RESEARCH ARTICLE

Differential cortical activation during the perception of moving objects along different trajectories

Finnegan J. Calabro[1,4] · Scott A. Beardsley[3] · Lucia M. Vaina[1,2]

Abstract

Detection of 3D object-motion trajectories depends on the integration of two distinct visual cues: translational displacement and looming. Electrophysiological studies have identified distinct neuronal populations, whose activity depends on the precise motion cues present in the stimulus. This distinction, however, has been less clear in humans, and it is confounded by differences in the behavioral task being performed. We analyzed whole-brain fMRI, while subjects performed a common time-to-arrival task for objects moving along three trajectories: moving directly towards the observer (collision course), with trajectories parallel to the line of sight (passage course), and with trajectories perpendicular to the line of sight (gap closure). We found that there was substantial overlap in the pattern of activation associated with each of the three tasks, with differences among conditions limited to the human motion area (hMT+), which showed greater activation extent in the gap closure condition than for either collision or passage courses. These results support a common substrate for temporal judgments of an object's time-to-arrival, wherein the special cases of object motion directly toward, or perpendicular to, the observer represent two extremes within the broader continuum of 3D passage trajectories relative to the observer.

Keywords fMRI · Object motion · hMT · Gap closure · TTP · TTC

Introduction

Many activities of daily living require the ability to detect how objects are moving around an observer. This ability is critical to safely negotiate the environment during walking, navigating a car in traffic, or to intercept and catch a moving object. In many such circumstances, it is not sufficient to detect solely an object's translation across the retina, since displacement across perceived visual space indicates only an object's frontoparallel motion (the motion within a plane perpendicular to the line of sight). Rather, it is critical that displacement be integrated with visual cues indicating motion-in-depth, such as looming. The combination of both translational and depth motion cues allow observers to infer a true 3D object trajectory, thus enabling them to accurately and effectively interact with their surroundings.

The need to integrate two disparate sources of visual information to determine an object's 3D trajectory raises the question of whether object motion is extracted by a single motion-responsive mechanism, or by a two-stage process in which all cues are first detected separately and then integrated. This is mirrored by a question of the underlying neuronal substrate, specifically, whether the same brain areas are used to estimate both translational and depth motion, or whether these cues are computed independently in the visual cortex, before being integrated, perhaps in yet another brain area. Here, we address this question using fMRI to determine whether separate neuronal substrates are recruited when subjects assess the motion of objects moving along different trajectories.

While a single object-motion detector would be neurologically efficient (the motion components are nearly always

✉ Finnegan J. Calabro
calabrofj2@upmc.edu

1 Brain and Vision Research Laboratory, Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

2 Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02214, USA

3 Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin, Milwaukee, WI, USA

4 Present Address: Departments of Psychiatry and Bioengineering, University of Pittsburgh, Pittsburgh, PA 15213, USA

co-present when viewing real world scenes, so processing them together and simultaneously seems logical), there are several lines of evidence against it. The first reason stems from known motion detection capacities of early visual areas. Motion detected by area V1 neurons is insufficient to characterize how an object is moving in space (3D). For example, a rightward moving edge is consistent with both an object moving to the right, or with the right edge of an expanding object (or the left edge of a contracting one). Determining the 3D trajectory of an object, therefore, depends on the integration of motion along the object boundary and a determination of how this boundary changes, either in its location across space or in its size and shape. While both position and size computations depend on the integration of motion along an object's boundary and/or surface, they constitute distinct computations: translational motion can be estimated when solving the aperture problem (Marr 1982; Hildreth 1983), since all points along the contour have the same direction of motion, while looming requires the integration of different directions of motion along the object's boundary. This is especially complex for non-colliding looming trajectories, because the expanding motion is non-symmetric. Thus, the differences intrinsic to computing translational and looming motion may be reflected in different computational and/or neuronal substrates.

Psychophysical evidence suggests a dissociation between translational motion and motion-in-depth when based on monocular and binocular (i.e., disparity) cues, respectively (Gray and Regan 1998), but less direct evidence exists when the different trajectories are based on monocular motion cues alone. We have shown that when estimating the arrival time for an object moving perpendicular to the line of sight ("gap closure"), observers relied on angular (translational) motion alone, and were not affected by the presence of looming motion, but when estimating the arrival time for objects on collision or passage courses, both translational and angular motion affected performance, and observers' performance was explained best by a weighted combination of the two (Calabro et al. 2011). This suggests the involvement of trajectory-specific mechanisms, and raises the question of whether there is a similar dissociation in the neuronal mechanisms.

