human Cortical Visual Area V6. *Journal of Neuroscience* 26:7962-7963.
- Ponce CR, Lomber SG, Born RT (2008) Integrating motion and depth via parallel pathways. *Nature Neuroscience* 11:216-223.
- Ponce CR, Hunter JN, Pack CC, Lomber SG, Born RT (2011) Contributions of indirect pathways to visual response properties in macaque middle temporal area MT. *Journal of Neuroscience* 31:3894–3903.
- Rees G, Friston K, Koch C (2000) A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience* 3:716-723.
- Salzman C, Murasugi C, Britten K, Newsome W (1992) Microstimulation in visual area MT: effects on direction discrimination performance. *Journal of Neuroscience* 12:2331–2355.
- Schwarzbach J (2011) A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behavior Research Methods* 43:1194-1201.
- Shipp S, Blanton M, Zeki S (1998) A visuo-somatomotor pathway through superior parietal cortex in the macaque monkey: cortical connections of areas V6 and V6A. *European Journal Of Neuroscience* 10:3171-3193.
- Smith AT, Curran W, Braddick OJ (1999) What motion distributions yield global transparency and spatial segmentation? *Vision Research* 39:1121-1132.
- Smith AT, Kossilo P, Williams AL (2011) The confounding effect of response amplitude on MVPA performance measures. *NeuroImage* 56:525-530.
- Smith AT, Singh KD, Williams AL, Greenlee MW (2001) Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex* 11:1182-1190.

- Smith AT, Wall MB, Williams AL, Singh KD (2006) Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience* 23:561-569.
- Smolyanskaya A, Haefner RM, Lomber SG, Born RT (2015) A Modality-Specific Feedforward Component of Choice-Related Activity in MT. *Neuron* 87:208-219.
- Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual Topography of Human Intraparietal Sulcus. *Journal of Neuroscience* 27:5326-5337.
- Van Essen DC, Zeki SM (1978) The topographic organization of rhesus monkey prestriate cortex. *Journal of Physiology* 277:193-226.
- Wall MB, Smith AT (2008) The representation of egomotion in the human brain. *Current Biology* 18:191-194.
- Williams DW, Sekuler R (1984) Coherent global motion percepts from stochastic local motions. *Vision Research* 24:55-62.
- Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology (London)* 236:549-573.
- Zeki SM (1978) Functional specialisation in the visual cortex of the rhesus monkey. *Nature* 274:423-428.

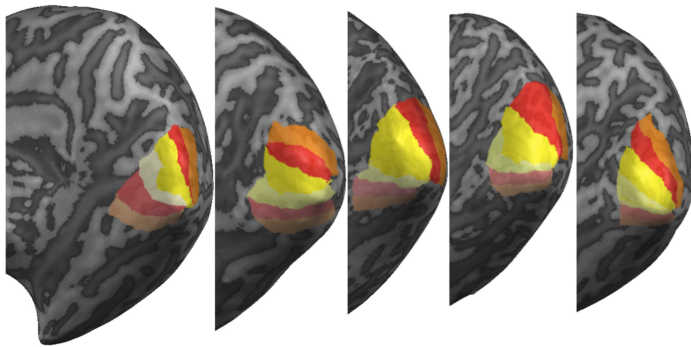
## FIGURES



**Figure 1**

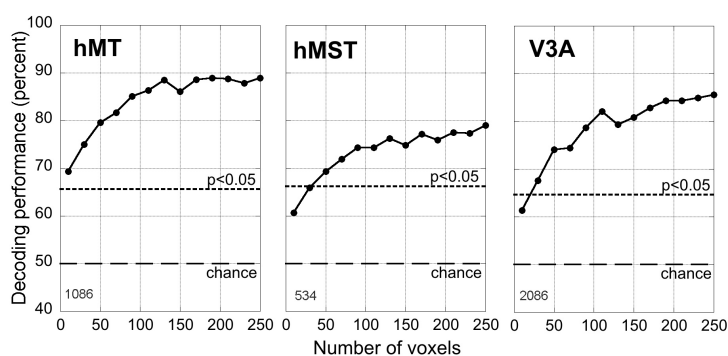
Diagrammatic representation of the construction of the stimuli. Each dot had a random position and a random axis of motion (double arrow). It oscillated back and forth along this axis, reversing every 500ms. All dots oscillated in synchrony, moving together in the directions shown by the solid arrows and then back in the directions shown by the broken arrows. Upper panel: Class 1. The dots all have “forward” directions (solid arrow) in the range upward-rightward-downward (0-180 deg) and “reverse” directions in the range downward-leftward-upward (180-360deg). The overall stimulus yields global motion alternately rightward and leftward. Each dot is labelled A or B according to whether its “forward” direction falls between 0 deg (up) and 90 deg (right) or between 90 deg and 180 deg (down). Lower panel: Class 2. The dots have the same positions and motion axes as in Class 1 but half of them differ in oscillation phase. All those marked ‘A’ have the same phase as in Class 1. Those marked ‘B’ have reversed phases (solid and broken arrows flipped). Again, dots move together in the direction of

the forward arrow and then reverse. This yields global motion that alternates between upward and downward. Thus, the two classes have exactly the same set of dots and are perfectly balanced at the local level but yield orthogonal global motions.



