Implicit representations of luminance and the temporal structure of moving stimuli in multiple regions of human visual cortex revealed by multivariate pattern classification analysis

Stephen T Hammett
Andrew T Smith
Matthew B Wall
Jonas Larsson

Dept of Psychology,
Royal Holloway University of London,
Surrey TW20 0EX UK.

Running Head: Motion and luminance encoding in human visual cortex

Corresponding author: Steve Hammett

Email: s.hammett@rhul.ac.uk

Abstract

The generation of a behaviorally relevant cue to the speed of objects around us is critical to our ability to navigate safely within our environment. However, our perception of speed is often distorted by prevailing conditions. For instance, as luminance is reduced, our perception of the speed of fast moving patterns can be increased by as much as 30%. In order to investigate how the cortical representation of speed may vary under such conditions we have measured the functional Magnetic Resonance Imaging Blood Oxygen Level Dependent (fMRI BOLD) response of visual cortex to drifting sine gratings at two very different luminances. The average BOLD response in all areas was band-pass with respect to speed (or, equivalently, temporal frequency) and thus contained no unambiguous speed information. However, a multivariate classifier was able to successfully predict grating speed in all cortical areas measured. Similarly, we find that a multivariate classifier can predict stimulus luminance. No differences in either the mean BOLD response or the multivariate classifier response with respect to speed were found as luminance changed. However, examination of the spatial distribution of speed preferences in V1 revealed that peri-foveal locations preferred slower speeds than peripheral locations at low, but not high, luminance. We conclude that although an explicit representation of perceived speed has yet to be demonstrated in the human brain, multiple visual regions encode both the temporal structure of moving stimuli and luminance implicitly.

Key words: speed, temporal frequency, visual cortex, human, fMRI

Introduction

The question of how an object's speed is encoded is critical to an understanding of how the visual system guides us in real-world situations. Whilst the cortical pathway responsible for motion processing is relatively well defined an understanding of the precise mechanisms involved in encoding the *speed* of a moving image has proven evasive. The earliest (retinal and thalamic) stages of visual processing are mediated by neurones whose responses are spatio-temporally separable (e.g. Tolhurst & Movshon 1975; Foster et al 1985). Such neurones do not provide an unambiguous code for speed. Various models of how the spatio-temporally separable signals generated in the retina are transformed to provide an unambiguous speed code have been proposed (e.g. Smith & Edgar 1994; Priebe & Lisberger 2004; Thompson et al, 2006; Stocker & Simoncelli 2006; Hammett et al 2007; Langley & Anderson 2007) but there is still no clear picture of exactly where in the pathway speed-tuning arises, nor how it is achieved. The speed tuning of many MT neurones (e.g. Perrone & Thiele 2001) and a direct link between their activity and speed perception (e.g. Rudolph & Pasternak 1999; Liu & Newsome 2005) raises the possibility that an explicit code for speed may be extracted from early spatially and temporally tuned responses in later cortical areas. However, other studies cast doubt upon a straightforward hierarchy of speed encoding across cortical areas and indicate that speed tuning in MT may, at least in part, be inherited from speed-tuned cells in V1 (Priebe et al 2003, 2006). Priebe et al (2006) report that around 25% of V1 and MT neurones have complete spatio-temporal separability, whilst another 25% show clear speed tuning. Thus both spatio-temporally separable responses and speed-tuned responses may contribute to the neural representation of speed. Indeed, Reisbeck and Gegenfurtner (1999) reported behavioral evidence consistent with such a scheme. In summary, previous models of how the brain encodes speed have been predicated upon the assumption that an explicit neural speed code exists. However,

electrophysiological evidence for such a code is mixed – some cells as early as V1 are speed tuned, but others are spatio-temporally separable. The same variability is seen in higher cortical areas such as MT. Thus, where such a code is generated is unclear. Moreover, how such a code is generated is equally unknown and recent electrophysiological evidence (Krekelberg et al 2006) has failed to find support for any extant model of how speed encoding may be accomplished.

The advent of fMRI has provided considerable clues as to how the visual system encodes object properties at the neural population level and a clear coupling of the BOLD response to fundamental image attributes such as contrast and motion coherence have been reported. For instance, Boynton et al (1999) have shown that the BOLD response increases monotonically as image contrast increases. Given the critical nature of encoding object speed, it is tempting to assume that its cortical representation would be similarly unambiguous. However, the few previous attempts to characterize speed encoding using fMRI have not been entirely consistent and have failed to reveal a straightforward relationship between image speed and BOLD magnitude. Two early studies (Chawla et al 1998, 1999) used spatially broadband stimuli (moving dots). They found heterogeneity in the BOLD response such that whilst later areas (V3 and MT) showed a band-pass response as a function of speed, V1 produced a low-pass response. However, interpretation of their results at high speeds is complicated by the severe under-sampling of motion inherent in their stimuli (each dot moved up to 8 times its diameter on each frame update), which may explain the reduced response. A subsequent study that used narrowband stimuli and a more limited speed range (Singh et al 2000) reported band-pass speed tuning in all cortical areas from V1 to MT. The tuning was both quantitatively and qualitatively similar across the entire visual cortex. A later study of flicker sensitivity (Hagenbeek et al 2002) also found bandpass tuning in V1. More recently, Lingnau et al (2009) have reported fMRI adaptation data

suggesting that the BOLD response is more sensitive to speed than temporal frequency, at least in MT and at high contrast. However, their results also provided evidence for temporal tuning across all cortical areas at lower contrast.

In summary, both electrophysiological and imaging studies have provided mixed evidence regarding the locus and nature of speed encoding. There is strong electrophysiological evidence of a direct link between speed perception and activity in MT but there is also evidence of speed tuning in V1 and spatio-temporally separable responses in MT. Unfortunately, psychophysical studies do not provide any greater illumination given the evidence that the substrates of perceived speed are not unequivocally tuned for either speed or temporal frequency (Thompson et al 2006). Computational approaches to speed encoding have proven equally unfruitful. Krekelberg et al (2006) have shown that Bayesian, labeled lines and ratio models (e.g. Ascher & Grzywacz, 2000; Priebe & Lisberger, 2004; Hammett et al 2007) of speed encoding are all inconsistent with the response characteristics of speed-tuned MT cells. Thus, to date, there is no clear account of the evolution of speed encoding, nor its locus within the visual pathways.

The absence of a consensus as to how a speed code may be computed renders any parametric study of the BOLD response prone to difficulties of interpretation. One potential way forward is to measure the BOLD response under conditions that are known to directly affect perceived speed. If the BOLD response is coupled to perceived (rather than physical) stimulus speed (see Lingnau et al's (2009) experiment 2 for some encouraging indirect evidence that this may be the case) then modulating perceived speed without changing physical speed should result in a concomitant modulation in BOLD response. The expectation of such a coupling, at least in MT, seems reasonable in light of evidence (Liu & Newsome 2005) that individual neurons in MT play a direct role in speed perception. Thus

examining the effect of a stimulus attribute that is known to affect perceived speed on BOLD responses may render a clearer picture of how and where cortical speed encoding occurs.

Recently, Hammett et al (2007) have shown that perceived speed is modulated by mean luminance such that low luminance stimuli appear significantly faster at high speeds. In any cortical region in which perceived speed is represented explicitly, the BOLD response measured under conditions in which both speed and luminance are varied should be affected by both variables such that it bears a consistent relation to perceived speed. We have therefore measured the BOLD response of various visual cortical areas to drifting sinusoidal gratings at a range of speeds and at two very different luminance levels. In addition, we use multi-voxel pattern classification analysis (MVPA) to check separately for sensitivity to physical speed and to luminance in each area.

Experiment 1

Methods

Subjects

All seven subjects were undergraduates or postgraduates at Royal Holloway. All had normal or corrected vision. The experiments were conducted in accordance with the Declaration of Helsinki, approved by a local ethics committee at Royal Holloway, University of London, and written informed consent was obtained. Standard MRI screening procedures were followed and volunteers were paid for their participation. Subjects were scanned on two occasions, usually separated by approximately a week. Additional scanning runs were also performed on other occasions to define regions of interest (ROIs; see below for details).

Data Acquisition

MRI images were acquired with a 3-Tesla Siemens Magnetom Trio scanner with an 8-channel array head coil. Anatomical (T1-weighted) images were obtained at the start of each scanning session (MP-RAGE, 160 axial slices, in-plane resolution 256 x 256, 1 mm isotropic voxels, TR = 1830 ms, TE = 4.43 ms, flip angle = 11°, bandwidth = 130 Hz/pixel). This was followed by six scanning runs of functional data acquisition with a gradient echo, echoplanar sequence (TR = 2s, 28 contiguous axial slices, interleaved acquisition order, 3 mm isotropic voxels, in-plane resolution of 64 x 64 voxels, flip angle = 90°, TE = 30 ms, bandwidth = 1396 Hz/pixel). Functional scanning runs consisted of 224 volumes and therefore lasted seven minutes 28 seconds.

