



OPEN

Distributed encoding of curvilinear self-motion across spiral optic flow patterns

Oliver W. Layton^{1,2} & Brett R. Fajen²

Self-motion along linear paths without eye movements creates optic flow that radiates from the direction of travel (heading). Optic flow-sensitive neurons in primate brain area MSTd have been linked to linear heading perception, but the neural basis of more general curvilinear self-motion perception is unknown. The optic flow in this case is more complex and depends on the gaze direction and curvature of the path. We investigated the extent to which signals decoded from a neural model of MSTd predict the observer's curvilinear self-motion. Specifically, we considered the contributions of MSTd-like units that were tuned to radial, spiral, and concentric optic flow patterns in "spiral space". Self-motion estimates decoded from units tuned to the full set of spiral space patterns were substantially more accurate and precise than those decoded from units tuned to radial expansion. Decoding only from units tuned to spiral subtypes closely approximated the performance of the full model. Only the full decoding model could account for human judgments when path curvature and gaze covaried in self-motion stimuli. The most predictive units exhibited bias in center-of-motion tuning toward the periphery, consistent with neurophysiology and prior modeling. Together, findings support a distributed encoding of curvilinear self-motion across spiral space.

Substantial progress has been made over the past several decades toward understanding how neurons in visual cortex encode information about the environment and state of the observer. The pioneering work of Hubel and Wiesel¹ exemplifies the successes that have been garnered in primary visual cortex and other areas early in the visual hierarchy². Tuning properties and the function of neurons in areas further along in extrastriate cortex, however, have remained more elusive.

One such area is the dorsal medial superior temporal area (MSTd), which has long been associated with self-motion perception. Much neurophysiological work has focused on translation, the self-motion scenario wherein the observer moves along a straight path without eye or head movements. Forward translation generates a pattern of motion on the eye (optic flow) that radiates outward from a singularity (radial expansion; Fig. 1A) known as the focus of expansion (FoE), which specifies the observer's direction of travel (heading)³. Many MSTd neurons exhibit tuning to the position of the FoE in the radial optic flow pattern^{4,5} and causal links have been established to heading perception^{6–8}.

Translation, together with another type of self-motion known as rotation, fully characterize the possible optic flow patterns (to a first-order approximation) that may arise during movement through a rigid environment⁹. Rotation arises during eye movements and has three primary components: pitch, yaw, and roll. Vertical and horizontal eye movements are examples of pitch and yaw rotation, respectively, and yield mostly uniform motion with some perspective distortion at the periphery (Fig. 1B,C). Clockwise (CW) and counterclockwise (CCW) rotation of the head about the line of sight are examples of roll rotation and generates concentric optic flow patterns (Fig. 1D,E). Combinations of translational and rotational self-motion yield spiral patterns, as Fig. 1F depicts for forward translation and roll. The continuum of patterns has been said to represent a "spiral space"^{10,11}. Neurons demonstrate systematic tuning to patterns that span the spiral space in MSTd^{10–12} and in areas 7a^{13,14} and VIP¹⁵, other areas at a similar depth along the dorsal stream¹⁶ with extensive optic flow sensitivity.

Much of what is known about the role of these brain areas in self-motion perception presupposes movement along a straight path. The neural basis of more general curvilinear self-motion is presently unknown. In the present study, we performed a computational analysis to determine whether populations of spiral space-tuned neurons could encode self-motion along curvilinear paths. Curvilinear self-motion commonly arises during naturalistic self-motion and yields optic flow (Fig. 1G–J) with both a translational component, due to observer movement, and a rotational component, due to rotation of the view about a point in the world (e.g. center of

¹Department of Computer Science, Colby College, Waterville, ME, USA. ²Department of Cognitive Science, Rensselaer Polytechnic Institute, Troy, NY, USA. email: oliver.layton@colby.edu

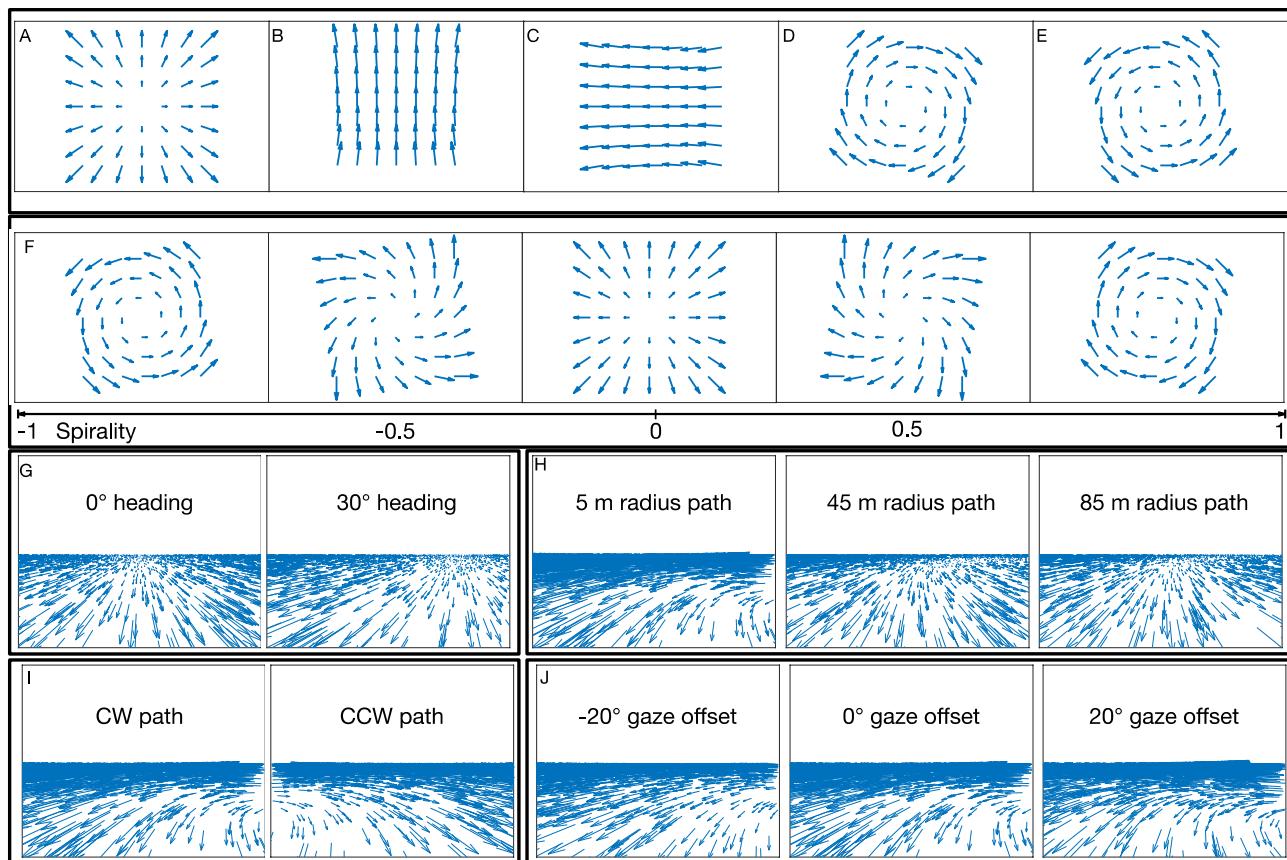


Figure 1. Sample optic flow vector fields and patterns in spiral space continuum. (A) Optic flow experienced by an observer moving straight-forward toward a frontoparallel wall. A stationary observer that performs a downward (B) or rightward (C) eye movement experiences approximately laminar optic flow due to pitch and yaw rotation, respectively. A stationary observer that rotates the head about the line of sight experiences CW (D) or CCW (E) concentric motion due to roll rotation. (F) Sample optic flow patterns in spiral space. The “spirality” metric defines the position in the spiral space and the proportion of roll rotation that is added to the radial translation field (center). Negative (positive) spirality yields CCW (CW) motion. (G–J) Optic flow fields generated by different types of linear and curvilinear self-motion over a ground plane. (G) Self-motion along straight-ahead (left) and rightward (right) linear trajectories. (H) Curvilinear self-motion along circular paths with different radii. (I) Curvilinear self-motion along CW (left) and CCW (right) 5 m radius circular paths. (J) Curvilinear self-motion along 5 m radius circular paths with the observer gaze offset from the path tangent (left: gaze directed outside the future path, center: gaze along path tangent as in (H,I); right: gaze directed inside the future path).

circular path). In the present study, we examine the case where the observer moves along a circular path of self-motion of constant radius. We focus on the neural encoding of path curvature, path sign, and gaze offset, three key parameters that describe the curvilinear self-motion of the observer. Path curvature is inversely related to the radius of the circular path traversed by the observer (Fig. 1H). Path sign indicates whether the observer moves along a CW or CCW path (Fig. 1I). Gaze offset refers to the horizontal offset in the observer’s gaze (i.e. bias in camera orientation) relative to the straight-ahead (Fig. 1J).

The present study addresses the following questions:

- Could units tuned to optic flow patterns in spiral space accurately encode the curvilinear self-motion of the observer? To what extent does the neural population plausibly account for human perceptual judgments?
- Do contributions predominately come from units tuned to specific patterns or is the model consistent with a distributed representation across spiral space? Are units tuned to radial expansion sufficient to decode accurate curvilinear self-motion estimates?
- What are the singularity tuning characteristics of the units that contribute to estimates of curvilinear self-motion? Specifically, are they uniformly distributed or biased toward the center or periphery?

