



# A proposed attention-based model for spatial memory formation and retrieval

Çağatay Soyer<sup>1</sup>

Received: 21 December 2021 / Accepted: 16 December 2022 / Published online: 28 December 2022  
© The Author(s), under exclusive licence to Marta Olivetti Belardinelli 2022

## Abstract

Animals use sensory information and memory to build internal representations of space. It has been shown that such representations extend beyond the geometry of an environment and also encode rich sensory experiences usually referred to as *context*. In mammals, contextual inputs from sensory cortices appear to be converging on the hippocampus as a key area for spatial representations and memory. How metric and external sensory inputs (e.g., visual context) are combined into a coherent and stable place representation is not fully understood. Here, I review the evidence of attentional effects along the ventral visual pathway and in the medial temporal lobe and propose an attention-based model for the integration of visual context in spatial representations. I further suggest that attention-based retrieval of spatial memories supports a feedback mechanism that allows consolidation of old memories and new sensory experiences related to the same place, thereby contributing to the stability of spatial representations. The resulting model has the potential to generate new hypotheses to explain complex responses of spatial cells such as place cells in the hippocampus.

**Keywords** Spatial memory · Hippocampus · Place cells · Remapping · Visual attention · Long-term memory

## Introduction

Starting from the 1970s, recordings from single cells in the hippocampal formation of the rodent brain resulted in the discovery of cells supporting spatial representations. The first one of these was the *place cell* — hippocampal neurons that increase their firing rates when the animal is at a specific location (O'Keefe and Dostrovsky 1971). This 'preferred' location of a place cell is represented by an area of increased activity called the *place field*. Another group of spatial cells, called *head direction cells*, indicate the direction in which the animal is facing. Head direction cells were first identified in the rat postsubiculum and later in several other areas of

the brain (Taube et al. 1990). Finally, *grid cells* — found in the rodent entorhinal cortex — produce a grid-like firing pattern with varying frequency and phase for different cells as the animal travels over long distances (Hafting et al. 2005). Populations of grid cells can be shown to provide a neural odometry function based on self-motion cues (Fiete et al. 2008; Rowland et al. 2016). Thus, grid cells, head direction cells and place cells together support a positioning and navigation system based on the measurement of the animal's own movements (Moser et al. 2015). The underlying process is sometimes called *path integration* as it relies on the integration of acceleration and velocity signals in order to keep track of the animal's position and orientation (Etienne 1992). Path integration inherently accumulates errors and causes drift in distance measurements. Mammals are thought to use landmarks in the environment to correct path integration errors and 'recalibrate' their navigation system (Etienne and Jeffery 2004; Etienne et al. 1996).

The discovery of place cells, head direction cells and grid cells supported the long-standing idea of a *cognitive map* — an allocentric spatial representation of the environment in the mammal brain (Tolman 1948; O'Keefe and Nadel 1978). In an allocentric representation, locations and relationships are expressed relative to an external

---

An early version of this paper was presented at the 8th International Conference on Spatial Cognition (ICSC 2021), September 13–17, 2021.

---

Editor: Benjamin Straube (University of Marburg); Reviewers: Kate Jeffery (University College London) and a second researcher who prefers to remain anonymous.

---

✉ Çağatay Soyer  
soyer@ieee.org

<sup>1</sup> Luxembourg, Grand Duchy of Luxembourg

coordinate system or reference. On the contrary, egocentric representations are relative to the observer and sensor (Klatzky 1998). Animals experience external stimuli in egocentric coordinates, and how these are referenced and converted into an allocentric representation remains an open question (Burgess 2006; Filimon 2015; Marchette et al. 2014; Epstein 2008).

Further studies on spatial representations revealed a wider network of spatially responsive cells, including, for example, *border cells* and *boundary vector cells*, which specifically responded to boundaries in the environment (Moser et al. 2015; Lever et al. 2009; Savelli et al. 2008). In this network, grid cells in the medial entorhinal cortex (MEC) are considered to be providing the primary inputs triggering the location-based responses of place cells. However, several studies have shown that place cell responses are also affected by external sensory information and behavior (Muller 1996; Moser et al. 2015; Eichenbaum et al. 1999). In the literature, such information came to be known as *context* or *contextual inputs*, while the information provided by grid cells is referred to as *metric* due to the underlying translational and angular measurements. The lateral entorhinal cortex (LEC) is considered to provide contextual inputs to the hippocampus as LEC neurons are modulated by olfactory and visual information as opposed to the spatial modulation of MEC neurons (Witter et al. 2017; Scaplen et al. 2017). Furthermore, the LEC seems to encode not only the existence of objects and landmarks, but also their egocentric locations (Knierim et al. 2014).

