

Integration and segregation of multiple motion signals by neurons in area MT of primate

J. Scott McDonald,^{3,4} Colin W. G. Clifford,^{1,3} Selina S. Solomon,^{1,2} Spencer C. Chen,^{1,2} and Samuel G. Solomon^{1,2,5}

¹ARC Centre of Excellence in Vision Science, The University of Sydney, New South Wales, Australia; ²Discipline of Physiology, School of Medical Sciences and Bosch Institute, The University of Sydney, New South Wales, Australia; ³School of Psychology, The University of Sydney, New South Wales, Australia; ⁴School of Psychology, University of New South Wales, Sydney, New South Wales, Australia; and ⁵Cognitive, Perceptual and Brain Sciences Research Department, University College, London, United Kingdom

Submitted 10 April 2013; accepted in final form 22 October 2013

McDonald JS, Clifford CW, Solomon SS, Chen SC, Solomon SG. Integration and segregation of multiple motion signals by neurons in area MT of primate. *J Neurophysiol* 111: 369–378, 2014. First published October 23, 2013; doi:10.1152/jn.00254.2013.—We used multielectrode arrays to measure the response of populations of neurons in primate middle temporal area to the transparent motion of two superimposed dot fields moving in different directions. The shape of the population response was well predicted by the sum of the responses to the constituent fields. However, the population response profile for transparent dot fields was similar to that for coherent plaid motion and hence an unreliable cue to transparency. We then used single-unit recording to characterize component and pattern cells from their response to drifting plaids. Unlike for plaids, component cells responded to the average direction of superimposed dot fields, whereas pattern cells could signal the constituent motions. This observation provides support for a strong prediction of the Simoncelli and Heeger (1998) model of motion analysis in area middle temporal, and suggests that pattern cells have a special status in the processing of superimposed dot fields.

vision; extrastriate cortex; population coding; transparency; plaid

A CENTRAL QUESTION IN SYSTEMS neuroscience is how neuronal populations represent multiple sensory values or motor outputs (Dayan and Abbott 2001). The analysis of visual motion in primate middle temporal area (MT) is a model system for understanding how sensory estimates can be derived from neural codes (Pouget et al. 1998, 2000; Shadlen and Newsome 1998). Previous work has inferred, from the direction tuning of single neurons, that the population response of area MT is a noisy “hill of activity.” A single motion direction can be estimated from this neural response by, for example, calculating a population vector. However, retinal image motion often arises from several overlapping sources. In these cases, the multiple directions of motion can be accompanied by the perception of transparent motion. The phenomenal segregation of motion into transparent surfaces provides two challenges for the brain and models of it (Snowden et al. 1991; Snowden and Verstraten 1999). First, models need to distinguish the number of motions present; second, they need to extract the appropriate directions.

It has been suggested (Treue et al. 2000) that the representation of transparent motion does not require two distinct hills of activity in the population response. Instead the breadth of the population response determines the number of surfaces that should be extracted (Mahani et al. 2005): broader population response indicates the presence of multiple motion directions. Using multielectrode recordings from area MT, we confirm that population response to superimposed dot fields conforms to these expectations.

The idea that population response width indicates the number of motions present cannot be applied to the perception of moving plaids, which are formed by the superposition of gratings with different motion directions. The population response of area MT is expected to be broader for plaids than it is for gratings (Rust et al. 2006), which we confirm here. If transparency were based on population width, plaids should always be seen as two transparent surfaces. Instead, plaids usually appear to move as a single moving surface.

A subset of neurons in area MT, “pattern cells,” show unimodal direction-tuning curves for plaids; their activity is therefore correlated with the single perceived direction of the plaid. Other neurons, “component cells,” show bimodal tuning curves for plaids, aligned to the motion directions of the component gratings. Which neurons are important for the representation of multiple directions during transparent motion? One possibility is that “pattern cells” represent the hypothesis that there is a single moving surface, and signal its overall direction, while “component cells” represent the hypothesis that there are multiple surfaces, and signal the multiple motion directions present. Whether one surface (as for superimposed gratings) or multiple surfaces (as for superimposed dot fields) were inferred depends on which neurons were used to make the decision. A second possibility is that whether the stimulus is a plaid or moving dot fields, only the response of pattern cells is used for motion estimation, such that the number and direction of the surfaces present must be inferred from the distribution of responses across them. Here we distinguish these hypotheses by a simple experiment: we compare the response of component and pattern cells to superimposed moving dot fields.

METHODS

Physiology. Nine adult marmosets (*Callithrix jacchus*, 340–430 g; 1 female) were obtained from the Australian National Health and

Address for reprint requests and other correspondence: S. G. Solomon, Cognitive, Perceptual and Brain Sciences Research Dept., 26 Bedford Way, London WC1P 0AH, UK (e-mail: s.solomon@ucl.ac.uk).

Medical Research Council (NHMRC) combined breeding facility. Procedures were approved by University of Sydney Animal Ethics Committee and conform to the Society for Neuroscience and NHMRC Code of Practice. Each animal was first anesthetized with an intramuscular injection of alfaxan (12 mg/kg) and diazepam (3 mg/kg). A tracheostomy was performed, and the animal artificially respirated with N₂O in carbogen (ratio 70:30). This supplemented postsurgical anesthesia maintained by continuous (via a tail vein) intravenous infusion of sufentanil citrate (6–12 $\mu\text{g}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$) in sodium lactate with added dexamethasone (0.4 $\text{mg}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$) and amino acids. Daily intramuscular injections provided antibiotic cover (Noricillin, 25 mg) and additional antiedematous cover (dexamethasone, 0.1 mg). The electrocardiogram, electroencephalogram and arterial O₂ saturation from pulse oximetry were monitored continuously. Muscular paralysis was then induced and maintained by continuous infusion of pancuronium bromide (0.3 $\text{mg}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$). At the end of the experiment (typically 72–96 h) the animal was killed with intravenous 160 mg/kg sodium pentobarbitone.

A craniotomy was made over area MT. In three animals, multi-channel recordings were made with a 96-channel array (Blackrock systems, 0.4 M Ω , 1.5 mm length, 0.4 mm separation), band-pass filtered (0.3–6 kHz), and sampled by a Tucker Davis Technologies RZ2 at 24 kHz. The array was inserted into the left hemisphere to a depth of \sim 1 mm using a high-speed pneumatic device; depth of penetration varied because of curvature of the underlying cortex. In six other animals single-unit extracellular recordings were made from the right hemisphere using tetrodes (Thomas Recordings, 2–5 M Ω). The analog signals from the electrodes were amplified, band-pass filtered (0.3–10 kHz) and sampled at 48 kHz by the computer that generated visual stimuli. The single-unit data included here comes from a subset of those neurons examined in Solomon et al. (2011). The craniotomy was sealed with agar (arrays; 2% in saline) or silicon elastomer (tetrodes).