Stimuli

All stimuli were back-projected onto a screen mounted in the rear of the scanner bore by a computer-controlled LCD projector (SANYO PLC XP40L) at resolution of 1024 x 768 pixels and a refresh rate of 60 Hz. Participants viewed the stimuli at a distance of 65cm via a

mirror mounted on the head coil which provided a horizontal viewing angle of approximately 30 degrees of visual angle. Stimuli were generated and delivered from Matlab (Mathworks, Natick, Massachusetts) using the MGL toolbox (www.justingardner.net; www.pc.rhul.ac.uk/staff/J.Larsson/dokuwiki/doku.php?id=software).

The stimuli consisted of single, horizontal, sinusoidal gratings, presented in the centre of the screen, and approximately 20 degrees in diameter. Sine gratings were chosen over broadband stimuli for comparability with previous work (Hammett et al, 2007). The background and central one degree of the stimulus was set at mean luminance and contained a small black fixation cross which was present throughout the experiment. The spatial frequency of the grating was 1 cycle/degree, and the grating drifted upwards throughout each trial. Four different speeds were used for the drifting grating, which defined the four conditions: 2, 4, 8 and 13.3 degrees/second. Because the stimuli were sine gratings and spatial frequency was invariant, the gratings can equivalently be regarded as drifting at 1, 4, 8 and 13.3Hz. The luminance of the stimulus was controlled by two cross-polarising filters mounted between the projector and the back-projection screen. By rotating one filter with respect to the other a continuous modulation of the output of the projector could be achieved, which enabled very precise adjustment with no change in contrast or gamma. Average luminance levels across the stimulus area were measured using a photometer to set the levels used in the experimental sessions. The high luminance condition's space-average luminance was 30 cd m^{-2}, and the low luminance condition was set at 1.5 cd m^{-2}, similar to the luminance levels found by Hammett et al (2007) to modulate perceived speed.

Procedure

The main experiment was a standard event-related design for fMRI. Each trial lasted 3 seconds and consisted of a continuous presentation of the drifting grating. The duration of inter-trial intervals (ITI) were randomly drawn from a Poisson distribution (Hagberg et al

2001) with a range of 2-10 seconds and a mean of 5.5 seconds. There were ten trials of each speed condition (forty in total) in a single scanning run. Each scanning run also included a 10 second buffer period at the beginning and end of the sequence. Each participant completed six scanning runs lasting approximately seven and a half minutes in each scanning session. The six sequences of trial orders and ITIs were the same for every participant, but each participant was exposed to them in a different random order. The scanning runs alternated between high and low luminance. For three of the participants the first scanning run was under high luminance conditions, for the remaining four the first scanning session was under low luminance conditions. A tightly-fitting black-out screen was placed over the window between the scanner room and the control room, and all other sources of light in the scanner room were either switched off or blocked in order that no stray light could influence the luminance levels used in the experiment.

Retinotopic areas V1-V3, V3A/B, V7 and hV4 (Larsson & Heeger 2006) were identified by a standard retinotopic mapping procedure (Sereno et al 1995), using a counter-phasing checkerboard 'wedge' stimulus (a 24-deg sector) of radius 12 degrees. MT/MST (henceforth refrred to as MT+) were identified on cortical flatmaps as the set of voxels that showed a significant response to the moving grating stimuli (defined arbitrarily as having a coefficient of determination $R^2>0.25$ in the event-related model fit described below) using retinotopic criteria (Huk et al 2002) to distinguish MT+ from nearby motion-responsive areas such as V1 and V3A. Contrast reversed at a rate of 8Hz. The wedge rotated clockwise at a rate of 64 seconds/cycle, and eight cycles were presented.

Data Analysis

Data were analyzed using bespoke software implemented in Matlab (The MathWorks Inc., Natick, MA, USA, 2000) and C/C++. Functional data were pre-processed to correct for head-motion using the MCFLIRT tool in FSL (Jenkinson et al 2002), and filtered with a

temporal high-pass filter with a cutoff of 0.025Hz. No spatial smoothing was performed on the functional data. Functional images from both scanning sessions were co-registered to a high-quality anatomical image (MDEFT; Deichmann et al 2004) - hereafter referred to as the reference anatomy - of the same participant, acquired in a separate scanning session. The reference anatomy was also used to extract cortical surface representations using the public domain software SurfRelax (Larsson 2001).

Retinotopic data were analyzed by fitting a sinusoid with a frequency corresponding to the period of the rotating wedge stimulus to the time course for each voxel (high-pass filtered with a cut-off of 0.025Hz and converted to percent signal change). For each voxel, this yielded a correlation (technically coherence), a phase and an amplitude. The phase of the fitted response corresponded to the visual field (polar angle) location of each voxel. The phases of voxels with a coherence >0.25 were visualized on computationally flattened cortical surfaces and boundaries between visual areas were identified as reversals of the direction of phase change across the cortical surface (Sereno et al 1995). ROIs were drawn by eye based on these boundaries.

Event-related analysis

fMRI data were analyzed by fitting a model of the mean fMRI response to the speed stimuli in two steps. First, the trial-triggered average fMRI response to each speed within a session was estimated using deconvolution (linear regression) (Burock & Dale 2000), yielding an individual estimate of the hemodynamic response function (HRF) for each subject and speed Estimating the HRF for each subject in this way, rather than using a model of the HRF, allowed us to better capture the inter-subject variability in hemodynamic responses and obviated the need for making a priori assumptions about the specific shape of the response. Second, the average fMRI response (or HRF) was used to estimate the amplitude

of the responses evoked by each individual stimulus presentation, using a general linear

model (GLM) analysis. Although the shape and delay of the trial-triggered responses

(HRFs) differed between subjects, it did not differ across speeds, ROIs, or luminance levels,

and hence HRFs were averaged across speed, luminance, and ROIs and normalized to unit

area individually for each subject. This analysis was carried out both for each voxel

separately (in order to identify MT/MST as described above) and for each ROI (by first

averaging the time-courses of all voxels within the ROI). In the GLM analysis, each trial

was modeled by a separate column in the design matrix, each column containing a copy of

the trial-triggered average fMRI response averaged across speeds and ROIs, time-shifted to

align it with the onset of the trial. The regression coefficients (obtained by standard multiple

regression methods) then corresponded to the fMRI response amplitudes for each trial. The

model was fitted separately to the fMRI time series (detrended by highpass filtering and

normalized to zero mean and percent signal change) for each ROI. From these amplitudes

we computed the mean responses (averaged across subjects and trials) for each speed and

luminance condition. We chose to use this method to estimate average response amplitudes

rather than using the peak amplitudes from the deconvolution analysis, as it allowed us to

directly compare the results of the ROI-based analysis with those of the multivariate analysis

described below.

Multivariate pattern classification analysis (MVPA)

In a complementary analysis, MVPA based on sparse regression techniques was applied to

investigate whether the multivariate fMRI response across all voxels in each area could accurately

predict the speed of the stimulus for individual trials. We used the LARS-LASSO algorithm for

sparse regression (Efron et al 2002) to automatically select the voxels that contributed most to

the decoding accuracy and overcome the problem of fitting a regression model with more

covariates than observations. For every voxel within each ROI, we estimated the response

amplitudes for individual speed trials using the same regression model described above. This yielded an [m x n] matrix of fMRI response amplitudes for each ROI, where m (number of rows, 240) is the number of trials (across all speeds) and n is the number of voxels in each ROI.

For each ROI, and separately for each subject and luminance, the fMRI response amplitudes (normalized to unit magnitude within each ROI) were used as covariates to predict the speed on each trial within each session by least angle regression (LARS) with the LASSO algorithm. A leave-one out procedure was used to validate the regression as follows. Data from five out the six scans in each session were used to fit the regression model, using the voxel response amplitudes as covariates and the speed on each trial as the dependent variable. Because comparisons of classification performance across ROIs are complicated by differences in size (i.e., number of voxels), response strength, and intrinsic response properties of the ROIs, we tried several criteria for selecting voxels to ensure our results were not due to the particular selection procedure used. Although the LARS-LASSO algorithm automatically selects the most informative voxels for classification, the algorithm can be constrained to use fewer voxels than the maximum possible (which is given by the number of trials). We explored the effect of varying the number of voxels on classification performance, and observed no qualitative difference in the results. However, in all ROIs we found that classification accuracy peaked at relatively small numbers of voxels (25), suggesting that inclusion of additional voxels provided largely redundant information and thus resulted in over-fitting. In the following, we show the results obtained using 25 voxels within each ROI, but emphasize that the results would have been qualitatively similar with a different number of voxels for all ROIs. The estimated regression coefficients were then applied to the data from the remaining scan to generate predicted speeds for each trial in that scan. This procedure was repeated for each of the six scans. The predicted speeds for all trials in a session were concatenated and the Pearson correlation computed between predicted and actual speeds on each trial to assess performance of the classification procedure. For each actual speed, the predicted speed across all trials was averaged within each subject and luminance. The overall

accuracy of the classification was then assessed by plotting the mean predicted speed for each subject and luminance against the actual speed and by computing the Pearson correlation between actual and predicted speeds across subjects. For areas in which the multivariate fMRI response across voxels encodes information about speed, this analysis would predict a monotonic increase in predicted speed with actual speed.

