Spatial precision of population activity in primate area MT

Spencer C. Chen,1,2 John W. Morley,3 and Samuel G. Solomon2,4

1Australian Research Council Centre of Excellence for Integrative Brain Function, The University of Sydney, New South Wales, Australia; 2School of Medical Sciences, The University of Sydney, New South Wales, Australia; 3School of Medicine, University of Western Sydney, Penrith, New South Wales, Australia; and 4Institute for Behavioural Neuroscience, University College London, London, United Kingdom

Submitted 16 February 2015; accepted in final form 1 June 2015

Chen SC, Morley JW, Solomon SG. Spatial precision of population activity in primate area MT. J Neurophysiol 114: 869–878, 2015. First published June 3, 2015; doi:10.1152/jn.00152.2015.—The middle temporal (MT) area is a cortical area integral to the “where” pathway of primate visual processing, signaling the movement and position of objects in the visual world. The receptive field of a single MT neuron is sensitive to the direction of object motion but is too large to signal precise spatial position. Here, we asked if the activity of MT neurons could be combined to support the high spatial precision required in the where pathway. With the use of multielectrode arrays, we recorded simultaneously neural activity at 24–65 sites in area MT of anesthetized marmoset monkeys. We found that although individual receptive fields span more than 5° of the visual field, the combined population response can support fine spatial discriminations (Δ<0.2°). This is because receptive fields at neighboring sites overlap substantially, and changes in spatial position are therefore projected onto neural activity in a large ensemble of neurons. This fine spatial discrimination is supported primarily by neurons with receptive fields flanking the target locations. Population performance is degraded (by 13–22%) when correlations in neural activity are ignored, further reflecting the contribution of population neural interactions. Our results show that population signals can provide high spatial precision despite large receptive fields, allowing area MT to represent both the motion and the position of objects in the visual world.

Copyright © 2015 the American Physiological Society

First published June 3, 2015; doi:10.1152/jn.00152.2015.

Address for reprint requests and other correspondence: S. Chen, The Univ. of Texas at Austin, Seidemann Lab, Center for Perceptual Systems, 1 University Station A8000, Austin, TX 78712-0187 (e-mail: spencer.chen@utexas.edu).

www.jn.org

0022-3077/15 Copyright © 2015 the American Physiological Society

869
MATERIALS AND METHODS

Subjects and electrophysiological recordings. The surgical and recording procedures have been detailed previously (McDonald et al. 2014; Solomon et al. 2014). All procedures were approved by the University of Sydney Animal Ethics Committee and conform to Australian National Health and Medical Research Council (NHMRC) policies on the use of animals in neuroscience research. Adult male marmosets (Callithrix jacchus, n = 3) were obtained from the NHMRC combined breeding facility. The animal was initially sedated with an intramuscular injection of 12 mg/kg Alfaxan (Jurox, New South Wales, Australia) and 3 mg/kg Diazepam (Roche, New South Wales, Australia). Postsurgical anesthesia was maintained by continuous intravenous infusion of sufentanil citrate (4–12 μg·kg⁻¹·h⁻¹; Sufenta Forte; Janssen Cilag, Beerse, Belgium). The animal was artificially ventilated with a 70:30 mix of N₂O and Carbogen. Rectal temperature was kept near 38°C with the use of a heating blanket.

Vital signs (ECG, EEG, SpO₂, and rectal temperature) 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 reduced level of anesthesia, the dose of sufentanil citrate was increased. To suppress eye movements, muscular paralysis was then induced and maintained by continuous infusion of pancuronium bromide (0.3 mg·kg⁻¹·h⁻¹; AstraZeneca, New South Wales, Australia). High-permeability contact lenses 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 an intravenous overdose of sodium pentobarbitone (500 mg/kg; Lethabarb; Verbac Australia, New South Wales, Australia).