Visual stimuli. Stimuli were presented on a calibrated cathode-ray-tube monitor (ViewSonic G810 or Sony G500, refreshed at 100 Hz, viewing distance 45 or 114 cm, mean luminance 45–55 cd/m^2). Stimuli were drifting sine-wave gratings, plaids, or fields of moving circular white dots (100% coherence, infinite lifetime, density of 0.3 $\text{dots}\cdot\text{s}^{-1}\cdot\text{deg}^{-1}$, dot width 0.4 $^\circ$) against a background of the mean luminance. Direction 0 $^\circ$ is motion to the right. For each single-unit, we established, with the preferred motion direction, the optimal speed of a moving dot field, and the optimal spatial frequency, temporal frequency, and size of a drifting grating of contrast 0.5. We measured direction tuning (30 $^\circ$ intervals) for this grating, and for plaids that were the sum of two gratings drifting at directions 120 $^\circ$ apart. We also measured direction tuning (15 or 30 $^\circ$ intervals) for two drifting dot fields of optimal speed and the same size as above, moving at directions 0, 30, 60 or 120 $^\circ$ apart. In each case, stimuli were presented in a pseudo-random sequence for 0.32 s, with no interstimulus interval; each stimulus was presented on average 35 times. For array recordings, we used large stimuli (diameter 40 $^\circ$) positioned to evoke responses from as many of the recording sites as possible. Response to gratings and plaids were obtained separately to that for dot fields. Grating spatial frequency was 0.4 cycles/ $^\circ$; temporal frequency was 5 Hz. Dot fields moved at 20 $^\circ$ /s. Stimuli were presented 50 times in a pseudo-random sequence for 0.5 s, with 0.1-s interstimulus interval.

Neural response. Multiunit activity was extracted from array recordings by finding all waveforms that exceeded 5 standard deviations of the mean. To ensure that the offset of the stimulus was not included, response was measured over a 0.4-s period starting 0.05 s after stimulus onset. Single units were extracted from tetrode recordings by principal components analysis of the waveforms (Solomon et al. 2011). Response was measured over the 0.32-s period that maximized the response variance over all stimulus conditions in the set (Smith et al. 2005; Solomon et al. 2011). Stimulus sets always included a blank screen of the mean luminance, from which we extracted the maintained discharge rate. The data sets include neurons

where 1) average response over the measurement period exceeded the maintained rate by 5 impulses/s; 2) response to single dot fields or drifting gratings was well tuned [circular variance < 0.5 (Solomon et al. 2011)]. For array recordings, these criteria permitted 53, 43 or 29 sites for further analysis of dot fields in the three animals. Off-line analysis was performed in the Matlab environment.

Indexes of motion integration and motion segregation. We used established methods to classify pattern and component cells (Movshon et al. 1985; Smith et al. 2005; Solomon et al. 2011). The partial correlations between the observed and ideal responses ($r_{\text{component}}$ and r_{pattern}) were transformed to z -scores (z_c and z_p) using Fisher's r -to- Z transformation. Neurons were classified as component because the correlation with the component prediction (z_c) exceeded 1.28, and the difference between the component and pattern correlations ($z_c - z_p$) also exceeded 1.28; the reverse criterion was applied to classify neurons as pattern selective. In two neurons classified as component cells in this way, direction tuning for plaids was single-peaked, but preferred motion direction was aligned to one of the component gratings, as if the neuron responded to one of the components but not the other. Four other component cells showed bimodal direction tuning to a single dot field. As their response is not consistent with that expected of component cells, these six neurons were excluded from further analysis. Of the 65 remaining neurons, we identified 15 component cells and 27 pattern cells; 23 were not classifiable as either pattern or component cells by these metrics.

We used the same method to generate an index of motion segregation, except that response to single dot fields (angular separation 0 $^\circ$) was used to generate predicted response to superimposed dot fields. We assumed that response to a single dot field was half the response to superimposed dot fields of angular separation 0 $^\circ$. The two predictions are as follows: 1) direction-tuning curves that are the same shape for single fields and superimposed fields (i.e. z_p); 2) direction-tuning curves that are the sum of responses to each component field (i.e. z_c). The motion segregation index is $z_c - z_p$.

Decoding. We implemented a linear support vector machine (SVM) to determine the capacity of population response to discriminate between two average motion directions, or discriminate the angular separation, of two superimposed dot fields. Response rate on each trial was measured over a 0.4-s period starting 0.05 s after stimulus onset. The SVMs were implemented in Matlab (Mathworks, Natick, MA) using the SVM-Light toolbox (Joachims 2002; <http://svmlight.joachims.org/>).

We used cross-validation to establish robustness of the SVM. For each discrimination, we removed one trial from each of the two sets of responses (each N trials). The decoder was trained on the $N - 1$ trials remaining in each set; this decoder was then required to identify the most likely stimulus associated with each of the two left-out trials. This process was repeated for all possible permutations of the training and test trials. To quantify performance of the decoder we calculated d' , as $[z(\text{hit rate}) - z(\text{false alarm rate})]$ (chance performance, $d' = 0$). To assess the impact of correlations in spike counts, we repeated this process, but shuffled the order of trials in the training data set, after removing the test trials. Performance was quantified as above, on the unshuffled "raw" test trials that had been left out of the shuffled training data set.

To characterize the weight that the SVM applied to different sets of neurons during the decoding, the weights associated with each iteration of the SVM were scaled to unit vector length. Weights from each recording site were then averaged across the cohort of cross-validations associated with each discrimination and aligned to the preferred motion direction at that site. Finally, we aligned the preferred motion direction to the average motion directions of the stimuli being discriminated.

RESULTS

Neural response to coherent and transparent motion. We first established the response of populations of neurons in area MT to the motion of a single dot field, and when two dot fields are superimposed. We did this by measuring spiking activity across a 10×10 array of electrodes implanted into area MT of anesthetized marmosets, a diurnal simian primate where area MT lies exposed on the cortical surface (Rosa and Elston 1998; Solomon et al. 2011). Figure 1A shows direction-tuning curves for spike response in one implant, for a large field of small dots that all moved coherently along a single direction. The array covered a large fraction of area MT, so spike responses show tuning curves that together cover all possible motion directions.

Having established population response to a single dot field, we then measured response to superimposed dot fields, where the individual fields moved along directions 120° apart. The response of two representative sites is shown in Fig. 1, D and E: response is shown as a function of the direction of a single dot field or as a function of the average direction of the superimposed dot fields. For the site in Fig. 1D, response is similar for both stimuli. For the site in Fig. 1E, response to superimposed dot fields is bilobed, with each lobe separated by 120° and corresponding to the motion direction where one of the dot fields is aligned with the preferred direction.

To characterize the responses of individual sites to superimposed dot fields, we calculated a motion segregation index. Response to a single dot field was used to generate two models. In the first model, neural activity is the same for both a single dot field, and two superimposed dot fields, expected when a neuron computes the average motion direction of the retinal image. In the second model, neural activity is the sum of responses to each individual field, so the predicted response to superimposed fields is bilobed. We computed the correlation between the actual response and that predicted by each model; the segregation index compares these two correlations, factoring out the correlation between the predictions themselves. Large positive indexes indicate sites with very bilobed-tuning curves for superimposed dot fields, and negative indexes indicate sites sensitive to the average direction of motion.

Figure 1F plots the distribution of the segregation index for the active sites on the implant in Fig. 1, A–E, and across three implants. The segregation index is widely distributed. To show what that variability represents, we aligned the response of each site to the preferred direction for a single dot field and ordered them by the segregation index. Figure 1B shows response of all sites to a single dot field, and Fig. 1C shows in the same order the response of each site to the superimposed dot fields. Fig. 1, C–F, shows that superimposed dot fields generate a wide continuum of responses in area MT: from neurons that seem to signal the average direction of the two dot fields, to those that are capable of representing the motion of each field.