As a control, the MVPA analysis was also carried out on the same data but with the speed labels randomly permuted on the training set only to estimate chance level performance of the classification algorithm. The classification procedure described above was repeated 1000 times for each ROI. On each randomization iteration, the speed labels for all trials were randomly permuted and the classification analysis re-run. For each actual speed and luminance, the chance-level predicted speed was computed as the average across randomization iterations. Assuming no bias in the randomization or the classification, this analysis would be expected to predict the same speed (equal to the average of presented speeds) for each actual speed presented. If the multivariate response across voxels in each area carried no systematic information about presented speed, the result of the original classification (using the correct speed labels) would have been expected to give the same result as the random classification, i.e. a flat speed prediction equal to the average of the speeds presented; conversely, a systematic deviation from a flat prediction would indicate that the multivariate response across voxels encoded information about presented speeds.

In a second control analysis we directly tested whether MVPA provided more information about stimulus speed than the original univariate (ROI) analysis. For this analysis we collapsed the (non-normalized) response amplitude data used for the MVPA analysis across all voxels within each ROI and then subjected the resulting univariate response to the same regression analysis used for the multivariate response. For univariate data, this reduces to a simple regression analysis with a single covariate, but the results had the same dimensions and format as the multivariate analysis, allowing a direct comparison to be made.

Results

In all our results, we use the term "speed" to describe the stimuli. However, it should be borne in mind that, in terms of tuning of neural mechanisms, our stimuli do not differentiate speed from temporal frequency. Thus, our speed tuning functions can equivalently be regarded as temporal frequency tuning profiles and speed decoding can equivalently be regarded as temporal frequency decoding. Fig 1 plots the tuning of the BOLD response as a function of speed for all areas measured. All areas show a clear band-pass tuning of the BOLD response at both low and high luminance. The functions in Fig 1 assume that subjects were able to maintain good fixation. It is possible that some degree of following eye movement may have occurred, despite the use of an unstimulated region immediately around the fixation spot. If so, this would tend to reduce retinal speed, which could distort the results. However, the results are both qualitatively and quantitatively consistent with those of Singh et al (2000). Importantly we observed no evidence of a consistent difference in speed response between the high and low luminance conditions.
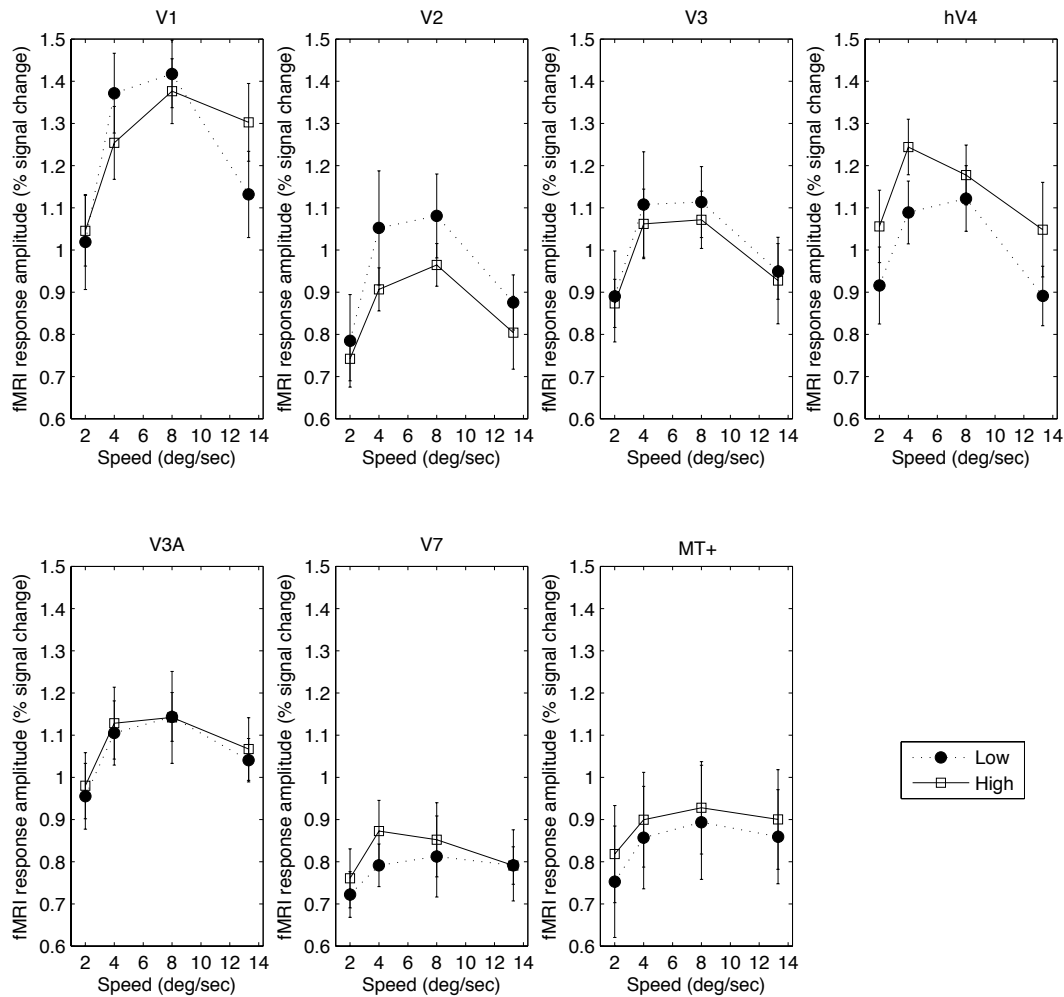
Fig 1 BOLD response as a function of the drift speed of a 1 c/deg sine grating for low (closed symbols) and high (open symbols) luminance in Experiment 1. Cortical area is indicated above each panel. Error bars represent ±1 SEM.

Given the absence of any reliable change in the BOLD response as a function of luminance, we wondered whether an explicit code for speed (or temporal frequency) might be found in a sub-population response. In order to investigate this we conducted an MVPA classification. Fig 2 plots the results of the MVPA analysis across all areas and at both luminance levels. The MVPA classifier could not distinguish the two slowest speeds measured but was able to distinguish them from higher speeds. All higher speeds (> 4 deg/sec) were unambiguously predicted in all cortical areas and at both luminance levels. At 2 deg/sec the classifier was unable to discriminate speed and the predicted speed