A craniotomy was made over area MT in the left hemisphere and the dura reflected. Multichannel recordings were made with a 96-channel array (Blackrock Microsystems, Salt Lake City, UT; 1.5 mm length, 0.4 mm separation, average impedance 0.265 MΩ), band-pass filtered (0.3–5 kHz), and sampled by a PZ2/RZ2 at 24 kHz (Tucker-Davis Technologies, Alachua, FL). The array was inserted to a depth of ~1 mm using a high-speed pneumatic device (Rousche and Normann 1992). The electrodes generally extended into or past layer 4, and slight curvature of the cortex means that the depth of the electrodes varies across the array. From the trajectory of receptive-field positions on each electrode (Rosa and Elston 1998), we were able to identify electrodes that were likely to be within area MT and others likely to be in area MTc [a thin area bordering the anterior of area MT; see Rosa and Elston (1998) and Solomon and Rosa (2014)]. For the current analyses, we have included neurons from electrodes in both areas.

Stimuli. Visual stimuli were generated by a Power Mac G4 computer using custom software (EXPO; Peter Lennie, University of Rochester, Rochester, NY) and presented on a cathode ray tube monitor (Sony G500, 100 Hz refresh rate, width 40 cm, height 30 cm, mean luminance 45–55 cd/m²), viewed directly at 45 cm. Supplementary lenses were used to focus the eyes, and the contralateral eye was occluded during measurements. In one animal, we made an additional measurement with the ipsilateral eye instead occluded. The primary stimulus was a white disk of diameter 3° and intensity twice that of the background, moving across a gray screen at 20°/s. The disk traveled in a straight line along one of five paths, length 40°, at angular —18°, —6°, 0°, 6°, and 18° from the horizontal [throughout, we use the term “degrees” (°) to refer to spatial separations on the monitor and “angular degrees” (angular °) to refer to the angle of motion], and all crossing at the center of the screen (see Fig. 1B). Both directions of motion were sampled for each path; each direction of motion is referred to as one trajectory. Each of the 10 trajectories was presented 80 times in pseudorandom order.

Multitunit spiking activity. Analyses were performed in MATLAB (R2012a; MathWorks, Natick, MA). The function “findpeaks” was used to identify waveforms with peak amplitude that exceeded 3 SD of the raw signal. Multunit spike count at each site was estimated in nonoverlapping time bins of width 0.05 s or 0.01 s and transformed into z-scores for subsequent analysis. z-Scores were calculated independently for each electrode, using the mean and the SD of the binned spike counts across all trials of all trajectories.

For explication in the following, we call the position and direction of motion associated with each time bin a single “stimulus.” Not every pair of stimulus positions fell within the spatial receptive fields of the recorded population, and discrimination performance between these pairs is difficult to interpret. To focus on the pairs of positions that elicited response from the recorded population, we first calculated the mean z-score across electrodes and trials at each stimulus position. Pairs of positions were included for analysis if mean z-scores both exceeded zero. When analyzing pairs of positions on different motion trajectories, we included only pairs separated by >0.5°; this means that the central point, where all trajectories cross, was excluded from analysis.

Estimating spatial receptive fields. Response latency (90 ms) was estimated as the delay that maximized the cross-correlation of each recording site’s response to opposite directions of motion along the same stimulus path, collapsed across recording sites and paths within an animal. Spatial receptive fields for each trajectory were characterized by finding the best predictions of a one-dimensional Gaussian model, with four parameters defining maximum response, center position, SD, and maintained activity. Preferred motion direction of each site was determined separately for each stimulus path. Some sites obviously did not respond to the stimulus at any position along the trajectory(ies) of interest, such that the peak amplitude of the Gaussian that best predicted response did not reach significance (P < 0.05). Our primary aim is to understand how receptive fields contribute to spatial discrimination, and therefore, we excluded these sites from the analyses. Counterpart analyses that nevertheless included these sites generally resulted in poorer discrimination performance (not shown).

Support vector machine analysis. Linear support vector machines (SVMs) (Graf et al. 2011; Joachims 1999; Vapnik 2000) were used to quantify the discriminability of population activity for pairs of stimuli. A leave-one-out cross-validation procedure was used: SVMs were trained on population activity, measured over 79 trials of each stimulus, and tested on the pair of left-out trials. The procedure was repeated for the 80 unique pairs of left-out trials. Support weights from each SVM (that is, for each pair of stimuli) were normalized to a unit vector and then averaged across the 80 cross-validated datasets. In each case, performance was estimated as d-prime (d’)

\[
d' = \text{norminv} \left( \frac{100 \cdot \text{HIT} + 1}{102} \right) - \text{norminv} \left( \frac{100 \cdot \text{FAR} + 1}{102} \right) (1)
\]

with the (100° k + 1)/102 adjustment to assign large, noninfinite values when the hit rate or false-alarm rate was 0 or 100%, and “norminv” is the inverse normal function with a mean of 0 and SD of 1. SVMs that were instead trained with 10 or 20 trials left out for cross-validation showed d’ performance and distribution of support weights that were consistent with that obtained in the leave-one-out approach (not shown), suggesting that the SVMs are identifying robust patterns of activity across the relevant populations.

