Selectivity for speed gradients in human area MT/V5

Julio C. Martinez-Trujillo,[1,CA] John K. Tsotsos,[2,3] Evgueni Simine,[2,3] Marc Pomplun,[4] Richard Wildes,[2,3] Stefan Treue,[5] Hans-Jochen Heinze[6] and Jens-Max Hopf[6]

[1]Department of Physiology, McGill University, Montreal, Canada; [2]Centre for Vision Research, York University, Toronto, Ontario, Canada; [3]Department of Computer Science York University, Toronto, Ontario, Canada; [4]Department of Computer Science, University of Massachusetts, Boston, Massachusetts, USA; [5]Cognitive Neuroscience Laboratory, German Primate Center, Goettingen, Germany; [6]Department of Neurology II, Otto-von-Guericke University, Magdeburg, Germany

[CA]Corresponding Author: julio.martinez@mcgill.ca

Received 23 November 2004; accepted 18 January 2005

Cortical area MT/V5 in the human occipito-temporal cortex is activated by visual motion. In this study, we use functional imaging to demonstrate that a subregion of MT/V5 is more strongly activated by unidirectional motion with speed gradients than by other motion patterns. Our results suggest that like the monkey homolog middle temporal area (MT), human MT/V5 contains neurons selective for the processing of speed gradients. Such neurons may constitute an intermediate stage of processing between neurons selective for the average speed of unidirectional motion and neurons selective for different combinations of speed gradient and different motion directions such as expanding optical flow patterns. *NeuroReport* 16:435–438 © 2005 Lippincott Williams & Wilkins.

Key words: Functional magnetic resonance imaging; Motion processing; Speed gradient; Speed selectivity; V5/MT

INTRODUCTION

Human area MT/V5, in the ascending limb of the inferior temporal sulcus, is specialized for the processing of visual motion [1–3] and it is considered to be the homolog of middle temporal area (MT) in the monkey [1]. Previous studies in monkeys have reported that some MT neurons are selective for the orientation of speed gradients in unidirectionally moving random dot patterns (RDPs) [4,5]. In these studies, neurons give different responses to RDPs moving in the same direction and at the same average speed, depending on whether the dots accelerate, move at a constant speed or decelerate. Currently, there is no evidence for the existence of such neurons in humans.

In the present study, we tested the hypothesis that human area MT/V5 contains neurons selective for unidirectional motion with speed gradients. We compared the blood oxygenation-level-dependent (BOLD) magnetic resonance imaging signal evoked by moving RDPs with speed gradients and other motion types in five human study participants. As a control, we determined the levels of activation evoked by the same stimuli within area V1.

MATERIALS AND METHODS

Study participants: Five healthy volunteers with normal or corrected vision (age 25–40 years) gave their informed consent and participated in the experiments. All the procedures were approved by the Ethics Committee of the Otto-von-Guericke University of Magdeburg, Germany.

Stimuli: The stimuli were RDPs [black dots (luminance=0.1 cd/m^2) against a white background (luminance=56 cd/m^2) within a virtual aperture of 45 deg^2)] generated on an Apple PC using custom-made software. They were displayed using an LCD projector on a translucent screen 30 cm away from the participants' eyes. The dots' lifetime was infinite and the dots' density was 10 dots/deg^2. Dots crossing one edge of the aperture reappeared at a new location within the aperture. The monitor refresh rate was 75 Hz.

Five different stimuli were used (Fig. 1a): (1) *Stationary* (S) were two RDPs with dot densities of 7 and 13 dots/deg^2 (average density=10 dots/deg^2). (2) *Random motion* (R) contained dots moving in different directions and at the same speed. Two patterns with speeds of 3.22°/s and 5.98°/s were used. (3) *Translation* (T) contained dots moving at the same speed and direction. Eight patterns were used (four cardinal directions × two different speeds: 3.22°/s and 5.98°/s). (4) *Gradient motion* (G) was similar to *translation* but the dots accelerated according to the formula (speed=k × distance from the border of the aperture in degrees). Eight patterns matching the direction and average speed of the *translation* patterns were used. (5) *Expansion* (E) contained dots moving radially. The speed of a dot was k × distance from the center of the aperture. Two patterns with average speeds of 3.22°/s and 5.98°/s were used. The shape of the aperture was square for *gradient motion* and circular for the other stimuli in order to simplify matching the average speed of the former to that of the other patterns.

