Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Steerable autoencoders underlying remapping, spatiotopy, and visual stability

Patrick Cavanagh[1,2,3] and David Melcher[4,5]

The encoding of locations in world coordinates is often called spatiotopy; it captures how we experience the layout of things around us. The visual system, on the other hand, appears to rely on retinotopic representations. Here we review earlier suggestions that replace spatiotopic maps with the tracking or remapping of a small number of attended targets to keep their positions updated on retinotopic maps, combined with suppression of motion responses to the shift of the retinal image. Importantly, we propose that each attended target's identity is connected to the locations of its features in early visual cortex through a steerable autoencoder network that shifts the feature activity to the target's new location with each saccade, explaining the many reports of trans-saccadic integration. This autoencoder process is part of the overall architecture of the visual system, acting to link target properties together, even as the target or the eyes move.

Addresses
[1] Centre for Vision Research, York University, Toronto, ON, M3J1P3, Canada
[2] Department of Psychology, Glendon College, Toronto, ON, M4N3M6, Canada
[3] Psychological and Brain Sciences, Dartmouth College, Hanover, NH, 03755, USA
[4] Psychology Program, New York University Abu Dhabi, United Arab Emirates
[5] Center for Brain & Health, New York University Abu Dhabi, United Arab Emirates

Corresponding author: Cavanagh, Patrick (patcav1@yorku.ca)

Current Opinion in Neurobiology 2026, **99**:103221

This review comes from a themed issue on **VSI: Computational Neuroscience 2026**

Edited by **Li Zhaoping** and **Carl Petersen**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.conb.2026.103221

Introduction

The world around us is reassuringly stable, aside from instances of earthquakes and rocking boats at sea. Things are mostly seen to remain where they are as we move around, avoiding obstacles and finding our way. The encoding of locations in world coordinates is often called spatiotopy and it captures how we experience the layout of things around us. However, the brain codes the representation of locations primarily in retinotopic coordinates—based on the locations of objects on the retina—and on corresponding retinotopic maps throughout most of the visual system. With every change in the direction of gaze due to an eye movement (a saccade) and/or head movement, the image on the retina shifts and the objects of interest are all at new locations. Helmholtz [1], among others, proposed that a reverse copy of the eye movement commands (an efference copy) could compensate for this shift and create a stable, spatiotopic representation. However, the search for that spatiotopic map has not met with much success. If not spatiotopic maps, then what? Spatiotopic maps offer two critical benefits. The first is the sense of stability. Because nothing moves on the spatiotopic map when the eyes move, there would be no sense of motion to ruin the appearance of stability. The second benefit is trans-saccadic integration. An object's representation would remain at the same location on a spatiotopic map so that the processing of its identity and properties could continue across saccades without having to start again from zero.

Here we review proposals for how these two critical properties are achieved in the absence of a spatiotopic map. The most successful proposal is one of partial compensation based on remapping [2, and reviews 3–5], where the retinotopic locations of a few attended objects are updated or "remapped" with each saccade on a retinotopic map while ignoring other scene elements. We will describe how this tracking of a few targets, combined with trans-saccadic integration and motion suppression, anchors our stable visual experience as the eyes move. We then focus on the central puzzle underlying the tracking of targets: how does the visual system connect the features of a target to its identity and how is this location-to-identity link updated with each eye movement? To do so, we introduce recent work on autoencoders (a type of deep neural net) already proposed to underlie object-based attention [6,7] and we review new approaches that predict future target positions (e. g., PredNet, [8,9]) that are currently of great interest in many areas including the development of autonomous vehicles (e.g., Ref. [10]).

Maps

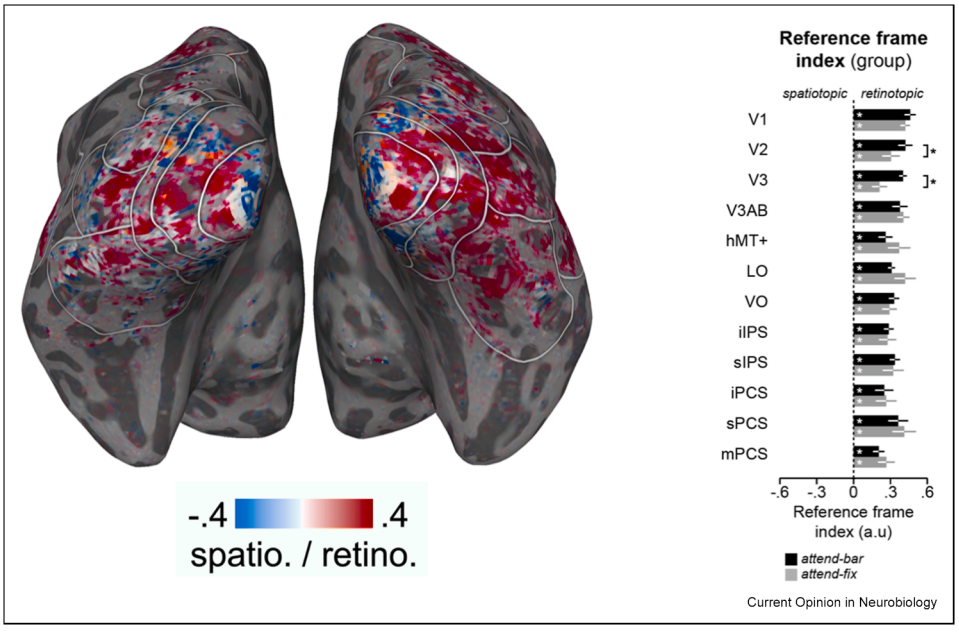
Surveys of spatial maps in visual areas of the brain have reported many topologically organized maps but all of them seem to be in the retinal coordinates (e.g., Ref. [11]; Figure 1). In 2007, (Ref. [12]) d'Avossa et al. reported that the medial temporal visual area (MT) coded locations on a spatiotopic map but these results have been challenged [13]. There has been evidence for spatiotopic representations for attended stimuli in several visual areas [14] (although see Ref. [11]) but only after some delay [15]. Because of the delay in updating this spatiotopic map, it cannot explain the experience of visual stability at the moment the eyes move—that requires a faster, predictive system like remapping [16]. If not spatiotopy, why would retinotopy, an awkward choice, be so widespread? First, retinotopic maps provide a common format across many visual areas and, combined with inter-area projections that respect this organization, they support a built-in processing pipeline where cross-module binding comes for free, based on co-localization. For example, the activity for an object of interest on a retinotopic saccade map projects to that object's features at the appropriate locations in retinotopic early visual cortex, and vice versa (e.g., Ref. [17]). Second, these retinotopic locations are, in practice, world locations. For example, microstimulation of a location on a retinotopic saccade map brings the corresponding location in the world onto the fovea (e.g., Refs. [18,19]). This demonstrates that the activity on a retinotopic map always "points to" a location in the world. All that is missing is something to keep track of a target's representation on the retinotopic map with changes in the direction of gaze, and somehow to link that activity to the target's identity that is held in basically non-retinotopic object areas [20,21].

