Title: Dynamic Population Codes of Multiplexed Stimulus Features in Primate Area MT.

Running Title: Dynamic Population Codes of Stimulus Features in Area MT.

Authors: Erin Goddard[i,ii] Samuel G. Solomon[iii] Thomas A. Carlson[i,ii]

Affiliations: (i) School of Psychology, University of Sydney, Sydney, NSW, 2006, Australia

(ii) ARC Centre of Excellence in Cognition and its Disorders (CCD), Macquarie University, Sydney, NSW, 2109, Australia

(iii) Department of Experimental Psychology, University College London, Gower Street, London, WC1E 6BT, United Kingdom

Corresponding Author: Erin Goddard
School of Psychology, University of Sydney,
Sydney, NSW, 2006, Australia
erin.goddard@sydney.edu.au

Conflict of interest: None to report

Abstract

The middle-temporal area (MT) of primate visual cortex is critical in the analysis of visual motion. Single-unit studies suggest that the response dynamics of neurons within area MT depend on stimulus features, but how these dynamics emerge at the population level, and how feature representations interact, is not clear. Here, we used multivariate classification analysis to study how stimulus features are represented in the spiking activity of populations of neurons in area MT of marmoset monkey. Using Representational Similarity Analysis (RSA) we distinguished the emerging representations of moving grating and dot field stimuli. We show that representations of stimulus orientation, spatial frequency and speed are evident near the onset of the population response, while the representation of stimulus direction is slower to emerge and sustained throughout the stimulus-evoked response. We further found a spatiotemporal asymmetry in the emergence of direction representations. Representations for high spatial frequencies and low temporal frequencies are initially orientation-dependent, while those for high temporal frequencies and low spatial frequencies are more sensitive to motion direction. Our analyses reveal a complex interplay of feature representations in area MT population response that may explain the stimulus-dependent dynamics of motion vision.

New & Noteworthy

Simultaneous multi-electrode recordings can measure population-level codes that previously were only inferred from single-electrode recordings. However, many multi-electrode recordings are analyzed using univariate single-electrode analysis approaches, which fail to fully utilize the population-level information. Here, we overcome these limitations by applying multivariate pattern classification analysis and Representational Similarity Analysis (RSA) to large-scale recordings from area MT in marmoset monkeys. Our analyses reveal a dynamic interplay of feature representations in area MT population response.

Introduction

Information about the external world is carried by signals that are distributed across populations of neurons within sensory brain areas. One of the most studied areas is the middle-temporal area (MT) of primate visual cortex. Area MT contains a high proportion of cells that are selective for motion direction and speed (Maunsell and van Essen, 1983; Albright, 1984; Movshon et al., 1985), whose activity correlates with perception of motion (Newsome et al., 1989; Salzman et al., 1990; Britten et al., 1996). While it is assumed that neurons in area MT provide a population code for motion (Simoncelli and Heeger, 1998; Jazayeri and Movshon, 2006), most electrophysiological work, like that conducted in other brain areas, has characterized the functional properties of individual neurons.

Our understanding of population codes in area MT has generally been established by extrapolating from single neuron responses, largely overlooking the diversity of neural tuning across the population and possible synergy in the signals of individual neurons. For example, a typical experiment will set the preferred size, spatial and temporal frequency for each neuron, before the responses to different grating directions are collected and averaged. By definition this approach excludes the possibility of uncovering interactions between selectivity for different stimulus features, and measuring multiplexed feature representations. Yet in some stimulus conditions motion signals evolve over 50-75ms after the onset of the stimulus-evoked response (Pack and Born, 2001; Borghuis et al., 2003; Smith et al., 2005), whereas in other conditions almost all information is present in the first few spikes (Osborne et al., 2004). Additionally, the temporal window of motion integration depends on stimulus speed, spatial frequency and contrast (Bair and Movshon, 2004). This work suggests that the response properties of MT neurons not only change over time, but that the nature and timing of these changes are stimulus dependent and carry stimulus-related information. To measure the stimulus-related information that is potentially available in the population response, we need to capture not only the response properties of single units, but also the dynamics of the information carried by a population of these units.

Modern techniques enable direct measurement of population activity, but we lack a unified framework for interpreting those measurements. One approach that holds promise is to combine multivariate pattern classifiers with Representational Similarity Analysis (RSA), which has been used to study high-level object representations and lower-level visual features in fMRI and MEG (Kriegeskorte et al., 2008; Cichy et al., 2013; Wardle et al., 2016). Classification analyses have been

employed to measure stimulus-related information present in the population response along single stimulus dimensions (McDonald et al., 2014; Chen et al., 2015; Zavitz et al., 2016). A multivariate approach affords the capacity to better understand how different stimulus features may be simultaneously represented in these populations (i.e. multiplexing), and how these representations dynamically change over time. By leveraging the response of multiple neurons these classification analyses make it possible to measure the stimulus-related information at different times after stimulus onset, and to explore the dynamics of population representations. Here we made multielectrode recordings from area MT of marmoset monkeys, and applied pattern classification to data from short time bins. We combine these classification analyses with RSA to test whether the population response in area MT to simple stimuli (moving gratings and dot fields) could multiplex motion properties and spatial features.

Materials and Methods

Experimental Preparation

Six adult marmosets (Callithrix jacchus; 5 males; weight 290-400 g) were obtained from the Australian National Health and Medical Research Council (NHMRC) combined breeding facility. Procedures took place at the University of Sydney and were approved by Institutional (University of Sydney) Animal Ethics Committee and conform to the Society for Neuroscience and NHMRC policies on the use of animals in neuroscience research. The data analyzed here were obtained during the course of other work, published previously (Solomon et al., 2015; McDonald et al., 2014), and the details of the experimental preparation can be found there.

Briefly, each animal was first sedated with an intramuscular (IM) injection of 12 mg/kg of Alfaxan and 3 mg/kg of Diazepam. Subsequent surgery was performed under supplemental local anesthesia (Lignocaine 2%; Astra Zaneca). A femoral vein was cannulated, the trachea exposed and an endotracheal tube inserted, and the animal placed in a stereotaxic frame. Post-surgical anesthesia was maintained by continuous intravenous infusion of sufentanil citrate (4-12 μg/kg/h; Sufenta Forte, Janssen Cilag) in physiological solution. The electrocardiogram, electroencephalogram (EEG), and SpO2 were monitored continuously. Dominance of low frequencies (1-5 Hz) in the EEG recording, and absence of EEG changes under noxious stimulus (tail pinch) were used as the chief sign of an adequate level of anesthesia. At any sign of the anesthesia becoming less effective, the

dose of sufentanil citrate was increased. Muscular paralysis was then induced and maintained by continuous infusion of pancuronium bromide (0.3 mg/kg/h; Astra Zaneca). The animal was artificially ventilated, with a 70:30 mix of N_2O and Carbogen, so as to keep end-tidal CO_2 near 33mm Hg. Rectal temperature was kept near 38°C with the use of a heating blanket. The corneas were protected with high-permeability contact lenses that remained in place for the duration of the experiment. No artificial pupils were used. At the end of the experiment, the animal was euthanized with intravenous 500 mg/kg sodium pentobarbitone (Lethobarb; Verbac Australia).

Electrophysiological recordings

In each animal, a craniotomy was made over area MT, a large durotomy was made and extracellular recordings were obtained using a 10 x 10 grid of parylene-coated platinum iridium microelectrodes (1.5mm in length, spacing 0.4mm; Blackrock Microsystems), pneumatically inserted to a depth of approximately 1mm (Rousche and Normann, 1992). Signals were band-pass filtered (0.3-6 kHz), and sampled by a Tucker Davis Technologies RZ2 at 24 kHz. For all implants, we identified electrodes that were likely to be located within area MT or MTc based on the directional-sensitivity of the multi-unit recordings, and using the trajectory of receptive field positions (Rosa and Elston, 1998), as described in detail by Solomon and colleagues (Solomon et al., 2015). Across animals, 59-96 of the possible 96 electrodes were located within area MT or MTc, and were included in the analyses below.

Visual stimuli

Visual stimuli were drawn at 8-bit resolution using commands to OpenGL, by custom software (EXPO; P. Lennie) running on a G5 Power Macintosh computer. Stimuli were displayed on a calibrated cathode ray tube monitor (Sony G520, refresh rate 100 Hz, mean luminance 45-55 cd/m^2, width 40cm and height 30cm). The monitor was viewed at a distance of 45cm. During measurements, one eye, usually the ipsilateral eye, was occluded.

In one set of stimuli, a large sine-wave grating (Michelson contrast 0.5) drifted within a circular window (diameter 30 degrees) with hard edges; outside the window, the luminance was held at the mean value. The spatial frequency was either 0.1, 0.32 or 1 cycles/degree, and temporal frequency was 2.5, 7.69 or 25 Hz. Gratings moved in one of 12 directions (30 degree steps), and each stimulus was presented for 500ms. The screen was held for 50ms at the mean luminance between each trial.

There were 20 repeats of each stimulus type (9 spatiotemporal frequencies x 12 directions), presented in a pseudo-random order, giving a total of 2160 trials of the grating stimuli.

In the second stimulus set, white circular dots (Weber contrast 1.0; diameter 0.4 degrees) moved across a quasi circular area (diameter 48 degrees; cropped at 37 degrees vertically) of the monitor; outside each dot, the monitor was held at the mean luminance. Dots were presented at a density of 0.3 dots/s/degree and moved with 100% coherence and infinite lifetime. The position of each dot at the beginning of a trial was specified by a random number generator; the same set of positions was used on every trial. There were 20 repeats of each stimulus type (7 speeds x 12 directions), presented in a pseudo-random order, giving a total of 1680 trials of the moving dot field stimuli. Both grating and dot field stimuli were moved at a constant speed from the time of stimulus onset (t=0ms).

For both stimulus sets, there were also blank trials on which no stimulus was displayed, and the screen remained at the mean luminance. There were 180 blank trials in the first stimulus set, and 140 in the second (1/13 of the total number of trials). These trials were pseudo-randomly interleaved with the stimulus trials, and the measured responses were used to estimate the spontaneous rate of spiking for each electrode in each dataset.

Preliminary data analysis

For each of the electrodes identified as being within area MT, we used the Matlab function *findpeaks* to identify candidate waveforms with peak amplitude that exceeded 3 standard deviations of the raw signal on that channel. We down-sampled the multi-unit activity of each channel to 500Hz, extracting the number of spike waveforms on that channel in each 2ms time bin. We did not sort spike waveforms into separate neuronal sources, so spike rates are expressed as the number of spikes per electrode. Data for each trial were extracted by taking the spike counts from 0ms to 600ms relative to stimulus onset. For each animal we characterised the dataset by plotting the average electrode spike rate as a function of different stimulus features (including direction, speed, and spatial and temporal frequency). We calculated direction index (*DI*) over time by finding the preferred stimulus direction for each electrode, and extracting the electrode's response to its preferred direction over time (*pref*) and to its antipreferred direction (*antipref*, preferred direction - 180 degrees). Both *pref* and *antipref* were normalised by dividing by the mean rate across stimuli for that electrode at each time bin. Then:

$$DI = (pref - antipref)/pref \quad (1)$$

Similarly, we also calculated the orientation index of each electrode. For the responses to grating stimuli, we averaged spike rates across gratings of the same orientation that were moving in opposite directions. For comparison, we also created a pseudo-orientation index for the responses to moving dot fields by averaging spike rates across dot fields that were moving in opposite directions. Then the normalised responses to the preferred orientation (*pref*) and antipreferred orientation (*antipref*, preferred orientation - 90 degrees) were used to calculate the orientation index *OI*:

$$OI = (pref - antipref)/pref \quad (2)$$

We then investigated the population tuning further using a classification analysis.

Classification analysis

We reduced the dataset by applying principal component analysis (PCA) to the entire dataset for each animal. Each animal's initial dataset comprised the entire 600ms of neuronal response following stimulus onset, for each of 2160 (gratings) or 1680 (dots) trials and up to 96 channels. Following PCA, the original channel data (up to 96) were reduced to spatial components, ordered by amount of variance for which they accounted (highest to lowest). Data from the first *n* components that accounted for 99% of the variance were retained; data from remaining components were discarded. Across animals and stimulus type (gratings and dots) *n* ranged from 54 to 93.