To address these questions, we fit linear decoders to predict gaze offset and path curvature parameters from the activations of model MSTd units and compared the accuracy of estimates with human judgments (Fig. 2). We used lasso regularization^{17,18} to create sparse decoding models that select patterns that are most predictive

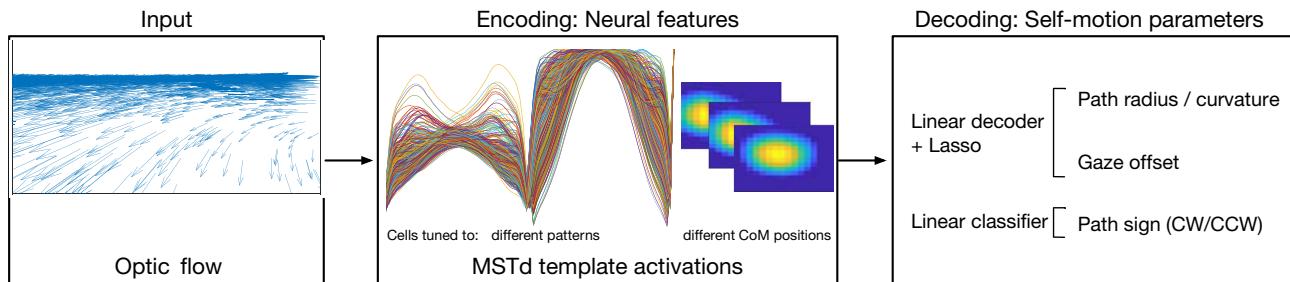


Figure 2. Overview of the neural encoder-decoder system. Optic flow generated by simulated self-motion (left) is passed through a neural model of areas MT and MSTd (center), which contains units tuned to optic flow patterns sampled from spiral space with different preferred centers of motion. The activation of the MSTd units is decoded (right) to recover parameters describing the observer's curvilinear self-motion (path curvature, gaze offset, and path sign).

when estimating each self-motion parameter. Models with sparse representations of optic flow capture important properties of MSTd neurons¹⁹ and human heading perception²⁰.

Results

We simulated a neural model that first processes the input optic flow (e.g. Figs. 1G–J and 2 left panel) with filters, which emulate the direction and speed tuning properties of the medial temporal (MT) area. We then matched the resulting activity distribution with template patterns that represent each unit's preferred spiral space pattern (Fig. 1F). This feedforward template matching paradigm is widely used to model optic flow selectivity in MSTd^{21–26}. We examined the encoding of curvilinear self-motion in a population of model units tuned to optic flow patterns ("MSTd") sampled from spiral space (henceforth "Full model"). Because one of our datasets consists of optic flow generated by simulated curvilinear self-motion over a ground plane (Fig. 1H–J), the Full model has units tuned to both full-field and hemi-field versions of spiral space patterns, the latter of which only included sensitivity to motion at or below the position of the singularity, which we refer to as the center of motion (CoM). For each preferred optic flow pattern, we simulated units tuned to different CoM positions sampled across the visual field.

We compared the Full model with a simpler model consisting only of units tuned to radial expansion (henceforth "Radial Only model"). We fit the linear decoders associated with the Full and Radial Only MSTd models using the activations produced from 900 optic flow samples that varied with respect to path curvature (radius: 5–200 m), gaze offset (-35 to $+35^\circ$), and path sign (CW, CCW). Path sign was predicted using a linear binary classifier. Unless noted otherwise, simulations involve optic flow from self-motion over a ground plane; we obtained similar results on a dataset consisting of simulated self-motion through a 3D dot cloud (see Supplementary Information).

The lasso yielded sparse models for decoding gaze offset and path curvature that included 758 (3.52%) and 718 (3.34%) units, respectively, of the simulated 21,504 MSTd units derived from the Full model (16×16 grid of CoM positions across the visual field tuned to 42 CW and 42 CCW spiral space patterns). The decoders derived from the Radial Only model to estimate gaze offset and path curvature included 83 (32.42%) and 87 (33.98%) units, respectively, of the 256 units (16×16 CoM) in the Radial Only model.

Figure 3 summarizes the accuracy of estimates of gaze offset (Fig. 3A,B) and path curvature (Fig. 3C,D) decoded from the full (top row) and radial only (bottom row) models on the 500 novel test optic flow samples. To quantify error across the dataset, we use mean absolute error (MAE). Gaze offset estimates decoded from the Full model (MAE: 0.76° , Fig. 3A) were substantially more accurate than those obtained from the Radial Only model (4.06° , Fig. 3B). Path curvature estimates decoded from the Full model (0.0027 1/m Fig. 3C) were also much more accurate than those obtained from the Radial Only model (0.0070 1/m, Fig. 3D). Although the predictions of the Full model were based on more units than the predictions of the Radial Only model, the Full model achieves more accurate predictions than the Radial Only model even when we constrain it to have the same number of nonzero regression weights (i.e. free parameters). This analysis is described in the Model Comparison section below.

To better appreciate the practical significance of the discrepancy between the models, we expressed the error in path curvature estimates with respect to the difference in visual angle subtended between the true and decoded paths ("path error"; Fig. 4A). Small angular differences imply close proximity between the two paths from the observer's vantage point. We measured this difference 10 m from the observer, a distance that has been used to obtain human curvilinear path judgments²⁷. This analysis revealed that the Full model estimated paths that deviated by 0.72° (Fig. 3E) compared to 1.89° (Fig. 3F) in the case of the Radial Only model. The Full model correctly classified the path sign of all test samples while the Radial Only model made 35/500 misclassifications (7%).

Comparison to human judgments. We compared the accuracy of model estimates to that of human judgments of curvilinear self-motion. Our aim was not to perform an exhaustive analysis of the existing psychophysical literature; rather we sought to gauge the sufficiency of model parameter estimates for capturing human performance under the simulated conditions.

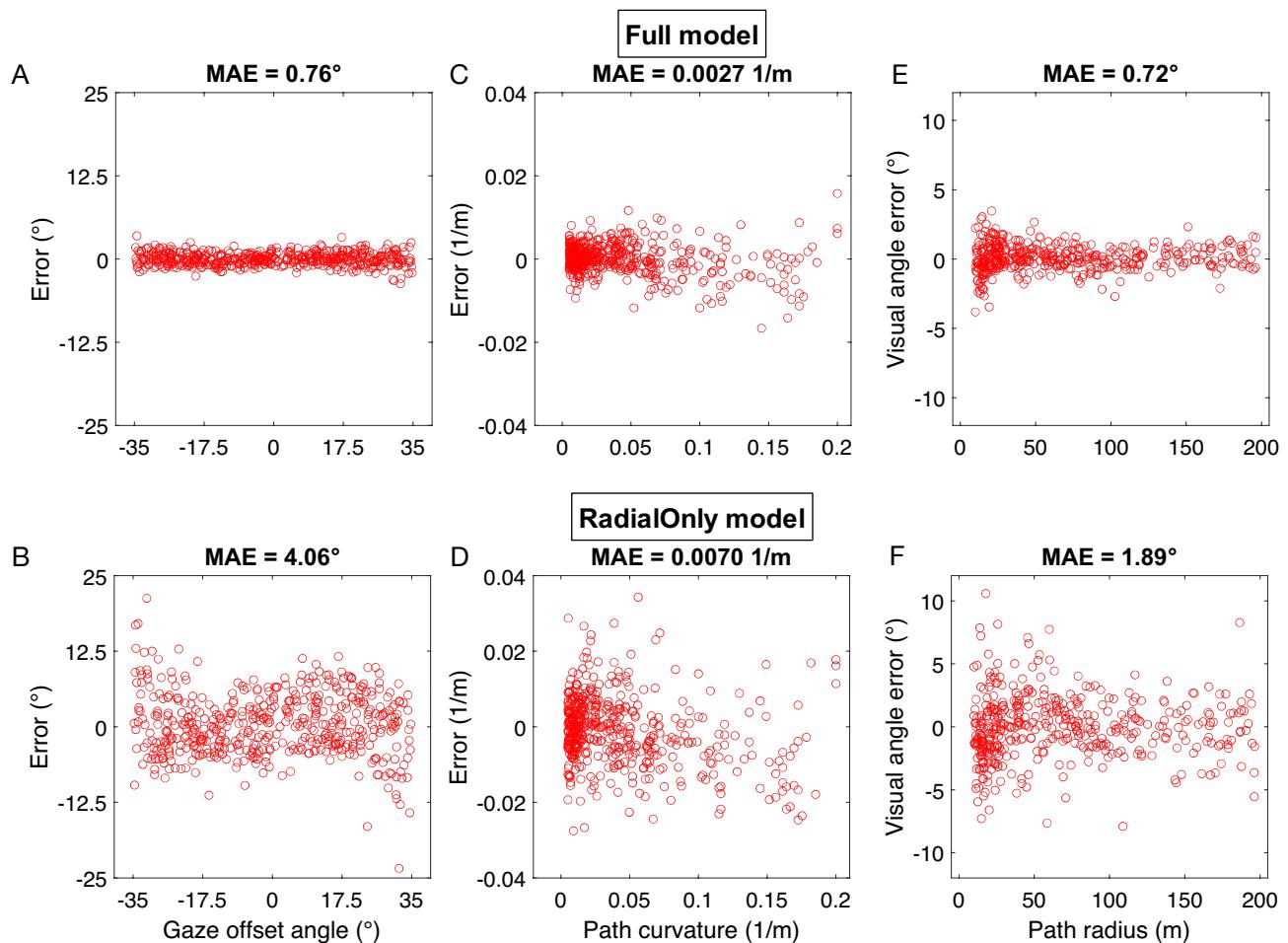


Figure 3. Error in the gaze offset (A,B) and path curvature (C–F) estimates produced by the Full model (top row) and Radial Only model (bottom row). (E,F) Error in path curvature is expressed with respect to the difference in visual angle between the true and predicted path measured at a depth of 10 m. The full model contains units tuned to heterogenous spiral space patterns, while the radial only model contains only units tuned to radial expansion.