Another piece of evidence that argues against the view that a single mechanism is used stems from psychophysical and functional imaging studies that have suggested that the detection of looming objects (those moving on a direct collision course with the observer) are associated with specialized collision–detection mechanisms in the human visual system. Psychophysical studies have shown that observers have a strong bias for detecting approaching objects, with detection rates significantly higher than for comparable receding objects (Georgeson and Harris 1978; Ball and Sekuler 1980). In fMRI, Field and Wann (2005)

have showed different cortical areas recruited depending on the trajectory of perceived motion, with sensorimotor areas specifically recruited in the case of approaching objects. However, their experimental conditions compared motion types across different tasks, and it remains unclear to what extent different activations are present across motion conditions when matched with a common, temporal arrival task. Billington et al. (2011) have shown that in addition to cortical regions, the superior colliculus and thalamus respond preferentially to approaching objects. However, they did not dissociate whether the responses were specific to objects on a colliding path with the observer, or simply to the presence of looming (an expanding object boundary, which may be present on an object moving on either a collision or passage course). Nevertheless, both psychophysical and neuroimaging evidences support the notion that specialized processing occurs for looming motion detection relative to translational motion.

Here, we used fMRI to characterize cortical responses to monocular cues resulting from a range of 3D object-motion trajectories. We operationalize 3D object motion as the implied movement of objects in 3D space, although we focus on only 2D cues for these trajectories based on combinations of retinal motion and changes in object size (looming). We identified the neuronal substrate of object-motion detection for objects containing translational motion only (gap closures), looming motion only (collision courses), or both (passage courses). We use a common task based on time-to-arrival estimation in different motion planes to identify differences in brain activation among these three motion conditions that cannot be attributed to differences in the task requirements. This allowed us to dissociate looming (i.e., expansion)-specific responses from collision-specific responses and to determine which brain areas respond to object motion independently of the specific visual information present. Our results show that despite computational differences employed to solve each task condition, the same cortical regions are in general recruited regardless of the 3D trajectory of the object. A single cluster consistent with left MT+ showed a difference in its magnitude of activation, indicating greater response in the gap closure condition, which may suggest specificity in motion-processing areas to the 3D trajectory of the object.

Materials and methods

Subjects

Ten right-handed neurologically healthy volunteer subjects from the Boston University student community (18–24 years, mean 20.3 years; 4 female) participated in the experiment. To ensure adequate performance, participants

practiced in the task in the Brain and Vision Research Laboratory at Boston University prior to the fMRI-scanning session. All participants had normal or corrected to normal vision and were screened for their ability to maintain fixation. Participants were screened for MRI contraindications and written informed consent was obtained according to the approved procedures of the Massachusetts General Hospital and Boston University Institutional Review Boards have been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Apparatus

Imaging data were collected at the Athinoula A. Martinos Center for Biomedical Imaging, using a 3T Siemens TrioTim 60 cm (RF coil ID) whole-body MRI. Stimuli were generated on and presented by a MacBook Pro running Matlab using in-house code based on the Psychophysical Toolbox (Brainard 1997; Pelli 1997) and OpenGL libraries, and projected by a Notevision6 LCD projector onto a translucent acrylic screen. Subjects viewed the screen through a Buhl Optical collimating lens placed on the head coil. A USB trigger code was sent from the imaging acquisition computer to the stimulus presentation computer to start each fMRI run and ensure accurate synchronization between the stimulus presentation and imaging systems. Subjects responded in a two alternative forced choice task using an MRI compatible button box with their right hand.

Stimulus and task

The stimulus was the same as in a previous psychophysical study (Calabro et al. 2011). Two blue spherical objects were positioned along the horizontal midline on either side of a red central fixation mark. On every trial, the initial eccentricity of each object was chosen randomly (between 2.8° and 9.5°) to randomize both the distances between the objects and their collision point as well as their speeds (by changing the distance traveled while maintaining a fixed time-to-arrival). Both objects had an initial size of ~1.5°, a mean luminance of 28 cd/m^2 on a low-luminance background (0.3 cd/m^2), and were labeled "1" and "2" throughout each trial. Object "1" was always to the left of fixation, and object "2" was always to the right of fixation; thus, subjects were not encouraged to saccade to the target to identify them (by number). The motion of each object was calculated, so that it crossed the medial plane (moving toward the fixation mark; gap closure, GC), or the frontal plane, either moving on a collision course directly toward the observer (time-to-contact, TTC), or on a passage course parallel to the observer's line of sight, (time-to-passage, TTP), as illustrated in Fig. 1. The mean expansion rate (TTC and TTP only; in GC, the objects did not expand) was 2.7° s^{-1}, and the mean speed during translational motion (in GC) was 5.4° s^{-1}.

During each trial, participants were instructed to fixate the central red fixation mark and determine which object would have crossed the medial plane (i.e., the fixation mark, for GC), or passed through the observer's frontal (depth) plane (TTC, TTP) first. Subjects responded with a keypress using an MR compatible button box after the object disappeared. Each imaging run lasted 300 s and consisted of 80 trials (0.5 s each) of a single trajectory condition (GC, TTC, or TTP). Two runs were acquired for each trajectory, for a total of six runs per subject. Subjects were told which experimental condition was being tested prior to the start of each acquisition to make the appropriate judgement (GC: arrival at the midline, TTC: arrival at the observer; and TTP: arrival at the observer's depth plane).