approached the value of the mean speed of those presented. Excluding the slowest speed, in all areas the classifier's predicted speed rose monotonically with stimulus speed. In a control condition we also trained the MVPA classifier with random assignment of stimulus speeds. As expected, under these conditions, the classifier was unable to predict any speed with all stimulus speeds classified around the mean of the four stimulus speeds, (shown by the broken line in Fig 2). Furthermore, to confirm that the performance of the classifier relied on information contained only in the multivariate response, we ran the classification analysis on the (non-normalized) multivariate data collapsed across voxels, reducing the analysis to a simple regression of speed against the average response across voxels. As expected, given the inherent ambiguity in the bandpass univariate response profile (Fig 1), the classifier failed to distinguish the different speeds from the average response (Fig 3). This result demonstrates that even though the univariate response does not unambiguously encode speed, information about stimulus speed above 2 deg s^{-1} is reliably encoded in the multivariate response.
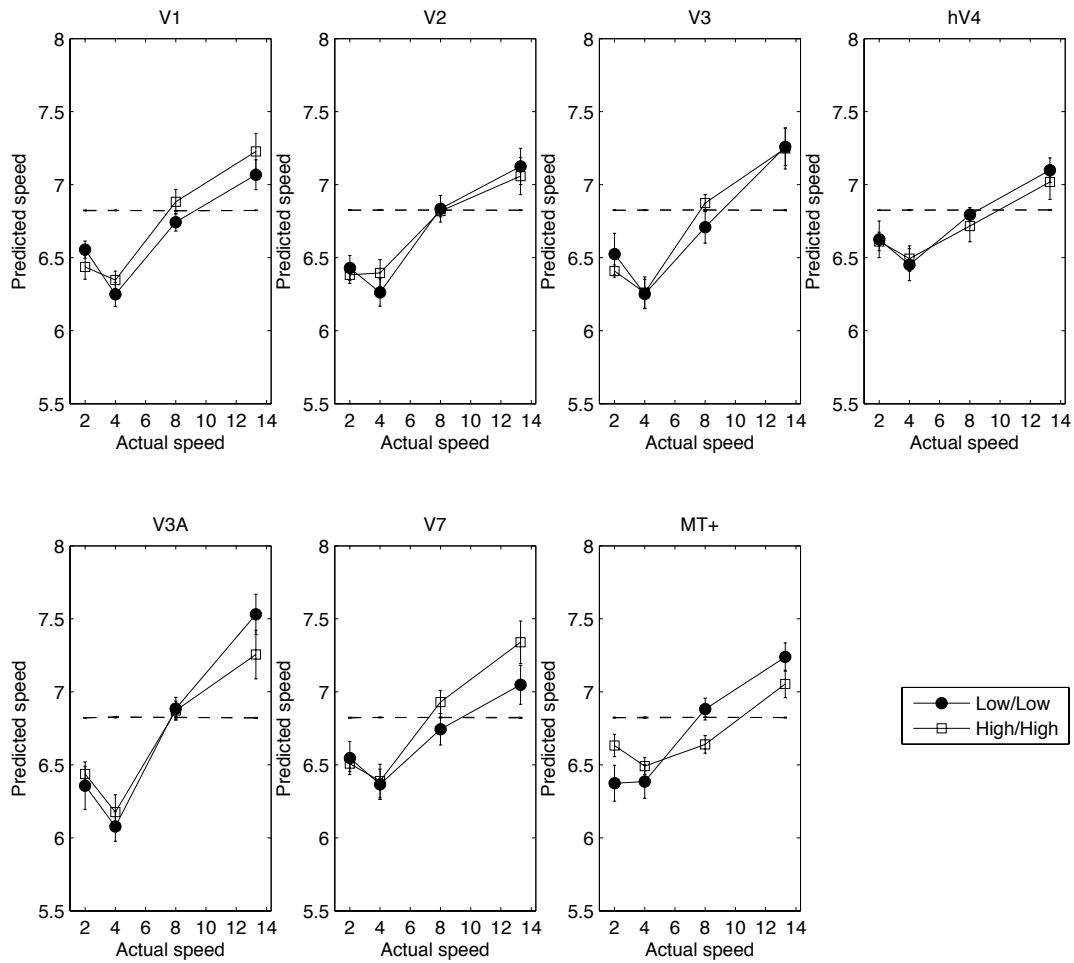
Fig 2 Predicted speed of the classifier as a function of stimulus speed is plotted for high (open symbols) and low (closed symbols) luminance in Experiment 1. Broken lines represent classifier performance for random assignment of speeds. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.
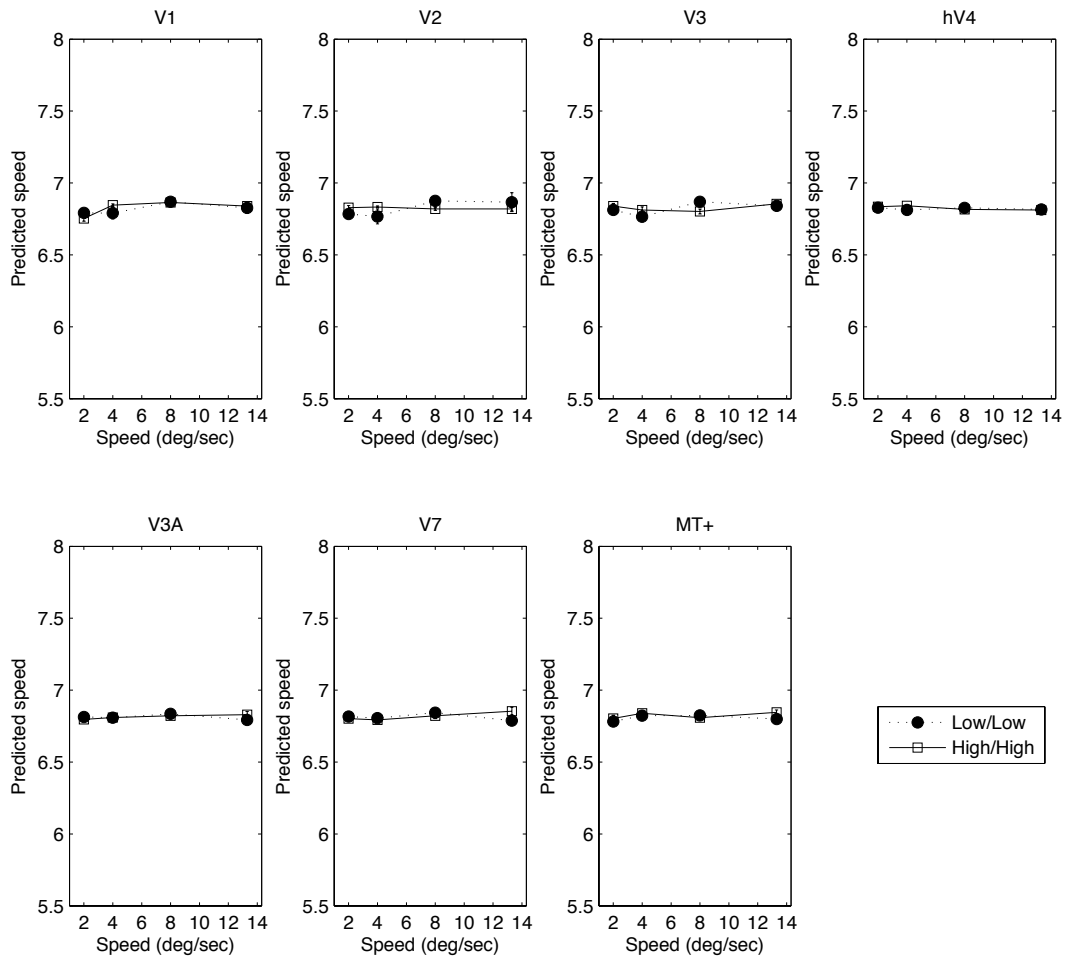
Fig 3. Results of univariate classification analysis. In each area, collapsing data across voxels reduces classifier performance to chance levels. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Experiment 2

Whilst the classifier could discriminate most speeds, its predictions did not differ between low and high luminance. We reasoned that any shift in the speed code with luminance should be revealed by training the classifier with low luminance and subsequently predicting the speed of the high luminance responses (or vice versa). If the classifier-predicted speed corresponded to perceived speed, we might expect a classifier trained at high luminance and tested on high luminance to show systematic deviations in predicted speed at low luminance that matched those observed perceptually. However, it was not tenable to examine the effect of such cross-classification since each luminance condition was measured in separate scans, thus any differences in performance between within- and across-luminance classification could be explained by differences in the exact voxel positions between scanning sessions. In order to investigate whether the speed code would reveal luminance dependency under cross-classification we repeated our measurement for both luminances within the same scanning session, reducing the number of speeds to three in order to retain an acceptable total scan time.

Methods

Subjects

Six subjects (age range 19-49, two males) took part in Experiment 2, one of whom had also been a subject in Experiment 1.

Stimuli and procedure

Motion stimuli were identical to those used in Experiment 1, except that only three speeds were used (3, 6 and 10 deg/second). The range of speeds was chosen to be within the range of speeds in Experiment 1. For some of the participants, slightly different procedures and stimuli were used to identify retinotopic visual areas (Larsson & Heeger 2006). For these

participants, stimuli consisted of radially moving checkerboards presented within a wedge 22.5 deg wide and extending from the centre of gaze to 13 degrees eccentricity, rotating clockwise or counter-clockwise stepwise by one wedge width every 1.5 seconds (synchronized with the scanner TR). Each stimulus cycle lasted 24 seconds and six cycles were presented. Data for clockwise and counter-clockwise runs were averaged as described previously (Larsson & Heeger 2006). For each participant, the experiment was divided into two back-to-back blocks, one for all the low-luminance conditions and the other for the high-luminance conditions. The order of the low and high luminance blocks was counterbalanced across participants. Hence, for three subjects the high luminance conditions were run first followed by the low luminance conditions, and for the remaining four subjects blocks were run in the reverse order. Each block was preceded by a short period of adaptation to the appropriate luminance level.