Spatial representations may also rely on distributed coding approaches that have no explicit maps. For example, gain fields combine eye position information with the retinotopically organized responses of cortical units. These gain fields have been reported in parietal areas, V1, and other early visual areas [22—24]. The information about eye position could allow the effects of direction of gaze to be discounted to stabilize a representation of the visual field. Or possibly, simply by virtue of the presence of the gaze direction signal, the location of targets is represented in a distributed fashion without further decoding [25]. This distributed coding will be important in our last section when we consider the examples from deep neural nets, in particular, autoencoders.

Remapping

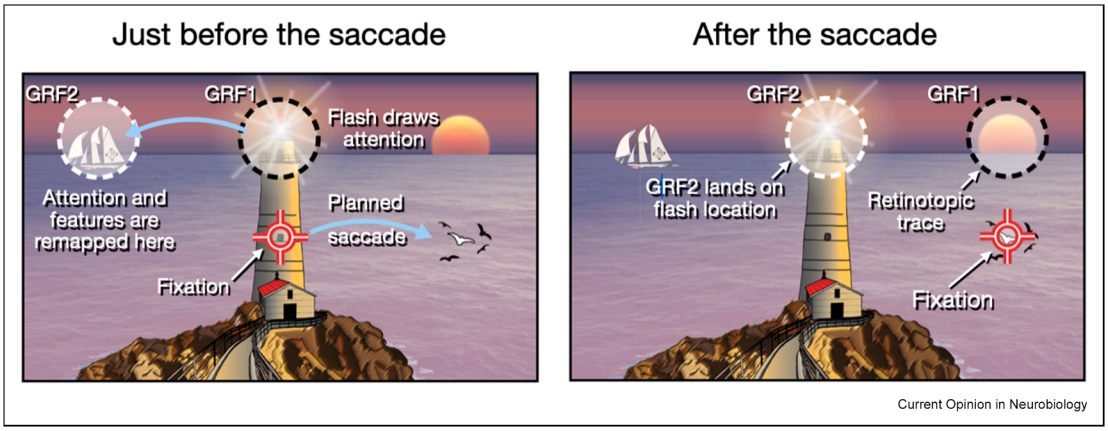
In the remapping proposal [2]; reviews, [3—5], only the locations of a few items are tracked with each eye movement, creating a manageable computational load. The evidence for this tracking, or remapping, was first reported by Duhamel, Colby and Goldberg [2]. Specif-

Figure 1



Maps in the visual system. There are many maps of the visual field across the visual system. The left panel shows a classification of local regions as retinotopic or spatiotopic during an eye movement task with attention to a bar target or the moving fixation point. Although there are hints of spatiotopically organized subregions, the analysis of individual visual areas (rightmost panel) shows only retinotopic organization (from Ref. [11]).

Figure 2



Remapping. **Left.** While a person is fixating on the lighthouse, they are also planning to shift their gaze to the flock of birds to the right. Just before this saccade, the beacon in the lighthouse flashes, triggering activity in a group of receptive fields GRF1 that cover that region of the retina, and drawing attention to that location. In preparation for the eye movement, the attention and feature activity from GRF1 are transferred to the group of receptive fields GRF2 to the left of the lighthouse where the beacon will fall after the saccade. Recent work (Kämmers et al., 2024; [31]) claims that the feature transfer happens principally for the saccade target (the birds here) while only attention is remapped for peripheral attended targets (the flash). **Right**. After the saccade, the preactivated units of GRF2 now align with the lighthouse beacon, providing a head start for the continued processing of its properties. The units of GRF1 are now centered over a different region and there is some evidence that there is a brief benefit from this retinotopic trace [32]. Images adapted from Ref. [33].

ically, just before a saccade, cells in the parietal and frontal areas (Figure 2) start responding to a stimulus far outside their receptive field if its location will fall in their receptive field after the eye movement. This suggests that some process is predicting the target's future location and somehow transferring its activity to that location from the cells that are currently responding to the target. This anomalous response is present only briefly before the saccade and, indeed, cells return to responding to their classical receptive fields once the saccade has landed. This momentary shift in response provides a remarkable glimpse into the inner workings of a tracking process that updates target locations. This finding has been contested [26] but appears to hold up for activity before and at the time of the saccade [27]. This updating applies not only to targets of eye movements but also to other attended stimuli in the visual field, although there is evidence that the saccade target itself is special and acts as the basic landmark for locations after the saccade [28]. Remapping and updating does not happen for unattended targets [29]. Instead, recordings in the lateral intraparietal area [30] suggests that attentional suppression of unattended targets is repositioned over their upcoming locations.