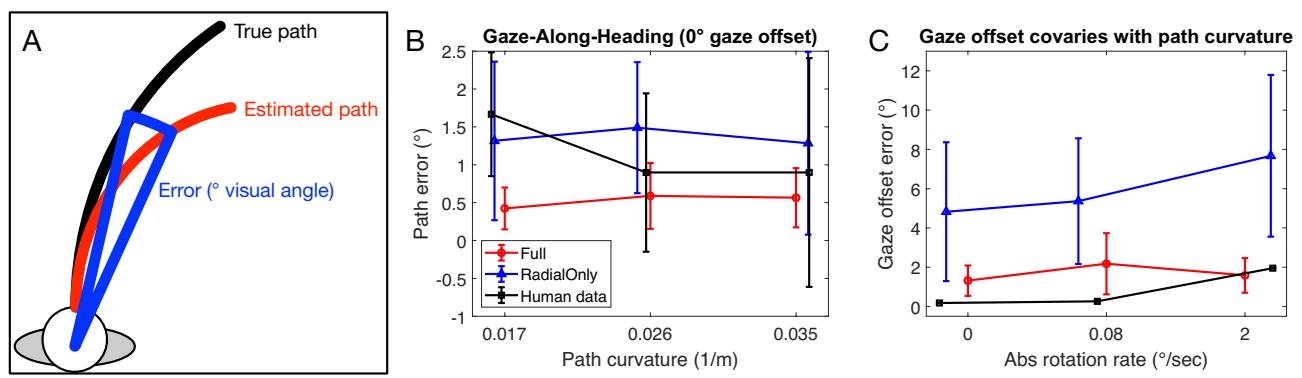


Figure 4. Comparisons between human judgments and estimates produced by the Full and Radial Only decoding models. (A) Schematic showing error in path curvature expressed as the angular difference in visual angle subtended by the estimated and true paths ("path error"). (B) Comparison between path the error produced by human subjects in the "gaze-along-heading" condition from Li and Cheng²⁷ and that produced by the decoding models. In this condition, the gaze offset is 0°. Positive (negative) path error corresponds to an overestimation (underestimation) of path curvature. (C) Comparison between the gaze offset error produced by human subjects in the "envelop" condition from Burlingham and Heeger²⁸ and that produced by the decoding models. In this condition, gaze offset covaried with path curvature. X values are staggered for visual clarity.

We simulated the “gaze-along-heading” condition from Li and Cheng²⁷ in which subjects judged path curvature while viewing simulated movement over a dot-defined ground plane. We used the previously fit decoders to estimate path curvature in this new condition. We expressed decoding error in terms of the previously used difference in visual angle (“path error”; Fig. 4A) measure to afford a direct comparison to the Li and Cheng data. Figure 4B suggests that both models produce human-like estimates when gaze is aligned with the instantaneous heading direction. Consistent with Li and Cheng, the positive error indicates an overestimation of the path curvature in both models. The difference between the path error predicted by the model and the path error based on the human data was significantly smaller for the Full model compared to the Radial Only model [$t(59) = 2.25$, $p = 0.03$]. However, inspection of Fig. 4B indicates that the advantage for the Full model is not consistent across path curvatures.

This prompted us to compare performance under more challenging conditions wherein gaze offset varied over a wider range. We simulated the “envelop” condition from Burlingham and Heeger²⁸ in which humans judged the gaze offset angle [−20, 20°] in displays simulating curvilinear self-motion along 43 m and 107 m radius paths. We expressed these path radii with respect to the rotation rate in the optic flow to afford direct comparison to their study (see “Methods”). Figure 4C shows that the Radial Only model does not plausibly account for human gaze offset judgments. The Full model yields comparable error to humans on the high curvature path (2°/s rotation) condition and slightly greater error compared to humans on the straight and 107 m paths, but the discrepancy (1–2°) is quite small. The difference between the gaze offset error predicted by the model and the gaze offset error based on the human data was significantly smaller for the Full model compared to the Radial Only model [$t(27) = 5.81$, $p < 0.001$]. This suggests that cells tuned to non-radial patterns in spiral space may play a role in estimating gaze offset with human-like accuracy.

Model position and pattern tuning characteristics. Given the improved accuracy of the Full model and its consistency with human data, we analyzed the tuning properties of the model units that contribute to self-motion parameter estimates. Figure 5 shows the number of units with different CoM preferences that contribute to gaze offset (Fig. 5A) and path curvature (Fig. 5B) estimates. Far more units tuned to peripheral CoM positions contributed to estimates than those tuned to central CoM positions. That is, the lasso set many more units with central CoM preferences to zero than those with peripheral preferences in both decoding models. As we explain in the Discussion, this finding is compatible with findings from both neurophysiological and computational studies.

Next, we considered how the relative contribution of units varied across pattern within the spiral space. This is quantified in terms of the relative strength of regression weights. Units that share a large proportion of the overall weight make more substantial contribution to parameter estimates than others with a smaller share of the weight. We use the “spirality” metric to express the fraction that the preferred spiral space pattern tuning depends on rotational optic flow. For example, 0 implies sensitivity to radial expansion, ±1 implies sensitivity to concentric motion, and values in between that approach ±1 imply sensitivity to spirals with increasing curvature (Fig. 1F). Figure 5C plots the proportion of the total weight that is shared among units tuned to different spiral space patterns in the gaze offset and path curvature decoders. Units tuned to concentric motion accounted for the most weight compared to units tuned to any other pattern type and therefore contribute most to estimates of gaze offset (21.7%; spirality 1) and path curvature (26.9%; spirality 1) in the Full model. The remaining proportion of the weight in both decoders is broadly shared across the other pattern types. This suggests a distributed representation across spiral space, which agrees with the findings of Xu and colleagues that show fairly even contributions among neurons tuned across spiral space toward monkey heading judgments¹⁰.

Figure 5D shows the proportion of units tuned to each pattern type included in each decoder. Taken together with Fig. 5C, units tuned to concentric motion represent the most numerous group and contribute most to estimates compared to any other single group.

Model comparison. The simulation results from Figs. 3 and 4 reveal the insufficiency of a model constructed with units tuned only to radial patterns. However, findings depicted in Fig. 5C raise the possibility that units tuned to other narrow ranges of the spiral space could primarily mediate the estimates produced by the Full model. We examined this possibility further by constructing models that constrain decoding from focused portions of spiral space. Specifically, we compared the accuracy of gaze offset and path curvature estimates garnered by the Full model with separate decoders fit to units tuned only to radial expansion (“Radial Only”); to spirals associated with the lowest, middle, and upper third of spiralities; and to concentric motion (“Concentric Only”). We also sought to address the extent to which the Full model’s superior accuracy to the Radial Only model may be attributed to the larger number of units that it includes. To that end, we included in the model comparison a version of the Full model constrained to possess the same number of nonzero weights as the Radial Only model (“Full Limited”).

Consistent with the analysis shown in Fig. 3 and Fig. 6 shows that the Radial Only model garnered considerably worse accuracy than the Full model when estimating gaze offset and path curvature. Decoding from any one group of spiral-tuned units or the concentric units produced gaze offset estimates that were almost as accurate as those produced by the Full model (Fig. 6A). In the case of path curvature, the accuracy garnered by the spiral-tuned decoders was approximately halfway between that achieved by the Radial Only and Full models (Fig. 6B). In contrast to gaze offset, the accuracy of the path curvature estimates decoded from the concentric units was poor and similar to that decoded from the Radial Only model. The discrepancy in performance between concentric unit models could stem from the larger number of concentric motion units that contribute to gaze offset predictions (Fig. 5D). It is possible that sensitivity to concentric motion on its own may not capture enough information that is essential for predicting path curvature across a range of gaze offsets (Fig. 1).