To assess discrimination performance of a single site and randomly drawn subpopulations of 2, 5, 10, or 20 sites, we repeated the SVM analyses above. In each case, 80 randomly selected subpopulations of responsive sites were drawn and d’ performance calculated. In further analyses, a Gaussian weighting function (SD 6° of the visual field), centered on the discrimination target, was used to bias the subpopulation toward the locality of the target position; 80 weighted subpopulations of 10 sites were drawn and the SVM analyses repeated. Linear regression was used to compare performance for weighted and randomly selected subpopulations of the same size.

Compensation for cortical footprint. The decoder performance depended on the fraction of the population recruited at each stimulus.
position. Two positions that are by larger sets of neurons or neurons with better signal to noise are inherently more discriminable than others. We used a general linear model (GLM), assuming normally distributed error, to estimate the variability in performance that could be attributed to stimulus position. Effectively, the GLM estimates $d'_{\text{POS1}, \text{POS}2}$ and $d'_{\text{POS2}}$

$$d' (\text{POS1, POS}2) = d'_{\text{POS1}} + d'_{\text{POS2}} + \hat{d'}$$

yielding residuals $\hat{d'}$, representing position-independent performance.

The position-independent performance was further modeled with an elliptical surface [parameters: $A$, $r$, spatial-direction ratio (SDR), and $d'_{\text{POS1}}$, with respect to separation in the stimulus position ($\Delta\text{POS}$) and motion direction ($\Delta\text{DIR}$)]

$$\hat{d'} = A \left( \exp \left( -\frac{1}{r^2} \frac{1}{\Delta\text{POS}^2 + \Delta\text{DIR}^2} \right) - 1 \right) + \hat{d'}_{\text{POS}}$$

from which we extract the SDR. SDR characterizes the difference in spatial position that is required to achieve the decoding performance observed for motion directions that are an angular $1^\circ$ apart. Finally, the offset value $\hat{d'}_{\text{POS}}$ was removed so that the final “estimated $d'$” (see Fig. 6, C and D) is zero when both positional and directional separation are zero.

In additional analyses, we asked if the discrimination surface could be better explained by alternative descriptive models. Linear, separable functions in $\Delta\text{POS}$ and $\Delta\text{DIR}$ failed to follow the elliptical contours in the data. Multiplicative, separable functions produced similar SDR values ($\sim 1.2 \times$ greater than the inseparable model used in the main analyses) but showed larger fitting error at small values of $\Delta\text{POS}$ and $\Delta\text{DIR}$.

Influence of noise correlations on decoding performance. To assess whether the performance of the SVM was dependent on having access to the structure of interneuronal correlations, we repeated the analyses above but after shuffling the order of the trials in the training dataset. As above, the decoder was cross-validated on “raw” trials that were left out of the training dataset; that is, the decoder is tested on actual brain activity, but training on shuffled data makes it incapable of using any correlations that may exist in the real dataset. Changes in the performance of the shuffled decoder therefore reflect the contribution of the information in the structure of interneuronal correlations to subsequent computations.

We recognized that there is a simple relationship between the performance from SVMs trained on the raw data (correlation-aware) and SVMs trained on shuffled data (correlation-blind) decoders. By definition, $d'$ measures the ratio between the difference of the mean responses to the SD of the noise. Since shuffling the training data does not change mean response, the impact of shuffling on $d'$ must be related to changes in SD of the noise in the population that is introduced by shuffling. The ratio of the $d'$ between the decoders can therefore be characterized by a simple gain/loss factor ($G$)

$$\frac{d'_{2}}{d'_{1}} = \frac{m}{\sigma_{2}} = \frac{\sigma_{1}}{m} = G$$

This allows us to use linear regression to characterize the change in performance ($G$) when using different decoders. In the relevant analyses (see Fig. 4), we used only $d'$ values $< 3.5$ in the regression to avoid ceiling effects of near-perfect discrimination.