**Figure 2**

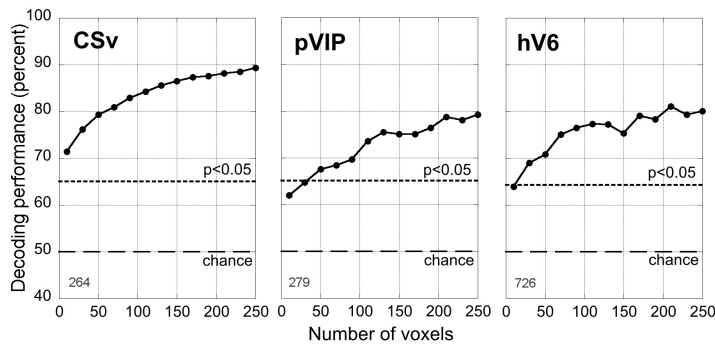
Medial posterior views of the ‘inflated’ right hemispheres of the participants showing the locations of V1 (yellow), V2 (red) and V3 (orange). V2d and V3d are distinguished from V2v and V3v by the use of more saturated colors. The regions marked are the regions of interest used in the study, as identified by retinotopic mapping, and exclude the representation of the far periphery of the visual field anteriorly and also the foveal representation at the occipital pole.



**Figure 3**

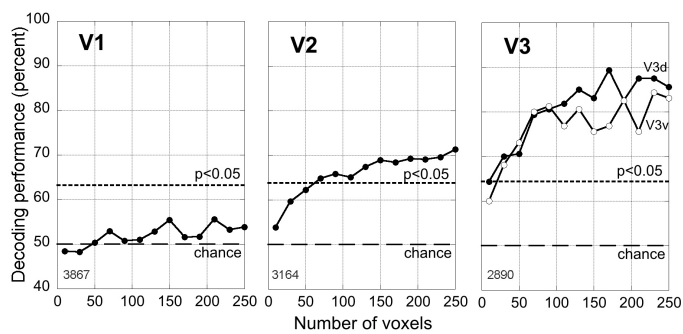
Decoding performance for three visual areas defined with independent localiser scans, hMT, hMST and V3A, as a function of the number of voxels (features) included in an MVPA analysis. Each point is the mean performance from 100 random voxel selections of the size shown on the abscissa. The total number of available voxels is shown in the

bottom left of each panel. Also shown are chance decoding performance (dashed line) and the performance level that is significantly above chance at  $p < 0.05$ , derived by permutation testing (dotted line).



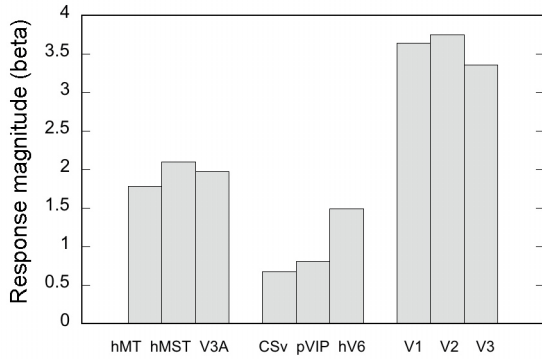
**Figure 4**

Decoding performance for visual areas CSv, pVIP and hV6 shown in the same way as in Figure 3.