The measurements above are for superimposed dot fields moving at $20^\circ/s$, near the average preferred speed of individual neurons in marmoset area MT (Solomon et al. 2011). The stimulus speed will often be different to that preferred by individual recording sites, and we, therefore, considered the possibility that the segregation index depends on the relationship between preferred speed and stimulus speed. In separate

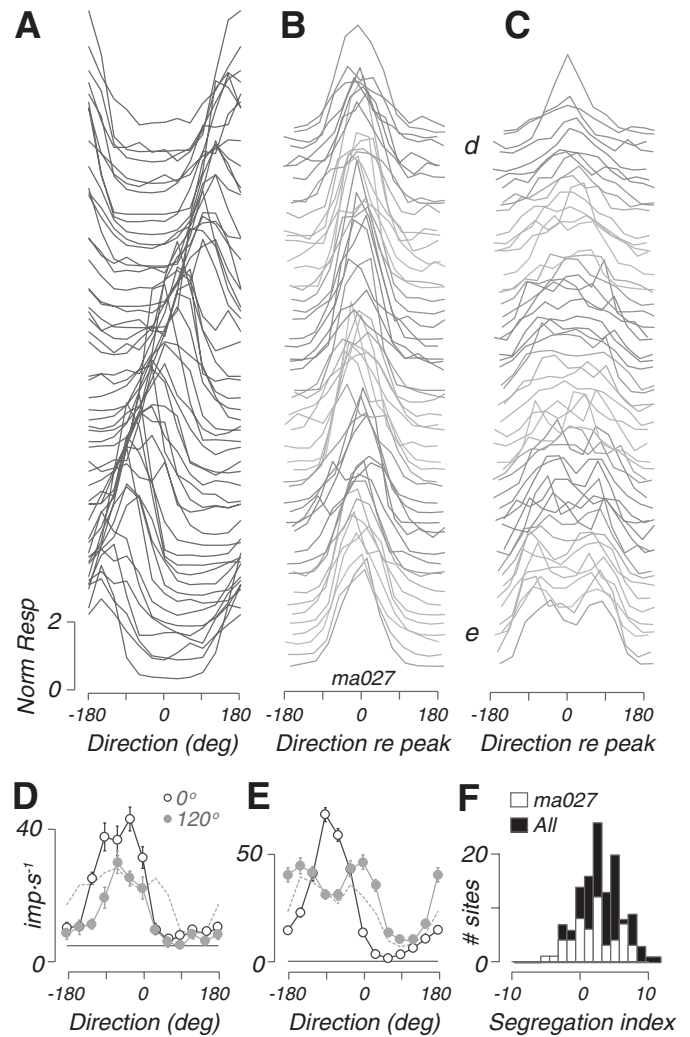


Fig. 1. Response of neurons in middle temporal area (MT) to moving dot fields. Simultaneous measurements from a 96-electrode array implanted into area MT of marmoset are shown. A: direction-tuning curves of 53 responsive, tuned sites to motion of a single dot field, arranged by preferred direction. Each curve is normalized to its mean across all stimuli. Data in A–E are from implant into left hemisphere of MA027. Sites were included if the range of responses exceeded 5 imp/s (81/96 sites), and response to a single dot field showed good directional tuning (circular variance of the tuning curve was less than 0.5; 53/81 sites). Average response range was 37 imp/s; average circular variance was 0.31. Dot speed was $20^\circ/s$. B: same as A, but aligned to the preferred direction (vector average) at each site. Tuning curves in B and C are arranged in the same order, by the motion segregation index described in RESULTS. Tuning curves at the bottom are those with the largest segregation index. C: direction-tuning curves during presentation of two superimposed dot fields moving 120° apart. Tuning curves are plotted as a function of the average motion direction of the two fields, aligned to the preferred direction for a single dot field. Conventions are as in A and B. D and E: comparison of response to single and superimposed dot fields. Average response is shown for single dot fields (open symbols) and superimposed dot fields (shaded symbols). Dashed line shows predicted tuning curve for superimposed dot fields, if response were the sum of that to the individual dot fields. Horizontal line shows maintained discharge rate. Error bars are 1 SE. D: a representative site showing unimodal tuning curve for superimposed dot fields, expected if response is aligned to the average motion of the two fields. E: a representative site showing bimodal tuning curve for superimposed dot fields, expected if response is the sum of that to individual dot fields. F: distribution of motion segregation index for MA027 (open bars), and across three implants (solid bars). Large positive values of the index indicate sites with bilobed response to motion of superimposed dot fields. Negative indexes indicate sites with unimodal response. The index is detailed in RESULTS and METHODS.

measurements, we obtained responses to single dot fields moving in each of 12 directions (30° intervals) and 7 speeds (5–80°/s in geometric steps). From responses along the optimum direction, we estimated preferred speed by finding the best predictions of a descriptive difference-of-exponentials function (Derrington and Lennie 1984). We grouped recording sites by their preferred speed: the motion segregation index for sites with preferred speed of 10° or less was on average 2.81 (SD 3.16; $n = 69$); for sites with preferred speed between 10 and 30° /s the index was on average 2.99 (SD 3.35; $n = 49$); for those with preferred speed greater than 30° /s, the index was on average 1.88 (SD 1.97; $n = 16$). Comparison of segregation index among groups yielded no significant differences (Students t -test, $P > 0.2$ in all cases). Motion segregation is therefore not simply dependent on the relationship between stimulus speed and the preferred speed of individual sites. We will return to this below, but, for the following analyses of population activity, we therefore included all recording sites.

Population response to coherent and transparent motion. As expected from the distribution of tuning curves in Fig. 1A, a single dot field moving in a particular direction evokes a “hill of activity” across the population (Fig. 2A, open symbols). The hill is centered on the neurons that prefer the motion direction presented and therefore respond vigorously to this stimulus. Figure 2A also shows population response for superimposed dot fields of the same average direction, but separated 120° (shaded symbols). Population response to motions 120° apart is broader than that for single dot fields, and there are two peaks. Each peak is centered on the neurons whose preferred directions are near the motion direction of one of the individual fields.

Figure 2B shows average population tuning curves across the three implants, during presentation of two dot fields whose directions of motion were separated by 0, 30, 60 or 120° . In each case, the dashed lines show the shape of population response expected if response was the sum of that for the component fields. The observed shape of the population response conforms to these expectations, confirming the inferences from earlier single-unit recordings (Treue et al. 2000): population response is unimodal for superimposed dot fields whose direction differs by 60° or less. The presence of multiple motions therefore does not necessarily bring about multiple peaks in area MT population response.

If multiple peaks are not present in the population response, what cues indicate the number of superimposed motion directions? One alternative decoding strategy is to use the breadth of the population response (Mahani et al. 2005). Figure 2, A and B, shows that the breadth of population response is, on average, larger for superimposed dot fields than single dot fields. These differences can be slight, however, and neural response to repeated presentations of the same stimulus is variable. Trial-to-trial variability might obscure the small differences in mean response.