Fig. 1 Object-motion trajectory conditions. **a** Screenshot of the test showing the two objects on either side of a red fixation dot. (Fixation dot enlarged for clarity). **b** Schematic view of the three trajectory conditions used in the time-to-arrival task (from left to right): head-on collision (time-to-contact, TTC), gap closure (GC) and passage (time-to-passage, TTP) courses. Conditions are illustrated from a zenithal view, with the viewer's position indicated by the small open circle, and the line of sight (as maintained by fixation on the red dot) shown with the solid vertical line
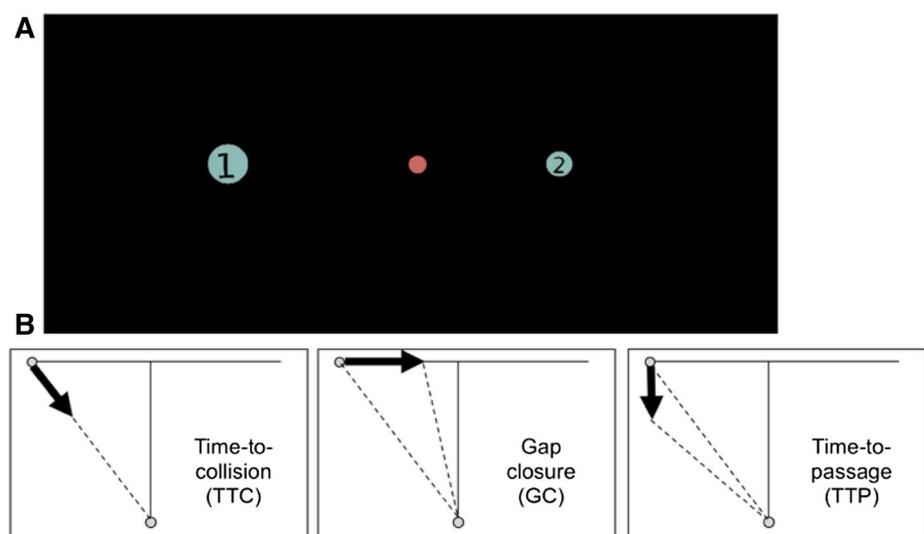
On each trial, subjects viewed the two objects simultaneously for 0.5 s. Object speeds were calculated, so that the first-arriving object would reach the collision point 1 s after stimulus onset. The later-arriving object's speed was determined, such that it reached the collision point 0.3 s later. Thus, the extrapolation times (time between when the objects disappeared and when they would have reached their target point) were 0.5 and 0.8 s, respectively. This time difference, chosen based on previously collected psychophysical data from each participant, made the task difficult enough, so that subjects had to actively engage, but not so hard that they would be unable to perform the task accurately. Furthermore, task difficulty (difference in arrival times) was chosen to produce approximately equal performance among subjects across the three task conditions.

Timing of task presentations was determined based on an optimized event-related imaging sequence (optseq) (Burock et al. 1998; Burock and Dale 2000). Events were coded specific to the task being performed and the resulting sequence was used to determine the stimulus presentation times. Each trial was shown for 500 ms, and subjects were given 1.5 s to respond, with the instruction to respond as quickly as possible. The interstimulus interval (ISI) was pseudo-random, ranging from 2 to 14 s, as determined by the optseq sequence. During this time, the screen was blank, except for the red fixation mark which subjects were instructed to fixate on.

MR scanning and data analysis

Two high resolution, T1-weighted (T1w) structural MRI scans were first obtained for registering the functional data using a 3D magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence (TR 2530 ms, TE 3.39 ms, inversion time 1100 ms, and flip angle 7°) with 128 slices of 1.33 mm thickness, and 256×256 in-plane sampling (1×1 mm resolution). Functional volumes were acquired using an interleaved, gradient echo EPI sequence every 2 s for each 5 min run (TR 2000 ms, 150 TRs; TE 30 ms, flip angle 90°). Six runs were obtained for each subject (two per task condition). Functional volumes were acquired using 33 axial slices of 3.6 mm thickness with in-plane sampling of 64×64 (resolution of 3.125×3.125 mm). Slice locations were positioned based on an AutoAlign sequence (van der Kouwe et al. 2005).

Preprocessing of functional images was based on an established pipeline designed to minimize the effects of head motion (Hallquist et al. 2013). Images were first simultaneously slice time and motion corrected, then underwent wavelet despiking to reduce motion related artifacts. We used an affine registration to the subject's T1w volume to align each volume and obtain motion estimates, together with a combined linear/non-linear (based on FSL's FLIRT and FNIRT) registration of the structural volume to MNI standard space. Functional data were smoothed using a 5 mm Gaussian kernel and high-pass filtered based on a 100 s cutoff ($f > 0.01$ Hz). Motion estimates were used to censor volumes with excessive head motion (framewise displacement > 0.9 mm or DVARS > 20) (Smyser et al. 2011).