Pearson’s correlation coefficient was calculated for all possible pairing of responsive sites. To factor out site-by-site variations in correlation magnitude, GLMs were used to estimate the correlations ($\hat{\rho}$) not attributable to the pairs of electrodes (MU1 and MU2) under consideration (applied as seen in Fig. 5)

$$\rho(\text{MU1, MU2}) = \rho + \rho_{\text{MU1}} + \rho_{\text{MU2}}$$

Selective correlation masking. We used synthetic data to investigate how decoding performance depends on noise correlations between pairs of electrode sites whose receptive fields are at different locations with respect to the stimulus. For each electrode, the position of each stimulus was expressed as distance from its receptive-field center in multiples of the SD of its receptive field. These normalized positions were then binned in five partitions (negative numbers imply the position is earlier in the trajectory than the receptive-field center): less than $-1.5$ SD; $-1.5$ to $-0.5$; $-0.5$ to $0.5$; $0.5$ to $1.5$; $>1.5$. Responses of each pair of electrodes were collated for each pair of normalized positions. This created 25 datasets; for example, in one dataset, we have the responses of pairs of neurons, where one of the receptive fields is centered on the stimulus ($-0.5$ to $0.5$ SD), and one is far ahead of the stimulus ($>1.5$ SD).

We calculated the correlation matrix for these datasets and then created two synthetic datasets (1,280 trials each), in which we preserved mean rate at each recording site but either retained or modified the covariances. We used MATLAB’s multivariate normal random number generator “mvmrnd” to generate the synthetic datasets, providing it with the mean activity for each electrode site and the preserved or modified covariance matrix. In the modified set, we halved (rather than removed) the covariances and used the nearest “symmetric-positive definite” matrix to minimize side effects that can arise when altering the covariance matrix. SVMs were trained on the real datasets and tested on the two synthetic datasets; performance gain for each dataset was calculated as in Eq. 4.

RESULTS

In the following, we characterize the precision of spatial representations in area MT by recording from populations of neurons with planar multielectrode arrays implanted into three anesthetized marmosets; each implant yielded between 24 and 65 responsive sites for analysis. The arrays covered a large fraction of area MT, such that the underlying receptive fields occupied a wide swathe of the visual field (Fig. 1A). As expected, the receptive fields in area MT were retinotopically organized and highly overlapping (Born and Bradley 2005; Britten 2003; Rosa and Elston 1998). To understand how populations of motion-sensitive neurons might represent spatial position, we analyzed multiunit activity evoked by a single moving object (diameter $3^\circ$) that traveled smoothly across the visual field at moderate speed ($20^\circ$/s; Fig. 1B). We analyzed four sessions obtained in three animals; in one animal, we had sufficient time to reposition the monitor and obtain a second recording.

The receptive fields of neurons in area MT are large, and the consequence is that neurons respond to a wide range of spatial positions (Fig. 1C). Receptive-field diameter (full width at half height) increased with eccentricity from a mean of $6.3^\circ$ ($\pm 3.7$ SD) in the parafoveal half of the stimulus trajectory to $11.7^\circ$ ($\pm 4.6$ SD) in the more peripheral half of the stimulus trajectory. Individual neurons will therefore be poorly capable of detecting small changes in position. By contrast, as an object moves across the screen, it evokes reliable responses in many neurons at each spatial position, and the moving object therefore creates a moving “hill” of activity in the MT population response. This distributed response is a potentially rich source of information for fine position judgements.