The influence of contextual inputs on place cell responses is more complex than just resetting the path integration system using landmarks. Several experiments discussed in the next section point to a more complex mechanism and demonstrate place cell responses that do not correlate directly with the movement of landmarks or other changes in the environment. Overall, the complexity of place cell firing supports the idea of a spatial representation that is “more than space” and more like rich memories and experiences that are connected to a place (Eichenbaum et al. 1999). However, the exact nature of contextual inputs and their contribution to place cell responses are not fully understood (Jeffrey 2007).

In this paper, I argue that attentional effects along the ventral visual pathway and in the parahippocampal region can explain non-spatial and unpredictable place cell responses. I further suggest that attentional control provides a feedback mechanism by simultaneously guiding the retrieval of spatial memories (*attention to memory*) and new sensory experience (*attention to stimuli*), thereby contributing to the stability of spatial representations despite the richness and variability of the sensory environment. For the purpose of the current discussion, I specifically focus on vision as it provides the richest sensory modality for humans and its

underlying attention mechanisms are extensively studied in humans and non-human primates.

In the next section, I summarize the findings on non-spatial place cell phenomena and the models proposed to explain them. Then, I briefly discuss attention mechanisms affecting the collection and processing of visual information with reference to the extensive literature on visual attention and memory. Finally, I review the evidence of attentional effects on areas processing and relaying visual information to spatial cells and develop a conceptual model for attention-based integration of visual context in spatial representations. I conclude with a discussion on future directions and potential experimental studies.

## Non-spatial place cell responses and remapping

In their early experiments, O’Keefe and colleagues already reported non-spatial and unpredictable responses of place cells (O’Keefe 1976; O’Keefe and Conway 1978; Muller R., 1996). For example, some place cells could have place fields in multiple environments; some fired in the presence of one or two cues and others were affected by more cues in a more complex way. On the other hand, significant landmarks in an environment could be changed or removed without having a big effect on place cell responses (O’Keefe and Conway 1978; Muller and Kubie 1987). One of the criteria determining the presence of a place field at a certain location was behavioral relevance. For example, place fields changed with the introduction of obstacles or changing boundaries in an environment, and they accumulated around reward or goal locations (Muller and Kubie 1987; O’Keefe and Burgess 1996). Other behavioral effects were observed in a virtual corridor experiment, where the direction of travel and anticipation of goal locations changed place cell responses (Battaglia et al. 2004). In another experiment, four rats developed significantly different place fields for the same environment, further demonstrating the variability and individuality of place cell responses (Skaggs and McNaughton 1998).

Changes in place cell responses usually result in a complete or partial reconfiguration of their place fields, called *remapping* (Muller and Kubie 1987). In a complete remapping, groups of interconnected cells (*place cell assemblies*) can switch from one set of place fields to another, indicating that these cells participate in multiple representations (Jeffery 2011; Moser et al. 2015; Jeffrey 2007). On the other hand, partial remapping, where some cells change their place fields and others do not, remains a more challenging problem.

Despite the apparent influence of the environment, place cells do not depend on sensory inputs to function. Place fields are sustained in the dark and without any visual or

olfactory cues, and they can occur almost instantaneously in a novel environment. However, they are initially not stable and gradually improve with experience in a new environment (Wilson and McNaughton 1993; Kim et al. 2020), indicating a temporal accumulation of sensory information contributing to spatial representations.