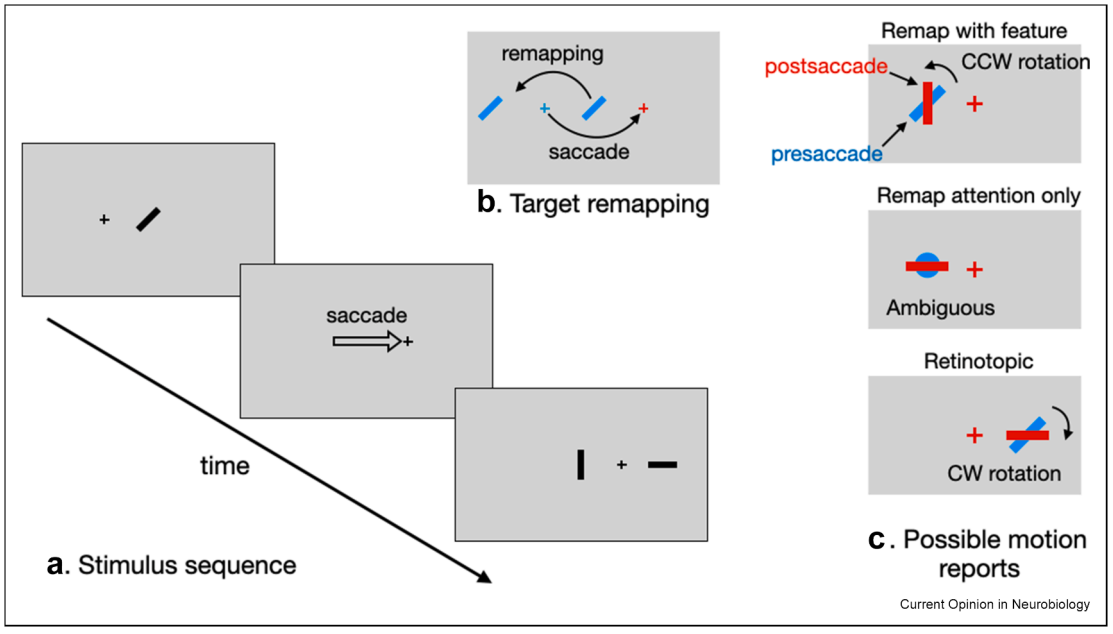
Next, we describe how this remapping process captures the two essential properties that simulate spatiotopy: suppression of motion responses and trans-saccadic integration.

Suppressing motion

Every change in gaze direction creates a massive shift of the visual field that would be expected to drive strong motion responses but does not. There are two parts to this motion silencing. First, much of the motion responses to the shift of scene elements should be signaled by the low-level motion system, based on direction-sensitive responses of units principally in areas V1 through V3A and MT [34]. Conveniently, these reflexive responses are suppressed (e.g., Refs. [35,36]) principally by masking from the onset and offset transients of the saccade [37,38]. Second, in contrast to the reflexive responses of low-level motion, the high-level motion system actively tracks targets with attention and generates an impression of motion when they move [39]. These high-level motion percepts can also be seen across saccades [40,41] as long as the displacement is large enough [42]. During a saccade, the expected location of an attended target is updated based on the retinal displacement that should be produced by the upcoming eye movement (e.g., Ref. [4]). If the target is then at or near its expected location after the eye movement (within an elliptical region equivalent to typical uncertainty of saccadic landing, [43]), no displacement of attention is detected, and consequently no motion is seen.

Finally, the shift of attention from the target's current location to its future location (driven by efference copy), is fully explained by the eye movement and so does not

Figure 3



Trans-saccadic feature remapping. Fracasso et al. [40] reported a compelling example of transfer of a target's feature, specifically orientation, during remapping. (**a**) Observers were presented with an obliquely oriented bar to the right of fixation and then they saccaded across the bar to the new fixation further to the right. The original bar disappeared just prior to the saccade onset. Once the saccade landed at the new fixation, two new bars were presented: one in the same spatial position as the original bar on the screen, rotated counterclockwise from it (in this example), and the other in the same retinal position with respect to fixation but rotated clockwise. Participants reported the direction of rotation of the bar. (**b**) Just before the saccade, remapping will move attention and possibly the feature (orientation) to the left of the fixation so their activity will line up with the same spatial location after the saccade. Blue indicates pre-saccdic items and red indicates post-saccadic items. (**c**) If the target is remapped with its feature (top panel), then the perceived rotation will be counterclockwise. If only attention is remapped, without orientation information (middle panel), rotation reports will be random. If the perceived motion is set by activity in receptive fields at the original retinal location, to the right of fixation, it will be clockwise. Observers consistently reported rotation of the bar according to the remapped, spatiotopic matching indicating that the orientation feature was remapped. Adapted from Ref. [40].

itself translate to any impression of motion in the world. As a result, nothing seems to have moved—no motion, no problem as Stalin might have said.