In order to measure the similarity or dissimilarity of population multi-unit activity in different stimulus conditions, we used linear discriminant analysis (LDA) to study the variability in the decodability of different stimulus conditions. We also repeated the analysis using a linear support vector machine (SVM) classifier, and obtained very similar results (data not shown). For each possible pair of the 108 unique grating stimuli we trained the classifier to discriminate between two stimulus conditions (for example, discriminating responses to a grating of 1 cycle/deg, moving at 25 cycles/s, to the right, from the responses to a grating of 0.1 cycles/deg, moving at 2.5 cycles/s, to the left), then tested the classifier accuracy using 10-fold cross-validation. The classification rule was learnt using 90% of trials (18 trials of each type), and the accuracy of this rule was tested on the remaining 10% of trials (2 trials of each type). This process was repeated for each of 10 partitions of

the data, such that all data were included in the test set once, and no data were ever used in both the training and test set (leave-one-out train-and-test).

Similarly, we trained classifiers to discriminate each pairing of the 84 unique dot field stimuli, and again tested the classifier accuracy using 10-fold cross-validation. To measure how classification accuracy evolved over time we repeated this process at each time point (every 2ms) in the 600ms window.

For both moving grating and moving dot field stimuli the entire classification analysis was performed separately for each animal, and the average classification accuracy was obtained by averaging classifier performance (% correct) across animals.

Representational Similarity Analysis (RSA)

We employed 'Representational Similarity Analyses' (RSA) to track the evolution of representations of image features in area MT. RSA analyses have been applied previously to analyses of object representations obtained by fMRI measurements and electrophysiological recordings (Kriegeskorte et al., 2008), as well as MEG (Cichy et al., 2014; Redcay and Carlson, 2015; Goddard et al., 2016). To do this we constructed representational dissimilarity matrices (RDMs), separately for grating and dot stimuli. Each RDM was a 108x108 matrix (for moving grating stimuli data) or 84x84 matrix (for moving dot field stimuli data), where each cell in the RDM is the classification accuracy for a single pair of stimuli. The diagonal axis of these matrices are nominally zeros, and the matrix is by definition symmetric about the diagonal axis, so for all correlation values calculated below we included only the triangular part of the matrix above the diagonal. We correlated each of these RDMs with model matrices, each describing the predicted dissimilarity of the pair of stimuli along a single feature dimension. For grating stimuli, the model matrices were direction, spatial frequency and temporal frequency; for dot stimuli, the features were direction and speed. Each model matrix is the same size as the data matrix, and each cell predicts the similarity (i.e. the decodability) of that pair of stimuli considering only that model's feature.

For grating stimuli, we also repeated the RSA within each spatiotemporal frequency. For each of the 9 spatiotemporal frequencies, we considered the subset of data where classifiers were trained to discriminate stimuli of that spatiotemporal frequency, constructing a 12x12 RDM in each case. These RDMs were correlated with model matrices based on stimulus direction and orientation.

We correlated each model with each animal's individual classification data at each time point using Spearman's rho (a rank correlation). At each time point we tested whether the between-dataset average was significantly positive using a one-sided *t*-test, using a False Discovery Rate (FDR) correction to control for multiple comparisons across time points (Genovese et al., 2002). We also estimated the maximum expected correlation at each time point given the noise in the data, following the methods of Nili and colleagues (Nili et al., 2014). This method is based on the fact that the maximum average correlation that can be achieved with any one model is limited by the variablity across datasets. To do this we defined a 'perfect model' using the average of the individual animal data, then correlated each animal's data with this perfect model using Spearman's rho. We calculated the upper bound of the 'noise ceiling' by taking the average of these correlation values. To calculate the lower bound of the 'noise ceiling' we repeated a similar process, except that when we correlated each animal's data with the 'perfect model' we excluded that animal's own data from the average. Where a single model achieves a correlation with the data that falls within the lower and upper bounds of the noise ceiling, there is good evidence that the model is providing a near-complete account of the 'true model' underlying the data.

Results

We tested how the representations of stimulus features evolve over time in the activity of populations of neurons in area MT. To do this we recorded spiking activity using a 10x10 planar electrode array implanted into area MT of anesthetized marmoset monkeys.

Functional properties of individual electrodes

In our analyses, we use electrode spike rates, without attempting to separate the responses of multiple units contributing to a single electrode. The functional properties of multiunit activity are, however, consistent with previous analyses of single-unit activity (cf. Solomon et al., 2015). Multiunit responses at each electrode were strongly tuned for the motion direction of a drifting grating (Figure 1A), or a moving dot field (Figure 1D). This is consistent with a columnar organization of direction selectivity in area MT, such that nearby neurons prefer similar motion directions (Maunsell and van Essen, 1983; Baker et al., 1981). In addition, Figure 1A shows a smaller secondary peak in activity for gratings moving 180 degrees from the preferred motion direction (ie. gratings with the same contour orientation but opposite motion direction). This is also

observed in single-unit work (Solomon et al., 2011; Maunsell and van Essen, 1983; Baker et al., 1981). The secondary peak is absent for dot fields, which lack oriented contours.

[Figure 1 about here]

Our analyses focused on the representation of stimulus features across populations of neurons. Individual electrodes showed tuning for spatial- and temporal frequency of a drifting grating, and each dataset included electrodes preferring a range of preferred spatial- and temporal frequencies (Figure 1B & C). The distribution of preferred spatial or temporal frequency differed slightly between datasets, but in each dataset we saw strong population responses for the range of stimuli tested. Similarly, individual electrodes showed tuning for the speed of a moving dot field, and each dataset included electrodes preferring a range of motion speeds (Figure 1E).

The average tuning curves mask pronounced dynamics in the activity at individual electrodes. To illustrate these dynamics, Figure 2A shows the average time-course of response at individual electrodes, during presentation of moving gratings. Here, and in all the analyses that follow, we exploited the temporal resolution of multiunit activity to analyze spiking activity in bins of duration 2ms. Responses for the following analyses were collapsed across spatial- and temporal frequency. These data show that the response to gratings moving in the anti-preferred direction is most prominent shortly after stimulus onset. To quantify this we calculated indices of direction selectivity (*DI*) and orientation selectivity (*OI*; see Methods), shown in Figure 2B. Orientation tuning emerges quickly, and is first significant 70ms after stimulus onset ($OI>0$, $p < 0.05$, one-sided *t*-test, FDR corrected for multiple comparisons across time bins). Direction tuning ($DI>0$) also emerges quickly and is first significant 78ms after stimulus onset. The *OI* and *DI* are similar over the first 50ms of response, but the *DI* is increasingly prominent at later time points. Figure 2C-D show counterpart plots of response to moving dot fields, collapsed across dot speed. The *DI* is first significant 86ms after stimulus onset, similar to that for gratings, but dynamics of the response are very different to that for gratings: responses to the preferred and 'flank' motion directions increase with time and responses to anti-preferred motion directions are weak throughout.

[Figure 2 about here]

Although moving dot fields do not have any oriented form, we considered the *OI* for moving dot fields (Figure 2D) to estimate the magnitude of the *OI* based on direction-tuned response alone. The *OI* was significant from 84ms, but was approximately half the magnitude of the *OI* for moving grating stimuli, and less than half as a proportion of the *DI* for the same stimuli. This demonstrates that the *OI* for moving gratings cannot be attributed to direction tuning alone. Overall, the data in Figure 2 show that the response of neurons in area MT is dynamic, and that these dynamics are stimulus-dependent, consistent with previous analyses of single-unit activity. We next sought to characterise the dynamics of stimulus-related information carried by populations of neurons in area MT.

Population-based measures of stimulus information

To quantify the amount of stimulus-related information that was present in the population response, and how this evolved over time, we performed multivariate pattern classification analyses on the electrode spike rates for each 2ms time bin. The classifier's ability to learn to discriminate two stimuli quantifies the difference in population response between two stimuli. To illustrate these analyses, consider the examples in Figure 3, which shows how well the decoder can discriminate motion direction for moving gratings (B) and moving dot fields (C). The decoder operated on population activity obtained in bins of 2ms. In both cases classifier performance is poorest when the classifiers were trained to discriminate directions with the smallest angular separation (30°), and performance increases with increasing angular separation up to 90° separation. This is consistent with the population response being most similar for pairs of stimuli with only a small angular separation in their motion direction (30°), then becoming increasingly different (and discriminable) for stimuli separated by 60° or 90°. Beyond 90° separation the discriminability of moving dot field pairs continued to increase, so that classifier performance was best for dot fields moving in opposite directions, as predicted. However, for moving grating stimuli classifier performance did not increase as angular separation increased from 90° to 180°, but instead decreased slightly, consistent with the responses in Figure 1A. We explore this in greater detail in subsequent analyses.

[Figure 3 about here]

Dynamics of feature representations

A drifting grating is defined by four features: its spatial frequency, temporal frequency, orientation and motion direction. We therefore asked if the population response in area MT could represent each of these features, and if so, whether those representations changed over time. Because the population response will likely depend on multiple features whose representations are multiplexed in the response, we needed a method to partition the population response into the appropriate dimensions. To do this we used Representational Similarity Analysis (RSA) (Kriegeskorte et al., 2008; Nili et al., 2014), which has previously been applied to recordings from other brain areas to explore population encoding for real-world objects. The method is illustrated in Figure 4, where four candidate models of the data (model 'representational dissimilarity matrices' or 'RDMs') are plotted along with the observed RDM (averaged across the duration of the stimulus evoked response). In the average RDM and the models the stimuli are sorted hierarchically by their temporal frequency, then by spatial frequency, then by motion direction. The candidate models predict the pattern of classifier performance that is expected if the discriminability of the population response were based on the stimulus dissimilarity along a single feature dimension. For the candidate models illustrated in Figure 4A, the predicted patterns of classifier performance based on temporal frequency, spatial frequency, and direction are orthogonal with one other, while the prediction based on stimulus orientation overlaps with that for direction. This is because we defined grating direction as being perpendicular to grating orientation (although grating stimuli are physically consistent with a range of interpretations, moving with different directions and speeds, in the absence of other cues they are perceived as moving orthogonal to their orientation, the interpretation of lowest speed). Gratings that are moving in opposite directions have the same orientation and those separated by 90° have orthogonal (maximally different) orientation. Visual comparison of the candidate models (Figure 4A) with the average RDM in Figure 4B reveals similarities between features of the observed data and each of the candidate models (Figure 4A), suggesting that the population response multiplexes information about these different stimulus features.

[Figure 4 about here]

To explore the temporal dynamics of these representations, we generated RDMs from the classification performance data for each 2ms time bin after stimulus onset. Figure 5B shows RDMs at each of four time-points, and illustrates that the qualitative (visual) similarity between the

candidate models and the data varies over time. Figures 5C & D provide a quantitative analysis of these dynamics, and plot the correlation between the observed data at each time point and each of the four candidate models shown in Figure 5A, along with a fifth model based on stimulus speed (the ratio of temporal to spatial frequency) which depends on both spatial and temporal frequency.

Figures 5C & D confirm that the way stimulus features are encoded by the population changes over time. The representation of spatial frequency in the population activity had emerged by 52ms after stimulus onset: this can be confirmed by the lineplot in Figure 5D, and by inspection of the left-most RDM plotted in Figure 5B, which resembles the spatial frequency model plotted above it in Figure 5A. At this early time, the classifier was above chance when discriminating pairs of stimuli that had different spatial frequency, but was poor at discriminating pairs within the same spatial frequency (that nevertheless varied in direction or temporal frequency). At 64ms after stimulus onset, the correlation with the spatial frequency model was higher (Figure 5D), and correlation with other stimulus feature models started to emerge (Figure 5C & D), demonstrating multiplexing of different stimulus features in the population response. Specifically, the lines along the negative diagonal, seen in the RDMs for 64ms and 316ms (Figure 5B), correspond to the models for orientation and direction (Figure 5A). At 64ms, the correlation with the orientation model was higher than the correlation with the direction model, but for the majority of the stimulus-evoked response this relationship was reversed, and the direction model provided a better account of the data than orientation (Figure 5C). This reversal can be seen by comparing the 64ms and 316ms RDMs (Figure 5B): at 64ms, the diagonal lines (for example in the highlighted square) are more closely spaced, as in the orientation model, whereas at 316ms (the time of peak correlation with the direction model) the lines are more widely spaced, as in the direction model. That is, at the onset of the population response to a moving grating, the response varied more with contour orientation than with the direction of movement. Over time, the population response shifted towards encoding stimulus direction (as opposed to contour orientation). Unlike SF, TF and speed, which showed transient peaks at the onset and offset of the population response, the correlations with the orientation and direction models were sustained throughout the response. We found that a similar pattern of results for each dataset, as shown in Figure 6 where the orientation index (*OI*), direction index (*DI*) and results of the RSA are plotted separately for individual datasets.

