Measuring response saturation in human MT and MST as a function of motion density

Szonya Durant\* and Michele Furlan

Department of Psychology,

Royal Holloway, University of London

Egham, TW20 0EX, UK

Szonya.durant@rhul.ac.uk

Michele.furlan@rhul.ac.uk

\*corresponding author

Phone: +44 (0)1784 276 522

The human brain areas MT and MST have been studied in great detail using fMRI with regards to their motion processing properties, however to what extent this corresponds with single cell recordings remains to be fully described. Average response over human MT+ has been shown to increase linearly with motion coherence, similar to single cell responses. In response to motion density some single cell data however suggest a rapid saturation. We ask how the combination of these responses is reflected in the population response. We measured the Blood Oxygen Level Dependent (BOLD) response function of MT and MST using a motion density signal, comparing with area V1. We used spatially fixed apertures containing motion stimuli to manipulate the area covered by motion. We found that MT and MST responded above baseline to a very minimal amount of motion and showed a rather flat response to motion density, indicative of saturation. We discuss how this may be related to the size of the receptive fields and inhibitory interactions, although necessarily residual attention effects also need to be considered. We then compared different types of motion and found no difference between coherent and random motion at any motion density, suggesting that when combining response over several motion stimuli covering the visual field, a linear relationship of MT and MST population response as a function of motion coherence might not hold.

Keywords: fMRI, motion density, motion coherence, response saturation

## Introduction

The primate visual cortex contains several defined visual areas which have a greater or lesser specialization for motion processing. Among these areas, a primary role in visual motion processing is played by the posterior middle temporal gyrus (MT) and the medial superior temporal cortex (MST). MT has been shown to contain a preponderance of neurons that are strongly sensitive to the direction of image motion ([Allman & Kaas, 1971](#_ENREF_1); [Dubner & Zeki, 1971](#_ENREF_7)). Like MT, MST has been show to contain neurons tuned for direction, which are located mostly in its ventral part ([Desimone & Ungerleider, 1986](#_ENREF_6)), and neurons that are tuned for more complex “optic flow” patterns, such as the expansion caused by forward motion, which are located mostly in its dorsal part ([Duffy & Wurtz, 1995](#_ENREF_8)). Because MT and MST share so many functional similarities, often they are referred to as one unique large motion-sensitive complex called MT+.

The homologues of MT+ have been found in humans ([Tootell et al., 1998](#_ENREF_21); [Zeki et al., 1991](#_ENREF_23)). Human MT has been shown to be sensitive to motion direction and motion coherence ([Braddick et al., 2001](#_ENREF_3); [Rees, Friston, & Koch, 2000](#_ENREF_17)), while MST has been shown to be sensitive to more complex motion components, such as optic flow ([Smith, Wall, Williams, & Singh, 2006](#_ENREF_19)). This functional specificity makes MT and MST two ideal candidates for comprehending how single neuron activity gives rise to population activity and the overall perception of motion.

Although non-human primate and human MT+ have been shown to share functional similarities, the parallel between the two complexes is still not fully established. Duffy and Wurtz ([1995](#_ENREF_8)) used motion patterns formed of randomly placed dots in order to investigate the response of neurons in monkey MST, and they found that increasing the dot density did not increase the response of individual neurons, suggesting that MST may reach a saturation point with just a weak motion signal. A similar result has been observed for monkey MT as well, when Snowden et al. ([1991](#_ENREF_20)) measured the response of MT for different levels of dot density, and found that the saturation of monkey MT cells usually occurred with just a few dots. This kind of saturation was not observed when MT and MST responses were measured as a function of motion coherence, i.e. the ratio between the number of dots moving in the same direction (signal) and the number of dots moving randomly (noise). The neural response in monkey MT has been shown to increase linearly with the amount of motion coherence ([Britten, Shadlen, Newsome, & Movshon, 1993](#_ENREF_5)), although interestingly this study also shows a large amount of heterogeneity in the amount of response to random motion, with most neurons showing a moderate response.