Population discrimination performance for coherent and transparent motion. To establish the capacity of area MT population response to represent one or more motion directions, we used linear SVMs, which operated on the trial-by-trial response of the population. These simple decoders were highly capable of discriminating between neighboring motion directions (Fig. 3A). To characterize the performance of the decoder, we calculated d' (METHODS: *Decoding*). In these ex-

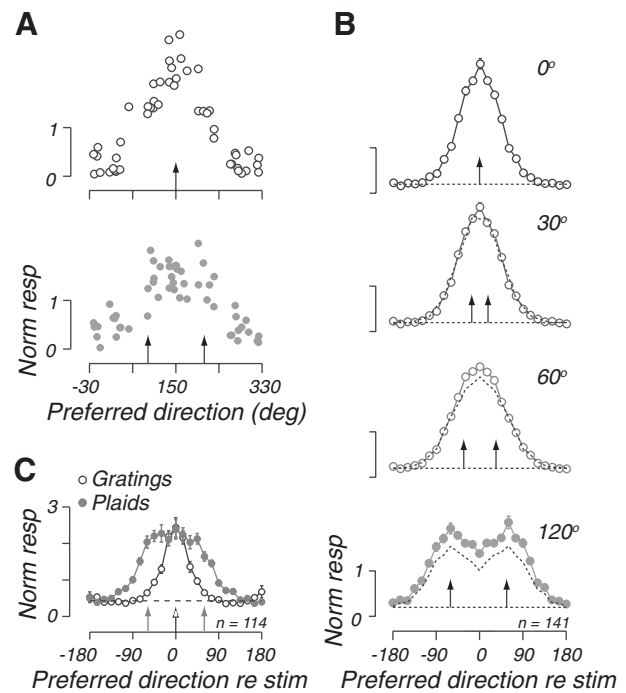


Fig. 2. Population response of area MT to coherent and transparent motion. **A**, top: population response to a single moving dot field, direction 150° (arrow). Average response at each of 53 sites (circles) is plotted as a function of preferred direction for that site. Same implant as in Fig. 1. Conventions are as in Fig. 1A. Bottom: population response to two superimposed dot fields, with direction 90° and 210° (arrows). Average response at each site (circles) is plotted as a function of preferred direction for a single dot field. Conventions as in top panel. **B**: population response across 3 implants. Each panel shows population response to superimposed dot fields of angular separation 0° (i.e., a single field), 30° , 60° or 120° . Arrows show motion directions of the dot fields. Dashed line shows average maintained discharge across all sites. Each datum plots average response across 12 motion directions (30° intervals), after the preferred direction for a single field was shifted to align to the average motion direction of the dot fields. The resultant population response was binned (15° bin width). Error bars, sometimes obscured, show 1 SE across the sites in each bin. Dashed lines show the predicted population response, if it were the sum of responses to the component dot fields. The predictions are lower than the measured activity. We return to this in the DISCUSSION. **C**: population response for gratings (open symbols) and plaids (shaded symbols). Plaids were the sum of two gratings with motion directions 120° apart. Responses were obtained from the same implants shown in B. Each datum plots average response across 12 motion directions (30° intervals), after the preferred direction for a grating was shifted to align to the average motion direction of the grating or plaid. Conventions are as in B.

periments, a d' value of 0.95 indicates the point at which an unbiased observer would make the correct decision on 75% of trials. Performance was best when the decoder needed to discriminate between two possible motion directions of a single dot field, but accuracy was also high when discriminating between the average directions of two superimposed dot fields. In a separate analysis, the decoders were required to discriminate the angular separation of two superimposed dot fields moving along a particular average motion direction. In this case, we expected that the decoder would regularly confound angular separations of 0, 30 and 60° , because the population response to these angular separations was so similar (Fig. 2B). Indeed, the decoder was largely incapable of discriminating angular separations of 0 and 30° , but population response was capable of discriminating larger angular separations (Fig. 3B).

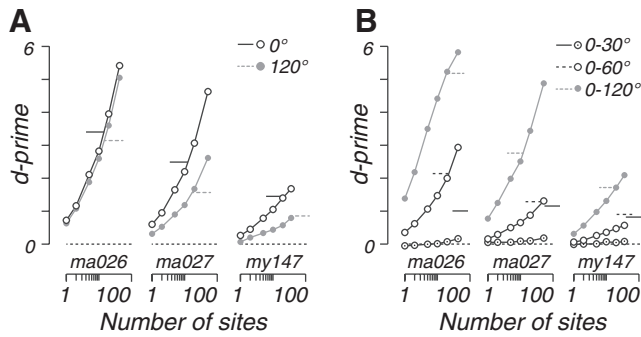


Fig. 3. Decoding motion direction and angular separation from population response of area MT. Linear support vector machines (SVMs) were used to establish capacity of the population of recording sites to represent angular separation and motion direction of superimposed dot fields. **A**: discrimination of average motion direction. The decoder was required to discriminate the average motion direction of single dot fields (open symbols), or of superimposed dot fields of angular separation 120° (shaded symbols). Symbols show average performance across motion directions. Performance was calculated for individual sites, for 100 random subsets of sites (with $n = 2, 5, 10$ or 20), and for the whole implant (rightward most point in each case). Horizontal lines show the best performance across individual sites, obtained from each of 12 motion directions, and then averaged. **B**: discrimination of the angular separation of two superimposed dot fields. The decoder was required to discriminate single dot fields from superimposed dot fields of angular separation 30° (dotted symbols), 60° (open symbols) or 120° (shaded symbols). Conventions are as in A.

How many neurons might be required to discriminate the angular separation, or two average motion directions? We provided the decoder the response of single sites, or randomly drawn samples of sites, and repeated the analyses. Individual sites were capable of correctly discriminating two average motion directions, for every angular separation, and performance improved rapidly as the number of neurons was increased (Fig. 3A). Individual sites could distinguish single fields from superimposed dot fields of direction difference 120°, but intermediate angular differences generally required larger groups (Fig. 3B).

Which elements of the population provide the capacity to discriminate between two average motion directions, or the number of motions directions present? Figure 4 shows the weight that the SVM applied to different neurons (see METHODS), as a function of the preferred motion direction of the neurons (cf. Fig. 2B). Each of the three columns shows the distributions from single animals, which were broadly in agreement.

Figure 4A shows the weights applied when discriminating between two average motion directions. Positive weights indicate the relevant neurons provided evidence that the motion was greater than (clockwise rotated from) the average, and negative weights indicate that the neurons provided evidence that the motion was less than (anti-clockwise rotated from) the average. For angular separations of 0–60°, the plots in Fig. 4A have the same basic shape. The arrows indicate the neurons given most weight in the discrimination: for a single moving field, most weight was given to neurons whose preferred direction lay 40° (case MA026), 58° (MA027) or 50° (MY147) away from the average motion direction. Similar values were obtained for angular separation of 30° (respectively 44, 61 and 61°), and angular separation 60° (respectively 64, 58 and 61°). For angular separation of 120°, the distributions are less clear,

presumably reflecting the presence of multiple peaks in population response.

Figure 4B shows similar plots, but during discrimination of single and multiple motions. Here, positive weights indicate evidence for the hypothesis that there is a single motion present. Where the distributions have coherent shape they are all in broad agreement: the decoder uses neurons with preferred directions near that of the average direction (which is that of the single field), and neurons with preferred direction far from the average. For 0–120° discrimination, the negative peaks are at –63, 115 and 108° respectively; for 0–60° discrimination these were 74, 96 and –55, respectively. Neurons with tuning curves aligned to the motion of the single field (and the average motion of the superimposed fields) are presumably useful