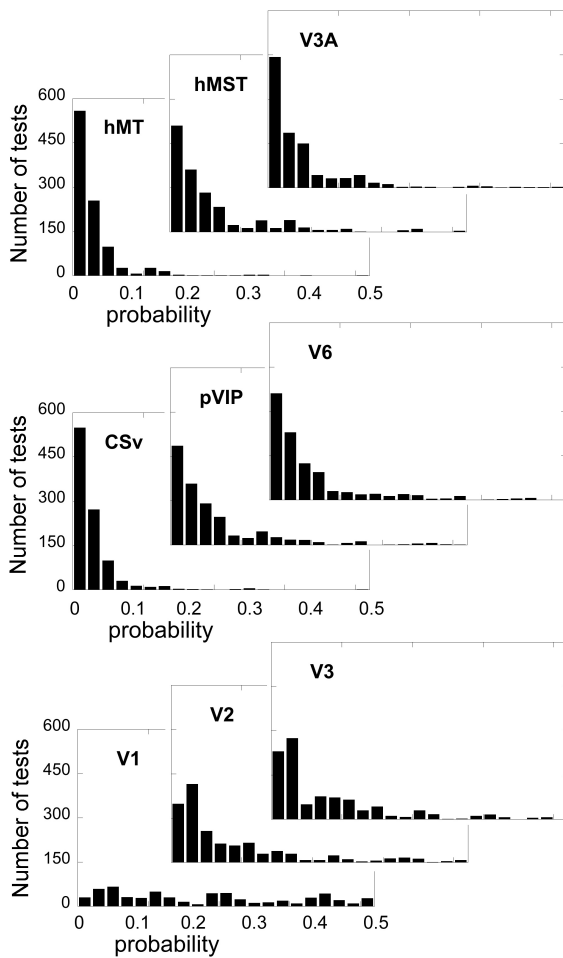
**Figure 5**

Decoding performance for visual areas V1, V2 and V3 shown in the same way as in Figures 3 and 4. In the case of V3, performance is shown separately for the dorsal (lower visual field) and ventral (upper visual field) portions.



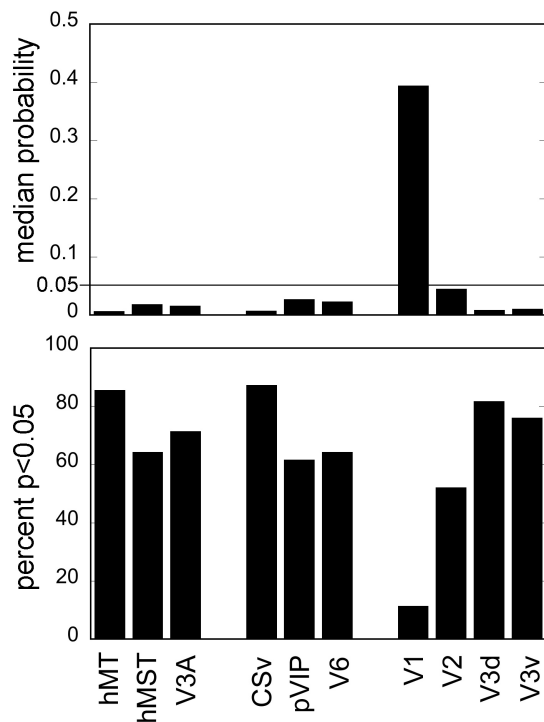
**Figure 6**

Mean amplitude, expressed as beta values from a standard (univariate) regression analysis, of the BOLD response to the stimuli (both classes combined), shown separately for each visual area examined. In each case, amplitude is the average from the 100 samples of 250 voxels used in the MVPA analysis shown in Figures 3-5.



**Figure 7**

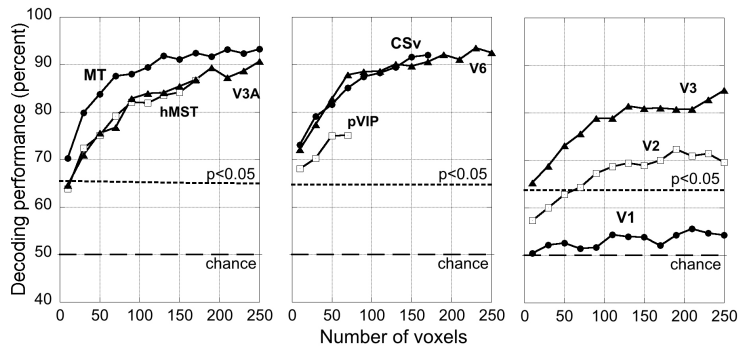
Results of an analysis in which decoding performance (based on 250 voxels) in each visual area was compared with performance in the same sample when the labels (vertical or horizontal global motion) associated with the exemplars (stimulus presentations) were randomly permuted. Each panel is a frequency plot showing binned  $p$  values (only values up to 0.5 are shown) from 1000  $t$ -tests; clustering at low values indicates strong sensitivity to global motion direction.



**Figure 8**

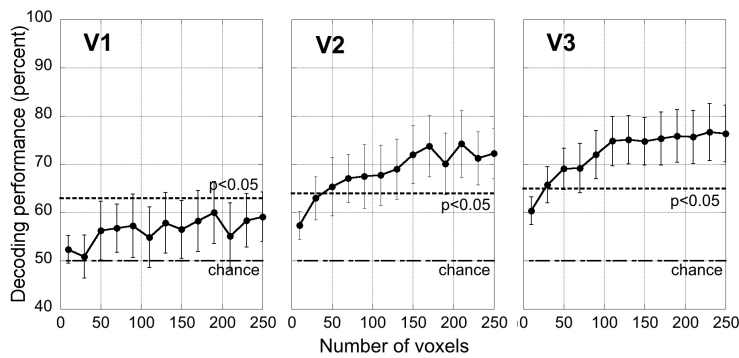
Summary statistics from the analysis shown in Figure 7. For each visual area, the median probability ( $p$ ) value from 1000  $t$ -tests is shown (top) together with the proportion of  $t$ -tests in which  $p$  was less than 0.05 (bottom).