Several studies have also documented the sensitivity of human MT+ to motion coherence ([Braddick et al., 2001](#_ENREF_3); [Rees et al., 2000](#_ENREF_17)). Particularly interesting is the study of Rees et al. ([2000](#_ENREF_17)), where the authors presented a random dot kinematogram in which dots moved either in a common direction (signal) or in different directions (noise), while they measured the BOLD response of MT+ for different levels of motion coherence. Their results showed a clear linear relationship between the BOLD signal and motion coherence (across the full motion coherence range from random motion to fully coherent motion), a result that was consistent with a simple summation of the activity found in neurophysiological results ([Britten et al., 1993](#_ENREF_5)) and therefore suggested that MT+ in monkey and human brains are comparable. However, this result is somewhat controversial. Some studies since have failed to find a clear difference between BOLD responses to coherent and uniform signals in human MT+ (Smith et al. 2006; McKeefrey et al. 2007), whilst some have suggested that this relationship is dependent on the attention paid to the stimulus (Kayser et al., 1997).

These results combined together raise an important question, which is how does the response in MT+ scale with the amount of pure motion signal, i.e. signal not embedded in noise? In other words, how does its response scale with coherent motion per se, i.e. how much information is necessary to extract coherent motion in a setting where the amount of coherent motion does not scale inversely with the amount of noise? Recent work using real-life movie clips that have been manipulated to reduce the amount of stimulus has suggested that with the amount of dynamic visual information there is a very rapid saturation of response ([Durant, Wall, & Zanker, 2011](#_ENREF_10)), but these natural image stimuli were not well controlled as to the amount of contrast, and the amount of and type of motion information available.

The aim of this study is to describe fully how the saturation function of human MT and MST varies with motion density. To this aim, we initially presented a series of circular apertures filled with coherently moving dots. We used this configuration in order to present motion that was spatially localized and thus we could be sure that smaller amounts of motion were stimulating a subset of the same neurons responding to larger amounts of motion and we could measure the response unambiguously to the area of visual field covered by motion in terms of number of apertures. All the dots shown in the visual field were moving either leftward or rightward. We then measured the hemodynamic response of MT and MST while varying the number of apertures. Initially we varied the amount of available uniform motion signal without varying motion coherence. We also measured the haemodynamic response in V1, which we compared with the results found in MT and MST. This stimulus then allowed us in a further experiment to consider the effect of different levels of motion coherence at different levels of motion density as our previous results suggested that there was no difference in activity caused by locally and globally coherent stimuli in real life dynamic stimuli ([Durant et al., 2011](#_ENREF_10)), so we compared these and also completely random motion.

As single cell responses saturate rapidly with motion density, a simple summation over the neurons should result in a similar linear dependence on the motion signal as observed with motion coherence. We investigate whether and how response saturation and other combination rules over visual space may play a role in shaping the population response.

## Materials and Methods

### Experiment 1

#### 2.1.1 Data acquisition

Data were acquired using a 3T Siemens Trio MR scanner with a 32 channel array head coil. Functional images were acquired with a *T2\**-weighted gradient-recalled echo-planar imaging (EPI) sequence (35 axial slices, TR 2500ms, TE 31ms, flip angle 85°, resolution 3 mm isotropic, echo-spacing: 1.42ms). Structural data were acquired using a *T1*-weighted 3D anatomical scan (MPRAGE, Siemens, TR 1830 ms, TE 5.56 ms, flip angle 11°, resolution 1x1x1).

#### 2.1.2 Functional localizers

We acquired two functional localizers which provided us with the regions of interest needed for the analysis of the data obtained from the main experiments (MT, MST, V1).

Human MT and MST were defined based on a standard method ([Dukelow et al., 2001](#_ENREF_9); [Huk, Dougherty, & Heeger, 2002](#_ENREF_12)). A circular patch of dots (8° diameter) was presented with its centre placed 10° to the left or right of ﬁxation. Blocks of 15 s in which the dots were static were alternated with blocks of 15 s in which the dots moved alternately inward and outward along the radial axes, creating alternating contraction and expansion. Sixteen blocks (8 static and 8 moving) were presented in each scan run; 1 run was completed with the stimulus on the left and another with it on the right. This procedure allowed us to differentiate two regions. MT was defined as the set of contiguous voxels that were contiguously active during contralateral stimulation alone, while MST was defined by the voxels that were responding to both ipsilateral and contralateral stimulation ([Smith et al., 2006](#_ENREF_19)). In addition, since previous research ([Huk et al., 2002](#_ENREF_12); [Smith et al., 2006](#_ENREF_19)) has shown that the centre of MST is located anterior to the centre of MT, any MT voxels situated further anterior than the median value of the MST ROI on the horizontal (axial) plane were removed from the MT ROI. Finally, it is likely that “MST” as we defined it comprises 2 or more regions that respond to motion and have large receptive ﬁelds, but further reﬁnement requires demanding high-resolution mapping techniques ([Amano, Wandell, & Dumoulin, 2009](#_ENREF_2); [Kolster, Peeters, & Orban, 2010](#_ENREF_15)) that are beyond the scope of this study. Although the localizers are spatially restricted this method captures the whole of MT and MST.