[Figures 5 & 6 about here]

Strikingly, despite variation between datasets in their tuning for spatial and temporal frequency (cf. Figures 1B & C), the results of the RSA were qualitatively similar across these datasets (as suggested by the 95% confidence intervals in Figure 5, and shown in Figure 6D & E), where data for individual animals are plot separately). This suggests that the dynamics of representations that the analyses are describing reflect consistent stimulus-related differences in neural responses, and not simply overall biases in the population response.

The thin black lines in Figure 5C & D show the lower and upper bounds of our estimate of the 'noise ceiling' (see Methods), that is, the maximum expected correlation of the data with any one model (Nili et al., 2014). Where a model approaches the estimated noise ceiling it suggests that the model is approaching the maximum possible correlation with the data, and is a good estimate of the true model underlying the data. Where a model's correlation with the data is below this noise ceiling it suggests that there is variance in the data not captured by that model alone, but which may be explained, for example, by a model based on another stimulus feature or a combination of features. The correlation of the data with the spatial frequency model (Figure 5D) approached the noise ceiling at the onset of the population response, suggesting that the spatial frequency of the stimuli provides a near complete account of the population response of MT neurons at this time. After this early peak the correlation between the spatial frequency model and the data was lower. This may reflect the presence of stimulus-direction-related variance in population activity, or a decrease in information about spatial frequency. Regardless, the pattern of results reveals that around the onset of the population response, spatial frequency is the dominant feature being represented in the population response.

We found weakest correlation between the data and the temporal frequency model (Figure 5D): there were no time points at which the correlation reached a value significantly above zero ($p < 0.05$, after FDR correction). Note that by applying false discovery rate correction for multiple comparisons across the 301 time points we are adopting a strict statistical threshold to avoid false positives, but lose statistical power, as shown by the fact that the temporal frequency model did not reach significance despite high local peaks. Inspection of the RDM at around stimulus offset (Figure 5B, 560ms) suggests that the local peak in correlation at this time is driven by poor classifier performance within stimuli of highest temporal frequency. This pattern of results suggests an earlier offset of population response for stimuli of high temporal frequency, consistent with earlier findings (Bair and Movshon, 2004) that neurons in MT have a shorter window of temporal integration for stimuli of high speed than for those of low speed. Interestingly, combining predictions from stimulus

spatial and temporal frequencies into a 'speed' model (Figure 5D) did not result in a better account of the data than that provided by spatial frequency alone. The spatial frequency model tended to have a higher correlation with the data, although at no time points was there a significant difference between the data's correlation with the speed and spatial frequency models (i.e. $p < 0.05$, paired two-tailed *t*-test, FDR correction for multiple comparisons across time points).

Dynamics of the population response to dot fields

A moving dot field has two significant features: the motion direction, and the motion speed. Dot fields are isotropic in orientation, so they allow analysis of direction tuning in the absence of strong orientation signals. As a validation of the methods, and to supply a point of comparison with population response to gratings, we therefore applied RSA analyses to the population response during moving dot fields.

We found pronounced dynamics in the population response to moving dot fields that were analogous to the dynamics we see for moving gratings. As for moving gratings, we correlated responses with candidate models of the RDMs based on single stimulus features (Figure 7A): speed and direction of motion. The speed and direction models correlated with the data from a similar onset. This is illustrated in Figure 7B, where even at 64ms (when classifier performance is low overall) there is evidence of structure in the RDM that follows the predictions of the speed and direction models. This observation is confirmed by correlation values plotted in Figure 7C, where the correlations with the speed and direction models show similar onsets. The speed and direction models have similar onsets in their correlation with the data. However, while the speed model correlates most strongly with the data at the onset and offset of the population response, the correlation with the model based on direction difference continues to increase for the first 20ms of the stimulus-evoked response, before reaching a plateau. This is qualitatively similar to the results for grating stimuli, where the correlation with the direction model was sustained throughout the population response, whereas the correlation with SF, TF and speed peaked around the onset and offset of the response. Again, although there was variation between datasets in their tuning for speed (seen in Figure 1E), the results of the RSA were qualitatively similar across datasets (data not shown).

[Figure 7 about here]

Our analyses also included a 'pseudo-orientation' model, which was constructed in an equivalent manner to the 'orientation' model that was correlated with the data for moving gratings. Specifically, the 'pseudo-orientation' model predicts that classification performance will be highest for moving dot fields that are moving in orthogonal directions, and lowest for those moving in opposite directions. We did not expect this model to provide a good account of the data. We included this model since it provided a way for us to estimate the expected correlation with an orientation model for a dataset where there was no orientation information in the stimulus. This was important since the direction and orientation models are non-orthogonal, so it is likely that the correlation between the moving grating data and the orientation model was driven in part by the direction-tuned response. The 'pseudo-orientation' model for the dot field data provided a natural way to measure this. In the absence of orientation information in the stimulus, we reasoned that any correlation between the moving dot field data and the pseudo-orientation model must be driven by the stimulus-direction-related responses, and the non-independence of the direction and orientation models.

As expected, the pseudo-orientation model does not correlate with the data as strongly as the direction model; the correlation between the moving dot field data and the direction model is approximately 7 times greater than the correlation between these data and the pseudo-orientation model. The pseudo-orientation model for dot fields performs much worse than the orientation model for gratings, both in absolute and relative terms. This provides strong evidence that the high correlation between the orientation model and the moving grating data cannot simply be attributed to stimulus-direction-related responses, but reflects the presence of an orientation-tuned signal in the population response.

Dynamics of orientation and direction signals

Using representational similarity analysis, we found prominent dynamics in the population response to moving gratings, and moving dot fields. To further characterize the dynamics of the orientation and direction signals in the population response, we collapsed the performance matrices according to stimulus direction difference. In Figures 8 and 9, we plot classifier performance over time for stimulus pairs of varying relative motion directions.

First, we considered the classification performance based on the population response to moving gratings, collapsed across spatial and temporal frequencies. If classifier performance simply

increased with direction difference this would imply that the neural population was encoding grating direction. If classifier performance were instead greatest for stimuli separated by 90° and the classifier could not discriminate stimuli that are moving in opposite directions, this would imply that the neural population was tuned to the orientation of the contour not its direction of motion. The result of this analysis, shown in Figure 8A, lies between these two extremes, suggesting that the population response depends upon both the grating orientation, and its direction of motion. We found the greatest dependence on grating orientation at the onset of above-chance classification performance. The peak classifier performance shifts gradually towards a direction tuned response over time, though never reaches it.

[Figure 8 about here]

For clarity, Figure 8B shows classifier performance in 3 illustrative time bins. If the population response depended only on the stimulus direction then classifier performance should peak for direction differences of 180°, when stimuli are maximally different in direction. Alternatively, if the population response depended only on contour orientation then the tuning curve should be centered on 90°, when the stimuli are orthogonal. For the earliest time bin in Figure 8B (64-74ms post stimulus-onset) the tuning curve peaked at 90°, suggesting that orientation, not direction, was the dominant feature determining the population response at this time. In later time bins peak performance shifts towards 180°.

Applying the same analyses to responses during moving dot fields produced a different pattern of results. Unlike for the gratings, classifier performance increased with increasing angular separation, and the most discriminable stimuli were those that were moving in opposite directions (180° separation; Figure 9).

[Figure 9 about here]

Spatiotemporal asymmetries in direction sensitivity

We next considered the possibility that the interaction of orientation and motion signals depends on where the gratings lie in the spatiotemporal spectrum, as previous work in single neurons (Bair and Movshon, 2004) shows that temporal kernels in area MT, and earlier areas, can depend on spatial

and temporal frequency. This suggests that the dynamics of the contour-orientation response may also vary with stimulus spatiotemporal frequency.

We therefore repeated the analyses above, for subsets of gratings of the same spatial and temporal frequency but varying motion direction. Figure 10 shows the classifier performance at each stimulus spatiotemporal frequency. We found that the early orientation-dependent response depends on stimulus spatiotemporal frequency: it is most marked for stimuli of high spatial frequency and low temporal frequency (i.e. slowest speed, bottom right plot in Figure 10). Indeed, at early time points the classifier performance for gratings separated by just 30° exceeds that for gratings with directions separated by 180°.

[Figure 10 about here]

To understand how population response dynamics vary with stimulus spatiotemporal frequency, we repeated the RSA analyses independently for each combination of spatial and temporal frequency (Figure 11). At each time point, we correlated each 12 direction x 12 direction RDM with model matrices based on stimulus direction and orientation. The results confirm that dependence on orientation is most marked at the onset of response and is strongest at low temporal frequency and high spatial frequency. The reader should note that the model RDMs for orientation and direction are not orthogonal (for direction differences smaller than 90° they are positively correlated, and for direction differences larger than 90° they are negatively correlated). This is why high positive correlation with the orientation model, seen at the highest spatial frequency and lowest temporal frequency, is accompanied by a slightly negative correlation with the direction model.

[Figure 11 about here]

Discussion

We measured the population response of area MT and applied multivariate pattern classifications in conjunction with Representational Similarity Analysis (RSA). We found that simultaneous multielectrode recordings combined with these analyses provide a powerful tool for exploring the relationship between stimulus features and the population response. Our analyses show that the population response can simultaneously represent multiple features (ie. multiplexing), and

that these features have distinct population dynamics. The analyses reveal temporal ordering of stimulus features in the population response of area MT. For both moving gratings and moving dot fields, the spatiotemporal frequency (for gratings) and speed of the stimuli were most evident in the population response around its onset and offset, whereas the representation of direction was sustained throughout the population response. For moving gratings, the spatial features of the stimulus were also better encoded around stimulus onset while encoding of motion direction emerged at slightly later time points. This temporal ordering arises because a representation of contour orientation precedes that of motion direction: a feature which is most pronounced for stimuli of high spatial frequency and low temporal frequency.

Our analyses are based on the population responses to a battery of stimuli. Because the classifiers have access to all neurons, and the RSA compares classifier performance across the entire battery of stimuli, these analyses can provide a sensitive characterization of stimulus-related information that is weak but consistent. For example, the RSA was sensitive enough to detect the presence of information about grating orientation 64ms after stimulus onset, before we could detect orientation selectivity by calculating indices independently for each electrode. Critically, this early time was also where our analyses revealed the largest qualitative shift in the population tuning for motion direction, from more to less dependent on contour orientation. This early qualitative shift would be largely missed if we had only considered the activity at individual electrodes.

Dynamics of population signals in area MT

Neurons in area MT play a central role in the analysis of visual motion (Newsome et al., 1989; Salzman et al., 1990; Albright, 1992; Britten et al., 1996). Motion analysis, however, is rarely independent of spatial form. Our analyses demonstrate that motion integration by populations of MT neurons varies with stimulus spatial and temporal frequency, and that responses depend not only on motion direction but also on stimulus orientation. Similarly, psychophysical work shows that motion detectors tuned to different spatial frequencies differ in their tuning for temporal frequency, and in their integration time (Burr et al., 1986). Perceived speed also increases with increasing spatial frequency of the stimulus (Brooks et al., 2011).