Several efforts were made to model and explain the complex influence of sensory context on place cell responses. Skaggs and McNaughton described a conjunction of spatial and non-spatial information to explain apparent partial remapping of place cell responses. According to their *conjunctive coding* idea, the hippocampal cells receive inputs from a spatial map layer as well as a context layer such that the changes in the set of active units in the context layer give rise to different activation patterns in the hippocampal units (Skaggs and McNaughton 1998). More recently, Hayman and Jeffery proposed the *contextual gating model*, where they simulated the combination of inputs from the MEC and the LEC at the dendrites of place cells (Jeffrey 2007; Jeffery 2011; Hayman and Jeffery 2008; Jeffrey et al. 2004). A key aspect of contextual gating is the fact that place cells continue to function in the absence of contextual inputs, but these inputs play a ‘gating’ role when they are present. In another modeling study, Burgess’ team focused on the combination of inherently different coordinate frames of contextual and metric inputs. They proposed a model of spatial memory, where egocentric representations (e.g., egocentric directions to landmarks) in the parietal cortex are transformed into allocentric representations in the medial temporal lobe, with the retrosplenial cortex (RSC) acting as a transformation circuit (Byrne et al. 2007; Bicanski and Burgess 2018). In humans, the role of the RSC and the parahippocampal place area (PPA) (Epstein and Kanwisher 1998) in encoding ‘spatial layout’ relative to environmental features has been demonstrated experimentally using fMRI (Epstein 2008; Marchette et al. 2014). Similarly, in tracing studies, the RSC was shown to be the likely ‘bridge’ between the parietal cortex and the hippocampus (Wilber et al. 2015).

The conceptual model proposed in this paper is consistent with previous proposals in terms of the joint and variable contributions of metric and contextual information. However, it further introduces attention as a key controlling function in multiple stages upstream of the hippocampus, modulating sensory inputs based on task goals, as well as the animal’s memory and expectations from its perceived location.

## Attention

Attention is one of the fundamental processes of the brain, playing a critical role in perception, cognition and action (Johnson and Proctor 2004). It can be defined as “the process

by which organisms select a subset of available information upon which to focus enhanced processing and integration” (Ward 2008). Attention involves mechanisms for orienting sensory receptors toward relevant stimuli, searching for a target, and filtering information (including both attenuation of unattended stimuli and enhancement of attended stimuli). Research on attention increased significantly from its early philosophical foundations, with multiple theories explaining its underlying mechanisms (Johnson and Proctor 2004; Ward 2008; Broadbent 1958; Treisman 1964; Wolfe 1994; Posner 1980). In the next section, I review attentional effects in vision as a key sensory contribution to spatial representations.

## Visual attention

Mammals have advanced eyes and wide visual fields, giving the impression of an almost complete perception of one’s visual environment at all times. While this may be possible in certain simple settings with few and highly salient stimuli, in most real-world situations, visual information is too rich and cluttered to be processed in its full detail across the visual field. Therefore, complex physical and mental attention mechanisms focus processing resources on important and task-relevant information, while avoiding irrelevant information and distractors (Carrasco 2011).

Attention plays a key role in how we perceive our visual environment. As an example, consider two individuals walking the exact same path in a city for the first time. At the end of their tour, if we could analyze their mental representations of the city, the results would be significantly different due to the different visual, auditory and other sensory information that they attended to during their experience. If both individuals were asked to draw and annotate a map of their path on paper, each map would have different levels of detail and accuracy in different parts, and the two individuals’ actual memory or experience of the city would be very different. Furthermore, the way they could recall different places would also be different based on how, or using which features and landmarks, these places were originally registered. Therefore, if some of these landmarks were changed or moved at a later time, the effect of these changes on the two individuals would again be different based on whether or not they had registered these landmarks in the first place.

Due to the complexity of such experiences and the underlying attention mechanisms, “where we look” and “what we see” have been key questions that kept researchers busy for more than half a century (Chun and Wolfe 2005; Schütz et al. 2011; Schall 2013; Carrasco 2011; Egeth and Yantis 1997; Yarbus 1967). Early models of attention were based on the idea of information filtering or attenuation of irrelevant information (Treisman 1964; Broadbent 1958).