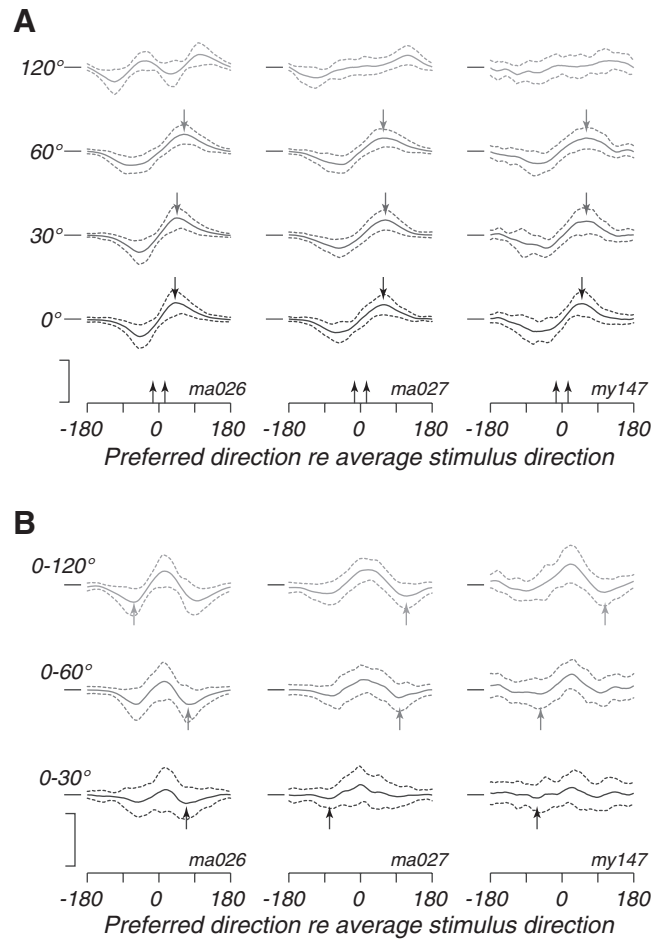


Fig. 4. Relationship between directional preference and utility in discrimination. Discrimination weights were recovered from the SVMs and shifted so that the preferred direction for a single field was aligned to the average motion direction of the dot fields being discriminated. Lines show means \pm SD of the weights across neurons and stimulus directions. **A**: weights for discrimination of average motion direction. Each row shows distribution of weights during discrimination of a 30° difference in the average motion direction of a single dot field (bottom), and superimposed dot fields of angular separation 30, 60 or 120°. Each column shows the distributions over a different animal. Vertical scale bar on left shows normalized weight of 0.5, calculated as described in METHODS. **B**: weights for discrimination of the angular separation of two superimposed dot fields. Each row shows distribution of weights during discrimination of a single dot field from superimposed dot fields of angular separation 30, 60 or 120°. The average motion direction of the superimposed dot fields was the same as that of the single field. Each column shows the distributions over a different animal. Vertical scale bar on left shows normalized weight of 0.5.

because the peak response amplitude can be different for the two stimuli (cf. Fig. 2*B*). Neurons with tuning curves that flank the motions of the superimposed dot fields are also informative, presumably because they can communicate the small differences in response amplitude that arise because superimposed dot fields generate broader direction-tuning curves. Figure 4*B* thus suggests that population-tuning width does indicate the spread of motion directions present and indicates that, where single fields and superimposed fields drive different response amplitude, this will also provide a potent cue.

Impact of neuronal correlations on discrimination performance. The variable firing rate of neurons in the cerebral cortex is often shared with other neurons, especially those that are nearby. Whether these correlations help or hinder the representation of sensory stimuli remains controversial (Averbeck et al. 2006; Graf et al. 2011; Latham and Nirenberg, 2005; Pouget et al. 2000). Neural correlations were present in the datasets analyzed here. To illustrate this, we calculated the Pearson's correlation coefficient between the spike counts at each pair of recording sites (50 trials per stimulus). Correlation coefficients were obtained for each stimulus, or after *z*-scoring the spike counts and collapsing them across motion direction, for each angular separation. Across all stimuli and datasets, spike count correlation was a mean 0.104 (SD 0.188, $n = 143,081$ comparisons; means in individual datasets ranged from 0.091–0.108). There was a weak tendency for the magnitude of correlation to increase with angular separation, increasing from a mean 0.098 for angular separations of 0°, to a mean of 0.108 for an angular separation of 120°. Application of a linear ANOVA, with dataset and angular separation as grouping variables, yielded main effects for angular separation [$F(3,11,954) = 5.88$; $P = 0.0005$], and dataset [$F(2,11,954) = 29.33$; $P < 0.0001$].

We therefore asked whether the performance of the SVM was dependent on having access to the structure of interneuronal correlations. To do this, we repeated the analyses above, but after shuffling the order of the trials in the training data set; as above, the decoder was cross-validated on “raw” trials that were left out of the training data set. That is, interneuronal correlations may be present in the test data, but the decoder is blind to them. If shuffling the training data changes the performance of the decoder on the test data, this indicates that there is information in the structure of interneuronal correlations that may be useful to subsequent computations (Averbeck et al. 2006). This manipulation had little impact on the distribution of weights (yielding plots indistinguishable from those in Fig. 4). Decoders that had access to the raw dataset performed better than those that had access to the shuffled datasets. That is, d' scores were improved by on average 6.7% (MA026), 13.2% (MA027) and 6.6% (MY147) for motion discrimination, and 5.5% (MA026), 14.5% (MA027) and 6.8% (MY147) for discrimination of angular separation. These performance improvements were not obviously dependent on the angular separation of the superimposed stimuli. We conclude that the performance of this simple decoder is improved by knowledge of correlations in spike rate, but that this improvement is mild.

In summary, population response is broadly tuned and is unimodal for superimposed dot fields whose direction differs by 60° or less, confirming the inferences from earlier single-unit recordings (Treue et al. 2000). Perceptually, angular separations of 30 or 60° are usually associated with transparent

motion (Braddick et al. 2002). The presence of multiple motions therefore cannot be inferred from the presence of multiple peaks in area MT population response (Jasinschi et al. 1992; Wilson and Kim 1994).

Our analyses show that for superimposed dot fields, the width of the population response may provide a cue to the number of motions present (Treue et al. 2000). We therefore asked if that cue remained robust if different stimuli were used. Figure 2*C* shows population response for gratings and plaids, obtained from the same implants. Plaids were the sum of two gratings, moving at directions 120° apart. The population response to plaids is broader than that to gratings. The breadth of the population response is therefore not a reliable indicator of motion transparency (Mahani et al. 2005), because plaids, which cohere into a single moving surface, produce a broad population response.

Role of component and pattern cells in motion segregation. Models of population coding often assume that neurons in area MT form a homogenous set of motion detectors, which differ only in their preferred stimulus. Neurons in area MT are not functionally homogenous (Born and Tootell 1992; Movshon et al. 1985; Smith et al. 2005), and Fig. 1 shows that the shape of direction tuning for superimposed motions is highly variable. In the following, we explore why this is the case. We will first show that different functional classes of neurons provide quite different signals during presentation of multiple motion directions.

Response to moving sinusoidal gratings, and plaids made by adding two gratings with different motion directions, reveals a spectrum of cell properties between two functional archetypes in area MT (Movshon et al. 1985; Rust et al. 2006). At one extreme, “component cells” show bimodal direction-tuning curves with peaks aligned to plaid's component directions (Fig. 5*A*); at the other extreme “pattern cells” show unimodal direction-tuning curves aligned to the average motion of the plaid (Fig. 5*C*). Classification of neurons as component or pattern or intermediate, “unclassified,” cells requires establishing how they respond to plaids and gratings. Because these are narrowband stimuli, this requires careful adjustment of the stimulus to that preferred by the neuron under study, and monitoring of individual spike waveforms over several hours of investigation at each site. This is not generally possible with recordings from multielectrode arrays so, in additional experiments, we isolated 65 single neurons in area MT using tetrodes, and categorized them as component, unclassified and pattern cells using standard methods (Movshon et al. 1985; Smith et al. 2005). For the sake of explication, we initially concentrate our attention on the extremes: the component and pattern cells.