We observed considerable dynamics in the population encoding of stimulus features. For moving gratings, classifier performance at response onset was best attributed to stimulus spatial frequency alone, then later it also depended on stimulus orientation, and then motion direction. We

672 considered the possibility that the lagged representation of direction simply reflects the fact that
673 defining motion requires at least two stimulus frames, whereas spatial frequency and orientation
674 require only the first frame of the stimulus. If this were the case, we would expect correlation with
675 the direction model to lag the correlation with the orientation model across all spatiotemporal
676 frequencies. By contrast, we found some spatiotemporal frequencies for which the onset of
677 correlations with orientation and direction models are simultaneous (Figure 11). This suggests
678 that the dynamics are not simply explained by the fact that computing direction requires a least 2
679 stimulus frames.
680
681 Across all spatiotemporal frequencies, directional information increases monotonically for 250ms
682 after stimulus onset. This is driven primarily by responses to high spatial frequency and low temporal
683 frequency. For gratings of low spatial and high temporal frequency directional information instead
684 rapidly reached a plateau (Figure 11). Longer dynamics at high spatial and low temporal frequencies
685 is consistent with previous results (Bair and Movshon, 2004) where single neurons in MT show
686 longer temporal integration for stimuli of lower speed, or higher spatial frequency. Our results extend
687 this by showing that the longer temporal integration is accompanied by a qualitative temporal change
688 in the population representation of direction, moving from contour-orientation dependent towards
689 orientation independent.
690
691 The population response to moving dot fields also showed strong directional dynamics, but, in this
692 case decoding performance was always highest for stimuli separated by 180°. This is consistent with
693 earlier work on single-units in area MT of macaque showing that, for moving dot fields, the first
694 few spikes already contain directional information (Osborne et al., 2004). Indeed, the dynamics of
695 population activity for moving dot fields appear similar in the two studies (cf. our Figure 2; their
696 Figure 2). That work also found that the information about stimulus direction decayed with time
697 following the first few spikes. Instead, we found that the decoding of stimulus direction improved
698 for at least 200ms. We cannot rule out an impact of stimulus differences - in Osborne et al. (2004)
699 the dot fields appeared and were then stationary for 256ms before moving, while in our study the
700 moving dot fields were preceded by a gray screen. However, the different dynamics of directional
701 information may reflect the fact that our analyses separate the dynamics of information about
702 stimulus speed from that about stimulus direction. Indeed, our analyses imply that information about
703 stimulus speed is rapid before subsequently decaying. The dynamics of speed-related information

were similar for dot fields and gratings, and both may be particularly driven by the dynamics of spatial frequency information.

Sensitivity to contour orientation

Previous studies have reported a proportion of direction selective cells in MT that are also selective for the orientation of a grating or line (Albright, 1984; Maunsell and van Essen, 1983). Similarly, we see a small secondary peak in the average direction tuning curve, in response to the anti-preferred direction (Figure 1A), similar to that reported in previous single-unit studies (Rodman and Albright, 1989; Albright, 1992; Solomon et al., 2011). These small responses have not been emphasized in previous work. They are emphasized here because the analyses that we have used are capable of extracting and highlighting the information present within very short time bins (2ms), in individual animals. This mitigates any impact of variation in response latency between animal (cf. Figure 6). In addition, our analyses show that this orientation dependence is pronounced only at some spatiotemporal frequencies. Previous work (Kumano and Uka, 2013; Gharaei et al., 2013) suggests that neurons in area MT are better able to encode the direction of moving patterns when those patterns contain multiple orientations. The pronounced orientation-dependence of the population response that we observe may be most prominent for stimuli where the motion direction is ambiguous, such as the gratings used here.

In some neurons in area MT (Type 1, 'component selective') the preferred orientation is a bar orthogonal to the path of preferred motion direction. In others (Type 2, 'pattern selective'), preferred orientation is parallel to the preferred direction (Rodman and Albright, 1989; Albright, 1984). The latter is consistent with Intersection-of-Constraints (IOC) frameworks for motion analysis, including the implementation by Simoncelli and Heeger (1998), though unlike those frameworks the response to static patterns is always weaker than response to moving ones (Nishimoto and Gallant, 2011). The motion integration that is performed by pattern-cells takes time to develop (Pack and Born, 2001; Smith et al., 2005; Solomon et al., 2011). We cannot distinguish 'pattern-like' activity from these data, but interestingly, the shift from 'component-like' to 'pattern-like' responsiveness in pattern cells shows a similar timecourse to the shift from strong to weaker orientation-dependence observed here. Regardless our observations reinforce the idea that there are qualitative changes in the population tuning for motion direction over the first 100ms of the stimulus-evoked response.

Mechanisms

Our results reveal the dynamic nature of population responses evoked by a simple moving stimulus. They show that the encoding of stimulus direction is shaped in the first hundred milliseconds of the stimulus-evoked population response. Yet while our results highlight the significance of dynamics in area MT response, we do not yet understand the processes that underlie these changes. Bair and Movshon (2004) extended the motion energy model by incorporating an integrate-and-fire model of spiking activity in area MT. They were, however, unable to fully account for the stimulus-dependence of temporal integration windows. Similarly, our results do not reveal whether these dynamics are present in the feedforward input to area MT, arise from computations within area MT, or require feedback from other visual areas. One possibility is that the dynamics reflect convergence of two visual pathways into area MT: an early, orientation selective pathway and a later, motion selective pathway. Alternatively, the motion signals conveyed to area MT may include strong orientation dependence that is subsequently countered by computations within MT. The timing of the contour-orientation dependence, along with the spatiotemporal tuning of the effect, may be used in future work to identify likely mechanisms that are driving the effect at the population level.

Acknowledgments

This project was funded under an Australian Research Council Future Fellowship (FT120100816), ARC Discovery Project (DP160101300), and a National Health and Medical Research Council of Australia Project Grant (APP1005427). We thank S.S. Solomon, S.K. Cheong, S.C. Chen and A.S. Pietersen for assistance with electrophysiological data collection.

References

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology* 52:1106–1130.

Albright TD (1992) Form-cue invariant motion processing in primate visual cortex. *Science* 255:1141–1143.

780 Bair W, Movshon JA (2004) Adaptive temporal integration of motion in direction-
781 selective neurons in macaque visual cortex. *Journal of Neuroscience* 24:7305-
782 7323.
783
784 Baker J, Petersen S, Newsome W, Allman J (1981) Visual response properties of neurons in four
785 extrastriate visual areas of the owl monkey (*aotus trivirgatus*): a quantitative comparison of
786 medial, dorsomedial, dorsolateral, and middle temporal areas. *Journal of Neurophysiology*
787 45:397–416.
788
789 Borghuis BG, Perge JA, Vajda I, van Wezel RJA, van de Grind WA, Lankheet MJM (2003) The
790 motion reverse correlation (MRC) method: a linear systems approach in the motion domain.
791 *Journal of Neuroscience Methods* 123:153–166.
792
793 Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship
794 between behavioral choice and the visual responses of neurons in macaque MT. *Visual*
795 *Neuroscience* 13:87–100.
796
797 Brooks KR, Morris T, Thompson P (2011) Contrast and stimulus complexity moderate the
798 relationship between spatial frequency and perceived speed: implications for MT models of
799 speed perception. *Journal of Vision* 11:19, 1–10.
800
801 Burr D, Ross J, Morrone M (1986) Seeing objects in motion. *Proceedings of the Royal Society of*
802 *London B - Biological Sciences* B227:249–265.
803
804
805 Chen SC, Morley JW, Solomon SG (2015) Spatial precision of population activity in primate area
806 MT. *Journal of Neurophysiology* 114:869–78.
807
808 Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time.
809 *Nature Neuroscience* 17:455–462.
810
811 Cichy RM, Sterzer P, Heinzle J, Elliott LT, Ramirez F, Haynes JD (2013) Probing principles of
812 large-scale object representation: category preference and location encoding. *Human Brain*
813 *Mapping* 34:1636–1651.
814
815 Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional
816 neuroimaging using the false discovery rate. *Neuroimage* 15:870–878.
817
818 Gharaei S, Tailby C, Solomon SS, Solomon SG (2013) Texture-dependent motion signals in
819 primate middle temporal area. *Journal of Physiology* 591:5671–5690.

Goddard E, Carlson TA, Dermody N, Woolgar A (2016) Representational dynamics of object recognition: Feedforward and feedback information flows. *Neuroimage* 128:385–397.

Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neural populations. *Nature Neuroscience* 9:690–696.

Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141.

Kumano H, Uka T (2013) Responses to random dot motion reveal prevalence of pattern-motion selectivity in area MT. *Journal of Neuroscience* 33:15161–15170.

Maunsell JH, van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. i. selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology* 49:1127–1147.

McDonald JS, Clifford CWG, Solomon SS, Chen SC, Solomon SG (2014) Integration and segregation of multiple motion signals by neurons in area MT of primate. *Journal of Neurophysiology* 111:369–378.

Movshon JA, Adelson EH, Gizzi M, Newsome WT (1985) The analysis of moving visual patterns., In: *Pattern Recognition Mechanisms*, eds Chagas, C, Gattass, R, Gross, C. Chapter 54, pp. 117–151 Vatican Press.

Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52–54.

Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A toolbox for representational similarity analysis. *PLoS Computational Biology* 10:e1003553.

Nishimoto S, Gallant JL (2011) A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. *Journal of Neuroscience* 31:14551–14564.

Osborne LC, Bialek W, Lisberger SG (2004) Time course of information about motion direction in visual area MT of macaque monkeys. *Journal of Neuroscience* 24:3210–3222.

Pack CC, Born RT (2001) Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* 409:1040–1042.

Redcay E, Carlson TA (2015) Rapid neural discrimination of communicative gestures. *Social Cognitive and Affective Neuroscience* 10:545–551.

Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research* 75:53–64.

Rosa MG, Elston GN (1998) Visuotopic organisation and neuronal response selectivity for direction of motion in visual areas of the caudal temporal lobe of the marmoset monkey (*Callithrix Jacchus*): middle temporal area, middle temporal crescent, and surrounding cortex. *Journal of Comparative Neurology* 393:505–527.

Rousche PJ, Normann RA (1992) A method for pneumatically inserting an array of penetrating electrodes into cortical tissue. *Annals of Biomedical Engineering* 20:413–422.

Salzman CD, Britten KH, Newsome WT (1990) Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346:174–177.

Simoncelli E, Heeger D (1998) A model of neuronal responses in visual area MT. *Vision Research* 38:743–761.

Smith MA, Majaj NJ, Movshon JA (2005) Dynamics of motion signaling by neurons in macaque area MT. *Nature Neuroscience* 8:220–228.

Solomon SS, Chen SC, Morley JW, Solomon SG (2015) Local and global correlations between neurons in the middle temporal area of primate visual cortex. *Cerebral Cortex* 25:3182–3196.

Solomon SS, Tailby C, Gharaei S, Camp AJ, Bourne JA, Solomon SG (2011) Visual motion integration by neurons in the middle temporal area of a new world monkey, the marmoset. *Journal of Physiology* 589:5741–5758.

Wardle SG, Kriegeskorte N, Grootswagers T, Khaligh-Razavi SM, Carlson TA (2016) Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *Neuroimage* 132:59–70.

Zavitz E, Yu HH, Rowe EG, Rosa MGP, Price NSC (2016) Rapid adaptation induces persistent biases in population codes for visual motion. *Journal of Neuroscience* 36:4579–4590.

Figure Legends

Figure 1: Summary of responses to moving grating stimuli (**A-C**) and moving dot field stimuli (**D-E**). **A-E**: The normalized spike rate averaged across electrodes for each of 6 datasets, expressed as a proportion of the mean spike rate for that electrode. Dashed black lines indicate the average spontaneous activity across datasets (normalized response during trials where the screen was blank). In **A** & **D** the spike rate to moving gratings and dot fields is plotted relative to the electrode's preferred direction. In **B** & **C** the normalized spike rate is plotted as a function of the actual grating spatial and temporal frequency, and in **E** as a function of the actual speed of the moving dot field. The legends indicate the animal from which the data were recorded (e.g. ma025), whether the stimuli were presented to the contralateral (c) or ipsilateral (i) eye, and the number of electrodes that were identified as located in MT.