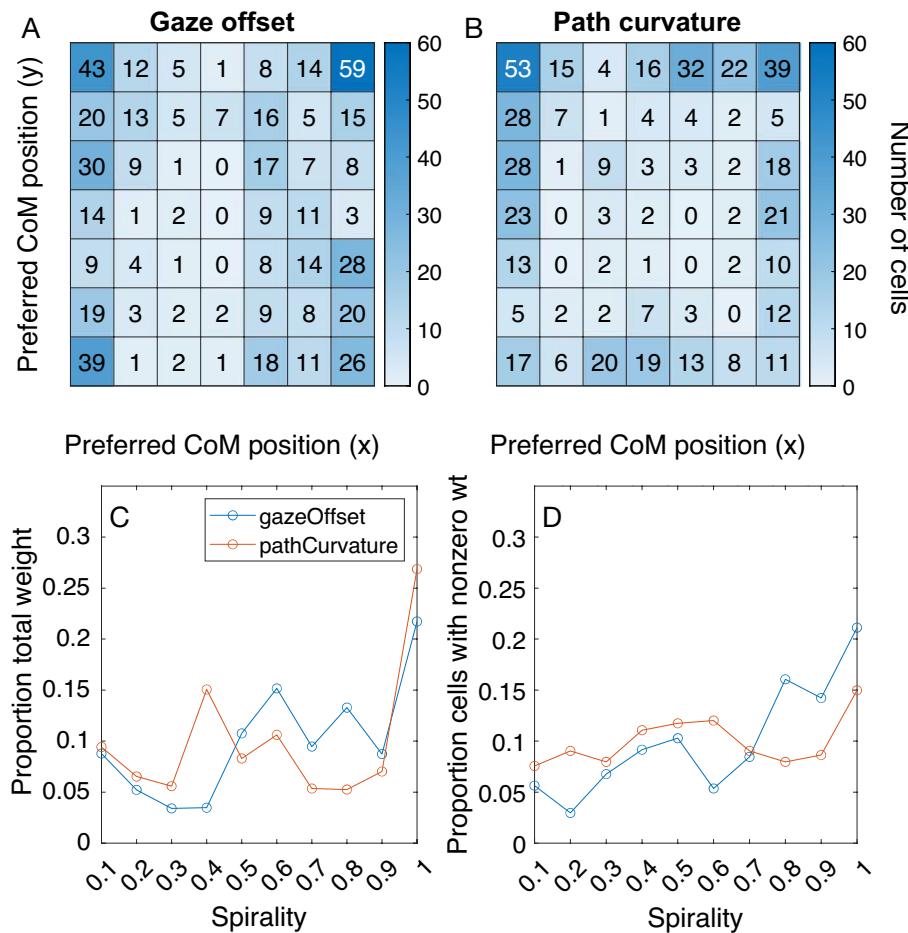


Figure 5. Spatial and pattern tuning properties of gaze offset and path curvature decoding models. (Top row) Histograms showing the number of cells tuned to CoM positions in the different regions of the visual field models included in the gaze offset (A) and path curvature (B) decoding models. (C) The proportion of the total regression weight accounted for by units tuned to different optic flow patterns (spirality) in each decoding model. (D) Proportion of cells tuned to different patterns included in each decoding model. Units are grouped in (C,D) based on the absolute value of spirality.

The accuracy of estimates produced by the Full Limited model was in between those achieved from Radial Only and Full models and comparable to the models constructed from thirds of the spiral space. The improved accuracy of the Full Limited model compared to the Radial Only model suggests an advantage of decoding gaze offset and path curvature from neurons tuned to a broad, distributed set of spiral patterns, even when controlling for the number of free parameters.

We examined the tradeoff between free parameters and accuracy more formally by computing the Akaike Information Criterion (AIC) for each model (Fig. 6C,D). This analysis also helps contextualize the accuracy of the other models, each of which contained different numbers of nonzero weights. For gaze offset, the numbers of nonzero weights were 385, 243, 83, and 73 for the Lowest Third, Middle Third, Upper Third, and Concentric Only models, respectively. For path curvature, these numbers were 268, 407, 342, and 171 for the Lowest Third, Middle Third, Upper Third, and Concentric Only models, respectively. Figure 6C,D reveals that the Full Limited model is favored according to the AIC metric for both gaze offset and path curvature estimates. Taken together, these results support a distributed representation of curvilinear self-motion across spiral space.

Discussion

While neurophysiology has linked the perception of linear self-motion to the activity of neurons in MSTd, the neural basis of more general curvilinear self-motion perception has been unclear. Our computational analysis suggests that tuning to combinations of translational and rotational optic flow in spiral space could support estimates of curvilinear self-motion. Decoding from units tuned to radial motion resulted in substantially less accurate and more variable estimates than the full spiral space model. While radial units in our simulations accounted for human judgments when gaze was aligned with the straight-ahead, they did not when gaze offset covaried with path curvature. The full spiral space model accounted for human judgments in both scenarios. Several of our analyses indicated a distributed representation of gaze offset and path curvature across spiral

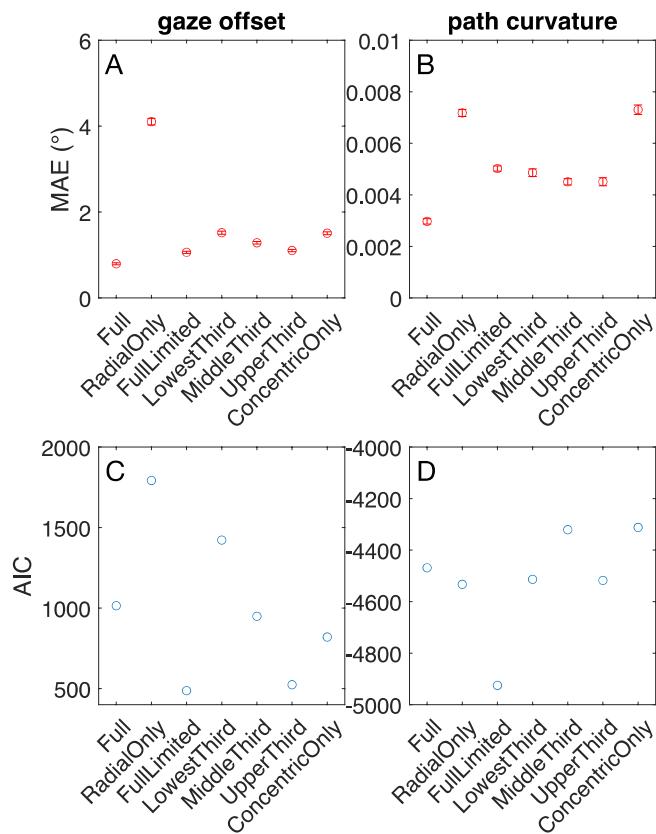


Figure 6. Comparison of decoding models. Top row: The MAE for gaze offset (**A**) and path curvature (**B**) parameter estimates. Bottom row: AIC computed for each of the models. In addition to the Full and Radial Only models, we considered a version of the Full model constrained to have the same number of nonzero weights as the Radial Only model (i.e. same number of free parameters; Full Limited). The Lowest Third, Middle Third, and Upper Third correspond to models that consist only of units tuned to lowest third (nearly radial), middle third, and upper third (nearly concentric) of spirality in the spiral space. Concentric Only indicates a model that consists only of concentric units. Error bars correspond to ± 1 standard deviation across MAE estimates obtained over 50 bootstraps. The small error indicates robustness to the specific set of optic flow stimuli included in the fitting process of each decoding model.

space. We found considerable redundancy in the representation of gaze offset—decoding from units tuned to non-radial patterns in spiral space yielded estimates that closely approximated those obtained by the Full model. This was mostly true of path curvature, albeit to a weaker extent and decoding from units tuned to concentric motion alone yielded poor estimates that were similar to those produced by the radial subpopulation. Overall, our findings support the possibility that MSTd could jointly signal the observer’s linear and curvilinear self-motion.

We used lasso regularization to create sparse linear decoding models that select the units that are most predictive of each self-motion parameter. In the resulting models, we found an overrepresentation of units tuned to peripheral centers of motion (Fig. 5) regardless of whether the optic flow simulated self-motion both over a ground plane and through a 3D dot cloud (see Supplementary Information). This agrees with the peripheral bias in the distribution of MSTd heading preferences reported by neurophysiological studies^{4,29}. Peripheral bias in the population tuning increases sensitivity to central headings³⁰ and enhances the overall accuracy of heading estimates²⁵. Both outcomes may arise because broad, Gaussian-like tuning curves of MSTd neurons tuned to peripheral headings change most rapidly for central headings^{30,31}. Our model is compatible with this hypothesis since the MSTd units are center-weighted about the preferred singularity position (Eq. 16). Beyeler and colleagues have proposed a computational model that explains many physiological properties of MSTd neurons, including the population peripheral bias, as an emergent outcome of a dimensionality reduction process¹⁹. The model produces a sparse representation of optic flow fields encountered during random combinations of translation and rotation. It is interesting that both their model and ours yield peripheral bias in CoM selectivity despite the difference in approach. The lasso regularization used in the present study also favors a model fit with a sparse set of predictors. This converging finding further supports the idea that peripheral bias in CoM selectivity may result from a process that promotes a sparse distributed representation of optic flow. It is noteworthy that the peripheral bias was found despite uniformity in the parameters used to generate the stimuli and regardless of whether the range of gaze offsets was restricted to ± 35 deg (in the present study) or spanned the entire 360° range (in Beyeler et al.). The relationship between the input statistics, tuning curving properties, and peripheral bias should be explored further in future research.

The use of a linear decoder in the present study also advances neural modeling of optic flow processing by expanding the range of self-motion properties that can be estimated. Within the literature on this topic^{21,23,24,32–35}, which spans almost 30 years, the standard approach to “reading out” the model’s heading estimate is to find the focus of expansion (FoE) location of the most active MSTd unit. This approach makes sense for pure translational self-motion because there is a simple one-to-one correspondence between self-motion direction and FoE location. However, when we expand the scope to include complex self-motion trajectories and the full gamut of MSTd cells types, it no longer makes sense to assume that the model’s best estimate of self-motion is reflected in the activity of a single unit. A more effective approach is to look at the global pattern of activity across units to extract self-motion estimates. The approach adopted in the present study offers a way to decode multiple self-motion parameters (e.g., gaze angle, path curvature) from the global activity across cells tuned to radial, spiral, and concentric flow patterns.