**Figure 9**

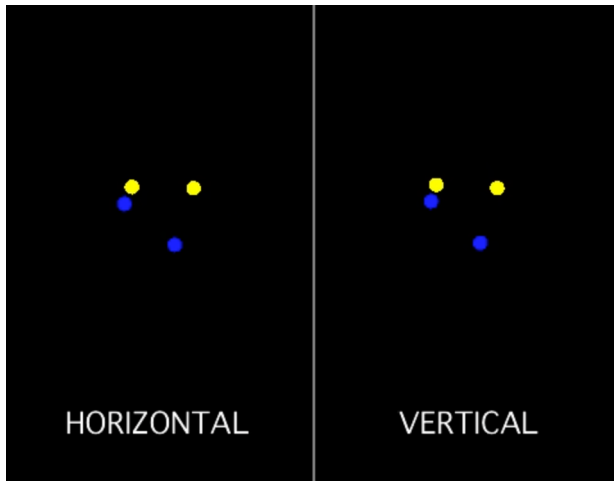
Decoding performance for all visual areas examined taken from a supplementary analysis in which the participants contributed equal numbers of voxels to any given analysis.



**Figure 10**

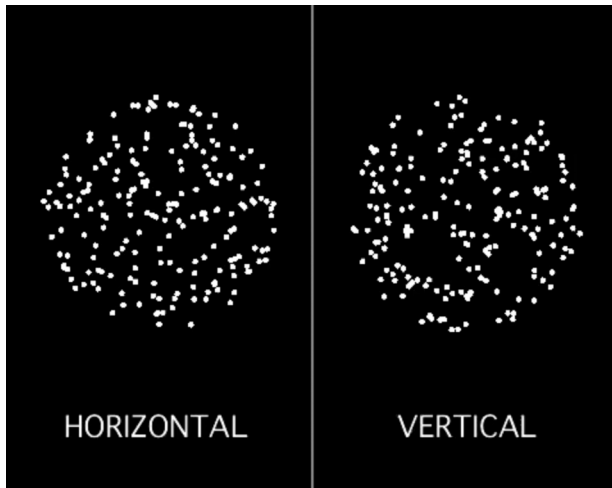
Decoding performance for V1, V2 and V3 from a supplementary analysis in which global motion was decoded separately in each participant and the results averaged. Error bars show  $\pm 1$  SEM.

## MULTIMEDIA



### Movie 1

Pictorial representation of the construction of the stimuli. Horizontal and vertical global motion are represented in the left and right panels respectively. In the left panel, 4 dots are shown. Each has a random start location and moves back and forth along a randomly chosen axis. Dot motion is timed such that the horizontal component of motion is the same (either leftward or rightward) for all dots at any given moment. The vertical component is not the same for all dots; two of the dots (shown in yellow) have upward components when the other two (blue) have downward components, and vice versa. The colors are for illustration only - in the experiment all dots were white. In the right panel, the first frame is identical to the first frame of the left panel and therefore contains the same 4 dots. Each dot moves along the same axis as in the left panel. Those shown in blue move in the same phase as in the left panel i.e. their movement is identical. However those shown in yellow move on the same axis as in the left panel but with the opposite temporal phase: when they have a downward component in the left panel they have an upward component in the right panel and vice versa. The result is that all 4 dots have the same vertical component (up or down) at any given moment but the horizontal components are opposite for blue and yellow dots. In summary, the two panels have exactly the same dots at the same locations, moving on the same axes, but differ in the temporal phase of half the dots. One temporal cycle is shown but the movie may be looped to give continuous back and forth motion.



### **Movie 2**

Examples of global motion stimuli constructed on the principal of Movie 1. Again the left and right panels show horizontal and vertical global motion respectively. When many dots with random motion axes are drawn and move according to the rule described, the overall (global) motion in the left panel is horizontal, even though local dot directions span the full 360deg range. In the right panel, global motion is vertical. The movie may be looped to give continuous back-and-forth global motion. Global motion is noisy and weak. When the movie is directly viewed, global motion may be difficult to see because of the distracting effect of individual dots: each panel contains both vertically and horizontally moving dots (along with all other axes). Global motion may become more obvious by fixating some way above or below the movie to place the movie in the visual periphery.