Figure 2: Dynamics of responses to moving grating (**A-B**) and moving dot field (**C-D**) stimuli. **A** and **C**: Evolution of direction tuning over the course of the stimulus presentation. In both plots, the spike rates are averaged across all electrodes from the 6 datasets, where each electrode's spike rate was normalized by its average spontaneous rate during the blank trials. These normalized spike rates are plotted as function of stimulus direction, relative to each electrode's preferred direction for each 2ms time bin from stimulus onset. **B** & **D**: Direction and orientation index values over time, with shaded error bars indicating the 95% confidence intervals of the between-dataset mean (n = 6). Hatched lines along the x-axis indicates when the stimulus was present (0 to 500ms). The three gray highlighted regions indicate fiduciary time points for later analyses and colored ovals at the bottom of the plot indicate time bins for which the corresponding indices were significantly above zero ($p <$ 0.05, one-sided t-test, FDR corrected for multiple comparisons across time bins).

Figure 3: Average discriminability of motion direction in area MT population responses. **A**: Schematic showing how direction difference varies across the pairs of stimuli that the classifier was trained to discriminate. Shading indicates direction difference. **B** & **C**: Discriminability of each grating direction pair (**B**) and each moving dot field direction pair (**C**). Classifiers were trained and tested on data within a 2ms time bin. Average

performance was calculated across each 2ms bin between 46ms and 596ms post-stimulus-onset.

Figure 4: Representational dissimilarity matrix (RDM) for the population response to grating stimuli. **A**: The model RDMs, based on grating stimulus spatial frequency (SF), orientation, direction and temporal frequency (TF) difference. **B**: The average RDM across time (46-596ms post-stimulus onset). The RDM shows the classifier's ability to discriminate each pair of unique exemplars (of given TF, SF and direction). As indicated by the axis label bars on the left of the RDMs in **A** and **B**, exemplars of the same TF are grouped together (black=lowest TF, white=highest TF), and within each TF the exemplars are grouped by SF (black=lowest SF, white=highest SF) and then direction (black=0° from downwards, white=330° from downwards, moving towards lower left). Exemplars are grouped in the same manner from left to right as from top to bottom, meaning that the negative diagonal is 50% (chance performance) by definition, and the matrix is symmetrical about this diagonal.

Figure 5: Representational Similarity Analysis for the population response to moving grating stimuli. **A**: The model RDMs from Figure 4**A** are replotted for comparison with the RDMs in **B**. **B**: RDMs for single time-points (those highlighted with inverted triangles in **C** & **D**). For t =64ms and t =316ms, the zoomed inset shows an example section of RDM where stimuli varied in direction and orientation but not spatial or temporal frequency. **C** and **D**: Rank correlations between the RDM and model matrices over time. For each 2ms time bin we correlated the observed RDM with model matrices based on direction and orientation (**C**), and spatial frequency (SF), temporal frequency (TF), and speed (TF/SF) (**D**). The upper and lower thin black lines are the upper and lower bounds of the maximum expected correlation with any one model (see text for details). Shaded error bars indicate 95% confidence intervals of the between-dataset mean (n=6), and the colored dots show points at which the corresponding curves were significantly greater than 0 (one-tailed t -test, $p < 0.05$, FDR corrected for multiple comparisons across time points). Hatched lines along the x-axes in **C** and **D** indicate when the stimulus was present (0 to 500ms). Data for individual datasets are plotted in Figure 6D&E.

Figure 6: Responses to grating direction and RSA within individual datasets. **A**: Average normalised response of each electrode as a function of preferred direction around response onset (blue: 46ms-100ms) and for the remainder of the stimulus-induced response (green: 102-550ms). Dashed black lines show the average spontaneous activity (normalized response during trials where the screen remained blank). **B**: Distribution of preferred directions across electrodes (0° = downwards, 330° = towards lower left). **C**: Direction index (blue) and orientation index (red) over time, averaged across electrodes. In both **A** & **C** the error bars are 95% confidence intervals of the mean over all electrodes within the dataset. **D**: (RSA): Correlation between data and direction and orientation models. **E**: (RSA): Correlation between data and the spatial frequency (SF), temporal frequency (TF) and speed models. Plotting conventions for **D** & **E** as in Figure 5C & D respectively. Hatched lines along the *x*-axes in **C-E** indicate when the stimulus was present (0 to 500ms). The datasets top to bottom are ma025c, ma025i, ma026i, ma026c, ma027c, my147c (using the dataset naming conventions from Figure 1).

Figure 7: Representational Similarity Analysis (RSA) for the population response to moving dot field stimuli, with plotting conventions as in Figure 5. **A**: The predicted (model) matrices based on stimulus speed (left) and direction (middle), along with a 'pseudo-orientation' model (right) that was used as a control analysis (see text for details). **B**: RDMs for a range of single time points (highlighted in **C**): in each case the exemplars are sorted by speed (slow to fast, with slowest speeds in the top left), then by direction (0-330° from vertical, with 0° in the top left). **C**: Rank correlations between the RDM and the model matrices over time.

Figure 8: Summary of classification of grating direction over time. **A**: Discriminability of grating direction over time, measured using classification performance (% correct). Stimulus pairs were grouped by their direction difference (*y*-axis), as indicated by the schematic. Average classification accuracy for these stimulus pairs across time (*x*-axis) is given by the color of the image. Overlaid on the plot are black contour lines joining points of the equal classifier performance. Hatched lines along the *x*-axis indicates when the stimulus was present (0 to 500ms). **B**: A subset of the classifier accuracy data in **A** is replotted here: the average classifier accuracy for three time bins (indicated by the shaded bars above the plot

in **A**). For individual animals, we took the average classification accuracy for each time bin averaged over stimulus pairs according to their direction difference, and expressed these as a proportion of the maximum accuracy in this time bin before averaging across animals. Shaded error bars indicate the 95% confidence intervals of the between-dataset mean (n = 6), and the peak of each curve is labeled in text of the same color.

Figure 9: Summary of classification of dot field direction over time. Conventions as in Figure 8, but here showing classification data based on the MT population response to moving dots. In **B**, the maximum classifier performance for every time point was for directions separated by 180°, consistent with a direction tuned population response.

Figure 10: Population tuning curves for grating stimuli over time, separated according to stimulus spatial and temporal frequency. Plotting conventions as in Figure 8B. Each of the nine plots contains a subset of the data in Figure 8B, showing the tuning curves when the training and test data are restricted to stimuli of a single spatiotemporal frequency. The spatial and temporal frequency of the stimuli in each plot are indicated across the bottom and along the left side of the figure.

Figure 11: Representational similarity analysis (RSA) for direction and orientation models, separated according to stimulus spatial and temporal frequency. Plotting conventions are as in Figure 5C. Each of the nine plots contains a subset of the data used in Figure 5C, showing RSA for direction (blue curves) and orientation (red curves) when the model matrices are correlated with classifier performance for stimuli of a single spatiotemporal frequency. The spatial and temporal frequency of the stimuli in each plot are indicated across the bottom and along the left side of the figure.

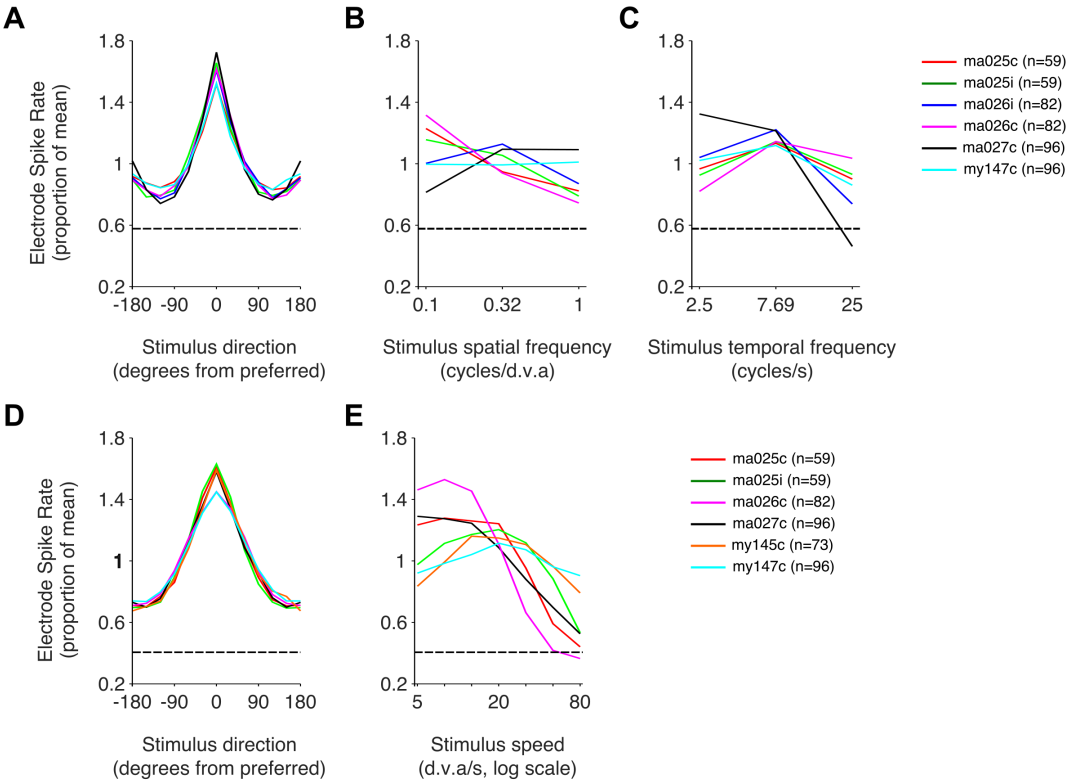
A
B
C
D
E
Electrode Spike Rate
(proportion of mean)
1.8
1.4
1
0.6
0.2
-180 -90 0 90 180
Stimulus direction
(degrees from preferred)
0.1 0.32 1
Stimulus spatial frequency
(cycles/d.v.a)
2.5 7.69 25
Stimulus temporal frequency
(cycles/s)
5 20 80
Stimulus speed
(d.v.a/s, log scale)
ma025c (n=59)
ma025i (n=59)
ma026i (n=82)
ma026c (n=82)
ma027c (n=96)
my147c (n=96)
my145c (n=73)