Trans-saccadic integration

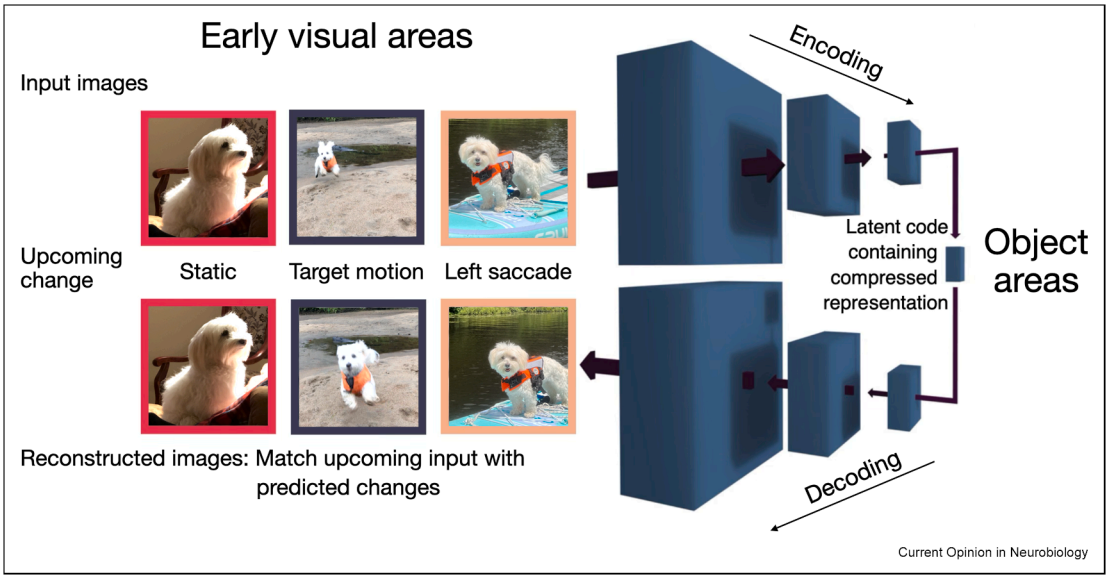
A central rationale for spatiotopy was to support the continuity of processing across a saccade, so that the target information accumulated before an eye movement could be linked to information accumulating after the eye movement. If this were not the case, the target's analysis would have to start again from zero after each saccade. Here we review the evidence that an attended target's information is integrated across the saccade, linking information from the current and future target locations, beginning just before the saccade.

Attention is remapped so that attention allocated to a target is passed on to its future location, even before the saccade [44,33]. A similar transfer of attentional benefits to the cued location is seen after the saccade (review, [45]; although see [46] and [32]; for evidence of retinotopic allocation in some conditions).

It has been proposed that this remapping of attention carries no feature information [3] but this claim has not held up as many studies have now shown transfer of feature information to the remapped location (Figure 3). For example, Kroell & Rolfs [31] showed pre-saccadic visual enhancement at the fovea for the orientation of the saccade target that was about to arrive at the fovea. Pre-saccadic effects are also seen at the remapped location for masking [47,48].

Other studies have shown that features presented before the saccade can influence the processing of features presented after the saccade. Examples include preview congruency effects (e.g., Refs. [49—51] and the spatial congruency bias (e.g., [52]). Trans-saccadic feature integration is also seen for visual features such as shape, color, and orientation of attended targets

Figure 4



Steerable autoencoder. An autoencoder takes an input pattern, here from early retinotopic cortex, passes it through several layers of encoding to generate a high-level, low-dimensional representation that we take to be the object areas of visual cortex. This then returns through decoding layers to produce an output pattern and the network's connections are adjusted until this output matches the input, a form of unsupervised learning. Versions of autoencoders and related architectures (e.g., Ref. [8]) use motion within the image to predict the subsequent images where an object in motion will be suitably displaced, represented here for the dog in motion. A steerable autoencoder would use distributed information about eye position or the upcoming saccade vector to predict object locations in the post-saccadic image, although only for the attended itemsand not the entire scene as depicted here. Adapted from Ref. [66].

(review, [53]). The integration follows the same maximum-likelihood principles as multisensory integration (review [54] as long as the pre- and post-saccadic information arecommensurate [55].

Further support for trans-saccadic integration has come from functional magnetic resonance imaging (fMRI) studies [56–58]. Importantly, feature information about the saccade target can be found at the fovea—the future location of the target—in the fMRI activation pattern [59] or with decoding techniques [60]. Electroencephalographic (EEG) and magnetoencephalographic (MEG) studies have also supported remapping of activity and the transfer of feature information [49,61,62].

Across all these studies, the evidence for trans-saccadic feature transfer is stronger for the saccade target than for other targets present when the saccade is made [31]. The evidence for trans-saccadic integration supports the remapping scheme where some updating processes link the pre-saccadic and post-saccadic locations of a few attended items, perhaps preferentially for the saccade target itself. These results and others showing spatiotopic effects suggest that there is a rapid, predictive, feature-specific mechanism such as remapping that simulates properties of spatiotopy for a few attended items [16]. However, labeling the process as "remapping" does not yet explain how it happens, how the target's features are linked to its identity and then transferred to the new location.

Autoencoders

Autoencoders [63,64] are a type of recurrent, deep neural net architecture with the ability to activate features at predicted locations, a property that we will exploit to suggest a mechanism for remapping (Figure 4). This has been used as a model of object-based attention [6,7,65] for predicting the next image in a sequence based on image motion [8,9]; and here we claim, for activating the features of a target at its remapped location at the time of a saccade.