Later, motivated by physiological evidence of functional specialization in the visual cortex (Zeki 1976), Treisman's *feature integration theory (FIT)* proposed a first stage of parallel feature processing followed by a second stage of combination via serial attention to target objects (Treisman and Gelade 1980). Duncan and Humphreys extended the FIT by describing attention as a competition between visual inputs to access limited perceptual resources based on target–nontarget similarity (Duncan and Humphreys 1989). For example, in search tasks, the difficulty and efficiency of the task depended on the similarity between the search target and the stimuli. In subsequent models, Wolfe and colleagues extended the idea of parallel preattentive processing of features to *guide* attention toward candidate targets (Wolfe 1994, 2007). Desimone and Duncan further developed the competition idea into a *biased competition* for processing resources, where bottom-up and top-down attention criteria are used to bias the competition toward task-relevant and important stimuli (Desimone and Duncan 1995). The biased competition idea is supported by more recent fMRI studies, which suggest that maintaining a search item in working memory enhances the processing of the matching visual input, while other non-relevant or “accessory” items hamper task performance (Peters et al. 2012).

Attention is usually categorized into different types based on how it is driven and deployed. Attention is called *bottom-up* or *exogenous* when it is driven by highly salient visual features in the environment (such as a sudden flash of light or a sudden motion), and *top-down* or *endogenous* when it is self-generated based on interests and task-related goals of the individual (such as searching for an object with a specific shape or color) (Chun and Wolfe 2005; Carrasco 2011; Egeth and Yantis 1997). Recently, effects of past viewing experience or *selection history* have also been identified as a separate mechanism, possibly operating via visual statistical learning, and distinct from goal-driven effects (Awh et al. 2012; Theeuwes et al. 2022; Theeuwes 2019).

Both low-level visual features (e.g., colors, edges, brightness, motion), discontinuities, irregularities, and complete objects or faces can attract attention (sometimes referred to as *feature-based* and *object-based* attention). While a sudden onset of motion or light can be hard to suppress, other salient features, such as color, brightness, and lines, can be ignored voluntarily, for example, based on the search task and their relationship to the search target (Chun and Wolfe 2005). When attention results in a location-based sampling of the visual environment around us, it is called *spatial attention*. Spatial attention can be deployed physically by moving the eyes and the head to the target (*overt attention*) or by mentally shifting attention to a peripheral region or object, while the eyes are fixated elsewhere (*covert attention*).

Finally, attention can also be directed internally, to memories, rather than to external stimuli. This *reflective* form of

attention (also called *attention to memory*) allows selective retrieval of memories of past experiences (Ciaramelli et al. 2008; Long et al. 2018; Cabeza et al. 2008).

## Are place cells attentive?

If attention can be defined as a bias toward certain stimuli based on relevance and goals, then the preceding account of place cell responses shows several characteristics of attention-based behavior. For example, place fields seem to be influenced by the animal's goals and activities in an environment, which essentially provide the underlying criteria or modalities of attention. Place fields are accumulated around environmental landmarks and behaviorally relevant locations, indicating a bias toward registering these types of cues. They also seem to be individualized based on past experience, such that complete firing patterns can be recalled or triggered without any change in the environment or the animal's position, which indicate a memory-based attentional state.

Attentional modulation of hippocampal representations has also been demonstrated in experimental studies. Fenton et al. showed that overdispersion of place fields reduced when rats were required to focus on a subset of stimuli in a spatial task (such as selectively using distal features for navigation) (Fenton et al. 2010). Kentros and colleagues analyzed place field stability in mice and suggested that long-term stability of place fields depends on the behavioral relevance of the spatial context, and attention-like cognitive processes play a role in establishing this link (Kentros et al. 2004; Muzzio et al. 2009). More recently, and in line with the above observations, Scaplen et al. showed that visual cues were processed differently in the hippocampus based on their relevance as navigational landmarks or as context (Scaplen et al. 2014). Such relevance was determined by the nature of the cue (e.g., distal landmark or small object on the floor), as well as its salience and the animal's experience. Further experiments identified LEC as the locus of this processing (Scaplen et al. 2017).

These results show that attentional effects are present at all levels during the collection and processing of sensory information and eventually, in the modulation of place cell response. Therefore, it follows that, especially in more complex environments, spatial memories must be influenced by the subject's interest and goals in the environment (or parts of it) as much as the environment itself.

Motivated by these observations, an attentional account of spatial representations may explain complex place cell responses and the interaction between attention and long-term spatial memories. In the next two sections, I follow visual information from the retina along the ventral visual pathway and into the MTL, reviewing at each step the

evidence of attentional effects on visual context contributing to spatial representations in the hippocampus. The first section below will focus on spatial attention and the spatiotemporal sampling of the visual context. The next section will review attentional effects in the primary (V1) and the extrastriate areas of the visual cortex and the inferior temporal (IT) cortex, as well as the parahippocampal cortex (in primates). In both sections, I will gradually incorporate these findings into a conceptual model describing the notion of attentional control in spatial representations.