Spatial precision of area MT population response. To characterize the spatial precision of population responses in area MT, we used linear SVMs, which allow us to estimate empir-
Fig. 1. Population response to a moving stimulus in the middle temporal area (MT). A: receptive fields of neurons in area MT are large and highly overlapping. Estimated receptive-field center (dots) and extent (shaded regions; 1 SD for multunit activity obtained with a 10 × 10 electrode array implanted into area MT in the left hemisphere of 1 animal. Estimated position of the fovea is marked with “F.” For clarity, the extent of a subset of receptive fields is illustrated. Color of receptive fields specifies position in the recording array. B: the stimulus was a disc moving across the screen. Response was obtained to both directions of movement, along each of 5 paths arranged around the horizontal axis. C: average response at each recording site to a disc moving along the horizontal axis from left to right (left) and from right to left (right). Note the x-axis on the right is flipped to preserve the true temporal order of neural response. Electrode sites were arranged by the preferred stimulus position. Electrode position in the recording array is identified by the color of the dots between panels. One recording site is highlighted by the thick lines in both left and right and illustrates the directional sensitivity of neural responses in area MT.

physically the discriminability of population response at two spatial positions without requiring assumptions about the correlation structure of population responses or how they might be used to discriminate changes in spatial position. Here, we trained the SVMs to discriminate a change in the spatial position of the moving object along a single trajectory of motion (Fig. 2A). For each pair of spatial positions, we trained a new set of SVMs, using a “leave-one-out” cross-validation approach to establish how well the population response could be discriminated on a single trial. We obtained similar results in each of the four recordings and therefore, pooled across them.

Application of the SVM showed that population response was capable of fine spatial precision on individual trials, with robust performance for both small and large changes in the spatial position of the moving object (Fig. 2B). For a 1° change in object position, average was 2.2 (±0.1 SE); that is, an unbiased observer, forced to choose between the two positions, would be correct on 86% of trials. Performance increased with spatial separation, and in all cases, the precision of population response exceeded that of the best individual site. To assess how performance depended on population size, we repeated the analysis, drawing on randomly sampled subpopulations (Fig. 2C). Performance increased with the number of sites included, and the performance of the 10 best sites approached that of an entire population of 24–65 sites.

The analyses above used a response window of 0.05 s, which parsed a trajectory into bins of 1°. To assess performance over finer spatial separations, we repeated the analyses with a response window of 0.01 s, thereby parsing the trajectories into bins of 0.2°. The smaller bin size naturally reduces spike count and in addition, may change covariances among neurons in the population, both of which may impair discrimination performance. Indeed, discrimination performance for spatial separations of 1° was impaired when using the narrower time windows to an average of 1.5 (77% correct). Performance nevertheless remained above chance (d = 0) for spatial separations of 0.2° (d = 0.35 ± 0.01, SE; 57% correct) and 0.4° (d = 0.74 ± 0.01, SE; 64% correct).

Neuronal contributions to a distributed code. How are the responses of neurons used to discriminate between alternative spatial positions? To address this question, we revisited the SVMs trained to discriminate pairs of positions (using time bins of 0.05 s) along a single motion trajectory and analyzed the weights that the SVMs applied to different recordings sites. The organization of our analyses means that the SVMs assign positive weights to neural activity that supports the hypothesis that the stimulus is at the target position and negative weights to neural activity that supports the hypothesis that the stimulus is at an earlier position in the trajectory, which can be 1°, 2°, 4°, or 8°. The SVMs give stronger weight to neurons that are more informative about the decision boundary, and the question here is how the weights relate to the spatial tuning of the underlying neurons.

We first investigated the coding strategy for fine position discriminations. Inspection of individual sites suggested that neurons were more likely to be informative when the peak of the receptive field was offset from the targets, such that the
positions to be discriminated lay on one of the flanks (Fig. 3A). To represent the distribution of weights across the population of neurons, we aligned the position of each stimulus to each tuning curve and expressed stimulus position in units of that tuning curve's width. Projection of SVM weights into the same space (Fig. 3B) confirms that neurons are more likely to be informative when their receptive field was offset from the discrimination. By contrast, large position differences—spatial separations larger than the size of receptive fields—activated largely separate populations of neurons, and neurons were more likely to be informative if one of the targets was near the center of the tuning curve (Fig. 3B). Analyses over a range of target separations showed smooth transition in the distribution of weights between the two coding schemes (not shown).

