

## Supplementary Notes

### Supplementary Note 1

In this Section, we study the relationship between the shape of the direction-tuning curve and fine discrimination thresholds. First, we computed an “average tuning curve” from all the 240 neurons and deduced the discrimination threshold at various points on this curve. We constructed this direction-tuning curve with the average background response, the average peak response and the average tuning-width. We computed direction-discrimination thresholds at 45 points on this curve starting from the peak and ending 90° off the peak. For each point, we derived the mean reference firing rate and the mean test firing rates for 16 direction differences between  $-3^\circ$  and  $+3^\circ$ . At each point on the tuning curve, the mean firing rate  $\mu$  was used to obtain the variance as  $\sigma^2 = 1.28\mu^{1.2}$  (these parameters were determined by non-linear regression of the mean firing rates on the variances for all the 240 neurons). Let  $\alpha = (\mu/\sigma)^2$  and  $\beta = \mu/\alpha$ . Then, the firing rate distribution of this MT neuron sampled at this point on the tuning curve is well approximated<sup>18</sup> by that of the scaled Poisson random variable  $p = \beta\text{Poisson}(\alpha)$ . Using this transformation, 20 trials were generated for each direction-difference, ROC analysis performed, a neurometric function fit, and neurometric threshold estimated. These calculations were repeated for all the 45 points (i.e., in steps of  $2^\circ$  on the tuning curve). The neurometric thresholds were converted into relative precisions using the average psychophysical threshold of the monkeys ( $1.7^\circ$ ). The curve fit to these relative precision values is shown as the grey line in **Fig 3b**.

Next, we directly estimated the average location of the steepest part of the tuning data relative to the preferred direction for all the 240 neurons in the sample and this gave a value of  $67^\circ$ , coinciding with the peak in **Fig. 3b**. Here, we investigate this result in

more detail. We fitted the Gaussian curve  $f(\theta) = r_0 + r_{\max} \exp(-(\theta - \theta_p)^2/2\sigma^2)$  to our direction-tuning data, where  $r_0$  is the background firing rate,  $r_{\max}$  is the maximum firing rate,  $\theta_p$  is the preferred direction and  $\sigma$  is the tuning bandwidth. The average value of tuning width  $\sigma$  for all the 240 neurons in our sample was  $47^\circ$ , close to the estimate reported in a previous study<sup>16</sup> that used stimuli (6 arc min diameter dots, 9 dots/square degree) that were very similar to our own. But this is slightly larger than the width inferred from the data reported in another study<sup>15</sup> that also used RDKs. However, this latter study used fairly larger dots (12 arc min –  $1^\circ$  diameter) and a much denser (40%) dot distribution compared to our study. This study also showed that tuning characteristics of MT neurons depend on the nature of the dot stimuli. Therefore it is reasonable that our result agrees with that of the former study<sup>16</sup> but is not quite close to that of the latter. Where does the steepest part of the Gaussian curve lie corresponding to  $\sigma = 47^\circ$ ? To a second-order approximation, it is easy to show that the location of the maxima of the first derivative of the Gaussian function lie at  $\theta_p \pm \sqrt{2} \sigma$  (or  $\theta_p \pm 1.414 \sigma$ ). Hence, the peak should be around  $66^\circ$  relative to the reference direction. This is very close to the peak in the data shown in Fig 3b as well as to our direct estimate of  $67^\circ$  for the average location of the steepest part of tuning curve relative to the preferred direction from our tuning data. But the exact location of the steepest part of the average tuning curve, calculated numerically, is  $58^\circ$ . All peaks in our data as well as the peak location of the precision deduced from the average tuning curve for the neural population are within  $2^\circ$ - $10^\circ$  of this estimate.

Finally, we used the measured responses at three sample directions closest to the vertical direction to estimate the average rate of change of the response with direction (all three sample directions were  $22.5^\circ$  of each other and within  $0^\circ$ - $40^\circ$  of the vertical direction). We compare these estimates of the slope with the neural precision in **Supp. Fig. 1**. There is substantial variance in the slope for the lowest precision neurons mostly

due to the decreased reliability in the slope estimates (the very shallow slopes make the estimates noisy for these neurons), but as both the slope and the precision become high, the variance decreases. The regression line (shown as the gray line) is statistically significant:  $r = 0.46$ , 99% CI is (0.40, 0.51),  $F=241.8$   $P<0.00001$ .