## Spatiotemporal sampling of the visual context

On the primate retina, the density of daytime photoreceptor cells (*cones*) is significantly higher in the central fovea region compared to the periphery (Curcio et al. 1990). This requires the eyes to be oriented in a way to capture task-relevant information in higher resolution, resulting in rapid eye movements called *saccades* (Yarbus 1967). Between saccades the eyes are fixated on the attended target (although not completely motionless), allowing the detailed processing of the retinal image. Despite frequent eye movements and the resulting changes in the retinal image, we perceive a clear and stable image of the world without any smear or motion blur. Interestingly, perception of both brief static stimuli and moving stimuli of low spatial frequency is possible during a saccade (Castet et al. 2002; Binda 2018; Burr and Ross 1982). However, complex neural mechanisms, including selective suppression of motion detection, predictive remapping of receptive fields (or *attentional pointers* as proposed in Cavanagh 2010), and transsaccadic memory, are employed to achieve visual stability across saccades (see reviews on neural mechanisms of visual stability (Melcher 2011; Wurtz 2008)).

The fact that saccadic eye movements are guided by attention has been established in several psychophysical and neurobiological experiments. In his early experiments with cueing tasks, Posner demonstrated that each saccade is preceded by first covertly shifting attention to the saccade target, followed by the saccade and fixation (Posner 1980). Later, the link between attention and saccades was established in dual-task experiments measuring both saccadic and perceptual performance. In these tasks, perceptual performance increased when the cued saccade target coincided with the location of the stimulus (Kowler et al. 1995; Deubel and Schneider 1996). More recently, Castet et al. investigated the dynamics of attentional deployment and observed the build-up of performance at the saccade target prior to a saccade (Castet et al. 2006). In neurobiological studies, key areas involved in controlling eye movements in the primate brain, such as the frontal eye fields (FEF) in the prefrontal cortex

(Schall 2004, 2013), the lateral intraparietal area (LIP) in the parietal cortex (Schall 2013; Gottlieb 2007) and the superior colliculus (SC) in the midbrain (Matsumoto et al. 2018), are shown to be influenced by attention through feature tuning and adjustment of receptive fields during spatial attention. There is growing evidence that “spatial attention shares spatial maps with saccade control centers in areas like SC, LIP, FEF” (Cavanagh 2010) and the activity in these maps not only points to the next saccade target but also enhances processing of information (e.g., detection of different features) from the saccade target (see reviews on neurobiology of attention Corbetta and Shulman 2002; Moore 2006)).

In free viewing conditions, spatial attention can be drawn to salient image features, such as color, brightness, and motion, representing a bottom-up (exogeneous) control mechanism. Saliency-based computational models based on the feature integration theory and Ullman and Koch’s early models have been vastly popular in attention and computer vision literature (Koch and Ullman 1985). However, purely bottom-up saliency has poor correlation with real human fixations unless modified by some top-down heuristics (Schütz et al. 2011; Boccignone et al. 2019; Tatler et al. 2011). It is now widely accepted that top-down attention criteria (or *attentional sets*) play a key role in guiding eye movements. Furthermore, during targeted search or other everyday tasks, eye movements are guided almost exclusively by top-down (endogenous) criteria in a highly purposive and task-oriented manner (Chun and Wolfe 2005; Schütz et al. 2011; Triesch et al. 2003; Land and Hayhoe 2001). In this case, bottom-up saliency (if present) may act as a distraction to the task, but allows the individual to detect unexpected and highly salient stimuli. The purposive nature of eye movements is also supported by the recently proposed role of visuospatial working memory (VSWM)<sup>1</sup> in guiding saccades and maintaining object and location continuity across saccades (Van der Stigchel and Hollingworth 2018). Equally, visuospatial memory retrieval is shown to be facilitated when subjects looked inside a square congruent with the spatial arrangement of the to-be-recalled objects during the encoding phase (Johansson and Johansson 2014).