Given open access to population response, fine position discrimination was primarily supported by those neurons with receptive fields within 2 SD (~6°) of the discrimination locus (Fig. 3B). This suggests that downstream areas may be able to recover spatial position with limited pooling over area MT output. To establish if limited pooling would be sufficient, we constructed SVMs where the probability of being included in the SVM was set by the centroid of the receptive field relative to the target positions. Specifically, subpopulations of 10 neurons were drawn from the population with probability given by a Gaussian distribution over visual space (SD 6°) and centered on the target positions. Discrimination performance of these localized subpopulations was, on average, 89% [±3%, 95% confidence interval (CI)] that of the random subpopulations shown in Fig. 2C. The reduction in performance implies that neurons farther away from the discrimination locus may provide useful signals, particularly when they have high signal to noise, and is consistent with the very large receptive fields of neurons in area MT. The high performance of the localized SVMs suggests that the spatial precision of population response can be “read out” with limited spatial pooling of its signals.

**Impact of interneuronal correlations on fine spatial discriminations.** Receptive fields of neurons in area MT are highly overlapping and are therefore likely to receive input from overlapping sets of neurons. Neurons in area MT are therefore unlikely to provide independent analyses of the retinal image—some of their activity will be shared with other neurons (Bair et al. 2001; Cohen and Kohn 2011; Cohen and Newsome 2008; Huang and Lisberger 2009; Solomon et al. 2014). These interneuronal correlations will impose motifs on area MT population response that might be important for its targets and be used by our decoding machines.

To assess whether the SVMs had identified population motifs, we further trained SVMs on artificial datasets, where the patterns of neural correlations were destroyed by shuffling the order of trials in each recording site. Shuffling allowed us to test an alternate decoding framework, where the brain is blind to its interneuronal correlations. We therefore refer to this set of SVMs as “correlation-blind” SVMs and the original set as “correlation-aware” SVMs. We cross validated the performance of the correlation-blind SVMs on real data to establish how those decoders interpret real neural activity. If the correlation-blind decoders showed the same performance as the correlation-aware decoders, then this would imply that the structure of interneuronal correlations is not useful for spatial discriminations.

The removal of correlations from the training set did not change the pattern of weights that the SVM attributed to individual sites (not shown). We used linear regression to estimate the change in decoder performance (G; Eq. 4) that is brought about by removing correlations from the training set. This analysis showed that removing correlations substantially reduced performance, by 13–22% (Fig. 4B). This provides evidence for the presence of reproducible population motifs in the spiking activity. The performance reduction was more pronounced for fine discriminations (1° separation: 22 ± 2%, 95% CI; Fig. 4C) than coarse discriminations (8° separation: 13 ± 3%), implying that population motifs are most useful when two stimuli elicit similar population responses.

To understand the structure of these motifs, we calculated the correlation in spike counts (“noise correlations” over time-scales of 0.05 s) between pairs of recording sites, using Pearson’s correlation coefficient. Preliminary analyses indicated that individual recording sites could show consistently higher or lower levels of correlation with other sites: individual sites contributed 9.5–23.7% of the variance in correlation estimates in each recording. This may reflect differences in the number of neurons contributing to multiunit activity at each site (Cohen and Kohn 2011). To compensate for the overall differences in correlation at individual sites, we removed them (see Eq. 5). The resulting, adjusted correlation coefficients are...
The population motifs that emerge in Fig. 5C are rich in structure, but they appear to be driven by two rules. The first rule is that correlation is highest in neurons that are near each other and therefore, have similar receptive fields (Huang and Lisberger 2009; Smith and Kohn 2008; Solomon et al. 2014); this provides a diagonal structure in the matrix (Fig. 5A). The second rule is that correlation depends on response amplitude. This provides a cross structure in the matrix (Fig. 5B). The shape of this cross arises from the inter-relationship of two factors: correlation increases with response amplitude when the stimulus is within the receptive fields of both neurons (Churchland et al. 2010; Gutnisky and Dragoi 2008; Kohn and Smith 2005), and correlation is also high when the stimulus is absent from both receptive fields (Smith and Kohn 2008). The result is that correlation is relatively low when the stimulus is within one receptive field but not the other, and this is what forms the centered vertical and horizontal bands of low correlation in the matrix.