The primary visual cortex (V1) was identified by standard retinotopic mapping procedure ([Sereno et al., 1995](#_ENREF_18)), involving a 8Hz counter-phasing checkerboard wedge stimulus (24° sector) of radius 12°. Check size was scaled by eccentricity in approximate accordance with the cortical magniﬁcation factor. The wedge rotated clockwise at a rate of 64 s/cycle, and 8 cycles were presented. This stimulus was presented twice to each participant, and the data from the 2 scan runs were averaged to give the final retinotopic maps. The temporal phase of the response to the rotating wedge at each voxel was obtained by ﬁtting a model to the time-series. Phases were superimposed as colors on a segmented and ﬂattened representation of the grey matter. Phase was taken as an indicator of visual ﬁeld position in terms of polar angle and the boundary of V1 was drawn by eye using conventional criteria.

#### 2.1.3 Participants

Seven healthy volunteers (4 females) took part in this experiment. The authors took part as well, but apart from the first author all other participants were naïve to the purpose of the experiment (the second author was not involved in the design of the experiment). All the participants had normal or corrected to normal vision. They were screened for MRI contraindications according to standard procedures and written consent was obtained. The experimental procedure was in accord with the Declaration of Helsinki and was approved by the appropriate local ethics committee.

#### 2.1.4 Stimuli and task

The stimuli were projected via Sanyo LCD projector at a refresh rate of 60Hz onto a rear-projection screen at the end of the scanner bore and were viewed via a mirror mounted on the head coil giving an image of 31.6° × 45.2° visual angle. The stimulus consisted of a field of white (90 cd/m2) dots presented on a grey background (44 cd/m2), 16 dots per degree squared. In between trials all dots were static and ISIs were of a random duration between 2s and 13s, mean 8s. The area covered by the stimulus was 24.5° × 33.9°, and the space between the edge of the stimulus and the end of the visual field was filled with grey background. On each trial the dots were positioned within areas defined by a set of notional circular apertures (diameter 0.8°), randomly positioned on the screen. The dots would move left or right at 2.5 deg/sfor 2s. Each aperture showed on average 3 moving dots (see Figure 1). The dots wrapped around within the aperture area and ended up in the same positions as initially. The number of circular areas containing dots could be 1, 2, 3, 6, 12, 64 or 128. All apertures contained the same type of uniform motion, and motion direction and number of apertures were randomized within each run and varied trial by trial. The random position of these apertures implies that the same proportion were spatially coincident with the MT+ localizer positions for each condition over all the trials for that condition. The single aperture appeared half of the time on the left visual field and half of the time on the right one. Therefore, the single aperture stimulated either the right or the left MT+ respectively. Averaging over all instances and sides provided us with a good estimate of the mean BOLD response in MT+ in the single aperture condition over all possible locations of the aperture (and the same is true in the conditions with more apertures). The 7 conditions per run each repeated 6 times, each run lasting 7m17s, with 6 runs in total. Each run started with the presentation of a central dark grey fixation spot that participants were required to fixate throughout the experiment and count how many times it flashed blue. Responses were recorded for double checking that participants were attending to the fixation point. After 10 seconds, the first trial was shown.

|  |
| --- |
| * Place Figure 1 approx. here -
 |

#### 2.1.5 Data analysis

Data were analyzed using BrainVoyager QX 1.4 (Brain Innovation, The Netherlands). The ﬁrst 4 volumes of each run were discarded. Three-dimensional motion correction and slice time correction were performed. The data were temporally high pass ﬁltered at 3 cycles/run (∼0.01 Hz). The preprocessed EPI scans were then coregistered with the anatomy. Finally, both functional and anatomical data were aligned into AC-PC space. The preprocessed data were analysed within the general linear model (GLM), separately for each participant. For the main experiment, each motion condition was modelled separately, with a regressor formed by convolving the stimulus time-course with a canonical hemodynamic impulse response function and then scaling to unity. Head motion regressors were also included. Having identified the regions of interest (ROIs), the strength of the response to motion signal in each region was assessed by calculating the mean activation averaged across all the voxels falling within each ROIs, which was expressed as percentage signal change (PSC).