Task: In a given trial, two RDPs were presented on both sides of a fixation cross (eccentricity=7.6°) for 500 ms. Using a two-alternative forced-choice design, participants reported which pattern moved faster. In trials containing *stationary* patterns, they reported which pattern contained more dots. Participants were instructed to maintain fixation during the trials. Although we did not measure eye movements, we tested two participants outside the scanner while monitoring eye movements using a video camera and they did not make saccades or apparent displacements in eye position toward either of the stimuli. They also reported that because of the short stimulus presentation time and bilateral visual field stimulation, it was not advantageous to look toward the stimuli during the trials.

The participants' performance was better than 85% correct with no significant difference between the different motion types, ruling out task difficulty as a source of difference in activation evoked by the different stimuli [6].

Functional magnetic resonance imaging measurements: A total of 309 images were acquired during each of the five runs on a neuro-optimized GE Signa LX 1.5 T system using a 5-inch surface coil beneath the participant's occipital pole. Functional images extended anteriorly from the occipital pole (23 contiguous coronal slices) and included posterior parts of the temporal and parietal lobe (Gradient-Echo-EPI sequence, slice thickness 3 mm, FOV 18 cm, matrix size 64 × 64, TR 2000 ms, TE 40 ms, flip angle 80°). Functional data were normalized to the Montreal Neurological Institute (MNI) template using a separate EPI image that covered about two-thirds of the whole brain. To superimpose the normalized statistical parametric maps (SPMs) onto individual anatomical structures, a high-resolution anatomical image (T1-weighted gradient-echo sequence, 124 slices, slice thickness 1.5 mm, FOV 25 cm, matrix size 256 × 256, TR 24 ms, TE 8 ms, flip angle 30°) was acquired for each participant and normalized to the MNI template.

Data analysis was conducted using SPM99 (Welcome Department of Cognitive Neurology, University College London, UK). First, functional volumes were phase shifted in time with reference to the first slice. Second, head-movement artifacts were corrected on the basis of an affine rigid body transformation with reference to the first image of the first run. Third, volumes were normalized to the MNI template, resliced to $2 \times 2 \times 2\,mm^3$ voxels, and spatially smoothed with a Gaussian kernel of 6 mm.

Statistical analysis was performed using a modeled hemodynamic response function for each experimental condition. Significant differences in hemodynamic responses were validated using the general linear model approach. For visualization, SPMs were superimposed onto the normalized high-resolution anatomical images. To quantify the observed effects, %-BOLD-signal change was estimated in regions of interest containing above threshold voxels ($p < 0.001$, corrected).

RESULTS

The middle panel in Fig. 1b shows an SPM superimposed on an axial anatomical slice of one participant. This shows the voxels with significant differences in the strength of the BOLD signal when contrasting the different motion types against *stationary* ($p < 0.001$). The positions of the local maxima for area MT/V5 (see the figure legend) are in agreement with previous functional magnetic resonance imaging studies [1–3,7–10].

The solid curves in the lateral panels of Fig. 1b show the average BOLD signal change corresponding to clusters of voxels activated by each motion type relative to *stationary* in a single participant. The five data points within each curve represent the average signal evoked by the different stimuli (S, R, T, G, E) within the corresponding group of voxels. Because there was little variance between the voxels activated by each motion type relative to *stationary* in both areas (MT/V5 and V1), we averaged across curves to obtain an estimate of the activation evoked by the different stimuli within each area (white dashed line). In both areas, *stationary* evoked the lowest activation, *random motion* evoked the strongest signal in V1 and *gradient motion* in MT/V5.