Data analysis

Data were analyzed using the same procedures as in Experiment 1. For the MVPA analysis, the LARS-LASSO classification algorithm was trained on data from high or low luminance runs separately and then tested on data from high or low luminance runs, resulting in four estimates of classification performance for each speed (train high, test high; train high, test low; train low, test low; train low, test high). In addition, we trained the algorithm on data from high and low luminance runs combined and tested on data for each luminance separately, yielding two additional estimates of classification performance (train all, test high or low).

In a complementary analysis we used the same multivariate analysis techniques to test whether stimulus luminance, rather than speed, could be decoded from the multivariate response. The analysis was identical to that for speed, except that the decoder was trained to distinguish luminance levels (coded as dummy variables with 0 indicating low luminance and 1 indicating high luminance), ignoring speed.

Estimating the spatial distribution of tuning for speed and luminance

To test whether responses to speed and/or luminance varied as a function of eccentricity in early visual areas we mapped the distribution of voxel-wise fMRI speed preferences at high and low luminance across the cortical surface in areas V1, V2, and V3 as follows. First, using the voxel-wise response amplitudes used in the MVPA analysis we computed a t-statistic for each speed and luminance for each voxel within these areas. Separately for each luminance, we then assigned each voxel a label corresponding to the speed that yielded the maximal t-statistic (i.e. the most significant response). For each area we subdivided the region corresponding to the visual stimulus (identified by voxels showing a significant increased response to the speed stimuli, defined as the voxels with an $R^2 > 0.25$) into ten bins along the cortical surface in the eccentricity dimension extending from the representation of the inner (peri-foveal) boundary of the stimulus to the outer (peripheral) boundary of the stimulus. The bins had equal width in terms of cortical distance. Data for the two most peripheral bins were excluded as only a small proportion (<10%) of voxels in these bins met the criteria for inclusion ($R^2 > 0.25$ and positive stimulus-evoked BOLD responses). Speed preferences were averaged across voxels within each bin and then averaged across subjects to yield a plot of average speed preference as a function of cortical distance from fovea to periphery for each luminance level. A resampling test was used to assess whether the difference in slopes between high and low luminance conditions was significant. For this test, a regression line was fit to the binwise difference between high and low luminance condition data to yield a regression coefficient b and a correlation coefficient r. Speed preference data for each subject, hemisphere, and luminance condition were then assigned randomly to one of two equal-sized sets. A regression line was then fit to the binwise difference between the two sets in the same way as for the real data. This procedure was repeated 10,000 times to provide

an estimate of the probability of obtaining a difference in slopes between the two sets by chance. The significance of the actual difference in slopes was estimated by computing the proportion of resampling iterations where the difference in slopes b and the regression coefficient r was greater than or equal to the observed ones.

Results

The speed tuning functions of the BOLD response were very similar to and consistent with those obtained in Experiment 1 (Fig 4). With the exception of V3A and V7, all areas showed band-pass tuning for speed. Since the maximum speed in Experiment 2 (10 deg/second) was lower than that in Experiment 1 (13.3 deg/second), the monotonic increase in BOLD response with speed in V3A and V7 is consistent with band-pass tuning with a peak response at or beyond 10 deg/seconds, as suggested by the results of Experiment 1 (Fig 1).
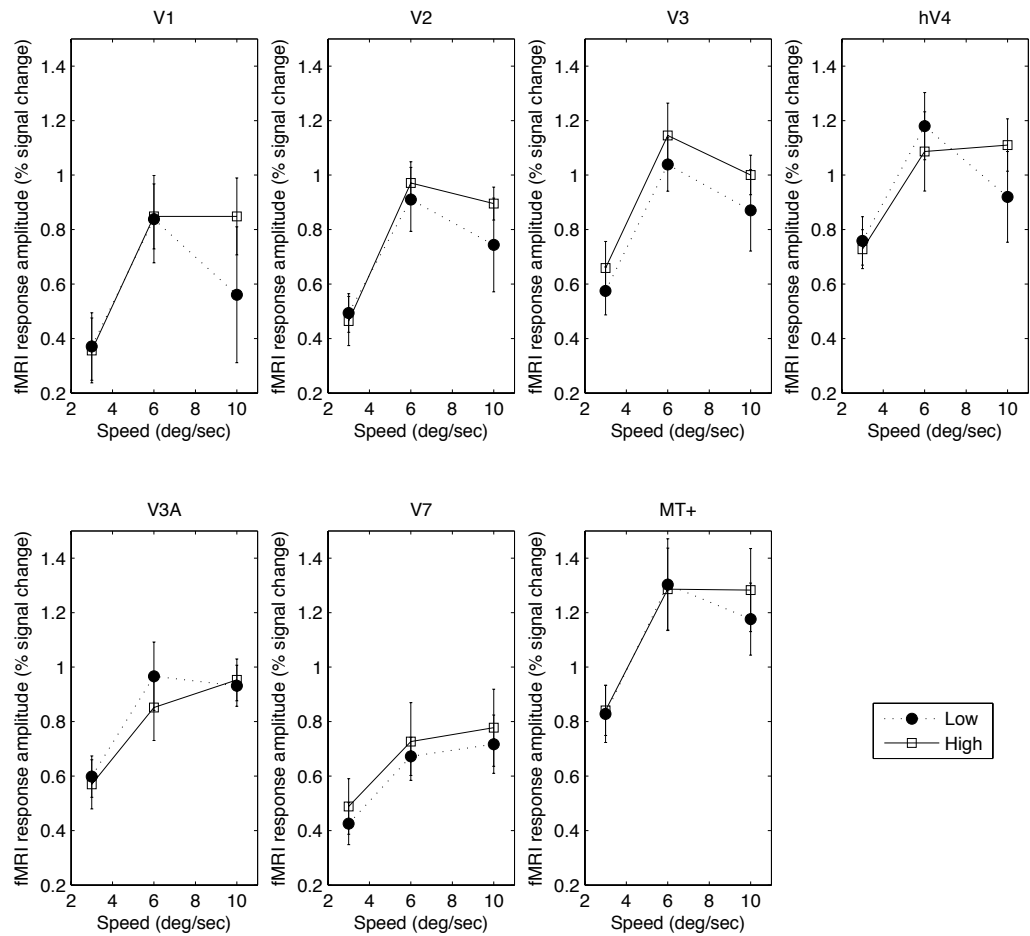
Fig 4 Mean BOLD response as a function of speed for low (closed symbols) and high (open symbols) luminance in Experiment 2. The areas are indicated above each panel. Error bars represent ± 1 SEM.

The results of the MVPA analysis for Experiment 2 are plotted in Fig 5. As in Experiment 1, in all areas the classifiers' predicted speeds rose monotonically with physical speed. However, contrary to our hypothesis, there was no discernible difference in the predicted speeds between the within- and across-luminance classification conditions.
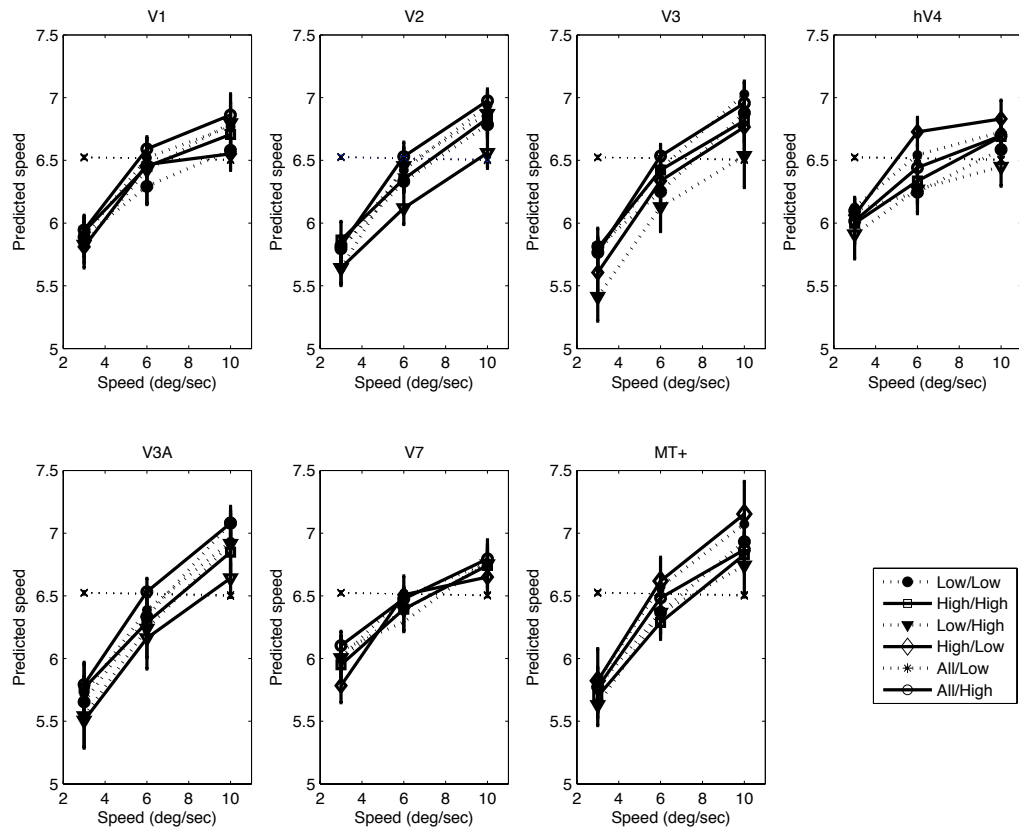
Fig 5 Predicted speed of the classifier as a function of stimulus speed is plotted for high (open symbols) and low (closed symbols) luminance and for low-high and high-low training regimes in Experiment 2. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