Functional analysis was performed in AFNI (Cox 1996). Level 1 analysis (per subject, per condition) used 3dDeconvolve to model the response timecourse (24 s TENT response) of the full hemodynamic response function (HRF) relative to each (TR-locked) trial. Previous studies have implicated both cortical (Field and Wann 2005) and sub-cortical (Billington et al. 2011) regions, which may have different patterns of HRF responses (Aguirre et al. 1998; Handwerker et al. 2004). Furthermore, differences in response latency across conditions could affect the shape of the HRF (Henson et al. 2002; Lewis et al. 2018). Thus, since it was unclear whether a canonical HRF fit would be appropriate for all conditions and brain regions, this modeling approach was agnostic as to the predicted HRF response, allowing us to estimate responses without making assumptions about the shape of the HRF. Analyses were limited only to the correct responses, since incorrect trials could represent lapses in attention or awareness (e.g., eyes-closed), and because correct response rates varied slightly among conditions (see "Behavioral performance" section). Incorrect responses were modeled with a separate set of regressors which were ignored in subsequent analyses.

Group analysis was performed based on the resulting response functions using a linear mixed effects model (with AFNI's 3dMVM). This approach performs a repeated measures ANOVA for each voxel with main effects of task condition and TR, as well as their interaction. A main effect of TR indicates at some TR which was common across conditions. A significant condition × TR interaction indicates that at some time(s) after stimulus onset, the response differed among conditions. This approach is highly sensitive to response differences in both amplitude and shape, and does not assume a canonical HRF (Chen et al. 2014).

Cluster correction was performed using the residuals of each subject's first-level analysis as described above. This was used to assess spatial correlation (FWHM) among nearby voxels, and smoothness was estimated based on the residuals using AFNI 3dFWHMx tool. We used a mixed-model correction for spatial autocorrelation (ACF) to account for non-normal smoothness which could otherwise lead to false-positve cluster results (Eklund et al. 2016). Smoothness parameters were averaged across subjects, and used as the basis for cluster simulation (based on AFNI's 3dClustSim) to assess clusterwise significance as a function of cluster size and voxelwise threshold. To determine overall activation independent of the experimental condition, we identified voxels with a significant main effect of TR (timecourse changed from

baseline) at $p < 0.001$ ($T > 7.2$), with cluster simulation performed within a gray matter mask to consider the entire cortex, with $n > 26$ contiguous voxels for a clusterwise $p < 0.05$. To determine brain regions that differed among conditions, we first computed task-specific activation maps by determining the main effect of TR independently for each condition. We used this to create a conjunction mask of voxels with FDR-corrected significant activation in any condition. We used this as a mask for the analysis of the TR × condition effect from the main mixed effects model. Within this mask, clusters were identified with a clusterwise $p < 0.05$, based on a voxelwise $p < 0.001$ threshold, and $n > 41$ contiguous voxels. To characterize differences among conditions, post hoc analyses were performed by visualizing the mean time course for each of the three task conditions extracted and averaged across all voxels within significant clusters, and performing post hoc statistical tests among pairs of conditions at each time course of the response.

Results

Behavioral performance

During the fMRI runs, there was a significant difference in performance by condition (repeated measures ANOVA, $F = 8.5$, $df = 2$, $p = 0.0025$). A post hoc test indicated that subjects performed slightly better on the GC task (91.6% ± 5.5% s.d. across subjects) than TTC (86.4% ± 5.4%) and TTP (80.7% ± 7.9%) conditions (two-tailed paired t test for performance across subjects showed GC significantly better than TTC, $b = 0.052$, $T(df = 9) = 2.35$, $p = 0.04$, and TTP, $b = 0.108$, $T(9) = 3.54$, $p = 0.006$, with the difference between TTC and TTP just failing to meet significance, $b = 0.056$, $T(9) = 2.21$, $p = 0.054$). Furthermore, a comparison of activation between correct and incorrect trials did not reveal any significant clusters of activation, and suggest that the differences in task accuracy are unlikely to produce significant differences in activation.