### 2.2 Experiment 2

#### 2.3.1 Participants

Seven healthy volunteers (5 females) took part in this experiment, and as in the first experiment the two authors were included in the sample, again all but one of the authors were naïve to the aims of the study at the time of being scanned.

#### 2.3.2 Stimuli and task

Stimuli were the same as in Experiment 1, except for the following two differences. The first difference was that there were three types of motion condition. The first type was random motion, in which each moving dot had a randomly chosen direction. The second type was locally coherent motion, in which dots inside each aperture were moving coherently in one direction, but this direction was randomly chosen for each aperture. The third type was global coherent motion, in which all the moving dots shown were moving coherently in one direction (randomly chosen on each trial). The second difference was the number of apertures, which were either 3 or 64. The aperture positions shown were the same across conditions. Each of the 6 conditions [motion (3) x density (2)] was repeated 6 times in each run, which lasted 6m26s.

#### 2.3.3 Data acquisition and analysis

Data were acquired and analysed as in Experiment 1. The analysis was run on the same regions of interests as in Experiment 1.

## Results

### -

### 3.1 Experiment 1 – Response as a function of number of moving apertures

- Place Figure 2 approx here -

We first checked performance on the attention task and found that all participants attended to fixation, with an average error of 12 blues counted on a range of number of blues that varied from 42 to 149. The response amplitudes were averaged across hemispheres for each participant and the mean over participants is plotted in Figure 2. There is some variability in amplitudes of BOLD responses across participants as shown by the error bars in Figure 2, but we can use repeated measures statistical tests to look for common patterns across conditions. Stimulus density affected the hemodynamic responses in all the ROIs, with a more dense stimulus producing a higher response, although in MST this is not quite significant (repeated measures ANOVA on percentage signal change main effect of “Density”, V1: F(6,36) = 16.1, p < 0.0005; MT: F(6,36) = 2.47, p < 0.05; MST: F(6,36)=2.53, p < 0.05 – however the MST data failed the test for sphericity and with Greenhouse-Geisler correction p=0.13). Although motion density affected area V1, the motion stimuli did not produce any response above baseline in V1 until we used 64 apertures, with a significant increase from the smallest to largest number of apertures (planned simple contrast: 1 vs 128 apertures F1.6=20.6, p<0.005, post-hoc univariate t-test 1 aperture against zero response t6=-2.18, p=0.07). In fact at low a number of apertures there appears to be a negative BOLD response, although this is not significant. By contrast, both area MT and MST showed a significant difference from the baseline even when a single aperture was presented and no significant difference between the smallest and largest number of apertures (MT planned simple contrast: 1 vs 128 apertures F1.6=3.11, p=0.13, post-hoc univariate t-test 1 aperture against zero response t6=5.56, p<0.005; MST planned simple contrast: 1 vs 128 apertures F1.6=0.40, p=0.55, post-hoc univariate t-test 1 aperture against zero response t6=3.98, p<0.01).

Whereas the increase in V1 response observed in the current experiment showed a clearly linear pattern (note – this shows as an exponential function on the log axes in Figure 2), the results shown by MT and MST appeared somewhat different. Indeed, both the regions showed a rather flat response as a function of stimulus density, suggesting signs of saturation. We fit the data with a simple saturating function (see Figure 2), taking the following form: *Rpop = kn/*(*c+n*), where *Rpop* is the overall population response, which is dependent on *n* the number of apertures stimulating the neural population, *c* is the semi-saturation constant and *k* is a scaling factor. This function implies a zero response difference when no apertures are present as is necessary, but as a function of number of apertures effectively steps up from 0 to 1 to near maximum response. Comparing the lowest and the highest levels of density in both MT and MST, even using one aperture produced a response which was 76% of the response observed using 128 apertures in the former and 88% in latter, making it difficult to characterize the saturation function of response in these areas, as it is already very close to saturating at the first point sampled. This flat response was not due to a lack of response as on average the response is in fact larger than in V1, where a marked increase can be seen. It cannot be due to visual contrast as this remains the same across all conditions.