We would like to know which aspects of the motifs in Fig. 5C might be exploited in decoding, but we cannot isolate them in real spiking activity. We therefore simulated population response using the original measurements of response amplitude and the correlation matrix in Fig. 5C to generate distributions of synthetic spiking activity. Selective attenuation of parts of the correlation matrix (Fig. 5D) allowed us to generate new distributions of spiking activity. We used changes in the performance of the SVM, implemented as above, to assess the importance of those correlations. The attenuation of all correlations reduced discrimination performance by, on average, 10% (±0.4%, 95% CI), similar to the reduction in performance that is brought about by shuffling training data above. Selective attenuation of parts of the correlation matrix revealed that this overall reduction in performance arises because the SVMs particularly rely on pairs of neurons in which the object lies on opposite flanks of the receptive field. In these pairs, changes in object position bring about an increase in spiking activity at one site and a decrease at the other, whereas interneuronal correlations either increase or decrease activity in both neurons. In these pairs, therefore, noise correlations distribute response along dimensions that are not aligned with the changes in response that are brought about by object motion. By contrast, the attenuation of correlations among neurons with similar receptive fields improved performance (Fig. 5E). This is because interneuronal correlations and changes in object position both increase or decrease activity in both neurons.

Comparison of spatial and direction discrimination. We established a benchmark for our estimates of spatial precision by measuring discrimination performance for pairs of positions along different motion trajectories (Fig. 6A). Each combination of position and motion direction has a different cortical footprint. That is, each position in visual space is sampled differently—both in terms of the number of electrode sites that are active at each position of the moving object in visual space and the signal to noise at those electrode sites. Combinations of positions that project onto larger ensembles of neurons or neurons with better signals are inherently more discriminable. To assess its impact, we created a surrogate measure of cortical footprint for each pair of positions: the average of z-scored spike rates across all recording sites and all trials at the two positions. This measure was a strong predictor of the d’ performance (not shown): as average z-score increased from −0.1 to 0.1, discrimination increased by 1.8 d’ units.

To establish how discrimination depends on spatial and direction separation, we need to establish the variation in performance that does not reflect variation in cortical footprint. We therefore estimated the impact of the cortical footprint by incorporating the actual positions of the two stimuli as predictors of discrimination performance in a GLM (see MATERIALS and METHODS, section on discrimination).
To determine the impact of retinal eccentricity on discrimination performance, we conducted the same analysis on the parafoveal and peripheral halves of the stimulus trajectories: estimated discrimination resolution for the two halves was, respectively, 0.10° and 0.14° (SDR of 0.049). The reduced resolution in peripheral receptive fields implies a significant decrease in the ability to discriminate between stimuli.

**DISCUSSION**

Visually guided behavior requires precise knowledge of both the position and motion of objects. The size of receptive fields of neurons in extrastriate areas of visual cortex, including area MT, means that individual neurons have limited capacity to signal the position of objects. By measuring the response of populations of these neurons, we have shown that the population signals of these areas allow high spatial precision and that these signals are accessible to simple decoders. Even our relatively small populations of neurons in area MT showed above-chance performance for targets separated by 0.2° of the field of moving dots (McDonald et al. 2014). Full-width, half-maximum directional tuning bandwidth was an angular 130° (±54 SD) and was similar for neurons with parafoveal and peripheral receptive fields (P = 0.10, t-test); a commonly used index of directional selectivity (1 - respref/respref) in the same recordings was 0.81 (±0.34 SD). These are comparable with that reported for single-unit activity in area MT of macaque and marmoset, in which tuning bandwidth is angular 100–120°, and the direction selectivity index is 0.85–1.0 (Britten 2003; Solomon et al. 2011).

In the analyses above, we excluded spatial discriminations for pairs of positions that lay on the same trajectory as a safeguard against effects arising from temporal correlations between these data points. We asked if the model could predict discrimination performance for these pairs of positions. We first factored out the impact of the cortical footprint (position-dependent performance) using the footprint that we estimated from the pairs of positions on different trajectories. The model’s predicted discrimination performance for positions on the same trajectory is very similar to that observed (Fig. 6D), with slight overestimation for small spatial separations.
null
aptation has its greatest impact on capacity to discriminate changes in pattern orientation, spatial frequency, or motion direction when the adapting stimulus is offset from the discrimination locus in the relevant dimension (Hol and Treue 2001; Regan and Beverley 1983, 1985; Tzvetanov and Womelsdorf 2008). By analogy, our observations predict that many aspects of spatial discrimination will be poorer during adaptation to displaced positions than during adaptation to overlapping positions. The most potent displacement should depend on the scale of the receptive fields engaged in the task, and this may help constrain which cortical regions are engaged in different forms of spatial vision.