Shared neuronal substrate of TTC, GC, and TTP tasks

We analyzed task and subject-specific hemodynamic responses as the basis for group analyses. Time courses were assessed in a repeated measures ANOVA (AFNI's 3dMVM) with effects of TR and condition. We first investigated parametric maps corresponding to the main effect of TR, indicating which voxels showed task-evoked responses, independent of the task condition (Fig. 2). This revealed widespread activation across the cortex, corresponding to a network of primarily occipital and parietal regions, as well as motor and
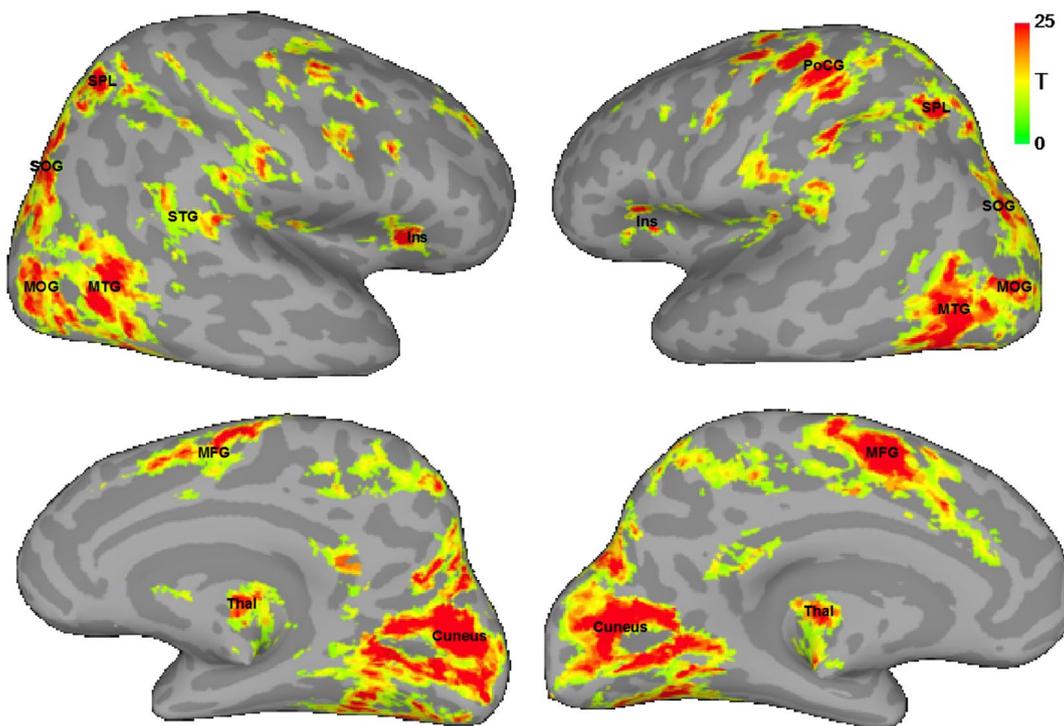
Fig. 2 Main effect of TR, indicating regions that showed a significant task-related activation, independent of the task condition. Regions of activation are labeled based on nearest anatomical location (*MOG* middle occipital gyrus, *MTG* middle temporal gyrus, *SOG* superior occipital gyrus, *SPL* superior parietal lobule, *STG* superior temporal gyrus, *Thal* thalamus, *Ins* insula, *PoCG* post-central gyrus)

motor planning areas (particularly in the LH) and insula. Individual regions were identified based on discrete peaks of activation ($T > 15.0$) separated by a minimum of 20 mm (using AFNI's 3dMaxima; Table 1).

Task-specific activation differences

To determine whether any regions within the main activation map exhibited task-specific activation, we examined the TR*condition interaction. Significant effects in this interaction would suggest that the response across conditions diverged for at least one TR during the response. A single region ($n = 42$ voxels) survived the cluster correction (see Fig. 3), and was situated in the left middle temporal sulcus (MNI center of mass 44, − 68,4; peak response $F = 7.21$). This location is commonly identified with area V5/hMT+. The visualize the sign of this effect, mean time courses extracted for each condition from this cluster revealed an elevated peak activation in the gap closure condition and nearly identical time courses of activation for the TTC and

TTP conditions (Fig. 3). A post hoc Wald Chi-square test among pairs of conditions revealed that this effect was driven by elevations in activation in the GC condition compared to both TTC ($\chi^2 = 4.26$, $p = 0.039$) and TTP ($\chi^2 = 6.21$, $p = 0.013$) occurring two TRs (4 s) after stimulus onset. No differences between TTC and TTP were observed at this point ($\chi^2 = 0.01$, $p > 0.9$). Importantly, no clusters were identified with activation that was unique to individual task conditions.

Discussion

We compared the activation elicited by three different object-motion conditions: gap closure (GC) containing angular motion only, time-to-collision (TTC) for objects on collision courses containing only looming motion, and time-to-passage (TTP) for objects moving along passage courses containing both looming and angular motion. We were interested in determining whether different brain regions mediate