Figure 5, *B* and *D*, plots average direction tuning of component and pattern cells to a single dot field; the responses are unimodal for both categories of neuron. The critical test is how component and pattern cells respond to superimposed dot fields. If component cells signal the presence of multiple surfaces, then we expect to see two lobes in the direction-tuning curves, imitating the response to plaids. If pattern cells signal average motion direction, we expect there to be a single lobe in the direction-tuning curve. We would expect such an arrangement if component and pattern cells are substrates of neural representation described by Hupé and Rubin (2003) that represent one and two surfaces and are engaged in mutual inhibition. Our data reveal instead the counterintuitive result

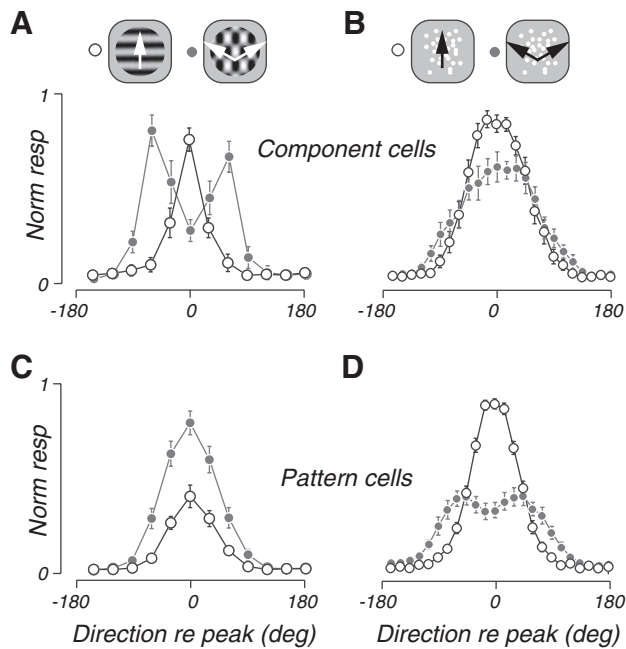


Fig. 5. Average response of component cells and pattern cells to coherent and transparent motion. *A* and *B*: component cells. *A*: average response of 15 component cells to gratings (open symbols) and plaids (shaded symbols). Component cells show characteristic bimodal response to plaids. Response to plaids is shown as a function of the average motion direction of the plaid. Response of each neuron was aligned to the preferred direction of a grating, and normalized to the maximum response to a grating, before averaging. Error bars show 1 SE. *B*: average response of the same component cells as in *A* to superimposed dot fields of angular separation 0° (i.e., a single field; open circles) or 120° (solid circles). Component cells do not show bimodal response for superimposed dot fields. Conventions are as *A*, *C* and *D*: pattern cells. Average response is shown of 27 pattern cells to gratings and plaids (*C*) or single fields and two superimposed fields (*D*). Pattern cells show unimodal response to plaids and bimodal response to superimposed dot fields. Conventions are as in *A* and *B*, respectively.

that, for superimposed dot fields, the direction tuning of component cells is single lobed (Fig. 5*B*), whereas the direction tuning of pattern cells is bilobed (Fig. 5*D*).

The average responses presented in Figure 5, *B* and *D*, do not reveal the diversity in the response of individual neurons. Figure 6, *A* and *B*, shows the response of 2 neurons to gratings and plaids (*upper* panels) and to one- and two-dot fields (*lower* panels). The response to gratings and plaids shows that the neuron in Fig. 6*A* is a component cell; for this neuron, response to superimposed dot fields is similar in shape to that to a single dot field. Figure 6*B* shows the response of a pattern cell: its response to superimposed dot fields is bimodal. Figure 6, *C* and *D*, shows histograms of the segregation index for component and pattern cells, respectively, where large values of the segregation index indicate bimodal direction-tuning curves for superimposed dot fields. Pattern cells, but not component cells, can show bimodal tuning curves for superimposed dot fields, but not all pattern cells show clearly bimodal response. More generally, our observations also reinforce the idea there is a continuum of response properties in area MT, between component cells and pattern cells.

Because of the average population response of the component and pattern cells to the dot fields, we expected that there would be a correlation between the difference in the component and pattern scores ($z_c - z_p$) with the motion segregation

index for all cells. However, we did not find a significant relationship between the indexes. We attribute this to the characteristics of the unclassified cells, whose index values are so uncorrelated across the two dimensions, taking negative and positive values, to render any relationship across the whole population insignificant.

In separate analyses, we explored among pattern cells the relationship between the motion segregation index and the stimulus preferences [established as described previously (Solomon et al. 2011)]. We saw no clear relationship between the index and preferred spatial frequency, the preferred size for a patch of drifting grating, or the strength of surround suppression as measured with drifting gratings. Large positive indexes, however, were associated with neurons that, when stimulated with gratings of optimal spatial frequency, preferred high temporal frequencies (correlation with log preferred temporal frequency, $r = 0.46$, $n = 19$, $P = 0.049$). The capacity of individual pattern cells to segregate the motion of superimposed dot fields, therefore, seems to require sensitivity to high temporal frequencies. This may reflect receptive field inhomogeneities similar to those reported in macaque (Nishimoto and Gallant

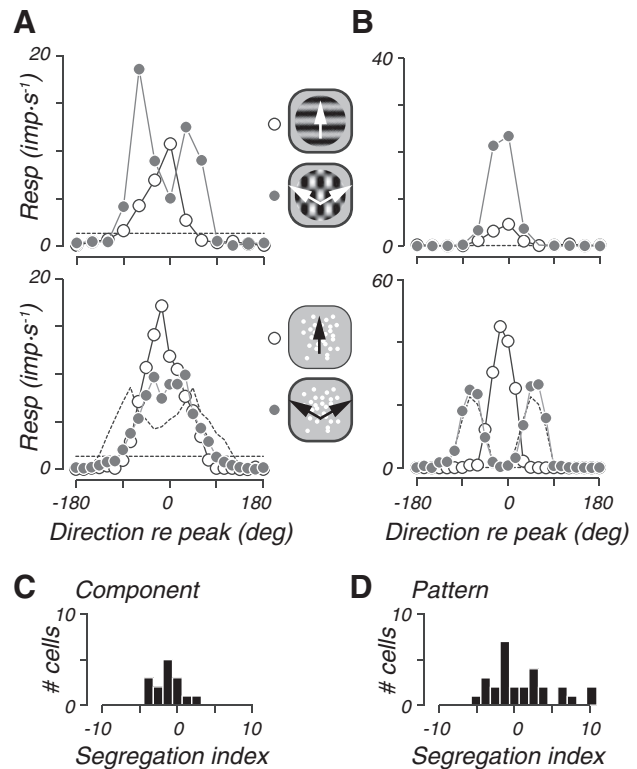


Fig. 6. Analysis of single-neuron responses. *A*: component cell. *Top*: response to gratings and plaids. Conventions are as in Fig. 5*A*. Response is plotted as a function of the preferred motion direction for gratings. *Bottom*: response to single dot field and two superimposed dot fields. Conventions are as in Fig. 5*B*. Dotted line in the *bottom* panel shows predicted response for superimposed dot fields, if it were the sum of responses to the individual dot fields. Response is plotted as a function of the preferred motion direction for single dot field. Horizontal line shows maintained discharge rate. For clarity, error bars are not plotted. *B*: pattern cell. Conventions are as in *A*. *C*: distribution of motion segregation index for component cells ($n = 15$). Segregation index was calculated for superimposed dot fields of angular separation 120°. Large negative values indicate unimodal response aligned to average direction of motion; large positive values indicate bimodal response aligned to the motion directions of the individual dot fields. *D*: distribution of motion segregation index for pattern cells ($n = 27$). Conventions are as in *C*.

2011), where neurons are better able to signal the motion carried by high temporal frequencies.