### Experiment 2 - Effect of number of apertures under different motion conditions

* Place Figure 3 approx. here -

We first checked performance on the attention task and found that all participants attended to fixation, with an average error of 5 blues counted on a range of number of blues that varied from 38 to 138. In the final experiment we set out to measure how the pattern between responses to low and high density motion stimuli varied across different types of motion, extending beyond the simple uniform motion in Experiment 1. We chose two points from the rising end and saturated end of the curves, 3 and 64 apertures, where the biggest difference in response was observed above (in fact a bigger difference than between 1 and 128 apertures). Figure 3 shows the mean response amplitude under different motion conditions for each cortical region separately with the response amplitudes averaged across hemispheres and participants. Again, there is some variability in amplitudes of BOLD responses across participants as shown by the error bars in Figure 3, but we can again use repeated measures statistical tests to look for common patterns across conditions. Stimulus density produced an increase of the hemodynamic response in all the regions, showing that the response is not actually fully saturated at 3 apertures (main effect of “Density”, V1: F(1,6) = 54.077, p < 0.0005; MT: F(1,6) = 17.233, p < 0.01; MST: F(1,6) = 39.049, p < 0.005). Quite surprisingly, varying the type of motion produced no significant change in response in any of the brain regions (main effect of “Motion”, V1: F(2,12) = 0.127, p = 0.882; MT: F(2,12) = 0.435, p =0.657; MST: F(2,12) = 0.018, p = 0.982). Different types of motion didn’t change the pattern of responses over 3 and 64 apertures (no significant interaction, V1: F(2,12) = 0.166, p = 0.849; MT: F(2,12) = 1.194, p =0.337; MST: F(2,12) = 1.522, p = 0.258), suggesting that the pattern of low to high density does not differ across different types of motion. If anything there appears to be slight trend of increased response to random motion in areas MT and MST with three apertures, but this is not significant.

## Discussion

Human brain regions MT and MST are well known to process motion information, particularly being sensitive to direction ([Zeki et al., 1991](#_ENREF_23)) and global motion ([Braddick et al., 2001](#_ENREF_3)). Although the functional selectivity of these two regions has been well documented, their response functions are still unclear, i.e. how their hemodynamic responses changes with stimulus intensity (in this case the amount of motion) and when their responses reach saturation point. Most of the studies so far have tried to describe MT and MST response as a function of motion coherence, a specific feature produced by manipulating the amount of motion signal embedded in noise. Whilst monkey evidence suggests that MT and MST are well described by a linear response to coherence (Britten et al. 1993), there has been some disagreement in the human fMRI literature: Rees et al. ([2000](#_ENREF_17)) have shown a linear increase as a function of coherence, whereas Smith et al. ([2006](#_ENREF_19)) found both MT and MST to be just as highly responsive to random motion as uniform. It may be the case that if MT and MST are almost equally driven by motion noise and motion signal, then it may require sensitive methods to detect any increasing effect. This suggested that a simpler way to investigate the saturation function of these two areas could be by using a coherent motion signal without noise while varying stimulus density. This also allowed a comparison between random and uniform motion at lower density levels, which may not result in such a high response.

To this aim, we set out first of all to measure how pure motion density, i.e. the amount of coherent global motion in the scene, affects the hemodynamic response in human motion processing areas MT and MST in comparison with primary striate cortex V1. We did this by increasing the number of localized motion patches in order to keep local motion density constant and restrict the visual area stimulated, so increased density would result in increased visual area, requiring combination of information across the visual scene. In V1 we observed a linear pattern of responses to the amount of motion, showing that motion contrast has a similar effect here to luminance contrast (Tootell et al. 1998). In MT and MST, however we observe a much flatter response, indicative of a rapid saturation with the amount of motion in these areas.