Table 1 Clusters defined by local maxima in overall activation analysis

MNI coordinate					
x	y	z	# Voxels	Anatomical location	Brodmann area
− 9.42	− 73.7	10.49	5639	Precuneus	BA 23
34.68	− 72.68	− 4.8	3372	Left inferior occipital gyrus	BA 19
18.65	− 67.42	45.86	2475	Left superior parietal lobule	BA 7
− 39.87	− 61.94	− 7.08	2918	Right inferior occipital gyrus	BA 19
− 30.5	− 49.26	51.56	1361	Right superior parietal lobule	BA 7
5.56	− 25.59	4.77	1758	Thalamus	
43.34	− 22.48	53.02	2440	Left precentral gyrus	BA 3
− 44.22	− 1.73	35.02	1449	Right precentral gyrus	BA 6
0.87	6.5	50.92	1852	Superior frontal gyrus/SMA	BA 32
38.46	17.24	13.79	323	Left insula	BA 13
− 27.93	42.56	25.95	329	Right superior frontal gyrus	BA 47
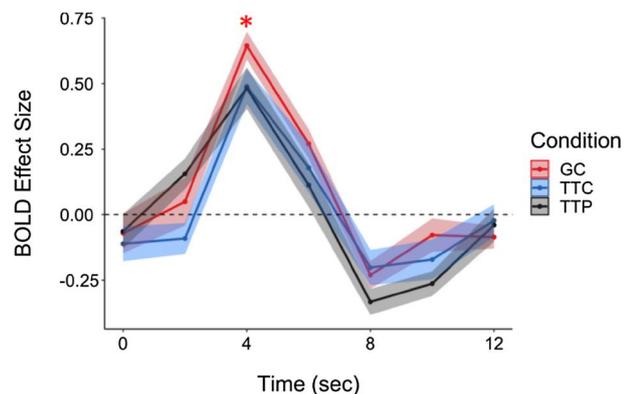
Fig. 3 Activation differences by task. Cluster-corrected voxels showing a significant TR × task condition interaction. Time courses showing the estimated HRF for each condition averaged across the cluster, indicate an elevated activation peak in the gap closure condition. Time points with significant post hoc differences among conditions ($p < 0.05$) are indicated by a star

object-motion perception based on the visual cues present in the stimulus. Our paradigm controlled for many variables among tasks (e.g., randomization of initial position meant there was no systematic difference in the location of the objects, extrapolation times were matched across conditions, etc.), so that differences among the conditions could be associated with the low level features (i.e., the speed of translational and/or looming motion) and high level interpretation (i.e., computation of arrival time in egocentric/allocentric coordinates) of performing each of the three tasks. We thus hypothesized that differences in the motion computations required could lead to task-specific activations. For example, translational motion detection (e.g., as present in the GC condition) has been extensively linked to area MT/V5 in the macaque monkey, since this area is one of the first shown to solve the aperture problem (Movshon and Newsome 1996; Pack and Born 2001). TTC, on the other hand, depends on the detection of looming, that is, each point along the surface of the object is moving in a different direction, which together form a radial pattern of motion. Perception of radial motion patterns such as these have been linked to area MST (Tanaka et al. 1986; Duffy and Wurtz 1991), though typically for large stimuli thought to be representative of self-motion (MSTd neurons have receptive fields on the order of 80° diameter, Tanaka and Saito 1989). Additional posterior parietal regions have been linked to approaching object-motion trajectories in both humans (Paradis et al. 2000; Vanduffel et al. 2002; Claeys et al. 2003) and non-human primates (Duhamel et al. 1998).

Consistent with the previous reports of activation on similar TTC and GC tasks (Field and Wann 2005), we found visual cortex activation in an inferior region of the lateral occipital complex, an area that has been found to be recruited during object and shape perception (Grill-Spector et al. 2001), and in MT+, a region critical for motion processing and perception (Tootell et al. 1995; Amano et al. 2009). While all three task conditions recruited activation in MT+, there was a difference in the magnitude of activation among task conditions, apparent as a change in the peak amplitude of response, with the GC condition having greater activation than TTC or TTP. Area MT/V5 in monkeys has typically been linked to translational, rather than looming, motion perception, while MSTd has been associated with perception of optic flow (wide field motion) rather than object motion. These areas are thought to constitute M+ in humans (Dukelow et al. 2001; Huk et al. 2002). Our results suggest that the MT complex overall is engaged by moving objects regardless of their trajectory, consistent with previous studies showing object representations in MT+ (Kourtzi et al. 2002), but that it is activated to a greater extent during the GC condition.

Consistent with the previous reports on TTC-driven activation (Field and Wann 2005; Billington et al. 2011), in all task conditions, we found significant bilateral activation of the anterior insula, a region which plays a role in judgments and comparison of short time intervals (Kosillo and Smith 2010). Contrary to the suggestion of a lateralization of temporal processing in the insula (Brunia and Damen 1988; Rao et al. 2001; Ferrandez et al. 2003), we found largely symmetric activation across hemispheres, consistent with reports of insular involvement in a bilateral timing circuit (Livesey et al. 2007; Stevens et al. 2007).