## Supplementary Note 2

We used a 2-interval task while previous studies used a single-interval task<sup>17,18</sup>. Since the reference presented in the first interval always moved up, it is possible that the monkeys ignored the reference and responded only to the test. If this is true, then the neurometric functions, computed using the deviation of the test direction relative to the reference, could have over-estimated the neurometric thresholds. Three controls rule out this possibility. First, **Fig. 4b** shows statistically significant CPs for the reference firing rates alone, showing that the reference firing rates were significantly correlated with monkeys' choices. This strongly indicates that the monkeys could not have completely ignored the reference stimuli. Second, during some training sessions, the reference and the test presentation order was randomised so that the monkeys learned to respond to direction differences rather than only to the direction of the second stimulus. During recording, we occasionally intermixed trials in which reference direction deviated 5° CW or CCW from the vertical with the normal trials. **Supp. Fig. 2** shows data pooled over several sessions for these control stimuli confirming that the monkeys made relative discriminations decisions. Finally, we re-computed neural thresholds under the assumption that the reference information does not come from the reference stimulus, but from an internalised standard. Instead of computing the neurometric values relative to the reference firing rates, we computed them relative to the rates for the test stimuli that moved in the same direction as the reference. Since the monkey's performance is at 50% when the test direction equals the reference direction, these neurometric functions give

the performance for all other test stimuli relative to this 50% performance level, as if they were computed relative to the internal standard used by the monkey. **Supp. Fig. 3** shows the threshold ratios computed in this manner. These ratios are not different from those in **Fig. 3a** (t-test,  $P > 0.8$ ).

Large fixational eye movements can influence the neural responses and the computed thresholds. But this cannot account for the difference between the two results. We aborted trials in which the eye moved out of a  $0.5^\circ$  square, which is comparable to, if not more stringent than the abort threshold used in other studies.

We used 1sec stimuli durations. Several psychophysical studies and our pilot data showed that fine direction discrimination thresholds continue to decrease with increasing stimulus duration and asymptote near 0.8 sec<sup>23,24</sup>. We chose 1sec duration so that we could compare the best-possible psychophysical performance to the performance limit of single neurons. Longer durations would give undue advantage to the *neuron over the monkey*, by reducing the response variance through longer averaging periods. Next, a recent study showed that in a reaction-time motion detection task, single neurons perform substantially poorer than the monkeys<sup>37</sup>. This study reported reaction times of 0.3-0.7 sec<sup>37</sup>. Suppose the monkey made the decision at 0.7 sec. Then shorter durations than 0.7 sec would put the neurons at a disadvantage compared to the monkeys by insufficiently integrating the neural responses. Therefore, we figured that 1 sec would be about the duration necessary and sufficient to allow a fair comparison of the neural and psychophysical thresholds. In order to rigorously test this hypothesis, we compared neural and psychophysical thresholds as a function of integration time. In one monkey, we measured psychophysical thresholds for exposure durations of 107, 160, 213, 400, 533, 667, 800 and 1000 ms. We took 8 neurons from our sample whose neuron-to-behaviour threshold ratios were about 3 or less and computed neural thresholds for

integration periods equal to the exposure durations listed above. This choice was guided only by the intuition that the less precise neurons will show a shallow relationship between threshold and integration time, making it difficult to discern the exact time at which threshold reached an asymptote. We excluded the latency of response (mean 87 ms, SEM=12) from the integration period. ROC analysis was performed on the firing rates obtained by averaging the spike count over this integration period. Neurometric functions were fit and thresholds were estimated in the same manner described in Methods. The results are shown in **Supp. Fig 4**. Psychophysical threshold value for 667 ms is statistically different from the value for 800 ms (t-test,  $p < 0.01$ ) but the threshold value for 800 ms is not different from that for 1000 ms ( $p > 0.6$ ). A Duncan multiple-range grouping test shows three distinct groups 107-160 ms, 213-667 ms, and 800-1000 ms ( $P = 0.01$ ). Therefore, the psychophysical threshold reaches an asymptote between 667 and 800 ms to a value of about  $1.5^\circ$ . The neural threshold shows two distinct groups, 107-533 ms and 667-1000 ms in the Duncan multiple-range grouping test ( $P = 0.01$ ). Again, the neural threshold value at 533 ms is different from the value at 667 ms (t-test,  $p < 0.01$ ). Hence, neural thresholds also reach an asymptote between 533 and 667 ms. In the asymptotic region, the neuron-to-behaviour threshold ratio is about 3, while at the shortest integration period of 107 ms, this ratio is about 5. Since the temporal dynamics of the psychophysical and neural threshold data are similar, neither the neuron nor the monkey is likely to have had an undue advantage in integration time.