We repeated the same analysis in each participant. In all cases we found voxels within MT/V5 and V1 with a significantly stronger signal change evoked by each of the four motion types relative to *stationary* in both hemispheres ($p < 0.001$). Figure 2a shows averaged data. Each of these curves represents the signal change (ordinate) evoked by the different stimuli (S, R, T, G, E) in voxels showing significant activation for a given motion-defining contrast (i.e. R–S, T–S, G–S, E–S) across participants.

We performed a two-factor ANOVA for repeated measurements (i.e. each motion-defining contrast was considered a measurement) for each area using hemisphere and stimulus type as factors. In MT/V5 we found no effect of hemisphere ($p = 0.31$) but a strong effect of stimulus type ($p < 0.0001$). In V1 we also found no effect of hemisphere ($p = 0.36$) but a strong effect of stimulus type ($p < 0.0001$).

Next, we compared the patterns of activation between MT/V5 and V1 by performing a two-factor ANOVA for repeated measurements using area and stimulus type as factors. We found a strong effect of both factors ($p < 0.0001$). The strongest activation in MT/V5 was evoked by *gradient motion* [$p < 0.0001$ in all comparisons (G vs. S, R, T, E), paired *t*-test after Bonferroni correction]. On the other hand, the strongest activation in V1 was evoked by *random motion* [$p < 0.0001$ in all comparisons (R vs. S, T, G, E), paired *t*-test after Bonferroni correction].

One explanation for this result is that some voxels within MT/V5 are more strongly activated by *gradient motion* than by other stimuli. If so, we should visualize such voxels when comparing *gradient motion* against the other stimuli. We deliberately used *translation* in this comparison because across all stimuli *translation* and *gradient motion* were completely matched for direction and average speed (see materials and methods). Figure 2b shows SPMs for the comparison *gradient motion–translation* (yellow) and *translation–gradient motion* (red) superimposed on an anatomical slice of the same participant appearing in Fig. 1b.

Translation evoked a significantly stronger activation in voxels corresponding to V1 (red), while *gradient motion* evoked a stronger activation in voxels corresponding to the right MT/V5 (yellow) ($p < 0.0001$). We did not find voxels in V1 or MT/V5 showing the reverse pattern. Of the remaining participants, three showed the same pattern ($p < 0.0001$, data not shown). The coordinates of the local maxima and the number of activated voxels corresponding to the right area MT/V5 for four participants were G–T (*S1*, #voxels=18, x=52, y=−58, z=6; *S2*, #voxels=138, x=58, y=−64, z=8; *S3*,