The lack of a consistent effect of luminance on response amplitude or decoding performance could indicate that luminance has little effect on visually evoked cortical responses, as suggested by results showing little or no modulation of fMRI BOLD response in V1 over a wide range of luminance levels (Hadjikhani and Tootell 2000). Alternatively, these results could indicate that luminance is encoded in cortical BOLD response but in a manner that does not systematically influence response magnitudes. To distinguish between these hypotheses we used the MVPA analysis to decode luminance instead of speed, training the classifier on data across low and high luminance runs and ignoring speed. The results of this analysis are plotted in Fig 6. With the possible exception of V7, the classifier successfully decoded stimulus luminance in every visual area examined, with early visual areas (V1-V3) showing most accurate decoding performance. These results show that luminance is indeed encoded in the multivariate BOLD response, although these effects appear not to bias speed responses or speed decoding.
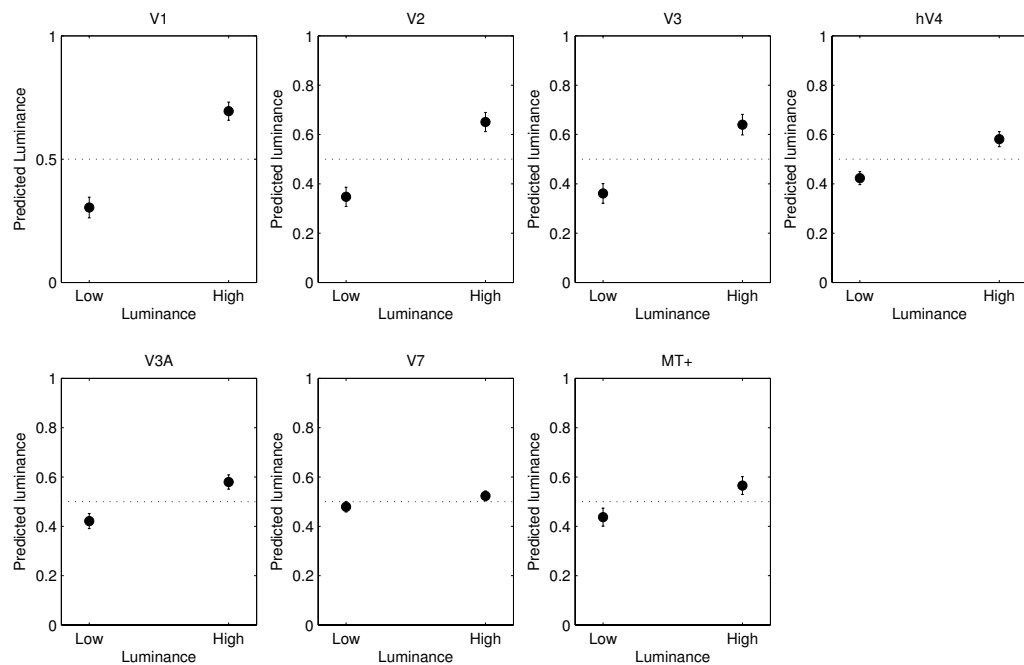

Fig 6 Predicted luminance as a function of actual luminance. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Spatial distribution of speed and luminance tuning

The observation that both speed and luminance could be accurately decoded from the multivariate BOLD response but not in the overall response averaged across all voxels indicated that voxels differed in their motion and luminance preferences. We wondered whether such differences would show a large-scale spatial distribution, specifically with regards to eccentricity. Magnocellular and parvocellular neurones differ in temporal frequency tuning, luminance and contrast response (Kaplan, Lee & Shapley, 1990; Purpura, Kaplan & Shapley, 1988) and are heterogeneously distributed across visual eccentricity (Dacey, 1994), thus the differences in speed/frequency and/or luminance preference could reflect such differences. To investigate this possibility we plotted speed preference across the cortical surface along the eccentricity dimension in V1 for high and low luminance conditions, averaged across subjects (Fig 7). For the low-luminance condition, a clear pattern was observed in that peri-foveal locations preferred slower speeds than peripheral locations. The high-luminance condition also shows an effect of eccentricity but the slope of the function is reduced, such that the two functions are well separated in the parafovea but converge in the periphery. Although we found no overall effect of luminance at any speed in Experiment 1 (Fig 1), the finding that the preferred speed of parafoveal voxels varies with luminance suggests that localized effects of luminance may exist (but see Discussion for consideration of the possible effects of measurement method).

The difference in slope between high and low luminance was statistically significant (resampling test, $P<0.05$). A similar but weaker and non-significant pattern was observed in V2 and V3 (data not shown). Hence, at least for V1, there appears to be large-scale variation in speed preference with eccentricity that differs between high and low luminance, suggesting that selectivity for speed and luminance interact in a systematic manner in this area.
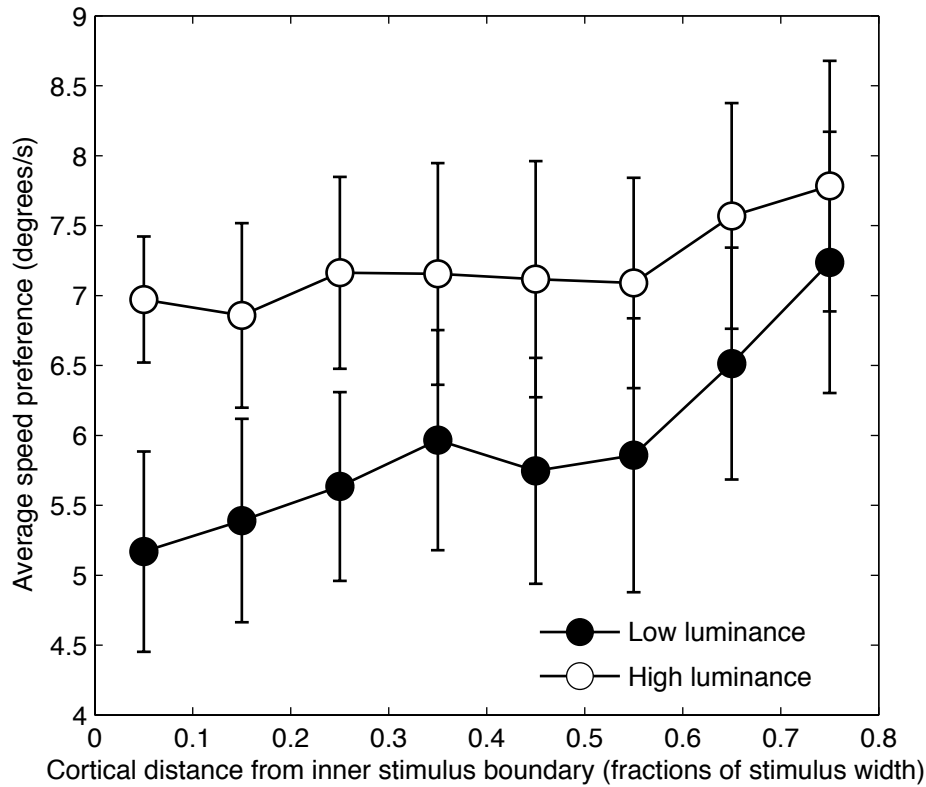
Fig 7 Preferred speed as a function of cortical distance from inner (foveal) stimulus boundary for high (open symbols) and low (closed symbols) luminance. Cortical distance corresponds broadly (but non-linearly) to eccentricity, which increases left-to-right. Error bars represent $\pm$ 1 SEM.