We emphasize that although neurons in MSTd, VIP, and 7a demonstrate spiral space pattern selectivity, no existing neurophysiological study has specifically linked this tuning to curvilinear self-motion perception. Therefore, it remains an open question whether these cortical areas possess specialized mechanisms that subserve curvilinear path perception. Cheng and Gu have shed some light on this issue, demonstrating that neurons in MSTd, VIP, and VPS exhibit tuning to translation, rotation, and the combinations thereof that arise during curvilinear movement on a 6 DOF platform. Interestingly, the analysis performed by Cheng and Gu revealed a distributed representation of translation and rotation, which supports our findings. The conditions in this study involve vestibular stimulation without accompanying optic flow, so it is unclear how these cells might relate to spiral space optic flow tuning. Nevertheless, the robust vestibular tuning and ~ 80% linear decoder accuracy suggests that vestibular signals could contribute significantly toward the neural basis of curvilinear self-motion perception when the information is available. Further neurophysiological work is required to clarify the connection between spiral space and vestibular tuning.

It should be noted that the present study does not address the influence of eye rotation, which have been shown to modulate MSTd signals during self-motion^{36–38} and human self-motion judgments^{39–41}. Eye rotations introduce an endogenous source of rotation into the optic flow field that is independent from the rotation caused by curvilinear self-motion. To estimate curvilinear self-motion during eye movements, the prevalent efference copy theory and related models propose that the visual system relies on nonvisual signals (e.g. proprioceptive or motor signals specifying eye rotation rate) to factor out the rotational flow due to eye movements, leaving on the rotational flow due curvilinear self-motion^{42,43}. Such a mechanism could be used to remove the effects of eye rotation prior to decoding from our model. However, the extent to which MSTd decomposes the optic flow field into distinct translational and rotational components visually⁴⁴ or in the presence of non-visual signals³⁶ has been questioned. The mechanism by which the visual system tolerates rotation in optic flow during self-motion remains unclear.

Some have examined linear heading perception when gaze remains fixed and eye movements are simulated within the virtual environment of the display (“simulated eye movement condition”)³⁹. In the absence of proprioceptive or motor signals that accompany real eye movements and when the scene lacks dense motion parallax⁴⁵, humans appear to treat the simulated rotation as curvilinear self-motion without eye movements^{46,47}. This suggests that rotation due to curvilinear self-motion is not factored out, decomposed, or removed, at least for the purposes of estimating curvilinear self-motion⁴⁸. The system presented here is consistent with this proposal, since neither the neural nor the decoding models explicitly remove rotation due to curvilinear self-motion. Indeed, pattern tuning in MSTd may emerge through the dimensionality reduction that neurons perform on their motion inputs¹⁹ rather than a process whose purpose is to segment translation from rotation.

Methods

Optic flow datasets. We created an optic flow dataset consisting of 10 frame digital sequences of simulated curvilinear self-motion over a ground plane (2000 dots) at 64×64 pixel resolution. In each video, the simulated observer translated at 3 m/s along a circular path. The virtual camera had a 90° horizontal and vertical field of view and a height of 1.61 m above the ground. On each video frame, we computed the optic flow using a pinhole camera model⁴⁹ and standard analytic equations⁹. We clipped and replaced dots that exited the field of view or valid relative depth range (1–50 m) to ensure that the same number of dots always remained visible. Figure 1G–J shows some example optic flow fields from the dataset.

Training sets. The training set used to fit the decoders consisted of 900 videos. We sampled the gaze offset parameter on a regular grid to ensure balanced coverage (-35 to 35° in steps of 8.7°). We sampled path radius differently than the other parameters due to the nonlinear relationship between circular path radius and curvature (Eq. 1). We accounted for this by sampling small radii more finely than large radii using the following recursive generating formula:

$$r_{\text{next}} = 1.078r_{\text{prev}}. \quad (1)$$

We used a starting radius of 5 m and selected the coefficient so that Eq. (1) would generate 50 radii between 5 and 200 m. We doubled the number of radii by including both CW and CCW versions of each path.

Test sets. We generated a 500-video test set to assess the quality of predictions made by the fit decoders. Gaze offset was sampled from a uniform distribution [-35 , 35°]. We used kernel density estimation (KDE) to randomly sampled path radii in proportion to the training set values. As in the training set, this biased values toward smaller radii that produce larger variation in path curvature. We fit the path radii used in the training set to an empirical probability density function via the MATLAB function *ksdensity*, which we then sampled

within the training set range with 0.1 m granularity. As in the training set, we also included CCW versions of these paths.

For Fig. 4, we generated optic flow sequences that emulated the gaze-along-heading condition from Li and Cheng⁴⁵: a gaze offset of 0° and 28, 38, and 58 m path radii. We simulated 20 repetitions of each path radius condition across which only the random placement of dots varied. The Li and Cheng test set consisted of 60 videos.

3D dot cloud environment. For the simulations described in Supporting Information, we used a dataset that satisfies the same train and test set parameters, except that we distributed the 2000 dots in a 3D cloud rather than on a ground plane. These simulations were the basis for our comparison with the Burlingham and Heeger²⁸ data. Burlingham and Heeger used simulated curvilinear self-motion with 1.5 m/s translation and 0, 0.8, or 2.0°/s rotation. The latter two curvilinear cases correspond to path radii of 107 m and 43 m, respectively. We compared corresponding human judgments of gaze offset with model predictions obtained on the 3D cloud test set stimuli that had a path radius within 2.5 m of the radii used by Burlingham and Heeger. To make a comparison to the straight path condition (0°/s rotation), we considered the predictions made for test stimuli with path radii ≥ 180 m. For such large path radii, the influence of rotation is small (~ 0.45 °/s rotation) and the optic flow appears effectively radial.

Decoding. We fit a linear model with lasso regularization to the neural activations produced by the neural model (described below) at the end of each video (frame 10) in the optic flow training set. The models for decoding path curvature and gaze offset made use of the *lasso* MATLAB function. We set the regularization parameter such that the mean squared error (MSE) of the fit was within 1 standard error of the minimum MSE computed with fivefold cross validation. To fit the Full Limited model, we used the *DFMax* argument of the *lasso* function to constrain the number of nonzero weights in the fit to those from the Radial Only model. We classified path sign according to a support vector machine binary classifier using the *fitcsvm* MATLAB function configured with the default linear kernel.

The values depicted in Fig. 6A,B reflect the outcome of bootstrap procedure performed 50 times. For each bootstrap, we simulated each model with optic flow sampled with replacement from the original dataset while preserving the total number of samples. Figure 6A,B shows the MAE and its standard deviation computed over each set of 50 bootstrap estimates.

Neural model of MT and MSTd. We processed the optic flow datasets with the Competitive Dynamics (CD) neural model of MT and MSTd (Fig. 7)^{22,35,50,51}. Populations of model neurons that emulate a range of physiologically supported properties transform the optic flow into neural signals from which we subsequently decode the parameters describing the observer's self-motion. The CD model accounts for human judgments of self-motion, including in the presence of independently moving objects^{22,52}, and object motion during self-motion^{50,51,53}. We made several key changes to the most recent version of the model^{50,51}:

- Given focus in the present work on self-motion estimation, we only simulated the requisite MT⁺-MSTd pathway of the model, excluding the MT-MSTv object motion pathway^{50,51}.
- We simulated units tuned to direction and speed only, omitting disparity tuning.
- Each MSTd unit received input from a random subset of the MT⁺ units²⁰. In prior versions of the model, MSTd units were fully connected to the constellation of MT⁺ units that compose the preferred pattern.
- For greater clarity and to better manage the greater diversity of cell types simulated (e.g. radial, spiral, concentric), we divided model MSTd into several processing stages (layers). The first stage integrates the match between each optic flow template over time. This is followed by spatial smoothing of the activation of units tuned to similar CoMs and patterns. Finally, units compete with one another in an on-center/off-surround contrast-enhancing network.

These changes and the complete model details will be fully described in the following sections. Figure 7 summarizes the architecture of the CD model simulated here and Table 1 specifies the values of parameters used to simulate the neural model.

Model area MT. We simulated the population of direction and speed tuned neurons in MT with excitatory surrounds (MT⁺), which project to MSTd^{16,54,55} and influence self-motion signals therein⁵⁶.

MT⁺ unit receptive fields (RFs) tuned to like speeds and directions were arranged in a regular 64×64 spatial grid across the visual field.

MT speed tuning. We simulated units tuned to 5 (N^s) different preferred speeds. Each unit's speed tuning followed a log-normal distribution:

$$S(s_{x,y}; v_{\text{pref}}, s_0, \sigma_v) = \exp\left(-\frac{\log\left(\frac{s_{x,y}+s_0}{v_{\text{pref}}+s_0}\right)^2}{2\sigma_v^2}\right), \quad (2)$$

where σ_v defines the speed tuning bandwidth; s_0 defines a non-negative offset parameter to prevent the singularity in the logarithm at 0; and v_{pref} defines the preferred speed of the model neuron. Given that MT neuron speed tuning varies considerably, we sampled values from probability distributions that approximate neurophysiological