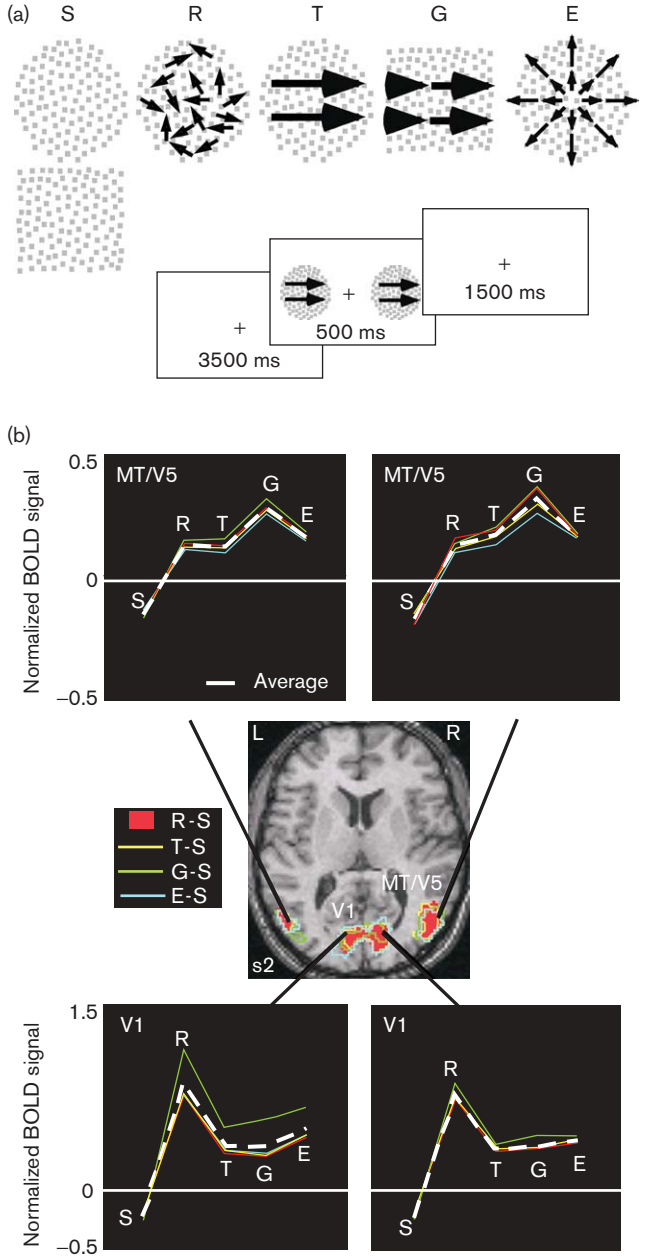

Fig. 1. (**a**) Stimuli, *stationary* (S), *random motion* (R), *translation* (T), *gradient motion* (G), *expansion* (E). Bottom: Event-related design. Intertrial intervals were optimized for blood oxygenation-level-dependent (BOLD) signal acquisition. (**b**) Center: Transversal anatomical images from one participant with statistical parametric maps superimposed. Red patches indicate areas significantly activated by R (left VI $x=-8$, $y=-84$, $z=10$; right VI $x=16$, $y=-78$, $z=12$; left MT/V5 $x=-50$, $y=-72$, $z=10$; right MT/V5 $x=48$, $y=-78$, $z=8$, $p<0.001$ *Z*-test). The yellow outline indicates areas significantly activated by T (left VI $x=-6$, $y=-84$, $z=10$; right VI $x=16$, $y=-80$, $z=12$; left MT/V5 $x=-52$, $y=-66$, $z=12$; right MT/V5 $x=48$, $y=-64$, $z=10$, $p<0.001$ *Z*-test), the green indicates those significantly activated by G (left VI $x=-8$, $y=-80$, $z=12$; right VI $x=16$, $y=-78$, $z=12$; left MT/V5 $x=-52$, $y=-72$, $z=8$; right MT/V5 $x=48$, $y=-76$, $z=6$, $p<0.001$ *Z*-test) and the cyan indicates those significantly activated by E (left VI $x=-8$, $y=-84$, $z=10$; right VI $x=16$, $y=-80$, $z=14$; left MT/V5 $x=-52$, $y=-66$, $z=12$; right MT/V5 $x=48$, $y=-76$, $z=8$, $p<0.001$ *Z*-test). Side panels: Average signal change (ordinate) evoked by the different stimuli (abscissa) in the regions significantly activated in the four motion-defining contrasts (see symbols). The white dashed line represents the mean.