A target's details—its identity, its features, and its location—are scattered across separate cortical areas and the puzzle of how they might be linked together—as an "object file" [67]—has a long history. A brief review of how this problem has been addressed for object-based attention will lay the groundwork for the proposed application of autoencoders to remapping. A partial linking of an object's properties is already achieved through the common format shared across retinotopic areas. The reciprocal and collimated connections between these areas links an object's activity on location-based priority or gain map [6] to the object's features

in early cortices. However, the link to the identity of the object is much less straightforward: How does an object's activity in areas like the lateral occipital complex (LOC), fusiform face area (FFA), and parahippocampal place area (PPA) project to that object's features in early cortex to boost the object's processing? Object areas and other ventral areas have at best only a rudimentary retinotopic format [20,21] and cells in these areas typically have very large receptive fields. It is difficult to see how these object-specific units could project activity to the precise location of the object's features in early visual cortex. Nevertheless, precise feedback to early areas is exactly what object-based attention does [6]. A solution based on autoencoders has been proposed [6,7,65] and a straightforward extension provides a solution for remapping: updating the link between an object's feature locations and its identity with each saccade. This proposal starts with the observation that even though individual units in object areas have very poor location information, position information can be decoded from their population responses in these areas (e.g., Ref. [68]).

The properties of autoencoder networks [64,66,69] provide the long-sought-after link between high-level identity and low-level location and feature information necessary to support object-based attention. In a neural implementation, we assume that the low-dimensional, latent values are the high-level, object representations of the ventral stream. Furthermore, to be comparable and high-dimensional, both input and output layers must both be in the same early visual areas of cortex. This generates a closed loop from early retinotopic areas up to object areas and then back to early retinotopic areas and embodies the basic property of object-based attention to enhance early processing of the attended object's features [6]. This architecture provides a compelling instance of linking a target's abstract identity to its early features and their locations using the diffusely distributed location information that is retained at the top, abstract object level. Can this autoencoder approach work to link identity and updated locations across saccades?

A partial step in that direction is the recent work on predictive autoencoders that has followed on the original demonstrations of PredNet [8]. These networks deal with video input and their goal is to match the output not the current input but the input in the next or subsequent frames. They manage very well, taking into account the motion information of each element in the scene to match its position in future frames. These networks have been adapted to many applications including autonomous vehicle control [10].

We propose that similar networks could base their predictions not on the motion of targets in the scene as in PredNet [8] but on distributed population information about gaze direction from gain fields (e.g., Ref. [24]) or efference copy gain fields (for saccade vectors rather than eye position) in object areas. This would enable predictions of where each attended target will be after a saccade and downward projections to that location would then activate the target's features there in advance. This "steerable" autoencoder would have continuous training to form these links based on the hundreds of thousands of saccades executed every day. The remapping function proposed here is not a separate system, it is a part of object-based attention that tracks targets and predicts their location as the target or the eyes move. We suggest that it is this neural architecture that activates the attended target's features at its post-saccadic location and that it is part of the overall architecture that links target identity with its low-level features, even as the target or the eyes move.

The proposal relies on distributed representations of gaze direction or upcoming saccade vectors in the object areas of cortex. There is no evidence for these, as yet, and to our knowledge, no studies have addressed the possibility. The gaze direction gain fields in parietal areas respond too slowly to make a contribution to visual continuity and visual stability [70]. The gaze direction gain fields in V1 [23] are in step with the eye movements but little or no remapping has been reported for units in V1 [71] or MT either (e.g., Ref. [72]. In contrast, pre-saccadic remapping is widespread in V2 [73]. An alternative would be that intrasaccadic motion [74], like the object motion used by PredNet [8], could support predictions of where the target would land. However, these predictions would only become available after the saccade.

Summary

We have followed on the earlier suggestions that in the absence of spatiotopic maps that can contribute effectively to visual stability at the time of each saccade, a remapping process fills in by tracking a small number of attended targets and updating their positions on retinotopic maps (e.g., Refs. [4,5]). This achieves a simulation of spatiotopy because the target locations are updated with each saccade, these target locations are directly connected to world coordinates (as shown by microstimulation in saccade areas), and motion responses to the shift of the retinal image are suppressed and discounted. Here we also propose that each attended target's identity is connected to the locations of its features in early visual cortex through a steerable autoencoder network that shifts the feature activity to the target's new location with each saccade. Unlike spatial attention (e.g., Ref. [17]), there is no salience map here, retinotopic or otherwise, the tracking of attended objects is instead based on the distributed coding of locations and gaze in a large scale, recurrent network.

Our proposal of steerable autoencoders to achieve the remapping step is highly speculative and we make it to provide a new option for understanding remapping and visual stability. It comes with several open questions. Will an autoencoder trained on a large image set generate object representations at its highest level? The studies by Refs. [7,69] suggest that it may as the higher-order, low-dimensional representations in these studies were specifically intended to be the set of familiar objects that the network can recognize. Are distributed codes for gaze or saccade vectors present in the loop between early visual areas and the object areas? As yet, there is no evidence for or against this possibility. Lastly, even if these codes are present, would they lead to accurate predictions of remapped locations? The success of the PredNet architecture based on motions within the image suggests that these remapping predictions are plausible within an autoencoder architecture although accurate remapping may be limited to the saccade targets which always remap to the fovea.

We have proposed that remapping is what object-based attention does when the eyes move, using a steerable autoencoder architecture that links early visual cortex to object areas and back. Of course, it is possible that several different architectures might simulate spatiotopic behavior and "solve" the problem of visual stability for specific tasks; for example, object tracking for autonomous vehicles. However, to directly explain human performance, it will be necessary to constrain these alternatives with neural and psychophysical data (e. g., Ref. [75]). One question, for example, could be the degree to which visual features are integrated across saccades and another could be the time course of position updating relative to the saccade. A working version of a steerable autoencoder could be tested for evidence that the pre-activation of the target's features may be detectable just prior to the saccade.