Finally, we note that the use of a short range of direction-differences necessitated the extrapolation of the neurometric functions to estimate the neural thresholds. The nature of the neurometric calculations performed in our study makes it problematic to use a wider range of direction difference. Suppose a neuron is fairly sharply tuned to  $85^\circ$  ( $5^\circ$  CW of the reference). Ideally, the response of this neuron for a  $95^\circ$  direction (CCW of the reference) will be smaller than its response to the  $90^\circ$  direction (reference) while its

response for the  $85^\circ$  direction (CW of the reference) will be greater than its reference response. Therefore, by comparing the response of this neuron for the  $85^\circ$  and the  $95^\circ$  stimuli with the response for the  $90^\circ$  stimulus, one can predict whether the stimulus is CW or CCW deviated from the reference direction. Now consider a stimulus that is moving along a direction CW of  $85^\circ$ , say at  $75^\circ$ . The response of the neuron to this stimulus will be lower than its response to the reference, thereby confusing the sign of the direction-difference. For this neuron, neurometric analysis can only be performed if the range of direction-differences is limited between  $-5^\circ$  and  $+5^\circ$ . If the direction-difference exceeds this range, then the “neurometric function” will start dipping after an initial increasing phase. A probability distribution function (used to fit neurometric data) should be a monotonic non-decreasing function. Hence when one fits a distribution function to such non-monotonic data, one would grossly over-estimate the threshold. **Supp. Fig. 5** shows three sample neurometric functions that illustrate the preceding discussion. The top row shows neurometric functions for the short and long ranges of direction differences for a neuron tuned  $3.5^\circ$  CW of the reference with a full-width at half-height of about  $18^\circ$ . As the range of direction differences is extended, the firing rates for larger CW differences start decreasing. The neurometric function, which is constrained to yield 50% CW answers when the test and reference histograms overlap for a direction difference of  $0^\circ$ , does not fit the data well. There is an over-estimation of the threshold by a factor of about 8. The second neuron is also a sharply tuned neuron ( $\sigma=19$ ) but with a preferred direction of about  $71^\circ$  CW of the reference direction. Therefore it has a fairly shallow slope near the stimuli directions, resulting in a large threshold. But the shallow slope also results in the two estimates of the neurometric threshold being nearly equal. The last neuron has a preferred direction of about  $67^\circ$  CW of the reference, with  $\sigma=43$ . Hence this neuron has the same, fairly steep slope near the stimuli directions for a large range of directions, yielding about the same, fairly low threshold value for both the short and long ranges of direction differences.

If we use a direction-difference range of  $-3^\circ$  to  $+3^\circ$ , then the above problem only affects a small fraction of neurons – those whose preferred directions fall in the range from  $93^\circ$  to  $87^\circ$ . Since the distribution of preferred directions among MT neurons is isotropic<sup>14</sup>, and we obtained an unbiased sample of neurons tuned from  $0^\circ$  to  $180^\circ$ , the *a priori* probability of this problem occurring is  $6/180 = 3.3\%$ , or approximately 6-7 neurons in the sample of 240 neurons. But if we use a broader range of direction-differences, say from  $-16^\circ$  to  $+16^\circ$ , then this problem will occur for neurons whose preferred directions fall within the range from  $106^\circ$  to  $74^\circ$ , i.e., approximately 18% of all neurons sampled, or about 47 out of the 240 neurons. Finally, note that other subtle factors may also influence the threshold estimated using a broad range of measurements. In particular, if the tuning curve were perfectly linear over a broad range near the reference direction then increasing the measurement range would not affect the threshold estimate. But the non-linear shape of the tuning curve, that has a compressive component to it, can itself cause an over-estimation of the threshold when the range of direction differences is extended. Therefore we chose to use a small range of direction differences to obtain as accurate an estimate of the best discrimination ability of single MT neurons as possible.