Spatial attention can also be deployed covertly without eye movements, by voluntarily shifting attention to peripheral targets while maintaining fixation at a central point (Carrasco 2011). As discussed above, attention can be deployed to saccade targets before eye movement. Shifts of attention (without a saccade) can also serve other goals such as evaluating potential saccade targets or monitoring

<sup>1</sup> Here VSWM refers to the combined system of visual and spatial working memories, whereas visual short-term memory (VSTM) in the next paragraph refers more specifically to short-term storage of visual information as part of the broader working memory system.

the environment. Therefore, any fixation target is essentially viewed and attended to at least once in the periphery before it is fixated. Despite this, studies on visual memory show that the reduction of receptor density (hence image resolution) in the periphery and the transient nature of raw sensory information significantly limit the contribution of peripheral vision in long-term scene representation (Henderson and Castelano 2005; Hollingworth 2006). Unless fixated, only high-level information (e.g., shape structure, parts, orientation) survives a saccade for a very short time (De Graef and Verfaillie 2002; Aagten-Murphy and Bays 2018; Deubel et al. 2002; Henderson and Castelano 2005). Precise visual information from the entire visual field is highly volatile and lasts for up to a few hundred milliseconds (generally 80 to 300 ms), while a more durable but limited visual short-term memory (VSTM) can hold abstract information for durations in the order of seconds (e.g., see Hollingworth's review of visual memory systems Hollingworth 2006).

Peripheral vision plays a role in guiding attention and may accelerate the processing of the saccade target (known as the *transsaccadic preview effect*). Recent reviews point to an intricate relationship, where foveal and peripheral vision serve opposing goals for focused, detailed processing vs. texture-like perception of a larger visual field, but are otherwise intertwined to optimize each goal (Stewart et al. 2020). Nevertheless, high acuity and contrast sensitivity of the fovea means that our understanding and long-term memory of a scene rely primarily on fixations and the processing of the attended target (Rensink et al. 1997; Hollingworth 2006; Hollingworth and Henderson 2002; Irwin and Andrews 1996; Stewart et al. 2020). A recent fMRI study provides a causal relationship between the number of fixations and memory recall and further supports the role of the attended target in visual memory tasks (Fehlmann, et al. 2020).

The prominence of the fovea image on the retina and the guidance of eye movements under top-down attentional control suggest a spatiotemporal and largely task-oriented sampling of the visual context during visual exploration. A model of scene perception and memory supporting this idea is described by Hollingworth and Henderson. Reviewing theories of visual short-term memory (VSTM) and object files (Kahneman et al. 1992), they describe a model where high-level abstract visual representations of attended objects are accumulated from each fixation, indexed to a spatial position and consolidated into the long-term memory (LTM) (Hollingworth and Henderson 2002; Henderson and Hollingworth 2003). Their model supports the notion of abstract visual information being retained across saccades and transferred to LTM under attentional control.

Overall, these results are consistent with the gradual stabilization of place fields in a new environment (Wilson and McNaughton 1993; Kim et al. 2020), and they suggest that