The pattern of activity here does not resemble a simple summation over neuronal response, where an increasing area containing motion would stimulate more and more neurons. It may be possible that most neurons are already fully saturated with the very small amount of motion covered by 1 aperture, but this would suggest very large receptive field sizes covering must of the visual field, not consistent with estimates of RF size of around 5° of visual angle in monkey MT ([Felleman & Kaas, 1984](#_ENREF_11)) and retinotopic maps in MT ([Kolster et al., 2010](#_ENREF_15)). There is some evidence from single-cell studies for inhibitory surrounds and divisive normalisation in area MT+, which would also lead to such a flat pattern of activity ([Britten & Heuer, 1999](#_ENREF_4); [Xiao, Raiguel, Marcar, & Orban, 1997](#_ENREF_22)). Area MST saturates somewhat earlier than area MT (shown by a flatter response), possibly reflecting the larger receptive field sizes in this area, that are hence more likely to respond to a small area of stimulus but also hence more likely to begin to suppress each other. This would suggest that the BOLD signal shows a relationship between receptive field size and the response saturation of these areas to a preferred stimulus. Our data is not able to distinguish between these two possibilities but on balance it seems more likely that the constant response across a wide range of motion densities is caused by inhibition between neurons, rather than a simple summation of responses.

Whilst we did attempt to rule out the effects of attention by using a central fixation task, even with a more demanding task it is impossible to rule out that the motion stimulus captured attention to some degree, as such making direct comparisons to anaesthetised monkey data tricky. To this extent when we are talking about response we mean the combined response due to the stimulus and residual attention from the task. The effect of attention on primate (including human) MT+ has been extensively studied (O'Craven, Rosen, Kwong, Treisman & Savoy, 1997; Treue & Maunsell, 1999; Martinez-Trujillo & Treue 2002). As we jittered stimulus onset times and there were always some dots present, we have ruled out the effects of a baseline shift in attention (i.e. a fixed additive increase in response to any given visual stimulus), as has been found in the case of induced stimulus expectation ([Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999](#_ENREF_13)). If, however the onset of the motion attracted attention to the stimulus then this could have the effect of increasing response and hence shifting of the response curve to the left. This effect of attention in MT+ has been shown for the contrast response function in macaques (Martinez-Truijllo & Treue, 2002) and we cannot rule out similar effects for a human motion density response. In this case it is possible that many apertures would cause response to saturate, masking this effect of attention, which could lead to a relative increase in response to the small amount of apertures due to attention. However, we would contend that it is still surprising that the bottom-up response to such a small amount of stimulus combined with the effects of the residual attention not demanded by the task brings MT/MST response close to saturation. Moreover, robust effects of attention have been previously also found in V1 (Somers, Dale, Seiffert & Tootell, 1999), yet here we find no evidence for such a boosting of the motion induced signal. Similarly we cannot rule out whether the cause of this may be higher level stimulus effects such as some sort of perceptual filling in or inductive self-motion effect, but clues to these processes may lie in this response that we have observed in MT/MST fMRI response. It would be interesting for instance to investigate in future work to what extent the BOLD response corresponds with a behavioural measure of a global sense of motion.

In our final experiment we find no difference in responses to the level of coherence of the motion stimulus, at low or high motion stimulus densities. Random stimuli elicit a significant response even with only three apertures. If anything there is a slight tendency with the smaller amount of apertures for there to be an increase in response to random motion, although this is not significant, but suggests we would not have found a decrease with more participants. The difference between our findings and those of Rees et al. (2000) may be down to attention differences, although Kayser et al. ([1997](#_ENREF_14)) suggest that motion coherence should have more of an effect when attention is directed away from the stimulus as in our task. This is not the only example in the literature where this lack of difference is found ([e.g., Smith et al., 2006](#_ENREF_19)) and it has even been found that incoherent motion produces a larger activation than coherent ([McKeefry, Watson, Frackowiak, Fong, & Zeki, 1997](#_ENREF_16)). We controlled fixation with the central fixation task, and there is no reason why different number of apertures should give rise to a different number of eye movements. It may be that these differences are caused by the difference between measuring responses based on a few spatially co-localized receptive fields, versus the combined response from several spatially separate receptive fields. In Reese et al. (2000), the response was measured to small 2° diameter apertures effectively in isolation. Although our apertures were smaller (0.8° diameter), even when we presented only three of them, they will have activated more spatially separate receptive fields, allowing for interactions between these to be observed and as shown, response saturation began to take place around the 3-6 aperture mark.

In conclusion, we have found that the motion specific areas MT and MST respond at close to their maximum response to a very small amount of motion stimulus and when motion stimuli are contrasted against static, we observe a very rapid saturation response, different from the linear increase found with motion coherence. The differences in saturation between the motion areas may reflect the larger receptive field sizes in MST and as such serve as a useful characterization. We find no differences in the pattern of response across low and high aperture numbers over three different types of motion from fully random, locally uniform, globally random to fully uniform, further suggesting population response is not simply a summing up of responses seen in single cell studies. Single cell effects of motion coherence do not necessarily emerge at the population level, when responses are a result of combining information over a large area of visual space.