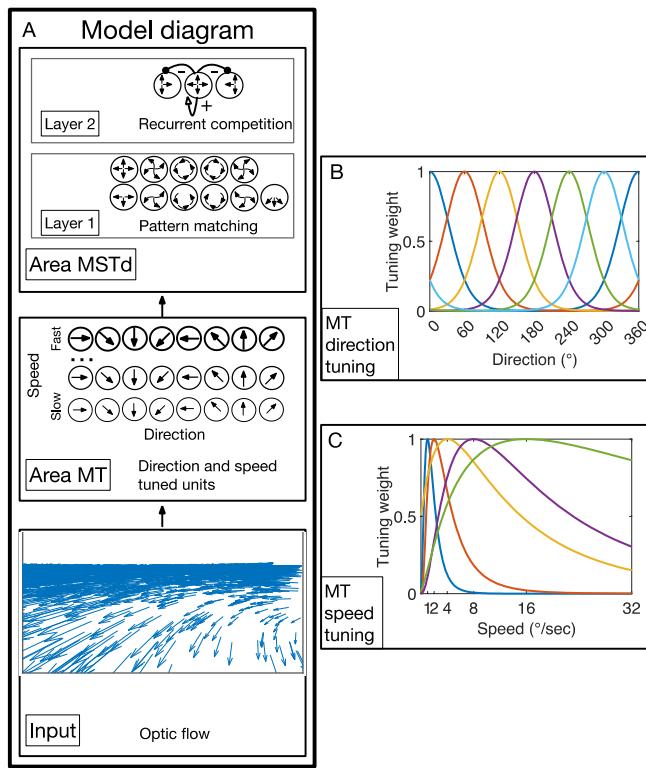


Figure 7. Overview of the Competitive Dynamics neural model of MT-MSTd. (A) Model area MT contains direction and speed tuned units. Units in model area MSTd integrate motion signals from MT consistent with the preferred motion pattern and compete with one another. MSTd units are tuned to full and hemi-field motion in a pattern continuum: radial, CCW spirals, CCW centers, CW spirals, CW centers, and ground flow. (B) Example MT direction tuning curves. (C) Example MT speed tuning curves.

fits to these parameters. Based on Fig. 4 of Ref.⁵⁷, we sampled σ_v from a Gaussian distribution (mean: 1.16° , SD: 0.5°) and s_0 from an exponential distribution ($\lambda: 0.25^\circ/\text{s}$). Consistent with Fig. 8 of Nover et al., we sampled v_{pref} from five uniform distributions with endpoints that yielded octave spaced bins (see Table 1).

We computed optic flow speed at retinotopic position (x, y) as:

$$s_{x,y} = \sqrt{u_{x,y}^2 + v_{x,y}^2}, \quad (3)$$

where $u_{x,y}$ and $v_{x,y}$ represent the horizontal and vertical components of the optic flow vector.

MT direction tuning. We simulated units tuned to $24 (N^d)$ motion directions (Fig. 7). We modeled each unit's direction tuning width using the following von Mises distribution

$$V(d; \mu_{\text{dir}}, \eta) = \frac{e^{\eta \cos(d - \mu_{\text{dir}})}}{e^\eta}, \quad (4)$$

where d represents the direction of the optic flow in the receptive field, μ_{dir} represents the unit's preferred direction, the shape parameter η controls the direction tuning width. We set $\eta = 3^\circ$ to approximate the 90° full-width at half-maximum bandwidth of MT neurons^{6,19}. We computed the direction of the optic flow vector $\vec{v}_{x,y} = (u_{x,y}, v_{x,y})$ at position (x, y) via the two-argument form of the arctangent function:

$$\Theta_{x,y} = \text{atan2}(v_{x,y}, u_{x,y}). \quad (5)$$

MT unit activation. The net input to the MT^+ unit tuned to speed s , direction d , and has its RF center at position (x, y) is the product of the directional $V(\Theta_{x,y}; d, \eta)$ and speed $S(s_{x,y}; s, s_0, \sigma_v)$ tuning curve outputs:

$$I_{s,x,y,d}^{\text{MT}} = V(\Theta_{x,y}; d, \eta)S(s_{x,y}; s, s_0, \sigma_v). \quad (6)$$

We modeled the dynamics of MT^+ unit $m_{s,x,y,d}$ as a simple leaky integrator:

Parameter	Description	Value
N^s	Number of preferred speeds	5
σ_v	Speed tuning bandwidth	Normal($\mu = 1.16^\circ$, $\sigma = 0.5^\circ$)
s_0	Non-negative offset parameter to prevent singularity in speed logarithm	Exp($\lambda = 0.25^\circ/s$)
v_{pref}	Preferred speed of MT neuron	U(0.5, 2.0)°/s, U(2.0, 4.3)°/s, U(4.3, 7.6)°/s, U(7.6, 12.7)°/s, U(12.7, 32.0)°/s
N^d	Number of preferred directions	24
—	Shape of MT ⁺ speed layers	$64 \times 64 \times 24 \times 5$
η	Von Mises direction tuning curve bandwidth parameter	3°
α^{MT}	MT ⁺ passive decay rate	1
τ^{MT}	Depressing synapse time constant	0.1
κ^{MT}	Rate at which the synaptic efficacy decreases	10
$N^{\text{MSTD,com}}$	Number of units tuned to CoM positions for each motion pattern (radial, spiral, etc.) arranged in regular across the visual field	$16 \times 16 = 256$
$\vec{\lambda}^{\text{spl}}$	Spirality of patterns to which spiral cells are tuned. Negative (positive) values produce CCW (CW) spirals	$\pm [0.0, 0.05, \dots, 1.0]$
N^{spl}	Number of CW and CCW spiral patterns to which units are tuned	21 (CW) + 21 (CCW) = 42
—	Indices of units in pattern tuning continuum. Units corresponding to the end-points exhibit tuning to radial expansion or CW/CCW center patterns	CCW full-field spiral (decreasing spirality): 1–21 CW full-field spiral (increasing spirality): 22–42 CW hemi-field spiral (decreasing spirality): 43–63 CCW hemi-field spiral (increasing spirality): 64–84
—	Shape of each MSTd layer	$16 \times 16 \times 84$
N^T	Number of MT ⁺ connections within each MSTd pattern template	200
b^{MSTD}	Rate at which MT ⁺ connection weights decrease with distance from the preferred CoM position	$1e-3$
$\alpha^{\text{MSTD,1a}}$	Layer 1a unit passive decay rate	1
$\vec{\sigma}^{\text{MSTD,1b,com}}$	Layer 1b Gaussian 2D spatial smoothing kernel horizontal and vertical standard deviations	[5, 5]
$r^{\text{MSTD,1b,com}}$	Diameter of Layer 1b Gaussian 2D spatial smoothing kernel	8
$\sigma^{\text{MSTD,1b,pat}}$	Layer 1b Gaussian 1D pattern smoothing standard deviation	1.5
$r^{\text{MSTD,1b,pat}}$	Diameter of Layer 1b Gaussian 1D pattern smoothing kernel	7
$\alpha^{\text{MSTD,1b}}$	Layer 1b unit passive decay rate	1
$\alpha^{\text{MSTD,c}}$	Layer 2 net input adaptive threshold passive decay rate	1
$\alpha^{\text{MSTD,2}}$	Layer 2 unit passive decay rate	10
$\beta^{\text{MSTD,2}}$	Layer 2 unit excitatory upper bound	3

Table 1. Parameters used in the neural model.

$$\frac{dm_{s,x,y,d}}{dt} = -\alpha^{\text{MT}} m_{s,x,y,d} + (1 - m_{s,x,y,d}) I_{s,x,y,d}^{\text{MT}}, \quad (7)$$

where α^{MT} is the passive decay rate of the cell.

MT output. Among models that explain MSTd responses based on their feedforward input from MT, those that include a nonlinearity that compresses MT signals perform best²⁶. The compressive nonlinearity could be explained by synaptic depression, the tendency for the same inputs to lose their efficacy over time. We modeled MT⁺ synaptic depression $h_{s,x,y,d}$ as follows:

$$\frac{1}{\tau^{\text{MT}}} \frac{dh_{s,x,y,d}}{dt} = 1 - h_{s,x,y,d} \left(1 + \kappa^{\text{MT}} h_{s,x,y,d} \right), \quad (8)$$

$$O_{s,x,y,d} = h_{s,x,y,d} m_{s,x,y,d}, \quad (9)$$

where $O_{s,x,y,d}$ denotes the output signal from MT⁺ units to MSTd, τ^{MT} is the synaptic time constant, and κ^{MT} represents the rate at which the efficacy of the input signal $m_{s,x,y,d}$ declines over time⁵⁸.

MSTd net input. We simulated MSTd cells tuned to radial expansion ($\vec{\Lambda}_{\text{rad},i,j,x,y}$) and spiral ($\vec{\Lambda}_{\text{spl},i,j,x,y}$) optic flow patterns that have been used to characterize MSTd selectivity in neurophysiological studies^{5,11}.

$$\vec{\Lambda}_{\text{rad},i,j,x,y} = (v_{\text{rad},dx}, v_{\text{rad},dy}) = (x - i), (y - j), \quad (10)$$

$$\vec{\Lambda}_{\text{cir},i,j,x,y} = (v_{\text{cir},dx}, v_{\text{cir},dy}) = (y - j), -(x - i), \quad (11)$$

$$\vec{\Lambda}_{\text{spl},i,j,x,y} = (v_{\text{spl},dx}, v_{\text{spl},dy}) = (1 - \lambda^{\text{spl}}) \vec{\Lambda}_{\text{rad},i,j,x,y} + \lambda^{\text{spl}} \zeta^{\text{spl}} \vec{\Lambda}_{\text{cir},i,j,x,y}. \quad (12)$$

In Eqs. (10–12), (i, j) indicates the location of the center of motion (CoM), which corresponds to the FoE in the case of radial expansion; (x, y) indicates visuotopic position of the local motion vector, ζ^{spl} is 1 for CW spirals and -1 for CCW spirals; and λ^{spl} is a scalar selected in the range $0–1$ that represents the spirality of the pattern. Setting $\lambda^{\text{spl}} = 0$ produces a radial expansion flow field, $\lambda^{\text{spl}} = 1$ produces a circular center flow field ($\vec{\Lambda}_{\text{cir},i,j,x,y}$), and $0 < \lambda^{\text{spl}} < 1$ creates spiral patterns spanning a continuum between these two extremes. We created MSTd units tuned to N^{spl} spiral patterns by sampling regularly spaced values of λ^{spl} (see Table 1). We simulated MSTd units tuned to both full-field and lower the hemifields. The latter was achieved by removing motion vectors above the CoM.