Testable hypothesis for future research

1. An autoencoder model should be able to reconstruct the attended targets at their new retinal location based on distributed eye position information at the latent value level (object areas).
2. The autoencoder will reconstruct only the attended items and not the entire input image.
3. The eye position information may be present as a gain field (or saccade vector field) at any location in the encoding stream, not necessarily at the highest, latent value (object area) level.
4. The proposed tight link between object-based attention and remapping predicts that object-based attention effects persist across eye movements.
5. Patients with lesions in object areas will have deficits in feature remapping.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Patrick Cavanagh reports financial support was provided by the Natural Sciences and Engineering Research Council of Canada. David Melcher reports financial support was provided by Tamkeen. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Josée Rivest, Martin Rolfs, Yalda Mohenzadeh, and Khac Vinh Nguyen for helpful comments. This research was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada grant RGPIN-2019-03938 (PC) and from the NYUAD Center for Brain and Health, funded by Tamkeen under the NYU Abu Dhabi Research Institute grant CG012 (DM).

Data availability

No data was used for the research described in the article.

References

1. Helmholtz H von (1867/1962). *Treatise on physiological optics*, vol. 3. New York: Dover; 1867. 1962); English translation by J. P. C. Southall for the Optical Society of America (1925) from the 3rd German edition of Handbuch der physiologischen Optik (first published in 1867, Leipzig: Voss).

2. Duhamel J-R, Colby CL, Goldberg ME: **The updating of the representation of visual space in parietal cortex by intended eye movements**. *Science* 1992, **255**:90–92.

3. Cavanagh P, Hunt A, Afraz A, Rolfs M: **Visual stability based on remapping of attention pointers**. *Trends Cognit Sci* 2010, **14**: 147–153.

4. Wurtz RH: **Corollary discharge contributions to perceptual continuity across saccades**. *Annu Rev Vision Sci* 2018, **4**: 215–237.

5. Golomb JD, Mazer JA: **Visual remapping**. *Annu Rev Vision Sci* 2021, **7**:257–277.

6. Cavanagh P, Caplovitz GP, Lytchecnko TK, Maechler MR, Tse PU, Sheinberg DR: **The architecture of object-based attention**. *Psychon Bull Rev* 2023, **30**:1643–1667, https://doi.org/10.3758/s13423-023-02281-7.
The authors propose that the targets of object-based attention are the object's features and locations in early visual cortices. These receive downward projections from object representations in high level, non-retinotopic, object areas of cortex mediated by autoencoder-like learning in the visual system.

7. Al-Tahan H, Mohsensadeh Y: **Reconstructing feedback representations in the ventral visual pathway with a generative adversarial autoencoder**. *PLoS Comput Biol* 2021, **17**, e1008775.

8. Lotter W, Kreiman G, Cox D: *Deep predictive coding networks for video prediction and unsupervised learning*. arXiv preprint; 2016. arXiv:1605.08104.

9. Zhou Y, Dong H, El Saddik A: **Deep learning in next-frame prediction: a benchmark review**. *IEEE Access* 2020, **8**: 69273–69283.

10. Riaz W, Gao C, Azeem A, Saifullah, Bux JA, Ullah A: **Traffic anomaly prediction system using predictive network**. *Remote Sens* 2022, **14**:447.

11. Szinte M, de Hollander G, Aqil M, Veríssimo I, Dumoulin S, Knapen T: **A retinotopic reference frame for space throughout human visual cortex**. *bioRxiv* 2024:2024. -02.
This study found that population receptive fields in all cortical visual field maps are indexed to the retina. A Bayesian decoding of stimulus location from the BOLD response patterns also ruled out the possibility of a spatiotopic reference frame independently computed at the population level.

12. d'Avossa G, *et al*.: **Spatiotopic selectivity of BOLD responses to visual motion in human area MT**. *Nat Neurosci* 2007, **10**: 249–255.

13. Gardner JL, Merriam EP, Movshon JA, Heeger DJ: **Maps of visual space in human occipital cortex are retinotopic, not spatiotopic**. *J Neurosci* 2008, **28**:3988–3999.

14. Crespi S, Biagi L, d'Avossa G, Burr DC, Tosetti M, Morrone MC: **Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention**. *PLoS One* 2011, **6**, e21661.

15. Zimmermann E, Morrone MC, Burr DC: **Buildup of spatial information over time and across eye-movements**. *Behav Brain Res* 2014, **275**:281–287.

16. Burr DC, Morrone MC: **Constructing stable spatial maps of the world**. *Perception* 2012, **41**(11):1355–1372.

17. Awh E, Armstrong KM, Moore T: **Visual and oculomotor selection: links, causes and implications for spatial attention**. *Trends Cognit Sci* 2006, **10**:124–130.

18. Robinson DA: **Eye movements evoked by collicular stimulation in the alert monkey**. *Vision Res* 1972, **12**:1795–1808.

19. Bruce CJ, Goldberg ME, Bushnell MC, Stanton GB: **Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements**. *J Neurophysiol* 1985, **54**:714–734.

20. Op De Beeck H, Vogels R: **Spatial sensitivity of macaque inferior temporal neurons**. *J Comp Neurol* 2000, **426**:505–518, https://doi.org/10.1002/1096-9861(20001030)426:4<505::AID-CNE1>3.0.CO;2-M.

21. Sereno AB, Lehky SR: **Population coding of visual space: Comparison of spatial representations in dorsal and ventral pathways**. *Front Comput Neurosci* 2011, **4**, https://doi.org/10.3389/fncom.2010.00159.

22. Andersen RA, Mountcastle VB: **The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex**. *J Neurosci* 1983, **3**:532–548.

23. Morris AP, Krekelberg B: **A stable visual world in primate primary visual cortex**. *Curr Biol* 2019, **29**:1471–1480.e6, https://doi.org/10.1016/j.cub.2019.03.069.

24. Merriam EP, Gardner JL, Movshon JA, Heeger DJ: **Modulation of visual responses by gaze direction in human visual cortex**. *J Neurosci* 2013, **33**:9879–9889, https://doi.org/10.1523/JNEUROSCI.0500-12.2013.