## Acknowledgements

The authors would like to thank Ari Lingeswaran for his help with the MRI data acquisition, Jaclyn Billington for providing help with data analysis and Johannes Zanker, Inci Ayhan and Andy Smith for useful discussions.

## References

Allman, J. M., & Kaas, J. H. (1971). A representation of the visual field in the caudal third of the middle tempral gyrus of the owl monkey (Aotus trivirgatus). *Brain Research, 31*(1), 85-105. doi: 10.1016/0006-8993(71)90635-4

Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology, 102*(5), 2704-2718. doi: 10.1152/jn.00102.2009

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception, 30*(1), 61-72. doi: 10.1068/p3048

Britten, K. H., & Heuer, H. W. (1999). Spatial summation in the receptive fields of MT neurons. *Journal of Neuroscience, 19*(12), 5074-5084.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience, 10*(6), 1157-1169. doi: 10.1017/S0952523800010269

Desimone, R., & Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *Journal of Comparative Neurology, 248*(2), 164-189. doi: 10.1002/cne.902480203

Dubner, R., & Zeki, S. M. (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research, 35*(2), 528-532.

Duffy, C. J., & Wurtz, R. H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience, 15*(7 Pt 2), 5192-5208.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology, 86*(4), 1991-2000.

Durant, S., Wall, M. B., & Zanker, J. M. (2011). Manipulating the content of dynamic natural scenes to characterize response in human MT/MST. *Journal Of Vision, 11*(10). doi: 10.1167/11.10.5

Felleman, D. J., & Kaas, J. H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys.*, 52*(1), 488–513. doi: 10.1016/0306-4522(83)90024-6

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience, 22*(16), 7195-7205.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751-761. doi: 10.1016/S0896-6273(00)80734-5

Kayser, A. S., Erickson, D. T., Buchsbaum, B. R., & D'Esposito, M. (1997). Neural representations of relevant and irrelevant features in perceptual decision making. *Journal of Neuroscience, 30*(47), 15778-15789. doi: 10.1523/JNEUROSCI.3163-10.2010

Kolster, H., Peeters, R., & Orban, G. A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *Journal of Neuroscience, 30*(29), 9801-9820. doi: 10.1523/JNEUROSCI.2069-10.2010

Martınez-Trujillo, J. C., & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, 35(2), 365-370. doi: 10.1016/S0896-6273(02)00778-X

McKeefry, D. J., Watson, J. D., Frackowiak, R. S., Fong, K., & Zeki, S. (1997). The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *Neuroimage, 5*(1), 1-12. doi: 10.1006/nimg.1996.0246

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT–MST. *Neuron*, 18(4), 591-598. doi: 10.1016/S0896-6273(00)80300-1

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience, 3*(7), 716-723. doi: 10.1038/76673

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*(5212), 889-893. doi: 10.1126/science.7754376

Smith, A. T., Wall, M. B., Williams, A. L., & Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience, 23*(2), 561-569. doi: 10.1111/j.1460-9568.2005.04526.x

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience, 11*(9), 2768-2785.

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences*, 96(4), 1663-1668.

Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences, 95*(3), 811-817.

Treue, S., & Maunsell, J. H. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience*, 19(17), 7591-7602.

Xiao, D. K., Raiguel, S., Marcar, V., & Orban, G. A. (1997). The spatial distribution of the antagonistic surround of MT/V5 neurons. *Cerebral Cortex, 7*(7), 662-677. doi: 10.1093/cercor/7.7.662

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience, 11*(3), 641-649.

## Figure legends

Figure 1. Illustration of the stimulus in Experiment 1. Arrows indicate direction of motion of the dots. Dotted lines show limits of the apertures, but did not appear on the screen.

Figure 2. Mean percentage BOLD signal change across participants for moving dots contrasted against static dots as a function of the number of apertures. See section 3.1 for details. V1 is fit by a linear function and MT and MST are fit by a saturating function. Note: x values are plotted on a log axis. Error bars are the standard error of means over participants, N=7.

Figure 3. Mean percentage signal change as a function of different types of motion and number of apertures. See section 3.2 for details. Error bars are the standard error of means over participants, N=7.