Figure 1

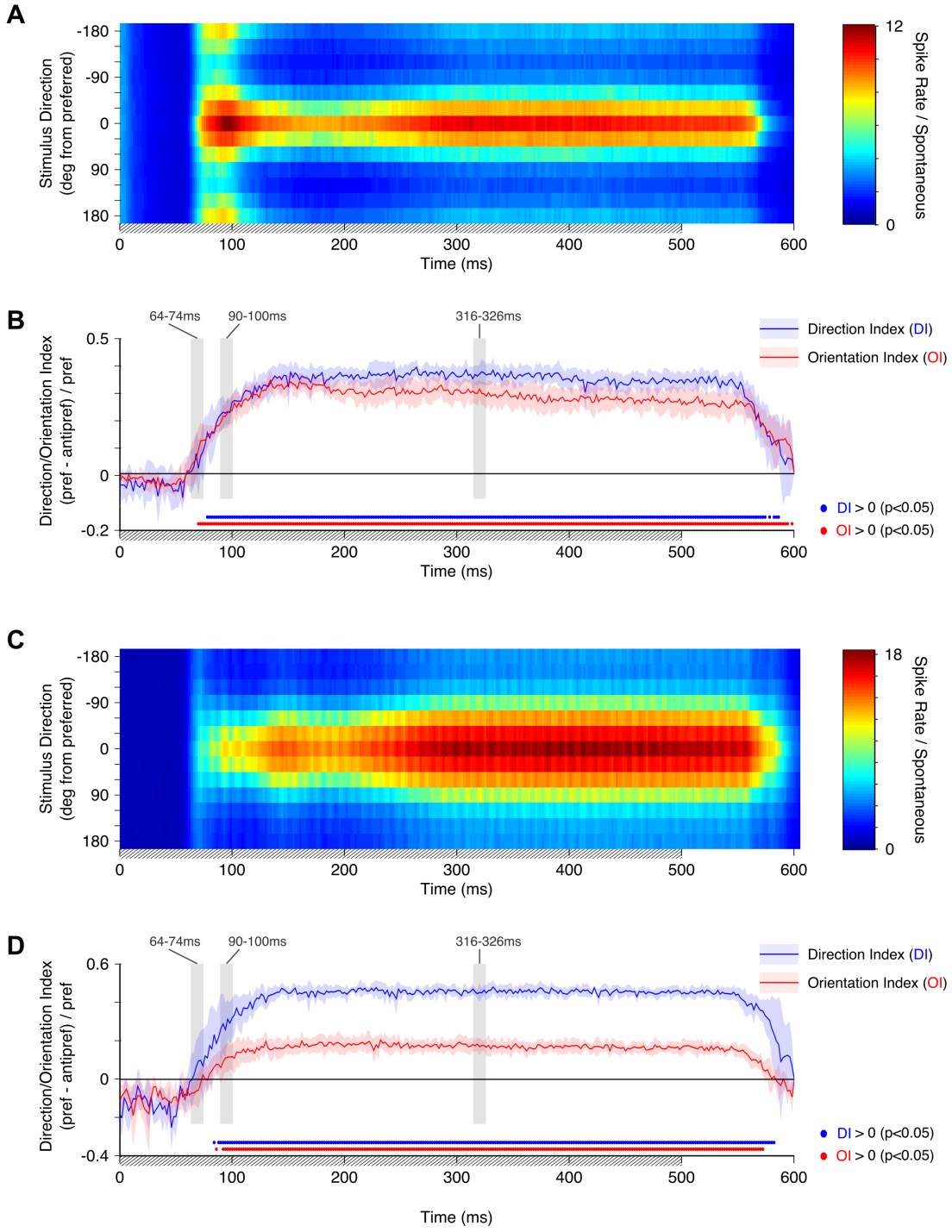
A
Stimulus Direction
(deg from preferred)
-180
-90
0
90
180
0
100
200
300
400
500
600
Time (ms)
12
0
Spike Rate / Spontaneous
B
64-74ms
90-100ms
316-326ms
Direction/Orientation Index
(pref - antipref) / pref
0.5
0
-0.2
Direction Index (DI)
Orientation Index (OI)
DI > 0 (p<0.05)
OI > 0 (p<0.05)
Time (ms)
C
Stimulus Direction
(deg from preferred)
18
0
Spike Rate / Spontaneous
Time (ms)
D
64-74ms
90-100ms
316-326ms
Direction/Orientation Index
(pref - antipref) / pref
0.6
0
-0.4
Direction Index (DI)
Orientation Index (OI)
DI > 0 (p<0.05)
OI > 0 (p<0.05)
Time (ms)

Figure 2

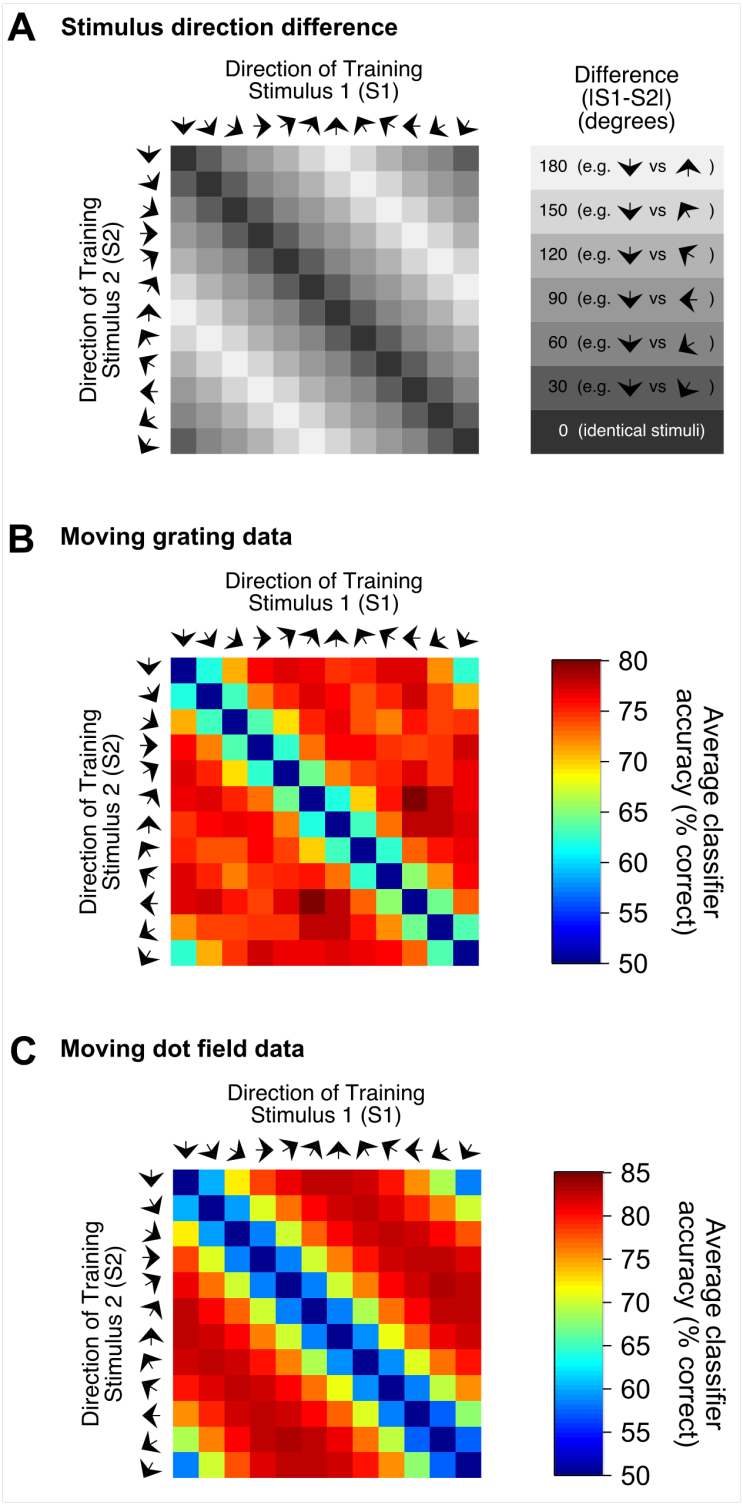
A
Stimulus direction difference
Direction of Training Stimulus 1 (S1)
Direction of Training Stimulus 2 (S2)
Difference (IS1-S2I) (degrees)
180 (e.g. vs)
150 (e.g. vs)
120 (e.g. vs)
90 (e.g. vs)
60 (e.g. vs)
30 (e.g. vs)
0 (identical stimuli)
B
Moving grating data
Direction of Training Stimulus 1 (S1)
Direction of Training Stimulus 2 (S2)
80
75
70
65
60
55
50
Average classifier accuracy (% correct)
C
Moving dot field data
Direction of Training Stimulus 1 (S1)
Direction of Training Stimulus 2 (S2)
85
80
75
70
65
60
55
50
Average classifier accuracy (% correct)

Figure 3

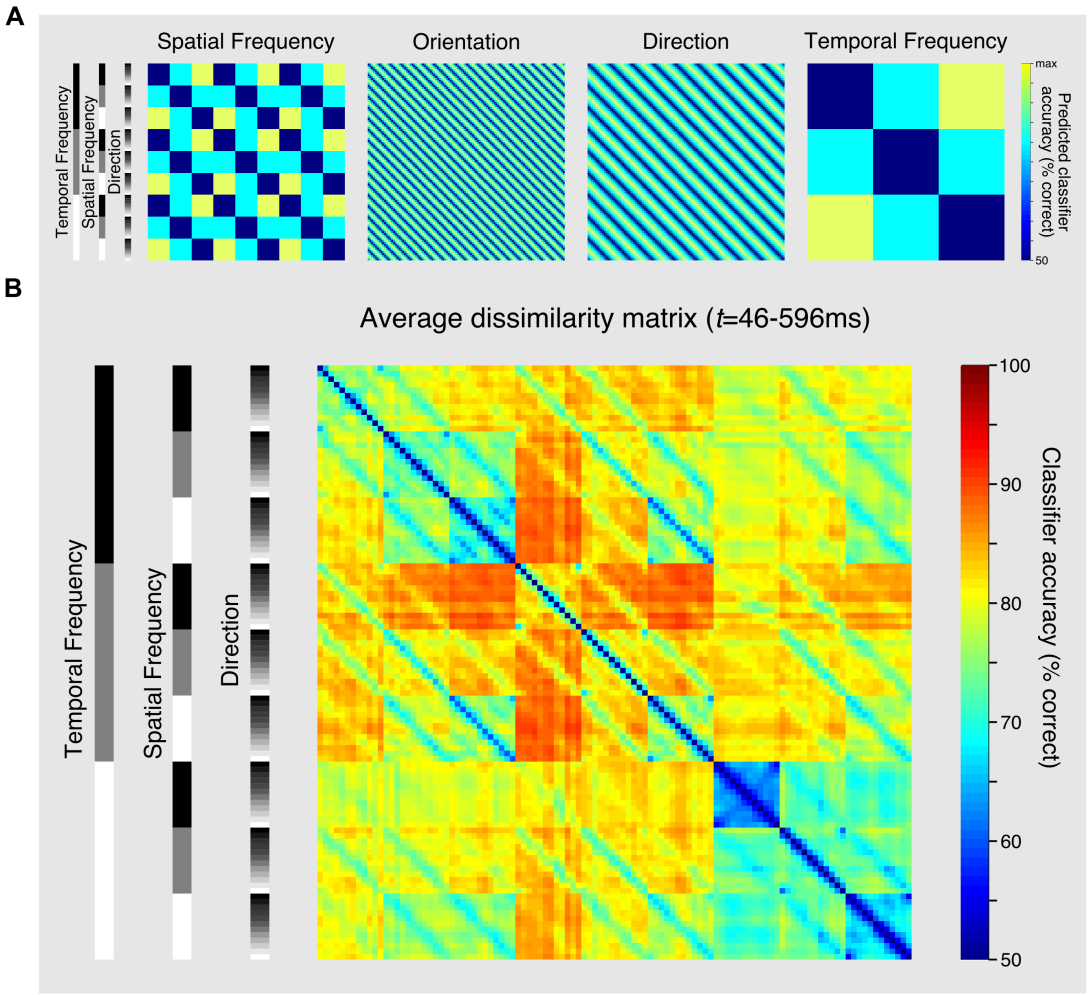
A
Spatial Frequency
Orientation
Direction
Temporal Frequency
Temporal Frequency
Spatial Frequency
Direction
max
Predicted classifier accuracy (% correct)
50
B
Average dissimilarity matrix (t=46-596ms)
Temporal Frequency
Spatial Frequency
Direction
100
90
80
70
60
50
Classifier accuracy (% correct)