25. Harrison WJ, Stead I, Wallis TS, Bex PJ, Mattingley JB: **A computational account of transsaccadic attentional allocation based on visual gain fields**. *Proc Natl Acad Sci* 2024, **121**, e2316608121.

26. Zirnsak M, Moore T: **Saccades and shifting receptive fields: anticipating consequences or selecting targets?** *Trends Cogn Sci* 2014, **18**:621–628.

27. Neupane S, Guitton D, Pack CC: **Dissociation of forward and convergent remapping in primate visual cortex**. *Curr Biol* 2016, **26**:R491–R492.

28. Deubel H, Bridgeman B, Schneider WX: **Immediate post-saccadic information mediates space constancy**. *Vis Res* 1998, **38**:3147–3159.

29. Gottlieb JP, Kusunoki M, Goldberg ME: **The representation of visual salience in monkey parietal cortex**. *Nature* 1998, **391**: 481–484.

30. Mirpour K, Bisley JW: **Anticipatory remapping of attentional priority across the entire visual field**. *J Neurosci* 2012, **32**: 16449–16457.

31. Kroell LM, Rolfs M: **Foveal vision anticipates defining features of eye movement targets**. *eLife* 2022, **11**, e78106.

32. Golomb JD, Chun MM, Mazer JA: **The native coordinate system of spatial attention is retinotopic**. *J Neurosci* 2008, **28**: 10654–10662.

33. Jonikaitis D, Szinte M, Rolfs M, Cavanagh P: **Allocation of attention across saccades**. *J Neurophysiol* 2013, **109**: 1416–1424.

34. Andersen RA: **Neural mechanisms of visual motion perception in primates**. *Neuron* 1997, **18**:865–872.

35. Burr DC, Holt J, Jonstone JR, Ross J: **Selective depression of motion sensitivity during saccades**. *J Physiol (Paris)* 1982, **333**:1–15. London.

36. Shioiri S, Cavanagh P: **Saccadic suppression of low-level motion**. *Vis Res* 1989, **29**:915–928.

37. Castet E, Jeanjean S, Masson GS: **Motion perception of saccade-induced retinal translation**. *Proc Natl Acad Sci* 2002, **99**:15159–15163.

38. Schweitzer R, Doering M, Seel T, Raisch J, Rolfs M: **Saccadic omission revisited: what saccade-induced smear looks like**. *Psychol Rev* 2025, https://doi.org/10.1037/rev0000574.

39. Cavanagh P: **Attention-based motion perception**. *Science* 1992, **257**:1563–1565.

40. Fracasso A, Caramazza A, Melcher D: **Continuous perception of motion and shape across saccadic eye movements**. *J Vis* 2010, **10**:14, https://doi.org/10.1167/10.13.14.

41. Szinte M, Cavanagh P: **Spatiotopic apparent motion reveals local variations in space constancy**. *J Vis* 2011, **11**:1–20. 4.

42. Bridgeman B, Hendry D, Stark L: **Failure to detect displacement of the visual world during saccadic eye movements**. *Vis Res* 1975, **15**:719–722.

43. Wexler M, Collins T: **Orthogonal steps relieve saccadic suppression**. *J Vis* 2014, **14**:13. -13.

44. Rolfs M, Jonikaitis D, Deubel H, Cavanagh P: **Predictive remapping of attention across eye movements**. *Nat Neurosci* 2011, **14**:252–256.

45. Mathôt S, Theeuwes J: **Visual attention and stability**. *Phil Trans Biol Sci* 2011, **366**:516–527.

46. Posner MI, Cohen Y: **Components of visual orienting**. *Attent Perfor X Contr Lang Proc* 1984, **32**:531–556.

47. De Pisapia N, Kaunitz L, Melcher D: **Backward masking and unmasking across saccadic eye movements**. *Curr Biol* 2010, **20**:613–617, https://doi.org/10.1016/j.cub.2010.01.056.

48. Hunt AR, Cavanagh P: **Remapped visual masking**. *J Vis* 2011, **11**:13, https://doi.org/10.1167/11.1.13.

49. Huber-Huber C, Buonocore A, Hickey C, Melcher D: **The peripheral preview effect with faces: combined EEG and eye-tracking suggest multiple stages of trans-saccadic predictive and non-predictive processing**. *Neuroimage* 2019, **200**:344–362.

50. Liu X, Melcher D, Carrasco M, Hanning N: **The extrafoveal preview effect is more pronounced where perception is poor**. *Proc Natl Acad Sci* 2024, **121**, e2411293121, https://doi.org/10.1073/pnas.2411293121.

This study provides further evidence that the visual system takes into account the precision of processing of the saccade target in determining to what extent the pre-saccadic input influences post-saccadic judgments. It builds on previous studies showing near optimal integration of information across saccades, this time by incorporating the natural

51. Huber-Huber C, Melcher D: **Saccade execution increases the preview effect with faces: an EEG and eye-tracking coregistration study**. *Atten Percept Psychophys* 2025, **87**(1):155–171.

52. Chui T-Y, Golomb JD: **The influence of saccade target status on the reference frame of object-location binding**. *J Exp Psychol Gen* 2025, **154**(5):1183–1200.

53. Melcher D, Morrone MC: **Nonretinotopic visual processing in the brain**. *Vis Neurosci* 2015, **32**, E017.

54. Stewart EE, Valsecchi M, Schütz AC: **A review of interactions between peripheral and foveal vision**. *J Vis* 2020, **20**:2. -2.

55. Herwig A, Schneider WX: **Predicting object features across saccades: evidence from object recognition and visual search**. *J Exp Psychol Gen* 2014, **143**:1903.