We used a single task for all three experimental conditions based on the comparison of arrival times of two objects placed symmetrically about a central fixation mark, so that any trajectory-dependent activation differences could not be attributed to differences in the temporal judgement task being performed by the observers. Despite this, we identified one notable difference in activation among conditions in a region corresponding to left V5/hMT+ (compared to probabilistic mapping, Malikovic et al. 2007). This region showed an elevated extent of activation in the GC condition. One possible explanation of why gap closure elicited stronger activation in this area is related to the relatively lower speed of movement of the object boundaries in the TTC or TTP conditions (e.g., the mean translation speed in GC, 5.4° s^{-1}, was four times higher than the speed of each edge in TTC, 1.3° s^{-1}). This would suggest that 2D features, rather than the 3D object speed or arrival time (which were matched among conditions), could contribute to differences in the extent of MT+ activation. Alternatively, the response may have been suppressed by the motions in opposite directions (i.e., in an expanding object one edge moves leftward and another moves rightward), creating a de facto translational object-motion detector, since the opposite directions of motion found along each edge of a moving object would prevent responses to motion-in-depth. Either case suggests that the response in the putative MT+ region is maximized by an object translating across space, rather than in depth relative to the observer. A possible counter to this interpretation is that we did not observe any differences between the TTC and TTP conditions, despite differences in their 2D motion cues. However, the differences between the motion cues of these conditions were small, and may be below the sensitivity of our imaging acquisition.

Interestingly, we did not identify any regions with activation that was unique to particular experimental conditions. Thus, activation across all higher level brain regions, likely associated with the interpretation of motion in three dimensions and temporal judgements, appears to be recruited similarly across conditions. These points to a shared functional–neuroanatomical substrate of the mechanisms involved in the detection of object motion across 3D motion trajectory conditions. This suggests that while many areas in the visual motion-processing stream undoubtedly prefer different motion features, at the level observable by fMRI,

these differences are diminished when carefully matching stimulus features, task instructions, and decision processes.

Our study has several notable limitations. First, while we attempted to match the task as closely as possible among conditions, the different trajectories necessitated slightly different instructions to the observers. Namely, in the GC condition, subjects assessed arrival at the midline, in TTC, they assessed arrival at the observer, and in TTP, they assessed arrival at their own depth plane. While the trajectory differences are themselves of interest, the need for slightly different instructions suggests that observers may employ different strategies among conditions. Relatedly, a second limitation is that eye movements were not assessed during the scan. While subjects had been trained with an eye tracker behaviorally, it is possible that their eyes were not successfully fixating during the scan, and instead could have fixated on one object or the other, or tracked objects with retinal motion. Recent reports have characterized the effect of free eye movements in fMRI (Son et al. 2019), but importantly, the pattern of activation that they have identified does not closely match the conditionwise differences that we report, so this is an unlikely source of the differences that we have identified.

Based on these results, we suggest that while the precise computations involved in each task may differ (Calabro et al. 2011), the overall similar nature of the three task conditions (all are discrete moving objects, whose motion must be computed and integrated along its edge) and task (a temporal judgement of motion trajectories) results in the recruitment of a largely overlapping neuronal substrate. This is compatible with a view of object-motion processing in which GC and TTC represent two special cases (containing exclusively horizontal or looming motion, respectively) of a broader continuum of possible trajectories (e.g., TTP).

References

Aguirre GK, Zarahn E, D'Esposito M (1998) The variability of human BOLD hemodynamic responses. Neuroimage 8:360–369

Amano K, Wandell BA, Dumoulin SO (2009) Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. J Neurophysiol 102:2704–2718. https://doi.org/10.1152/jn.00102.2009

Ball K, Sekuler R (1980) Models of stimulus uncertainty in motion perception. Psychol Rev 87:435–469

Billington J, Wilkie RM, Field DT, Wann JP (2011) Neural processing of imminent collision in humans. Proc Biol Sci 278:1476–1481. https://doi.org/10.1098/rspb.2010.1895

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10:433–436

Brunia CH, Damen EJ (1988) Distribution of slow brain potentials related to motor preparation and stimulus anticipation in a time estimation task. Electroencephalogr Clin Neurophysiol 69:234–243

Burock MA, Dale AM (2000) Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. Hum Brain Mapp 11:249–260

Burock MA, Buckner RL, Woldorff MG et al (1998) Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. NeuroReport 9:3735–3739

Calabro FJ, Beardsley SA, Vaina LM (2011) Different motion cues are used to estimate time-to-arrival for frontoparallel and looming trajectories. Vision Res 51:2378–2385. https://doi.org/10.1016/j.visres.2011.09.016

Chen G, Adleman NE, Saad ZS et al (2014) Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. NeuroImage 99:571–588. https://doi.org/10.1016/j.neuroimage.2014.06.027

Claeys KG, Lindsey DT, De Schutter E, Orban GA (2003) A higher order motion region in human inferior parietal lobule: evidence from fMRI. Neuron 40:631–642

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173. https://doi.org/10.1006/cbmr.1996.0014

Duffy CJ, Wurtz RH (1991) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J Neurophysiol 65:1329–1345. https://doi.org/10.1152/jn.1991.65.6.1329

Duhamel JR, Colby CL, Goldberg ME (1998) Ventral intraparietal area of the macaque: congruent visual and somatic response properties. J Neurophysiol 79:126–136. https://doi.org/10.1152/jn.1998.79.1.126