Figure 4

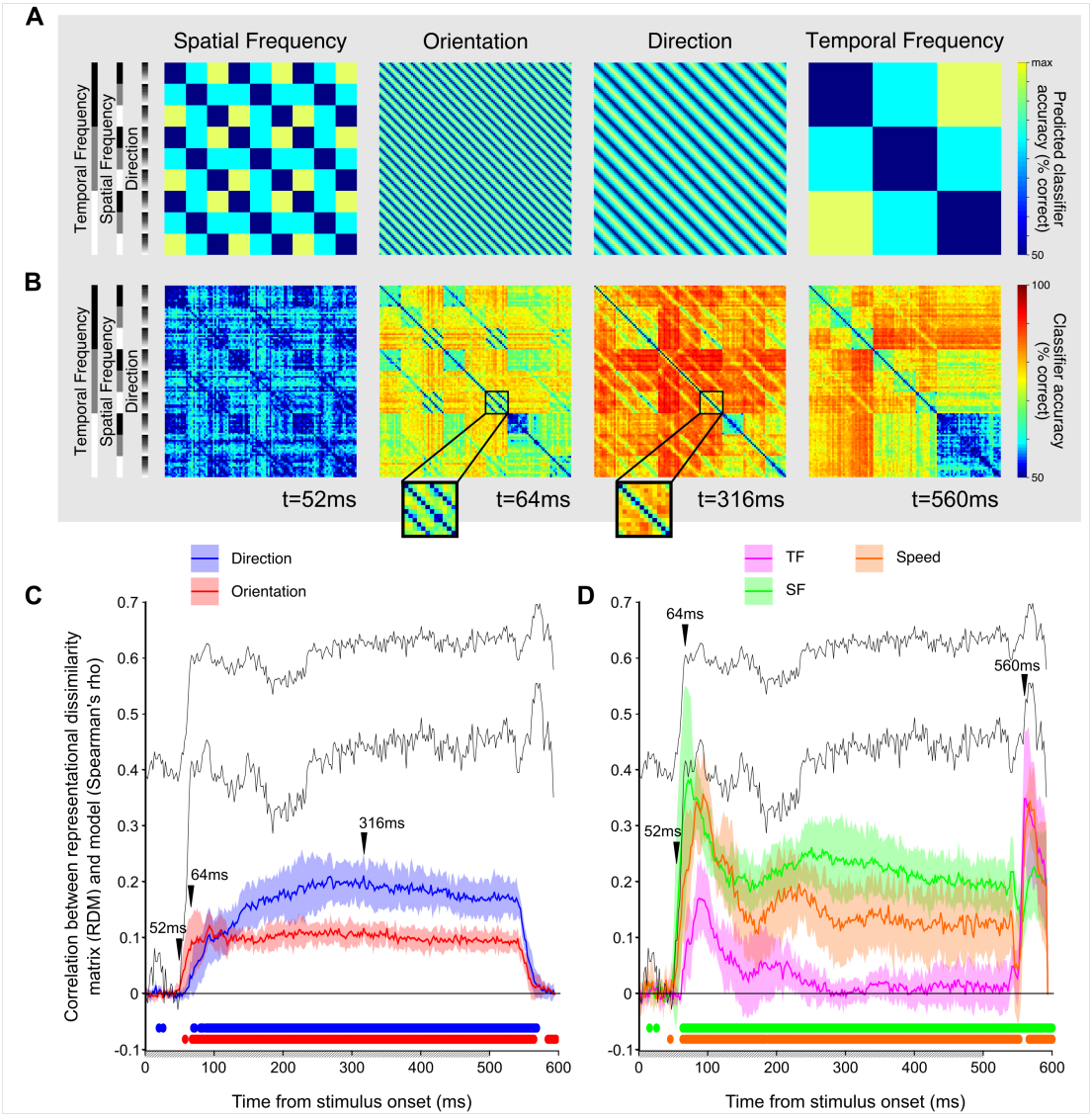
A
Spatial Frequency
Orientation
Direction
Temporal Frequency
Temporal Frequency
Spatial Frequency
Direction
max
Predicted classifier accuracy (% correct)
50
B
Temporal Frequency
Spatial Frequency
Direction
100
Classifier accuracy (% correct)
50
t=52ms
t=64ms
t=316ms
t=560ms
C
Direction
Orientation
Correlation between representational dissimilarity matrix (RDM) and model (Spearman's rho)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
316ms
64ms
52ms
0
100
200
300
400
500
600
Time from stimulus onset (ms)
D
TF
Speed
SF
64ms
560ms
52ms
Time from stimulus onset (ms)

Figure 5

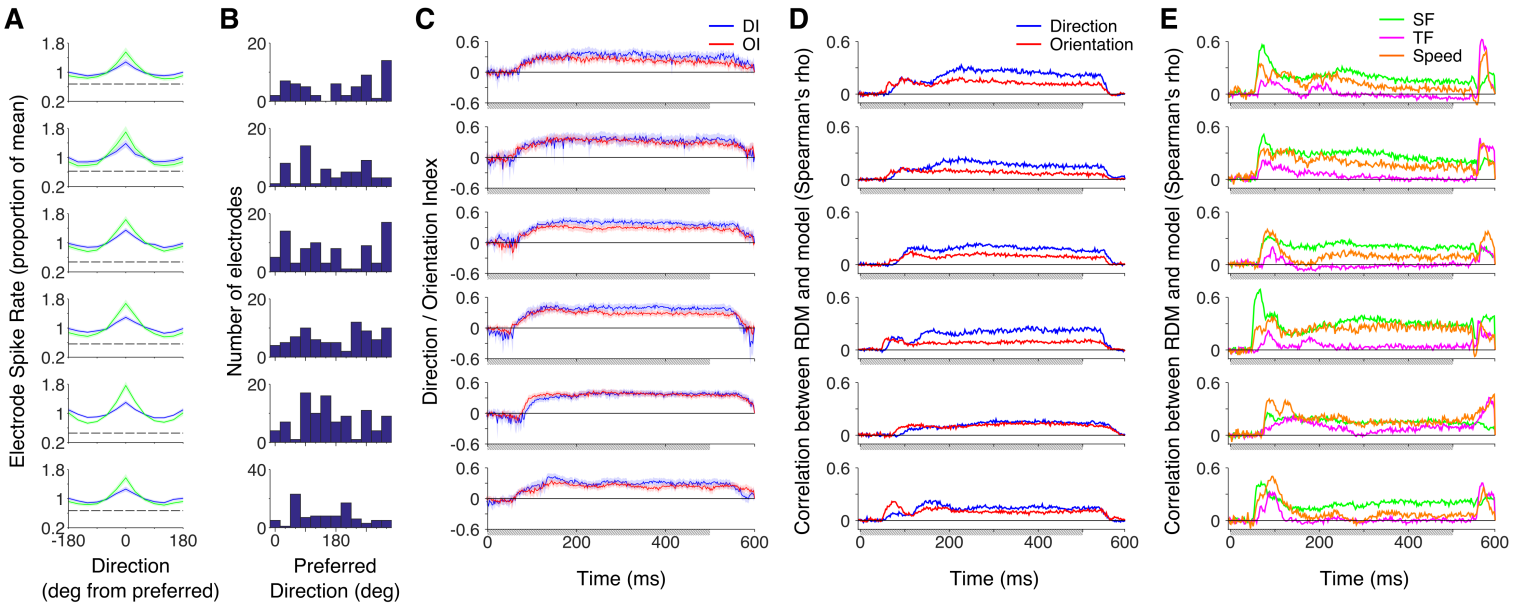
A
B
C
D
E
Electrode Spike Rate (proportion of mean)
Direction
(deg from preferred)
Number of electrodes
Preferred
Direction (deg)
Direction / Orientation Index
DI
OI
Time (ms)
Correlation between RDM and model (Spearman's rho)
Direction
Orientation
SF
TF
Speed

Figure 6

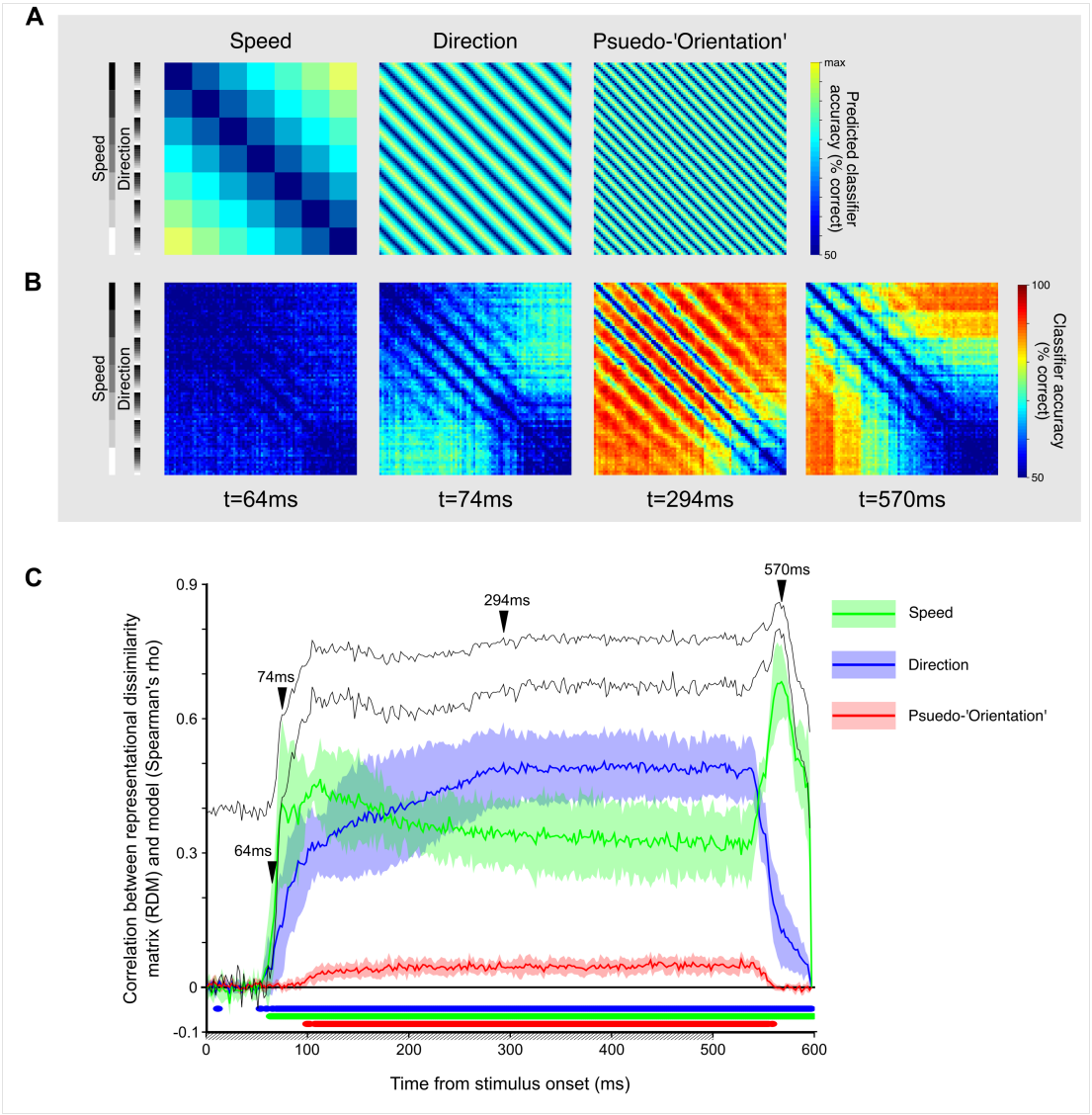
A
Speed
Direction
Psuedo-'Orientation'
Speed
Direction
max
Predicted classifier accuracy (% correct)
50
B
Speed
Direction
100
Classifier accuracy (% correct)
50
t=64ms
t=74ms
t=294ms
t=570ms
C
570ms
294ms
74ms
64ms
0.9
0.6
0.3
0
-0.1
Correlation between representational dissimilarity matrix (RDM) and model (Spearman's rho)
Speed
Direction
Psuedo-'Orientation'
0
100
200
300
400
500
600
Time from stimulus onset (ms)