We implemented MSTd tuning to the preferred pattern using direction templates that select MT^+ signals when they appear in appropriate spatial locations. For example, in the case of a cell tuned to radial expansion with a centrally positioned FoE, the rightward direction template pools the responses of MT^+ cells tuned to rightward motion when their receptive fields coincide with the right side of the visual field. The following equations define the RF template $T_{\psi,d,i,j,x,y}$ for pattern ψ that integrates MT^+ cells tuned to direction d , normalized by the total number of pooled cells ($\hat{x}\hat{y}$):

$$\chi_{\psi,d,i,j,x,y} = \text{atan2}(\Lambda_{\psi,i,j,x,y,dy}, \Lambda_{\psi,i,j,x,y,dx}), \quad (13)$$

$$\hat{T}_{\psi,d,i,j,x,y} = \begin{cases} 1 & \frac{2\pi(d-1-\hat{d})}{\hat{d}} < \chi_{\psi,d,i,j,x,y} < \frac{2\pi(d-\hat{d})}{\hat{d}}, \\ 0 & \text{otherwise} \end{cases}, \quad (14)$$

$$T_{\psi,d,i,j,x,y} = \frac{1}{\hat{x}\hat{y}} \hat{T}_{\psi,d,i,j,x,y}. \quad (15)$$

In Eq. (14), \hat{d} indexes the N^d MT preferred directions.

Within each pattern template, we sampled a small subset of the possible connections from MT^+ units (e.g. 200 instead of $64 \times 64 \times 24 = 98304$). Sparse connectivity is more biologically plausible and substantially improves the algorithmic efficiency of the model. We implemented this by randomly selecting only N^T connections to MT^+ units within each template tuned to pattern ψ with CoM positioned at (i, j) . After drawing N^T pairs of (x, y) indices per template, we selected directions at each spatial location according to the uniform distribution $U[1, N^d]$. We denote the set of N^T three-tuple (x, y, d) indices of MT^+ units sampled within each template Ω . We repeated this process for each of the 5 (N^S) preferred speeds. This transformed the original $64 \times 64 \times 24 \times 5$ signal into one with shape $N^d \times N^T$.

The following equation matches the direction templates with the output signals from $MT^+(O_{s,x,y,d})$ at the set of sampled spatial and directional indices $(x, y, d) \in \Omega$ to compute the net input ($R_{\psi,i,j}$) to the MSTd unit that prefers the pattern ψ with a CoM positioned at (i, j) :

$$R_{\psi,i,j} = \frac{1}{N^T} \sum_s \sum_{(x,y,d) \in \Omega} e^{-b^{\text{MSTd}}((x-i)^2 + (y-j)^2)} \frac{T_{\psi,d,i,j,x,y} O_{s,x,y,d}}{\max_d O_{s,x,y,d}}. \quad (16)$$

In Eq. (16), the exponential function makes MSTd units more sensitive to motion nearby the preferred CoM position and the parameter b^{MSTd} modulates how sensitivity decreases with distance^{22,32,35}. Dividing by the maximum directional input at each position ensures that dominant inputs make comparable contributions across visuotopic space before the exponential distance weighting from the CoM is applied.

MSTd layer 1a. We arranged MSTd units into separate layers for organizational clarity. Note that these layers do not correspond to anatomical laminae. Layer 1 corresponds to MSTd units that are driven by their pattern tuning and perform spatiotemporal averaging of their input $R_{\psi,i,j}$. We divided the temporal (Layer 1a) and spatial (Layer 1b) averaging into two sequential stages. Layer 2 includes recurrent competitive connections that normalize the total activation while enhancing the contrast between strongly and weakly active MSTd units. We rely on the signals produced by Layer 2 for decoding the parameters specifying the observer's self-motion.

Dynamics in Layer 1a average each MSTd unit's net input (Eq. 16) over time:

$$\frac{da_{\psi,i,j}}{dt} = -\alpha^{\text{MSTd,1a}} a_{\psi,i,j} + (1 - a_{\psi,i,j}) R_{\psi,i,j}. \quad (17)$$

MSTd layer 1b. Units in MSTd Layer 1b pool the activation among units tuned to similar CoM locations and patterns along the spiral continuum (Fig. 1F). We achieved this by convolving the spatial grid of signals emanating from Layer 1a units with a two-dimensional Gaussian $G_2(x, y; \vec{\mu}, \vec{\sigma})$

$$G_2(x, y; \vec{\mu}, \vec{\sigma}) = \frac{1}{2\psi_x\sigma_y} e^{-\left(\left(\frac{x-\mu_x}{\sqrt{2}\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sqrt{2}\sigma_y}\right)^2\right)}, \quad (18)$$

$$I_{\psi,i,j}^{\text{MSTd,1b}} = \sum_n \sum_m G_2(n, m; (i, j), \vec{\sigma}^{\text{MSTd,1b,com}}) a_{\psi,i-n,j-m}, \quad (19)$$

and convolving the pattern signals at each CoM position with a one-dimensional Gaussian:

$$G_1(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)^2}, \quad (20)$$

$$J_{\psi,i,j}^{\text{MSTd,1b}} = \sum_n G_1(n; \psi, \sigma^{\text{MSTd,1b,pat}}) I_{\psi,i,j}^{\text{MSTd,1b}}. \quad (21)$$

Because we arranged units tuned to similar patterns along a circular spiral pattern continuum (see Table 1), we wrapped the convolution in Eq. (21) around the boundaries.

The dynamics of Layer 1b units obey:

$$\frac{db_{\psi,i,j}}{dt} = -\alpha^{\text{MSTd,1b}} b_{\psi,i,j} + (1 - b_{\psi,i,j}) J_{\psi,i,j}^{\text{MSTd,1b}}. \quad (22)$$

MSTd layer 2. The input to Layer 2 undergoes an adaptive threshold to suppress signals weaker than the mean activation value $c_{\psi,i,j}$ achieved among units tuned to the same pattern over the recent time history

$$\frac{dc_{\psi}}{dt} = -\alpha^{\text{MSTd,c}} c_{\psi} + (1 - c_{\psi}) \left(\frac{1}{N^{\text{MSTd,com}}} \sum_m \sum_n b_{\psi,m,n} \right), \quad (23)$$

where $N^{\text{MSTd,com}}$ represents the total number of units tuned to the same pattern with differing CoM positions. The net input to Layer 2 is the Layer 1b output signal with the threshold applied:

$$I_{\psi,i,j}^{\text{MSTd,2}} = \lceil b_{\psi,i,j} - c_{\psi} \rceil^+, \quad (24)$$

where $\lceil \cdot \rceil^+$ represents the half-wave rectification $\max(\cdot, 0)$.

Layer 2 integrates signals from Layer 1b and contains recurrent connections within the layer that normalize activation, facilitate competition, and stabilize signals under changing conditions²². Units obey the recurrent competitive field with a contrast-enhancing feedback transfer function $f(w) = w^2$, where w represent the unit activation⁵⁹:

$$\begin{aligned} \frac{dz_{\psi,i,j}}{dt} &= -\alpha^{\text{MSTd,2}} z_{\psi,i,j} + (\beta^{\text{MSTd,2}} - z_{\psi,i,j})(f(z_{\psi,i,j}) + D_{\psi,i,j}E) \\ &\quad - z_{\psi,i,j} \left(\sum_{m \neq i} \sum_{n \neq j} f(z_{\psi,m,n}) + \sum_{p \neq \psi} \sum_{m \neq i} \sum_{n \neq j} D_{p,m,n} \right). \end{aligned} \quad (25)$$

In Eq. (25), $\beta^{\text{MSTd,2}}$ represents the excitatory upper bound of each unit and the terms involving $f(\cdot)$ implement competition among units tuned to the same pattern. The terms $D_{\psi,i,j}E$ and $\sum_{p \neq \psi} \sum_{m \neq i} \sum_{n \neq j} D_{p,m,n}$ describe excitatory and inhibitory recurrent connections, respectively, that subserve activity normalization across the layer⁶⁰. In the former term, $D_{\psi,i,j}$ denotes the difference between each unit's current activation and the feedforward input signal:

$$D_{\psi,i,j} = \lceil z_{\psi,i,j} - I_{\psi,i,j}^{\text{MSTd,2}} \rceil^+. \quad (26)$$

The variable E defines the discrepancy between input and activation across the entire layer:

$$E = \sum_p \sum_m \sum_n D_{p,m,n}. \quad (27)$$

This means that the network attempts to conserve the total activation across the entire layer, not just among units tuned to the same pattern.

Data availability

Competitive Dynamics model activations and the decoding model code are available on Github at <https://github.com/owlayton/R-Curvilinear-Path-Encoding>. The ground optic flow dataset is available on OSF, https://osf.io/su8h3/?view_only=0a18371f0828412f91895512a6656ea4.