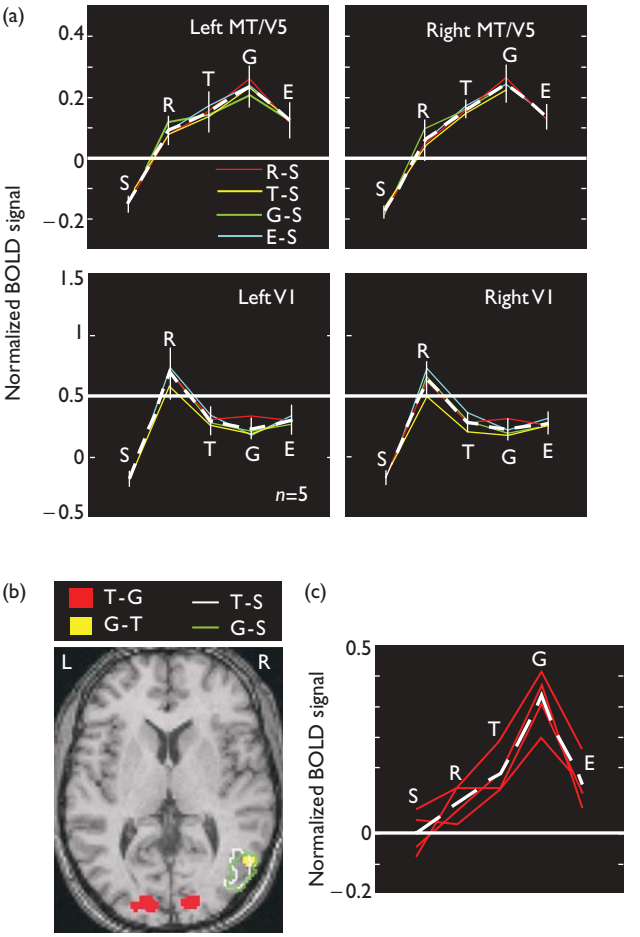

Fig. 2. (**a**) Average signal change evoked by the different stimuli across participants for all motion-defining contrasts in MT/V5 (upper panels) and VI (lower panels). The white dashed line indicates the average across contrasts. The error bars are standard errors. (**b**) Anatomical magnetic resonance image from a participant with statistical parametric maps for the motion-defined contrasts T–G and G–T superimposed. The contours indicate the regions activated by T and G relative to S on the right MT/V5. (**c**) Average signal change evoked by the different motion types within the right MT/V5 voxels significantly activated when contrasting G–T. The red lines indicate individual data and the white dashed line indicates the mean across participants. The error bars are standard errors.

#voxels=57, $x=52$, $y=-74$, $z=16$; *S4*, #voxels=66, $x=50$, $y=-62$, $z=0$), T–S (*S1*, #voxels=118, $x=48$, $y=-64$, $z=6$; *S2*, #voxels=180, $x=48$, $y=-64$, $z=10$; *S3*, #voxels=167, $x=48$, $y=-66$, $z=4$; *S4*, #voxels=494, $x=48$, $y=-72$, $z=14$), G–S (*S1*, #voxels=210, $x=50$, $y=-62$, $z=6$; *S2*, #voxels=156, $x=48$, $y=-76$, $z=6$; *S3*, #voxels=245, $x=52$, $y=-68$, $z=6$; *S4*, #voxels=627, $x=54$, $y=-70$, $z=8$).

The number of voxels activated by *gradient motion* relative to *translation* within the right area MT/V5 decreased significantly relative to the voxels activated by both kinds of motion relative to *stationary* ($p=0.03$ relative to T–S and $p=0.007$ relative to G–S), suggesting that indeed only a subset of voxels were more strongly activated by *gradient motion*. Although the effect was statistically significant in only the right area MT/V5 in four participants, three of them showed a similar pattern on the left hemisphere at $p<0.01$. We did not find any apparent cause for this interhemispheric difference; however, the pattern was consistent across participants.

When examining the activation evoked by the different motion types in the voxels activated by *gradient motion* relative to *translation* (Fig. 2c), we found a strong activation evoked by *gradient motion* relative to the other stimuli ($p<0.01$, Wilcoxon rank sum test after Bonferroni correction). This increase in activation evoked by *gradient motion* may be because of the contribution of a subpopulation of neurons within MT/V5.

DISCUSSION

This study provides evidence in favor of the existence of speed-gradient-selective neurons within a subregion of human area MT/V5. Our findings suggest that MT/V5 is anatomically and functionally divided into different subregions containing units specialized in the processing of different motion patterns. This is consistent with previous reports of a similar specialization in monkeys [4,5] and illustrates the similarities between the architecture of the motion processing systems in humans and monkeys [i.e. selectivity for the average direction and speed of translational motion in MT/V5 [7], for speed gradients in the same area (this study) and for optical flow patterns in a neighboring area, thought to be the homolog of area MST in the monkey [11–13]].

Our results agree with predictions of computational models of V1 and MT direction-selective neurons. In these models V1 units strongly respond to the presence of preferred motion inside their receptive fields independently of the presence of other motion directions. MT neurons, in comparison, perform a spatial integration of velocity signals inside their receptive fields [14]. We have recently incorporated speed-gradient-selective units into a computational model that mimics the architecture of the primate motion processing system and successfully classified different types of moving patterns generated by different stimulation conditions [15].