Figure 7

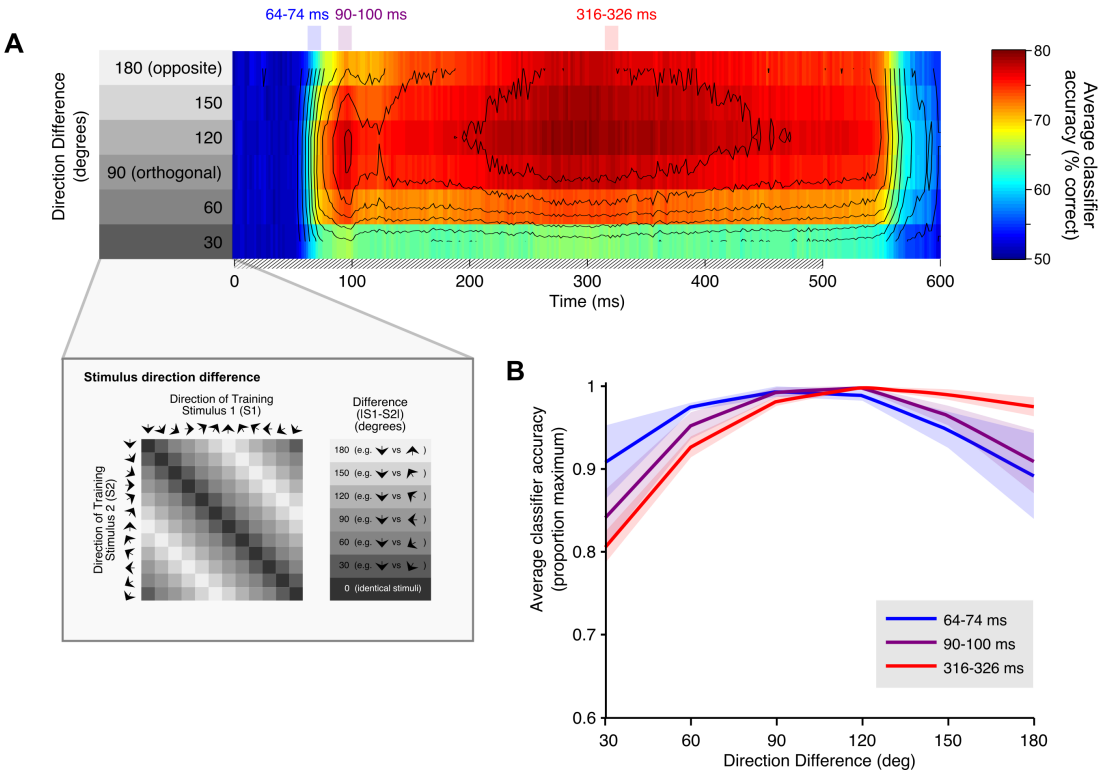
A
64-74 ms
90-100 ms
316-326 ms
180 (opposite)
150
120
90 (orthogonal)
60
30
Direction Difference (degrees)
0
100
200
300
400
500
600
Time (ms)
80
70
60
50
Average classifier accuracy (% correct)
Stimulus direction difference
Direction of Training Stimulus 1 (S1)
Direction of Training Stimulus 2 (S2)
Difference (|S1-S2|) (degrees)
180 (e.g. vs)
150 (e.g. vs)
120 (e.g. vs)
90 (e.g. vs)
60 (e.g. vs)
30 (e.g. vs)
0 (identical stimuli)
B
Average classifier accuracy (proportion maximum)
1
0.9
0.8
0.7
0.6
30
60
90
120
150
180
Direction Difference (deg)
64-74 ms
90-100 ms
316-326 ms

Figure 8

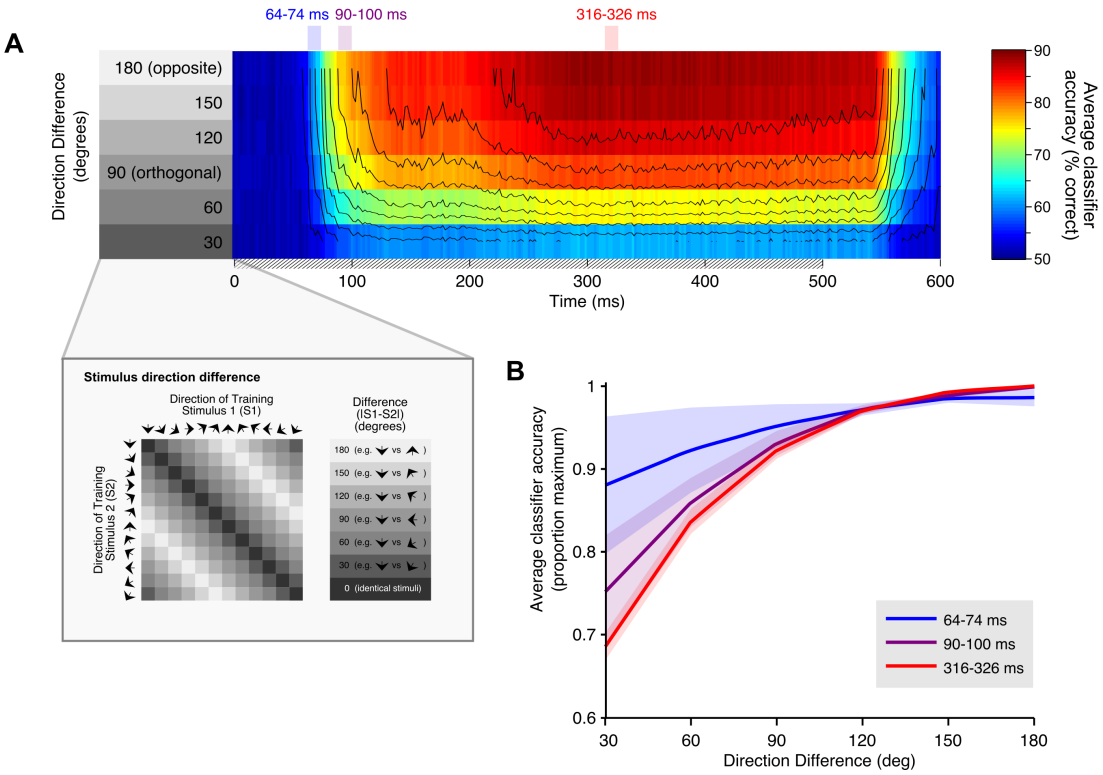
A
64-74 ms
90-100 ms
316-326 ms
180 (opposite)
150
120
90 (orthogonal)
60
30
Direction Difference (degrees)
0
100
200
300
400
500
600
Time (ms)
90
80
70
60
50
Average classifier accuracy (% correct)
Stimulus direction difference
Direction of Training Stimulus 1 (S1)
Direction of Training Stimulus 2 (S2)
Difference (|S1-S2|) (degrees)
180 (e.g. vs)
150 (e.g. vs)
120 (e.g. vs)
90 (e.g. vs)
60 (e.g. vs)
30 (e.g. vs)
0 (identical stimuli)
B
Average classifier accuracy (proportion maximum)
1
0.9
0.8
0.7
0.6
30
60
90
120
150
180
Direction Difference (deg)
64-74 ms
90-100 ms
316-326 ms

Figure 9

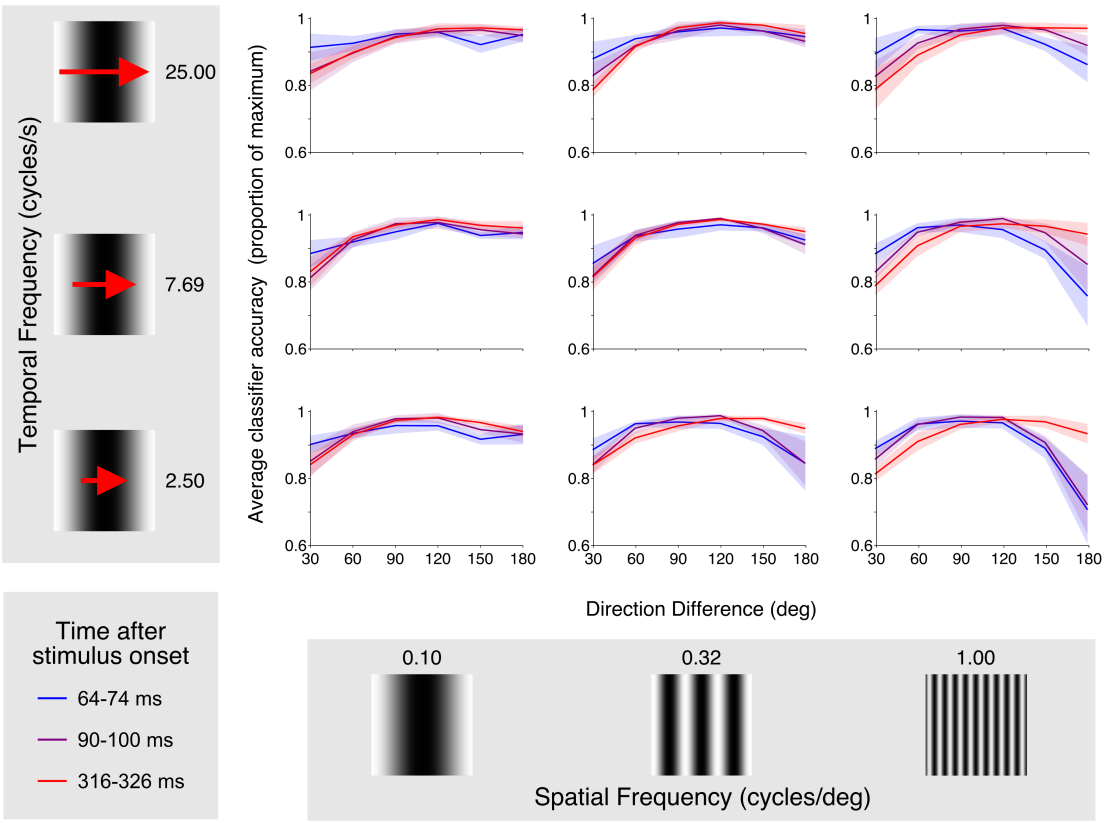
Temporal Frequency (cycles/s)
25.00
7.69
2.50
Average classifier accuracy (proportion of maximum)
1
0.8
0.6
30
60
90
120
150
180
Direction Difference (deg)
Time after stimulus onset
64-74 ms
90-100 ms
316-326 ms
0.10
0.32
1.00
Spatial Frequency (cycles/deg)

Figure 10

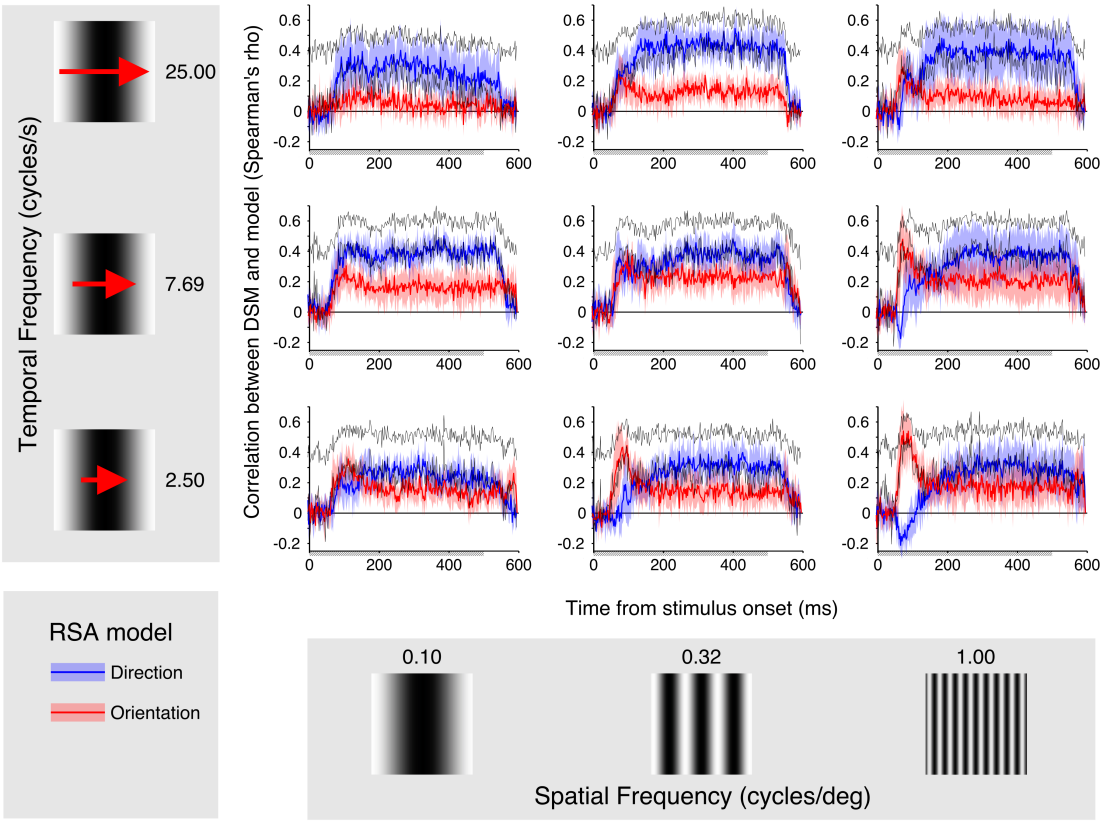
Temporal Frequency (cycles/s)
25.00
7.69
2.50
Correlation between DSM and model (Spearman's rho)
0.6
0.4
0.2
0
-0.2
0
200
400
600
Time from stimulus onset (ms)
RSA model
Direction
Orientation
0.10
0.32
1.00
Spatial Frequency (cycles/deg)

Figure 11