Discussion

Our results demonstrate that, at least for narrowband stimuli, the BOLD response to speed is band-pass, peaking between 4 and 10 deg/sec, in all cortical areas measured and at both luminances measured. Comparison of our results with those obtained for high luminance by Singh et al (2000) indicates good quantitative agreement over the speed range tested, in all areas. Thus the global population response does not render an unambiguous speed signal. This is perhaps unsurprising given the large body of electrophysiological evidence that both speed and spatio-temporally separable responses are found in the visual cortex. The BOLD response to a drifting grating in a given cortical region must be assumed to reflect the sum of activities of all mechanisms that exist in that region. Thus the band-pass tuning of the average

BOLD response reflects the composite activity of speed and temporal frequency tuned neurones. There is considerable psychophysical evidence for the existence of two (or possibly three) non-directional temporally tuned mechanisms (e.g. Kulikowski & Tolhurst 1973; Hammett & Smith 1992; Hess & Snowden 1992) whose overall response envelope is band-pass, peaking around 6-8Hz. Whilst the precise relation of these mechanisms to underlying neural physiology is far from clear, they do share temporal tuning characteristics that are broadly similar to the temporal tuning of Magno- and Parvo-cellular cells, the outputs of which presumably form the substrate for speed-tuned neurones (see e.g. DeValois et al 2000). Regardless of its precise neural substrate, our results are consistent with previous studies (e.g. Singh et al, 2000) that indicate that the average BOLD signal contains no unambiguous information regarding speed. It should again be noted that because we, like previous studies, have used only one spatial frequency we cannot disambiguate speed and temporal frequency and thus our result may reflect coding of flicker rather than speed *per se.*

However, the results of the MVPA classification indicate that although speed information is not unambiguously encoded in the global response, the multivariate response across voxels within each visual area does contain information about speed. Whilst the MVPA classifier could not distinguish the two slowest speeds measured in Experiment 1, it could distinguish them from higher speeds each of which was also unambiguously predicted. Thus, whilst the average BOLD response conveys no unambiguous speed information, the multivariate response in all cortical areas measured yields clear information about the speed of the stimulus over most of the range tested. However, we found no difference in either the mean BOLD response or the multivariate response across luminance and no difference in the classifier predicted speed applied within or across luminance conditions. This is surprising given the known variation in perceived speed with luminance (Hammett et al 2007; Vaziri-Pashkam and Cavanagh 2008) and the known changes in response of retinal ganglion cells with luminance

(Purpura et al 1988). We know of few studies that have directly measured the fMRI BOLD response as a function of luminance and it may be that early (retinal) luminance gain control effectively renders the BOLD response relatively immune to luminance modulation. Indeed, Hadjikhani and Tootell (2000) have shown that a million-fold reduction in luminance yielded little change in activity in areas from V1 to V4. Moreover, Cornelissen et al (2006) report that neither luminance nor perceived luminance modulations are accompanied by commensurate changes in the BOLD signal in V1 and V2 and Leonards et al (2003) report a cortical region tuned for luminosity perception that does not increase its activity with increasing luminance.

Whilst we found no systematic change in overall BOLD response with luminance, consistent with previous studies, MVPA analysis could successfully decode luminance in almost every visual area, indicating that, just like for speed, information about stimulus luminance is encoded in the multivariate response across voxels. The lack of effect of luminance on speed decoding would thus suggest that speed and luminance are encoded independently in visual cortex. Such a conclusion is however at odds with the effect of luminance on variations in speed preference as a function of eccentricity that we observed in V1, which indicated an *interaction* between speed and luminance. How can these two seemingly contradictory observations be reconciled? It is likely that the difference in results is at least in part due to differences in analysis methods. The variation in speed preference with eccentricity was computed by assigning to each voxel the speed label corresponding to the speed that evoked the most significant (rather than strongest) response. This is essentially a "winner-takes-all" algorithm, which being inherently non-linear can potentially exaggerate differences in responses to different speeds, even if the underlying response differences are small. Moreover, the computation of speed preferences was performed only on voxels showing a positive BOLD response (increase) to the speed stimuli, whereas the classification analysis was performed on all voxels showing a significant stimulus-evoked response, whether positive

(increase) or negative (decrease). Indeed, the observation that speed classification was equally or more accurate in V2 and V3 than in V1, even though there was much less evidence of a systematic eccentricity effect on speed preference in these areas than in V1, suggests that the classifier was not primarily driven by eccentricity-specific variations in speed preference. Similarly, for the high luminance condition speed preference did not vary with eccentricity in V1, yet speed classification performance was no worse for the high luminance than the low luminance conditions. Together this suggests that the effect of eccentricity on speed preference did not bias classifier performance, even if the classifier may in part have relied on information contained in these large scale biases. Although there is controversy about the spatial scale of BOLD responses underlying MVPA (Freeman et al. 2011, Kamitani & Sawahat 2010), it is likely that information in the BOLD response may be found at several spatial scales, all of which may drive classification performance (Chaimow et al., 2011). The fact that the speed classification analysis did not show a sensitivity to luminance could thus reflect the existence of smaller scale speed biases in the multivariate response which might have been invariant to luminance differences. While the classification results indicate that information about speed and luminance are encoded independently at the coarse scale of the multivariate BOLD response, these results do not rule out that the two stimulus features interact at the single neuron level, which could potentially account for the effect of luminance on perceived speed.

In summary, our results are broadly consistent with a scheme whereby the average BOLD response reflects the envelope of neural responses, many of which may not be speed tuned, but whose multivariate response contains unambiguous speed (or temporal frequency) information in all visual areas measured. We propose that the most parsimonious account of our results is that speed encoding is only implicitly represented in visual cortex and is estimated from the multivariate population response. How could such an implicit code for

speed be realized? One possibility is that, rather than a speed code being extracted upon the basis of the tuning of speed sensitive cells, either the responses of discrete neural sub-populations or the multivariate response of entire regions is used to render a speed-related signal. Our results do not speak directly to any specific model of speed perception but they do indicate that speed or temporal frequency tuning may occur as early as V1 and that any model must reflect either the response properties of sub-populations of neurons within motion-sensitive areas or the multivariate population response of those areas.

References

Adelson E H, Bergen J R Spatiotemporal energy models for the perception of motion.*Journal of the Optical Society of America A* 284-99, 1985.

Ascher D, Grzywacz N M. A bayesian model for the measurement of visual velocity. *Vision Res*40: 3427-34, 2000.

Bardy C, Huang J Y, Wang C, FitzGibbon T, Dreher B. 'Simplification' of responses of complex cells in cat striate cortex: suppressive surrounds and 'feedback' inactivation. *Journal of Physiology*574: 731-50, 2006.

Braddick O. A short-range process in apparent motion.*Vision Res***14**: 519-27, 1974.

Boynton G M, Demb J B, Glover G H, Heeger, D J. Neuronal basis of contrast discrimination. Vision Res 39: 257 – 269, 1999.

Burock MA, Dale AM. Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. *Hum Brain Mapping*11:249-60, 2000.

Castet E, Lorenceau, J, Shiffrar M, Bonnet C. Perceived speed of moving lines depends on orientation, length, speed and luminance . *Vision Res*33: 1921-1936, 1993.

Chaimow D, Yacoub E, Ugurbil K, Shmuel A. Modeling and analysis of mechanisms underlying fMRI-based decoding of information conveyed in cortical columns. *NeuroImage* 56:

627-42, 2011.

Chawla D, Phillips J, Buechel C, Edwards R, Friston K J, Speed-dependent motion-sensitive responses in V5: an fMRI study. *NeuroImage*7: 86-96, 1998.

Chawla D, Buechel C, Edwards R, Howseman, A, Josephs O, Ashburner J, Friston K J. Speed-dependent responses in V5: A replication study. *NeuroImage* 9: 508-15, 1999.

Cornelissen FW, Wade AR, Vladusich T, Dougherty RF, Wandell BA. No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *Journal of Neuroscience* 26: 3634–41, 2006.

Dacey D. Physiology, morphology and spatial densities of identified ganglion cell types in primate retina. In: *Higher-Order Processing in the Visual System*, CIBA Foundation. New York: Wiley, 1994, p. 12–28.

Deichmann, R, Schwarzbauer, C. & Turner, R. Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. *Neuroimage* 21: 757-67, 2004.

DeLange H. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light.*Journal of the Optical Society of America*48, 777-784, 1958.

De Valois RL, Cottaris NP, Mahon LE, Elfar SD, Wilson J a. Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision research* 40: 3685–702, 2000.

Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. *Ann. Statist.*32: 407-499, 2004.

Foster K H, Gaska J P, Nagler M, Pollen D A. Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology* 365: 331-63,1985.