Dukelow SP, DeSouza JF, Culham JC et al (2001) Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. J Neurophysiol 86:1991–2000. https://doi.org/10.1152/jn.2001.86.4.1991

Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci USA 113:7900–7905. https://doi.org/10.1073/pnas.1602413113

Ferrandez AM, Hugueville L, Lehéricy S et al (2003) Basal ganglia and supplementary motor area subtend duration perception: an fMRI study. NeuroImage 19:1532–1544

Field DT, Wann JP (2005) Perceiving time to collision activates the sensorimotor cortex. Curr Biol CB 15:453–458. https://doi.org/10.1016/j.cub.2004.12.081

Georgeson MA, Harris MG (1978) Apparent foveofugal drift of counterphase gratings. Perception 7:527–536. https://doi.org/10.1068/p070527

Gray R, Regan D (1998) Accuracy of estimating time to collision using binocular and monocular information. Vision Res 38:499–512

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vision Res 41(10–11):1409–1422

Hallquist MN, Hwang K, Luna B (2013) The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures

functional connectivity. NeuroImage 82:208–225. https://doi.org/10.1016/j.neuroimage.2013.05.116

Handwerker DA, Ollinger JM, D'Esposito M (2004) Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage 21:1639–1651

Henson RNA, Price CJ, Rugg MD et al (2002) Detecting latency differences in event-related BOLD repsonses Application to words versus nonwords and initial versus repeated face presentations. NeuroImage 15:83–97

Hildreth EC (1983) The detection of intensity changes by computer and biological vision systems. Comput Vis Graph Image Process 22:1–27. https://doi.org/10.1016/0734-189X(83)90093-2

Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. J Neurosci 22:7195–7205. https://doi.org/10.1523/JNEUROSCI.22-16-07195.2002

Kosillo P, Smith AT (2010) The role of the human anterior insular cortex in time processing. Brain Struct Funct 214:623–628. https://doi.org/10.1007/s00429-010-0267-8

Kourtzi Z, Bülthoff HH, Erb M, Grodd W (2002) Object-selective responses in the human motion area MT/MST. Nat Neurosci 5:17–18. https://doi.org/10.1038/nn780

Lewis LD, Setsompop K, Rosen BR, Polimeni JR (2018) Stimulus-dependent hemodynamic response timing across the human subcortical-cortical visual pathway identified through high spatiotemporal resolution 7T fMRI. NeuroImage 181:279–291. https://doi.org/10.1016/j.neuroimage.2018.06.056

Livesey AC, Wall MB, Smith AT (2007) Time perception: manipulation of task difficulty dissociates clock functions from other cognitive demands. Neuropsychologia 45:321–331. https://doi.org/10.1016/j.neuropsychologia.2006.06.033

Malikovic A, Amunts K, Schleicher A et al (2007) Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: a probabilistic, stereotaxic map of area hOc5. Cereb Cortex 17:562–574. https://doi.org/10.1093/cercor/bhj181

Marr D (1982) Vision: the philosophy and the approach. Freeman, San Francisco

Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J Neurosci 16:7733–7741

Pack CC, Born RT (2001) Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature 409:1040–1042. https://doi.org/10.1038/35059085

Paradis AL, Cornilleau-Pérès V, Droulez J et al (2000) Visual perception of motion and 3-D structure from motion: an fMRI study. Cereb Cortex 10:772–783

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442

Rao SM, Mayer AR, Harrington DL (2001) The evolution of brain activation during temporal processing. Nat Neurosci 4:317–323. https://doi.org/10.1038/85191

Smyser CD, Snyder AZ, Neil JJ (2011) Functional connectivity MRI in infants: exploration of the functional organization of the developing brain. NeuroImage 56:1437–1452. https://doi.org/10.1016/j.neuroimage.2011.02.073

Son J, Ai L, Lim R et al (2019) Evaluating fMRI-Based Estimation of Eye Gaze during Naturalistic Viewing. bioRxiv. https://doi.org/10.1101/347765

Stevens MC, Kiehl KA, Pearlson G, Calhoun VD (2007) Functional neural circuits for mental timekeeping. Hum Brain Mapp 28:394–408. https://doi.org/10.1002/hbm.20285

Tanaka K, Saito H (1989) 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 62:626–641. https://doi.org/10.1152/jn.1989.62.3.626

Tanaka K, Hikosaka K, Saito H et al (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J Neurosci 6:134–144

Tootell RB, Reppas JB, Dale AM et al (1995) Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. Nature 375:139–141. https://doi.org/10.1038/375139a0

van der Kouwe AJW, Benner T, Fischl B et al (2005) On-line automatic slice positioning for brain MR imaging. NeuroImage 27:222–230. https://doi.org/10.1016/j.neuroimage.2005.03.035

Vanduffel W, Fize D, Peuskens H et al (2002) Extracting 3D from motion: differences in human and monkey intraparietal cortex. Science 298:413–415