56. Golomb JD, Albrecht AR, Park S, Chun MM: **Eye movements help link different views in scene-selective cortex**. *Cerebr Cortex* 2011, **21**:2094–2102.

57. Zimmermann E, Weidner R, Abdollahi RO, Fink GR: **Spatiotopic adaptation in visual areas**. *J Neurosci* 2016, **36**:9526–9534.

58. Baltaretu BR, Dunkley BT, Stevens WD, Crawford JD: **Occipital cortex is modulated by transsaccadic changes in spatial frequency: an fMRI study**. *Sci Rep* 2021, **11**:8611.

59. Knapen T, Swisher JD, Tong F, Cavanagh P: **Oculomotor remapping of visual information to foveal retinotopic cortex**. *Front Syst Neurosci* 2016, **10**, https://doi.org/10.3389/fnsys.2016.00054.

60. Kämmer L, Kroell LM, Knapen T, Rolfs M, Hebart MN: **Decoding peripheral saccade targets from foveal retinotopic cortex**. *bioRxiv* 2025, **2025–02**.
This gaze-contingent fMRI study shows that information about peripherally presented visual stimuli is represented in the foveal visual cortex, and extends it by demonstrating that these representations are similar to those evoked by foveally presented stimuli. These findings reveal that foveal cortex predicts the features of incoming stimuli through feedback from higher cortical areas, which offers a candidate mechanism underlying stable perception and may facilitate object recognition across saccades.

61. Parks NA, Corballis PM: **Electrophysiological correlates of presaccadic remapping in humans**. *Psychophysiology* 2008, **45**:776–783.

62. Moran C, Johnson PA, Landau AN, Hogendoorn H: **Decoding remapped spatial information in the peri-saccadic period**. *J Neurosci* 2024, **44**, e2134232024.
This study uses EEG to show that stimulus location information can be decoded at its post-saccadic (remapped) retinotopic position, with this effect being strongest when stimuli are presented shortly before the saccade.

63. Hinton GE: **Learning multiple layers of representation**. *Trends Cognit Sci* 2007, **11**:428–434.

64. Hinton GE, Salakhutdinov RR: **Reducing the dimensionality of data with neural networks**. *Science* 2006, **313**:504–507, https://doi.org/10.1126/science.1127647.

65. Ahn S, Adeli H, Zelinsky GJ: **The attentive reconstruction of objects facilitates robust object recognition**. *PLoS Comput Biol* 2024, **20**, e1012159.
Building on auto-encoder neural networks, the model reconstructs an object representation from incomplete visual input. It then uses this reconstruction as a template to bias feedforward processing to align with the most plausible object hypothesis.

66. Storrs KR, Anderson BL, Fleming RW: **Unsupervised learning predicts human perception and misperception of gloss**. *Nat Hum Behav* 2021, **5**:1402–1417, https://doi.org/10.1038/s41562-021-01097-6.

67. Kahneman D, Treisman A, Gibbs DJ: **The reviewing of object files: object-specific integration of information**. *Cogn Psychol* 1992, **24**:175–219.

68. Carlson T, Hogendoorn H, Fonteijn H, Verstraten FAJ: **Spatial coding and invariance in object-selective cortex**. *Cortex* 2011, **47**:14–22, https://doi.org/10.1016/j.cortex.2009.08.015.

69. Han K, Wen H, Shi J, Lu K-H, Zhang Y, Fu D, Liu Z: **Variational autoencoder: an unsupervised model for encoding and decoding fMRI activity in visual cortex**. *Neuroimage* 2019, **198**:125–136, https://doi.org/10.1016/j.neuroimage.2019.05.039.

70. Xu BY, Karachi C, Goldberg ME: **The postsaccadic unreliability of gain fields renders it unlikely that the motor system can use them to calculate target position in space**. *Neuron* 2012, **76**:1201–1209.

71. Nakamura K, Colby CL: **Updating of the visual representation in monkey striate and extrastriate cortex during saccades**. *Proc Natl Acad Sci* 2002, **99**:4026–4031.

72. Inaba N, Kawano K: **Neurons in cortical area MST remap the memory trace of visual motion across saccadic eye movements**. *Proc Natl Acad Sci* 2014, **111**:7825–7830.

73. Denagamage S, Morton MP, Hudson NV, Nandy AS: **Widespread receptive field remapping in early primate visual cortex**. *Cell Rep* 2024, **43**, 114557.
This study reports that pre-saccadic remapping is widespread in primate visual area V2, with up to 70% of neurons remapping around 30 ms before saccade onset. The findings reinforce the theory that forward predictive remapping is a general mechanism for maintaining perceptual stability across saccades, also involving activity in relatively early visual areas like V2.

74. Schweitzer R, Rolfs M: **Intrasaccadic motion streaks jumpstart gaze correction**. *Sci Adv* 2021, **7**:eabf2218.

75. Zirnsak M, Lappe M, Hamker FH: **The spatial distribution of receptive field changes in a model of peri-saccadic perception: predictive remapping and shifts towards the saccade target**. *Vis Res* 2010, **50**:1328–1337.

76. Wang X, Zhang C, Yang L, Jin M, Goldberg ME, Zhang M, Qian N: **Perisaccadic and attentional remapping of receptive fields in lateral intraparietal area and frontal eye fields**. *Cell Rep* 2024, **43**.
This study addresses whether remapping involves mainly convergent remapping (toward the saccade target) during preparation/attention or forward remapping (toward future receptive field locations). They report that that there are quite distinct remapping dynamics in lateral intraparietal area (LIP) and frontal eye fields (FEF) in monkeys, and propose a mechanism that combines both attention modulation and corollary discharge signals to explain transsaccadic updating.