DISCUSSION

Our observations show that the response of small populations of neurons in area MT of primate is sufficient to distinguish transparent motion from that of a single surface. From a decoding perspective, this capacity does not require there to be two distinct “hills of activity” in the population response. The population response will be most robust if it is drawn from units that are capable of encoding the individual motions, and our observations show that pattern cells, but not component cells, are capable of encoding the motion direction of superimposed dot fields. Thus the same population of neurons may support the phenomenal coherence of moving plaids, and the segregation of superimposed dot fields.

Our single-unit measurements are largely in accord with an unpublished report by Bradley, Goyal and Scott (2005) [unpublished manuscript available at <http://sintn-seminars.stanford.edu/reprints/Bradley1.pdf>], that pattern cells in area MT of awake macaque respond to the individual motions present in a superimposed dot field. That work, however, also found component cells that responded to the individual motions, whereas those in our sample responded primarily to the vector average of the two motions. The discrepancy might reflect species differences or unknown impact of anesthesia. We discuss an alternate, functional, explanation below. Regardless, the two studies are in good agreement on the crucial issue: responses of pattern cells appear to closely reflect the perceptual experiences of the stimuli.

The unimodal response of component cells to superimposed dot fields is counterintuitive but is expected if the receptive fields of these neurons are tuned to motion energy within a localized range of spatial and temporal frequencies. This is best thought of in three-dimensional Fourier space. A dot field is represented by a plane in Fourier space (Fig. 7A). When two dot fields are spatially superimposed, local motion energy is greatest along the intersection of the two planes in Fourier space, which here corresponds to the average motion of the two fields (Fig. 7B). A component cell receptive field is a localized region of Fourier space (Fig. 7C). Component cell response is therefore maximal when the velocities of the two dot fields are such that in Fourier space the intersection passes through the region that corresponds to the component cell's receptive field (Fig. 7E). Component cells therefore have a unimodal direction tuning for dot fields, because they respond best to their average motion, where local motion energy is strongest. If the speeds of the superimposed dot fields are significantly different to that preferred by the component cell, then the component cell spectral receptive field will be further from line that is the vector average, and responses will align more with the motion direction of the individual fields (Fig. 7G).

Many pattern cells are capable of signaling the motion direction of each of two superimposed dot fields: these results

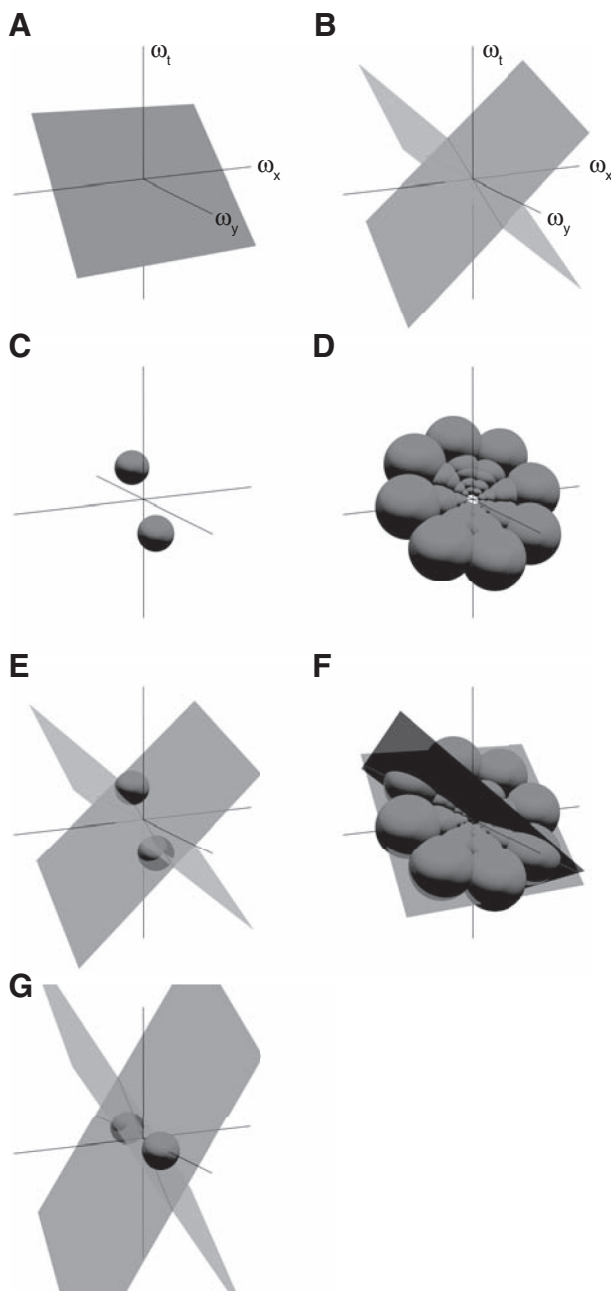


Fig. 7. Representations of dot fields and component cells in Fourier space. *A*: the power spectrum of a single, moving dot field. For translational motion, all of the power lies on a plane through the origin in frequency space. In a random dot stimulus, the power is distributed evenly over the plane. The tilt of the plane relative to the (ω_x, ω_y) plane corresponds to the speed of the dot field. *B*: the power spectra of two overlaid moving dot fields, moving in directions separated by 120° . The power spectra of the two dot fields intersect on a line in frequency space. This is the line of steepest ascent on the plane corresponding to the vector average of their motions. *C*: the localized spectral sensitivity of a model component cell. Azimuth indicates preferred direction, distance from the (ω_x, ω_y) plane indicates the preferred temporal frequency, and distance from the temporal frequency axis indicates the preferred spatial frequency. *D*: a model pattern cell, as formulated by Simoncelli and Heeger (1998). Notice that the pattern cell can be constructed by combining the responses of multiple component cells to tile a tilted plane through the origin in frequency space. *E*: two dot fields and a model component cell. The motions of the dot fields are tailored to the cell's spatiotemporal tuning, such that the intersection of the dot fields crosses the model component cell's region of sensitivity. *F*: two dot fields and a model pattern cell. The pattern cell will respond optimally when one or other of the dot fields are aligned with it, as illustrated. *G*: two dot fields and a model component cell. In contrast with *E*, the dot fields are no longer tailored to the cell's tuning properties. The intersection of the dot fields does not cross the model component cell's region of sensitivity. When the direction of motion is varied, the model component cell will respond when it intersects with the dot fields individually.

are compatible with a previously proposed model (Simoncelli and Heeger 1998), where pattern cells have receptive fields that are distributed across a plane in Fourier space (Fig. 7D). Because these neurons integrate motion energy across a range of spatiotemporal frequencies, they respond best to superimposed dot fields when the Fourier representation of either dot field matches that of their receptive field (Fig. 7F). Thus a strong prediction of this model of pattern cells, confirmed here, is that the response to superimposed dot fields is often multimodal. In our experiments, the measured response of neurons to superimposed dot fields is larger than the response predicted by summation of responses to a single field (Fig. 2B). Predictions were based on the assumption that each of the dot fields produced population response one-half that for superimposed dot fields of 0° angular separation. If the real response to a single field was greater than one-half that to superimposed dot fields (i.e., if a compressive nonlinearity was necessary to relate response to dot density), then we would underestimate response to angular separations greater than 0.

Simoncelli and Heeger's model (1998) is not the only one that can emulate pattern cells' responses to plaids. For instance, end-stopping V1 neurons can be exploited to determine the average direction of plaid (Pack et al. 2003; Tsui et al. 2010). Alternatively, the cascade model, in which V1 neurons are subject to gain control and their responses summed by MT neurons, with an accelerating nonlinearity output, can also predict pattern cells' responses to plaids (Rust et al. 2006). Our results introduce a new benchmark that must be passed by models for them to remain credible; they must also be able to emulate the responses of pattern cells to overlaid dot fields.