Freeman J, Brouwer G J, Heeger D J, Merriam E P. Orientation decoding depends on maps, not columns. *Journal of Neuroscience*: 4792-804, 2011.

Gegenfurtner K R, Mayser H, Sharpe L T. Seeing movement in the dark. *Nature*398: 475-6, 1999.

Gegenfurtner, K. R., Mayser, H. M. & Sharpe, L. T. Motion perception at scotopic light levels. *Journal of the Optical Society of America A*17: 1505-15, 2000.

Hadjikhani N, Tootell R.B.H. Projection of rods and cones within human visual cortex. *Human Brain Mapping* 9:51-63, 2000.

Hagberg G, Zito G, Patria F, Sanes J. Improved detection of event-related functional MRI signals using probability functions.*Neuroimage*14: 1193-1205, 2001.

Hagenbeek R E, Rombouts S A,RB, Van Dijk B W, Barkhof F. Determination of individual stimulus--response curves in the visual cortex. *Human brain mapping*17: 244-50, 2002.

Hammett S T, Smith A T. Two temporal channels or three? A Re-evaluation. *Vision Res*, 32: 285-291, 1992.

Hammett S T, Champion R A, Morland A B, Thompson P G. A ratio model of perceived speed in the human visual system.*Proceedings of the Royal Society (B)* 272: 2351-6, 2005.

Hammett S T, Champion RA, Thompson P G, Morland A B. Perceptual distortions of speed at low luminance: Evidence inconsistent with a Bayesian account of speed encoding. *Vision Res* 47: 564-568,2007.

Harris M G. Velocity specificity of the flicker to pattern sensitivity ratio in human vision.*Vision Res*20: 687-91, 1980.

Hess R F, Snowden R J. Temporal properties of human visual filters: number, shapes and spatial covariation. *Vision Res*32: 47-59, 1992.

Huk A C, Dougherty R F, Heeger D J. Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience* 22: 7195-205, 2002.

Jenkinson M, Bannister PR, Brady JM, Smith S M. Improved optimisation for the robust and accurate linear registration and motion correction of brain images.*NeuroImage*, 17:825-841, 2002.

Kamitani Y, Sawahata Y. Spatial smoothing hurts localization but not information: pitfalls for brain mappers. *Neuroimage* 49:1949-52, 2010.

Kaplan E, Lee B, Shapley R. New Views of Primate Retinal Function. In: Osborn N.N and Chader GJ (Eds) *Progress in Retinal Research*. 1990, p. 273–336.

Krekelberg, B, Van Wezel R J A, Albright T D. Interactions between speed and contrast

tuning in the middle temporal area: implications for the neural code for speed. *Journal of Neuroscience* 26: 8988-98, 2006.

Kulikowski J J, Tolhurst D J. Psychophysical evidence for sustained and transient detectors in human vision.*Journal of Physiology* 232: 149-62, 1973.

Larsson J. Imaging vision: Functional mapping of intermediate visual processes in man. Karolinska Institutet. 2001. http://diss.kib.ki.se/2001/91-7349-090-3/

Larsson J, Heeger D. Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience* 26: 13128–13142, 2006.

Langley, K. & Anderson, S. J. Subtractive and divisive adaptation in visual motion computations.*Vision Res* 47: 673-86, 2007.

Lee B B, Pokorny J, Smith V C, Martin P R, Valberg A. Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A***7:** 2223-36,1990.

Leonards U, Troscianko T, Lazeyras F, Ibanez V. Cortical distinction between the neural encoding of objects that appear to glow and those that do not. *Brain research. Cognitive brain research* 24: 173–6, 2005.

Lingnau A, Ashida H, Wall M B, Smith A T. Speed encoding in human visual cortex revealed by fMRI adaptation. Journal of Vision 9: 3 1-14.2009.

Liu J, Newsome W T. Correlation between speed perception and neural activity in the middle temporal visual area. Journal of Neuroscience 25: 711-22, 2005.

Lui L L, Bourne J A, Rosa M G. Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrixjacchus). *European Journal of*

Neuroscience 25: 1780-92, 2007.

Mullen K T, Baker C L Jr. A motion aftereffect from an isoluminant stimulus. *Vision Res* 25: 685-8, 1985.

Naka K I, Rushton W A. S-potentials from colour units in the retina of fish (Cyprinidae). *Journal of Physiology* 185: 536-55, 1966.

Olman C A, Ugurbil K, Schrater P, Kersten D. BOLD fMRI and psychophysical measurements of contrast response to broadband images. *Vision Res* 44: 669-683, 2004.

Perrone J A. Economy of scale: a motion sensor with variable speed tuning. *Journal of Vision* 5: 28-33, 2005.

Perrone J A, Thiele A. Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience* 4: 526-32, 2001.

Priebe N J, Cassanello C R, Lisberger S G. The neural representation of speed in macaque area MT/V5. *Journal of Neuroscience* 23: 5650-61, 2003.

Priebe N J, Lisberger S G. Estimating target speed from the population response in visual area MT. *Journal of Neuroscience* 24: 1907-16, 2004.

Priebe N J, Lisberger S G. Movshon J A. Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *Journal of Neuroscience* 26: 2941-50, 2006.

Purpura K, Kaplan E, Shapley R M. Background light and the contrast gain of primate P and M retinal ganglion cells. *Proc Natl Acad Sci U S A* 85: 4534-7, 1988.

Reisbeck T E, Gegenfurtner K R. Velocity tuned mechanisms in human motion processing. *Vision Res* 39: 3267-85, 1999.

RosaM G, Tweedale R, Elston G N. Visual responses of neurons in the middle temporal area of new world monkeys after lesions of striate cortex. *Journal of Neuroscience* 20: 5552-63, 2000.

Rudolph K, Pasternak T. Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey.*Cerebral Cortex*9: 90-100, 1999.

Sereno M I, Dale A M, Reppas J B, Kwong K K, Belliveau J W, BradyT J, Rosen B R, Tootell R B. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*268: 889-93, 1995.

Singh K D, Smith A T, Greenlee M W. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. *Neuroimage*12: 550-64, 2000.

Smith, A T, Edgar,G K. Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception.*Vision Res*34: 253-65, 1994.

Smith, A. T., Wall, M. B., Williams, A. L. & Singh, K. D. Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience* 23: 561-9, 2006.

Snowden R J, Hess R F, Waugh S J. The processing of temporal modulation at different levels of retinal illuminance. *Vision Res* 35: 775-789, 1995.

Stocker A A, Simoncelli E P. Noise characteristics and prior expectations in human visual speed perception.*Nature Neuroscience* 9: 578-85, 2006.

Sun P, Ueno K, Waggoner R A, Gardner J L, Tanaka K, Cheng K. A temporal frequency-dependent functional architecture in human V1 revealed by high-resolution fMRI. *Nature Neuroscience* 10: 1404-6, 2007.

Thompson P, Brooks K, Hammett S T. Speed can go up as well as down at low contrast: Implications for models of motion perception. *Vision Res* 46: 782-786, 2006.

Tolhurst D J, Movshon J A. Spatial and temporal contrast sensitivity of striate cortical neurones.*Nature*257: 674-5, 1975.

Vaziri-pashkam M, Cavanagh P. Apparent speed increases at low luminance. *Journal of Vision*8: 1-12, 2008.

Wall M B, Lingnau A, Ashida H, Smith A T. Selective visual responses to expansion and rotation in the human MT complex revealed by functional magnetic resonance imaging adaptation. *European Journal of Neuroscience* 27: 2747-57, 2008.

Figure Legends

Fig 1 BOLD response as a function of speed for low (closed symbols) and high (open symbols) luminance in Experiment 1. Cortical area is indicated above each panel. Error bars represent ±1 SEM.

Fig 2 Predicted speed of the classifier as a function of stimulus speed is plotted for high (open symbols) and low (closed symbols) luminance in Experiment 1. Broken lines represent classifier performance for random assignment of speeds. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Fig 3 Results of univariate classification analysis. In each area, collapsing data across voxels reduces classifier performance to chance levels. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Fig 4 Mean BOLD response as a function of speed for low (closed symbols) and high (open symbols) luminance in Experiment 2. The areas are indicated above each panel. Error bars represent ± 1 SEM.

Fig 5 Predicted speed of the classifier as a function of stimulus speed is plotted for high (open symbols) and low (closed symbols) luminance and for low-high and high-low training regimes in Experiment 2. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Fig 6 Predicted luminance as a function of actual luminance. Cortical area is indicated above each panel. Error bars represent ± 1 SEM.

Fig 7 Preferred speed as a function of cortical distance from stimulus boundary for high (open symbols) and low (closed symbols) luminance. Error bars represent ± 1 SEM.