One alternative explanation for our results could be that presenting many motion types within a circular aperture caused response adaptation in neurons with receptive fields covering the aperture's area. When presenting the *gradient motion* within a square aperture, nonadapted neurons with receptive fields covering the angular borders of the aperture would respond strongly causing a relative increase in activation. However, we consider this explanation unlikely because (1) it predicts a higher activation by *gradient motion* relative to the other stimuli in V1 (this was not the case), (2) we used *stationary* square-shaped stimuli, in addition to the circular ones, to contrast *gradient motion–stationary* and found no effect of aperture shape on the number of activated voxels, or in the levels of activation in both areas, (3) trials with different stimuli were randomly intermixed, and (4) we used short stimulus presentation times (500 ms) and longer intertrial intervals (e.g. 1.5, 3.5 s).

One issue that needs clarification is why we did not find voxels within MT/V5 in which *translation* evoked a stronger activation than *gradient motion*. This could be explained by considering that these two stimuli had the same average speed and direction; therefore, neurons selective for the average speed of translational motion would give the same responses to both. On the other hand, neurons selective for speed gradients would give a stronger response to *gradient motion* than to *translation*.

CONCLUSION

In summary, our results suggest that a subregion of human area MT/V5 specializes in the processing of speed gradients in optical flow patterns. This subregion may constitute an intermediate step between regions specialized in the processing of the average speed and direction of translational motion and regions selective for complex optical flow patterns (a suggestion regarding the computation of affine motion parameters made first by Longuet-Higgins and Prazdny, 1980) [16]. Our findings impose new constraints to existing models of the human motion processing system.

REFERENCES

1. Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS. A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 1991; **11**:641–649.
2. Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC *et al*. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex* 1993; **3**:79–94.
3. Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ *et al*. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 1995; **15**:3215–3230.
4. Treue S, Andersen RA. Neural responses to velocity gradients in macaque cortical area MT. *Vis Neurosci* 1996; **13**:797–804.
5. Xiao DK, Marcar VL, Raiguel SE, Orban GA. Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *Eur J Neurosci* 1997; **9**:956–964.
6. Beauchamp MS, Cox RW, DeYoe EA. Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J Neurophysiol* 1997; **78**:516–520.
7. Rees G, Friston K, Koch C. A direct quantitative relationship between the functional properties of human and macaque V5. *Nat Neurosci* 2000; **3**:716–723.
8. Sunaert S, Van Hecke P, Marchal G, Orban GA. Motion-responsive regions of the human brain. *Exp Brain Res* 1999; **127**:355–370.
9. Dumoulin SO, Bittar RG, Kabani NJ, Baker CL, Le Goualher G, Bruce PG *et al*. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex* 2000; **10**:454–463.
10. Dukelow SP, DeSouza JF, Culham JC, van den Berg AV, Menon RS, Vilis T. Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J Neurophysiol* 2001; **86**:1991–2000.
11. Morrone MC, Tosetti M, Montanaro D, Fiorentini A, Cioni G, Burr DC. A cortical area that responds specifically to optic flow, revealed by fMRI. *Nat Neurosci* 2000; **3**:1322–1328.
12. Graziano MS, Andersen RA, Snowden RJ. Tuning of MST neurons to spiral motions. *J Neurosci* 1994; **14**:54–67.
13. Tanaka K, Saito H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 1989; **62**:626–641.
14. Heeger DJ, Simoncelli EP, Movshon JA. Computational models of cortical visual processing. *Proc Natl Acad Sci USA* 1996; **93**:623–627.
15. Tsotsos JK, Pomplun M, Liu Y, Martinez-Trujillo JC, Simine E. *Attending to motion: localizing and labeling simple motion patterns in image sequences. Conference on Biologically-Motivated Computer Vision*, Tuebingen, Germany; 2002.
16. Longuet-Higgins HC, Prazdny K. The interpretation of a moving retinal image. *Proc R Soc Lond B Biol Sci* 1980; **208**:385–397.