But the use of a narrow range of direction differences could have affected our neurometric threshold estimates because of the need to extrapolate the neurometric function beyond the measurement range. To evaluate the impact of this extrapolation on the estimated thresholds, in 132/240 neurons we used two ranges of direction-differences, one from  $-3^\circ$  to  $+3^\circ$  and the other from  $-16^\circ$  to  $+16^\circ$ . In order to prevent the monkeys from selectively ignoring stimuli with the smaller direction differences, the direction differences were sampled more frequently near  $0^\circ$  than near the larger values (a logarithmic distribution of differences was used). In addition, the larger direction differences were presented less frequently (we accumulated more trials for the smaller

direction differences than we required). We analysed the data for the two ranges of direction-differences and compared the neurometric and psychometric thresholds. First, **Supp. Fig. 6** shows a comparison of the psychophysical thresholds obtained for the wider range of direction differences with those obtained for the smaller range. Results are shown for 31 sessions in which these 132 neurons were recorded from. Except for the one outlier, psychophysical thresholds for both ranges were about the same. Hence our precautions seem to have prevented the monkeys from slacking off when the large direction-deviation tests were shown randomly interleaved with the smaller direction-deviation tests. Next, **Supp. Fig. 7** shows the ratio of the neurometric threshold estimated using direction-differences in the range from  $-16^\circ$  to  $+16^\circ$  to that estimated using the range from  $-3^\circ$  to  $+3^\circ$ . The peak is near 3 and the average value of this ratio was about 4, indicating that the use of a narrow range of direction differences yielded lower threshold estimates. Finally, we investigated whether the magnitude of over-estimation is reasonable or not, in greater detail, by analysing the tuning properties of the 132 neurons on which we used both ranges of direction-differences. We selected 2 broad groups from these 132 neurons: the “very broadly” tuned neurons (with slope  $< 0.03$ ) and the “very narrowly” tuned neurons (with slope  $> 0.1$ , **Supp. Fig. 1**). The average ratio for the broadly tuned neurons whose preferred directions were within  $(-16^\circ, 16^\circ)$  degrees of the stimuli directions was 0.97. The average ratio for the sharply tuned neurons with the same range of preferred directions was 11.8. The average ratio for the broadly tuned and sharply tuned neurons with preferred directions outside this range was 1.2 and 1.4, respectively. Overall, the two factors explained in the previous paragraph were responsible for the ratio peaking around 3.

### Supplementary Note 3

The action potentials (APs) from individual neurons were isolated offline using the “Offline Sorter” software (Plexon, Inc. TX). The sorting was done by plotting all recorded APs from all the electrodes for the entire duration of the experiment in a 3-dimensional feature-space in which the APs are seen as clouds of points that cluster together depending upon similarities in their shape. The three feature-axes along which the APs were plotted were chosen from a set of 15 features including the first three principal components, the amplitude of the AP at any given slice in time, the peak, the valley and the waveform energy. All combinations of these features taken 3 at a time were tried. The 3 features that gave the highest clustering quality (which is a measure of the average normalized distance between the clusters) were chosen. The clouds of APs were then delineated with each cluster representing one neuron. This method also enabled us to check that a single neuron whose APs were recorded by more than one electrode (channels) was not counted more than once. Since the MAP system has a separate time trigger for each channel, an AP that is simultaneously captured by two or more channels will have a 100% temporal correlation between these channels. Additionally, waveforms captured from the same neuron by two or more electrodes were almost always found to have the same shape (except for amplitude scaling). These two indicators were powerful enough to eliminate the duplication of neurons in our database. Finally, note that the electrode tips were spaced about 250 microns apart, thereby reducing the probability of different electrodes picking up the same neuron.