Multiple attribute values have been shown to coexist in other visual and motor domains. For example, in dorsal premotor cortex, multiple potential reach directions are encoded prior to the decision of which target to reach for (Cisek and Kalaska 2002). Although numerous schemes have been proposed for decoding a single-valued stimulus attribute from the response of a neuronal population, the problem of coding one or more values depending on the stimulus characteristics remains challenging (Snowden and Verstraten 1999; Zemel et al. 1998). Our measurements show that the shape of population response to coherent plaid motion can be remarkably similar to that for transparent dot motion. Thus the presence of multiple motions cannot be inferred from the breadth of area MT population response or the presence of multiple peaks (Jasinschi et al. 1992; Wilson and Kim 1994). Instead, the perceptual interpretation of coherent vs. transparent motion correlates well with the response of a subset of neurons that we find to behave in accord with the model of pattern cells proposed by Simoncelli and Heeger (1998). Restricting the population to pattern cells will not generate multiple peaks for small angular differences, but may allow the width of the population response to be a reliable cue to the number of motion directions present. Our measurements do not imply that cues to motion transparency have a selective impact on the response of pattern cells, as how pattern and component cells response changes with manipulation of cues to transparency will depend on the spatiotemporal structure of the stimulus (Stoner and Albright 1992).

Our analyses show that the performance of simple decoders improves when they have knowledge of the correlations between neurons, but the performance gain was modest. We think that the decoder's performance improves because the relative

activity of neurons is less variable than the absolute level of activity. As yet, we have only attempted to include correlations in overall spike rate; additional cues might arise in synchronous activity within local populations of neurons representing each of the component motions (Castelo-Branco et al. 2000, 2002; but see Thiele and Stoner 2003).

ACKNOWLEDGMENTS

We thank S. K. Cheong, S. Gharaei, P. R. Martin and S. Pietersen for help in experiments.

GRANTS

Supported by National Health and Medical Research Council of Australia (NHMRC) Grant 1005427, an NHMRC Career Development Award to S. G. Solomon, an Australian Research Council (ARC) Future Fellowship to C. W. G. Clifford, and the ARC Centre of Excellence in Vision Science.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: J.S.M., C.W.G.C., S.C.C., and S.G.S. conception and design of research; J.S.M., C.W.G.C., S.S.S., S.C.C., and S.G.S. performed experiments; J.S.M., C.W.G.C., S.S.S., S.C.C., and S.G.S. analyzed data; J.S.M., C.W.G.C., and S.G.S. interpreted results of experiments; J.S.M., C.W.G.C., and S.G.S. prepared figures; J.S.M., C.W.G.C., and S.G.S. drafted manuscript; J.S.M., C.W.G.C., S.C.C., and S.G.S. edited and revised manuscript; J.S.M., C.W.G.C., and S.G.S. approved final version of manuscript.

REFERENCES

- Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci* 7: 358–366, 2006.
- Born RT, Tootell RB. Segregation of global and local motion processing in primate middle temporal visual area. *Nature* 357: 497–499, 1992.
- Braddick OJ, Wishart KA, Curran W. Directional performance in motion transparency. *Vision Res* 42: 1237–1248, 2002.
- Castelo-Branco M, Formisano E, Backes W, Zanella F, Neuenschwander S, Singer W, Goebel R. Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc Natl Acad Sci U S A* 99: 13914–13919, 2002.
- Castelo-Branco M, Goebel R, Neuenschwander S, Singer W. Neural synchrony correlates with surface segregation rules. *Nature* 405: 685–689, 2000.
- Cisek P, Kalaska JF. Simultaneous encoding of multiple potential reach directions in dorsal premotor cortex. *J Neurophysiol* 87: 1149–1154, 2002.
- Dayan P, Abbott LF. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: Massachusetts Institute of Technology Press, 2001.
- Derrington AM, Lennie P. Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J Physiol* 357: 219–240, 1984.
- Graf ABA, Kohn A, Jazayeri M, Movshon JA. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14: 239–245, 2011.
- Hupé JM, Rubin N. The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. *Vision Res* 43: 531–548, 2003.
- Jasinschi R, Rosenfeld A, Sumi K. Perceptual motion transparency: the role of geometrical information. *J Opt Soc Am A* 9: 1865–1879, 1992.
- Joachims T. *Learning to Classify Text Using Support Vector Machines*. Boston, MA: Kluwer Academic, 2002.
- Latham PE, Nirenberg S. Synergy, redundancy, and independence in population codes, revisited. *J Neurosci* 25: 5195–5206, 2005.
- Mahani AS, Carlsson AE, Wessel R. Motion repulsion arises from stimulus statistics when analyzed with a clustering algorithm. *Biol Cybern* 92: 288–291, 2005.

- Movshon JA, Adelson EH, Gizzi MS, Newsome WT.** Pattern recognition mechanisms. In: *Pattern Recognition Mechanisms*. Rome: Vatican Press, 1985, p. 117–151.
- Nishimoto S, Gallant JL.** A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. *J Neurosci* 31: 14551–14564, 2011.
- Pack CC, Livingstone MS, Duffy KR, Born RT.** End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron* 39: 671–680, 2003.
- Pouget A, Dayan P, Zemel R.** Information processing with population codes. *Nat Rev Neurosci* 1: 125–132, 2000.
- Pouget A, Zhang K, Deneve S, Latham PE.** Statistically efficient estimation using population coding. *Neural Comput* 10: 373–401, 1998.
- Rosa MG, Elston GN.** 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. *J Comp Neurol* 393: 505–527, 1998.
- Rust NC, Mante V, Simoncelli EP, Movshon JA.** How MT cells analyze the motion of visual patterns. *Nat Neurosci* 9: 1421–1431, 2006.
- Shadlen MN, Newsome WT.** The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18: 3870–3896, 1998.
- Simoncelli EP, Heeger DJ.** A model of neuronal responses in visual area MT. *Vision Res* 38: 743–761, 1998.
- Smith MA, Majaj NJ, Movshon JA.** Dynamics of motion signaling by neurons in macaque area MT. *Nat Neurosci* 8: 220–228, 2005.
- Snowden RJ, Treue S, Erickson RG, Andersen RA.** The response of area MT and V1 neurons to transparent motion. *J Neurosci* 11: 2768–2785, 1991.
- Snowden RJ, Verstraten FA.** Motion transparency: making models of motion perception transparent. *Trends Cogn Sci* 3: 369–377, 1999.
- Solomon SS, Tailby C, Gharaei S, Camp AJ, Bourne JA, Solomon SG.** Visual motion integration by neurons in the middle temporal area of a New World monkey, the marmoset. *J Physiol* 589: 5741–5758, 2011.
- Stoner GR, Albright TD.** Neural correlates of perceptual motion coherence. *Nature* 358: 412–414, 1992.
- Thiele A, Stoner G.** Neuronal synchrony does not correlate with motion coherence in cortical area MT. *Nature* 421: 366–370, 2003.
- Treue S, Hol K, Rauber HJ.** Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci* 3: 270–276, 2000.
- Tsui JMG, Hunter JN, Born RT, Pack CC.** The role of V1 surround suppression in MT motion integration. *J Neurophysiol* 103: 3123–3138, 2010.
- Wilson HR, Kim J.** A model for motion coherence and transparency. *Vis Neurosci* 11: 1205–1220, 1994.
- Zemel RS, Dayan P, Pouget A.** Probabilistic interpretation of population codes. *Neural Comput* 10: 403–430, 1998.