**Knowledge of noise correlations improves discrimination.**

Our analyses show that moving objects impose a rich pattern of interneuronal correlations, which may be useful for spatial discriminations. Because a small moving object moves into and out of individual receptive fields, the structure of correlations is richer than that previously revealed during visual stimulation with large surfaces, which always cover the receptive fields of the neurons under study [e.g., Bair et al. (2001), Cohen and Kohn (2011), Rosa and Elston (1998), Smith and Kohn (2008), and Solomon et al. (2014)]. We were able to reveal this structure because we made measurements from large populations of neurons simultaneously: measurements from pairs of neurons would only provide single slices through the matrix in Fig. 5C. This rich structure emerges from simple principles. First, interneuronal correlations in the absence of a visual stimulus are likely to reflect intrinsic rhythms of cortical networks (Smith and Kohn 2008; Vidne et al. 2012), and introduction of a stimulus reduces the impact of these rhythms in neurons whose receptive fields cover the stimulus position (Churchland et al. 2010; Gutnisky and Dragoi 2008; Kohn and Smith 2005). Second, the presence of a stimulus brings neurons away from spike threshold and allows more of the shared membrane-potential fluctuations to be visible in spiking activity (Cohen and Kohn 2011)—spike correlations are stronger when the stimulus is within the receptive fields of the relevant neurons. Third, neurons with similar receptive fields, those close together in the cortical sheet, show stronger correlations [e.g., McDonald et al. (2014) and Smith and Kohn (2008)].

How interneuronal correlations affect population performance will depend on the functional properties of the relevant neurons. The most important neurons in fine spatial discriminations are those whose receptive fields flank the discrimination locus (cf. Fig. 3). In pairs of similarly tuned neurons, where the object lies on the same flank of the receptive field, noise correlations increase overlap in firing rates evoked by each of the two positions and impair neural performance. By contrast, where the object lies on opposing flanks of receptive fields, noise correlations can reduce overlap in responses to two positions and can thereby improve neural performance. This is because the change in stimulus position brings about opposite changes in mean response. Thus interneuronal correlations can help or hinder neural computations, as implied by theoretical work (Averbeck et al. 2006; Averbeck and Lee 2006; Cohen and Kohn 2011; Latham and Nirenberg 2005; Nirenberg and Latham 2003; Romo et al. 2003; Sompolinsky et al. 2001). Our observations further suggest that these different impacts of noise correlations do not “cancel out” when considered across populations of neurons. When we implemented decoders that ignored correlations, performance dropped by 13–22%, consistent with other work (Graf et al. 2011; Pillow et al. 2008). The performance loss is especially acute in fine spatial discriminations, where two stimuli elicit similar population responses.

Our observations predict that in spatial discrimination tasks that rely on the activity of similarly tuned neurons, reduced noise correlations should be associated with increased behavioral performance (Fig. 5C). This is in agreement with previous work that shows that attention-related improvements in a change-detection task are associated with reduction in correlations between neurons, in which the visual change generally increased mean firing rate (Cohen and Maunsell 2009). For tasks that instead rely on neurons with opposite changes in mean response, our observations predict that improvements in performance may be associated with increased neural correlations. This is in agreement with recent work that shows that attention-dependent improvements in a contrast-discrimination task can be associated with increased noise correlations between neurons that provide evidence for opposite choices (Ruff and Cohen 2014).

**ACKNOWLEDGMENTS**

The authors thank S.-K. Cheong, A. S. Pietersen, S. S. Solomon, and P. R. Martin for assistance in experiments. The authors also thank K. Jeffery, P. R. Martin, and H. J. Spiers for comments on previous drafts of the manuscript.

**GRANTS**

Support for this work was provided by National Health and Medical Research Council of Australia Grant 10005427 and the Australian Research Council Centre of Excellence in Vision Science.

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the authors.

**AUTHOR CONTRIBUTIONS**


**REFERENCES**


Ruff DA, Cohen MR. Attention can either increase or decrease spike count correlations in visual cortex. Nat Neurosci 17: 1591–1597, 2014.