Received: 7 April 2022; Accepted: 8 July 2022

Published online: 04 August 2022

References

1. Hubel, D. H. Exploration of the primary visual cortex, 1955–78. *Nature* **299**, 515–524 (1982).
2. Hubel, D. & Wiesel, T. David Hubel and Torsten Wiesel. *Neuron* **75**, 182–184 (2012).
3. Gibson, J. J. *The Perception of the Visual World*. (Houghton Mifflin, 1950).
4. Gu, Y., Watkins, P. V., Angelaki, D. E. & DeAngelis, G. C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J. Neurosci.* **26**, 73–85 (2006).
5. Duffy, C. J. & Wurtz, R. H. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J. Neurosci.* **15**, 5192–5208 (1995).
6. Britten, K. H. & van Wezel, R. J. A. Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nat. Neurosci.* **1**, 59 (1998).
7. Britten, K. H. & van Wezel, R. J. A. Area MST and heading perception in macaque monkeys. *Cereb. Cortex* **12**, 692–701 (2002).
8. Gu, Y., DeAngelis, G. C. & Angelaki, D. E. Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. *J. Neurosci.* **32**, 2299–2313 (2012).
9. Longuet-Higgins, H. C. & Prazdny, K. The interpretation of a moving retinal image. *Proc. R. Soc. Lond. B* **208**, 385–397 (1980).
10. Xu, H., Wallisch, P. & Bradley, D. C. Spiral motion selective neurons in area MSTd contribute to judgments of heading. *J. Neurophysiol.* **111**, 2332–2342 (2014).
11. Graziano, M. S., Andersen, R. A. & Snowden, R. J. Tuning of MST neurons to spiral motions. *J. Neurosci.* **14**, 54–67 (1994).
12. Duffy, C. J. & Wurtz, R. H. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* **65**, 1329–1345 (1991).
13. Siegel, R. M. & Read, H. L. Analysis of optic flow in the monkey parietal area 7a. *Cereb. Cortex* **7**, 327–346 (1997).
14. Raffi, M. & Siegel, R. M. A functional architecture of optic flow in the inferior parietal lobule of the behaving monkey. *PLoS One* **2**, e200 (2007).
15. Schaafsma, S. J. & Duysens, J. Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their. *J. Neurophysiol.* **75**, 4056–4068 (1996).
16. Maunsell, J. H. & van Essen, D. C. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586 (1983).
17. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **58**, 267–288 (1996).
18. Hastie, T., Tibshirani, R., Friedman, J. H. & Friedman, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009).
19. Beyeler, M., Dutt, N. & Krichmar, J. L. 3D visual response properties of MSTd emerge from an efficient, sparse population code. *J. Neurosci.* **36**, 8399–8415 (2016).
20. Layton, O. W., Steinmetz, S., Powell, N. & Fajen, B. R. Computational investigation of sparse MT-MSTd connectivity and heading perception. *J. Vis.* **19**, 237a (2019).
21. Perrone, J. A. Model for the computation of self-motion in biological systems. *JOSA A* **9**, 177–194 (1992).
22. Layton, O. W. & Fajen, B. R. Competitive dynamics in MSTd: A mechanism for robust heading perception based on optic flow. *PLoS Comput. Biol.* **12**, e1004942 (2016).
23. Royden, C. S. Computing heading in the presence of moving objects: A model that uses motion-opponent operators. *Vis. Res.* **42**, 3043–3058 (2002).
24. Browning, N. A., Grossberg, S. & Mingolla, E. A neural model of how the brain computes heading from optic flow in realistic scenes. *Cogn. Psychol.* **59**, 320–356 (2009).
25. Yumurtaci, S. & Layton, O. W. Modeling physiological sources of heading bias from optic flow. *eNeuro* **8**, (2021).
26. Mineault, P. J., Khawaja, F. A., Butts, D. A. & Pack, C. C. Hierarchical processing of complex motion along the primate dorsal visual pathway. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E972–E980 (2012).
27. Li, L. & Cheng, J. C. K. Perceiving path from optic flow. *J. Vis.* **11**, 22–22 (2011).
28. Burlingham, C. S. & Heeger, D. J. Heading perception depends on time-varying evolution of optic flow. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33161–33169 (2020).
29. Takahashi, K. *et al.* Multimodal coding of three-dimensional rotation and translation in area MSTd: Comparison of visual and vestibular selectivity. *J. Neurosci.* **27**, 9742–9756 (2007).
30. Gu, Y., Fetsch, C. R., Adeyemo, B., Deangelis, G. C. & Angelaki, D. E. Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron* **66**, 596–609 (2010).
31. Laughlin, S. A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung c* **36**, 910–912 (1981).
32. Warren, W. H. & Saunders, J. A. Perceiving heading in the presence of moving objects. *Perception* **24**, 315–331 (1995).
33. Lappe, M. & Rauschecker, J. P. A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Comput.* **5**, 374–391 (1993).
34. Wang, R. A network model for the optic flow computation of the MST neurons. *Neural Netw.* **9**, 411–426 (1996).
35. Layton, O. W., Mingolla, E. & Browning, N. A. A motion pooling model of visually guided navigation explains human behavior in the presence of independently moving objects. *J. Vis.* **12**, 20–20 (2012).
36. Danz, A. D., Angelaki, D. E. & DeAngelis, G. C. The effects of depth cues and vestibular translation signals on the rotation tolerance of heading tuning in macaque area MSTd. *eNeuro* **7**, (2020).
37. Shenoy, K. V., Bradley, D. C. & Andersen, R. A. Influence of gaze rotation on the visual response of primate MSTd neurons. *J. Neurophysiol.* **81**, 2764–2786 (1999).
38. Inaba, N., Shinomoto, S., Yamane, S., Takemura, A. & Kawano, K. MST neurons code for visual motion in space independent of pursuit eye movements. *J. Neurophysiol.* **97**, 3473–3483 (2007).
39. Royden, C. S., Crowell, J. A. & Banks, M. S. Estimating heading during eye movements. *Vis. Res.* **34**, 3197–3214 (1994).
40. Warren, W. H. & Hannon, D. J. Eye movements and optical flow. *JOSA A* **7**, 160–169 (1990).
41. Cheng, J. C. & Li, L. Effects of reference objects and extra-retinal information about pursuit eye movements on curvilinear path perception from retinal flow. *J. Vis.* **12**, 12 (2012).
42. Perrone, J. A. Visual-vestibular estimation of the body's curvilinear motion through the world: A computational model. *J. Vis.* **18**, 1 (2018).
43. Perrone, J. A. & Krauzlis, R. J. Vector subtraction using visual and extraretinal motion signals: A new look at efference copy and corollary discharge theories. *J. Vis.* **8**, 24.1–2414 (2008).
44. Orban, G. A. *et al.* First-order analysis of optical flow in monkey brain. *Proc. Natl. Acad. Sci.* **89**, 2595–2599 (1992).
45. Li, L. & Warren, W. H. Jr. Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vis. Res.* **40**, 3873–3894 (2000).
46. Royden, C. S. Analysis of misperceived observer motion during simulated eye rotations. *Vis. Res.* **34**, 3215–3222 (1994).
47. Banks, M. S., Ehrlich, S. M., Backus, B. T. & Crowell, J. A. Estimating heading during real and simulated eye movements. *Vis. Res.* **36**, 431–443 (1996).
48. Saunders, J. A. & Ma, K. Y. Can observers judge future circular path relative to a target from retinal flow. *J. Vis.* **11**, 16 (2011).

49. Raudies, F. & Neumann, H. A review and evaluation of methods estimating ego-motion. *Comput. Vis. Image Underst.* **116**, 606–633 (2012).
50. Layton, O. W. & Fajen, B. R. Computational mechanisms for perceptual stability using disparity and motion parallax. *J. Neurosci.* **40**, 996–1014 (2020).
51. Layton, O. W. & Niehorster, D. C. A model of how depth facilitates scene-relative object motion perception. *PLoS Comput. Biol.* **15**, e1007397 (2019).
52. Layton, O. W. & Fajen, B. R. Possible role for recurrent interactions between expansion and contraction cells in MSTd during self-motion perception in dynamic environments. *J. Vis.* **17**, 5 (2017).
53. Layton, O. W. & Fajen, B. R. A neural model of MST and MT explains perceived object motion during self-motion. *J. Neurosci.* **36**, 8093–8102 (2016).
54. Beregovskii, V. K. & Born, R. T. Specificity of projections from wide-field and local motion-processing regions within the middle temporal visual area of the owl monkey. *J. Neurosci.* **20**, 1157–1169 (2000).
55. Born, R. T. & Tootell, R. B. H. Segregation of global and local motion processing in primate middle temporal visual area. *Nature* **357**, 497 (1992).
56. Yu, X., Hou, H., Spillmann, L. & Gu, Y. Causal evidence of motion signals in macaque middle temporal area weighted-pooled for global heading perception. *Cereb. Cortex* **28**, 612–624 (2018).
57. Nover, H., Anderson, C. H. & DeAngelis, G. C. A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *J. Neurosci.* **25**, 10049–10060 (2005).
58. Grossberg, S. Intracellular mechanisms of adaptation and self-regulation in self-organizing networks: the role of chemical transducers. *Bull. Math. Biol.* **42**, 365–396 (1980).
59. Grossberg, S. Contour enhancement, short term memory, and constancies in reverberating neural networks. *Stud. Appl. Math.* **52**, 213–257 (1973).
60. Boardman, I. & Bullock, D. A neural network model of serial order recall from short-term memory. *Int. Jt. Conf. Neural Netw.* **2**, 879–884 (1991).

Author contributions

O.W.L. and B.R.F. conceptualized the research; O.W.L. curated the data; O.W.L. programmed the model; O.W.L. and B.R.F. analyzed the results; O.W.L. wrote the manuscript; O.W.L. and B.R.F. revised and reviewed the manuscript.

Funding

This study was funded by Office of Naval Research, United States (Grant no. ONR N00014-18-1-2283).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16371-4>.

Correspondence and requests for materials should be addressed to O.W.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022