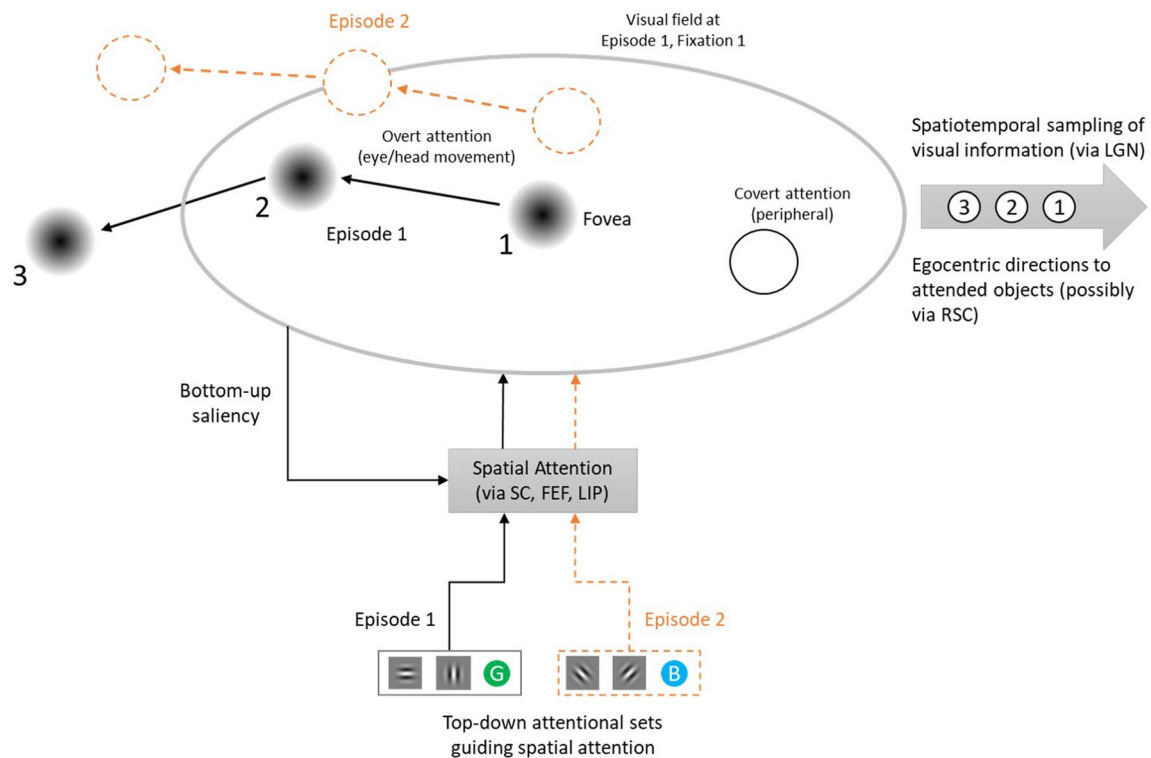
this process takes place in a task-oriented manner and largely based on attended targets.

To summarize the effects of early spatial attention mechanisms on spatial representations, the illustration in Fig. 1 shows an animal visiting the same physical location twice with different sets of top-down, task-oriented attention criteria. During each 'episode,' although the environment, the animal's position, and bottom-up saliency remain the same, saccades are guided differently due to the two attentional sets (shown by two separate sequences of fixations). This results in two different streams of retinal information being fed into the visual pathways and eventually into the spatial representations. Considering the significance of the fovea in the formation of long-term visual memories, this information is simplified as a stream of foveal images (1, 2, 3,...) guided by the corresponding attention criteria in each episode. At each fixation, head direction and the saccade vector further provide directions to attended targets and their spatial relationships in egocentric coordinates. This information is assumed to be provided via the parietal cortex (PC) and the retrosplenial cortex (RSC) as discussed above (Wilber et al. 2015; Marchette et al. 2014; Epstein 2008). Note that the model equally accounts for covert spatial attention, where attended information would be collected from the periphery albeit at a significantly lower spatial resolution.

## Attentional tuning and modulation in visual areas and the parahippocampal cortex

Visual information from the retina is relayed to the visual cortex via the lateral geniculate nucleus (LGN) of the thalamus. Besides being a relay center, LGN has been shown to be affected by attention so that neural responses to attended stimuli are enhanced, while responses to unattended stimuli are attenuated (O'Connor et al. 2002). In the visual cortex, a large part of early processing and feature detection continues to be dedicated to the fovea due to the topographic organization of the primary visual cortex (area V1). V1 neurons have small receptive fields and respond to simple features such as lines of different orientation, spatial frequency and color. Along the ventral visual pathway, the processing continues with areas V2, V4 and the inferior temporal (IT) cortex, where the receptive fields of cells get bigger and the features they respond to become more high-level (Kandel and Mason 1995; Zeki 1976).

Attentional effects in the visual cortex are traditionally studied by controlling covert attention to stimuli which fall either inside or outside of a cell's receptive field (Moran and Desimone 1985). In primates, the modulatory effects of top-down attention have been observed in all areas of the visual pathways during both overt and covert shifts of attention



**Fig. 1** Spatial attention guides fixations (1, 2, 3,...) and covert attention, resulting in spatiotemporal sampling of visual information and directions to attended objects based on top-down, task-oriented attention criteria in each episode

(Motter 1993; Reynolds and Chelazzi 2004; Freiwald and Kanwisher 2004).

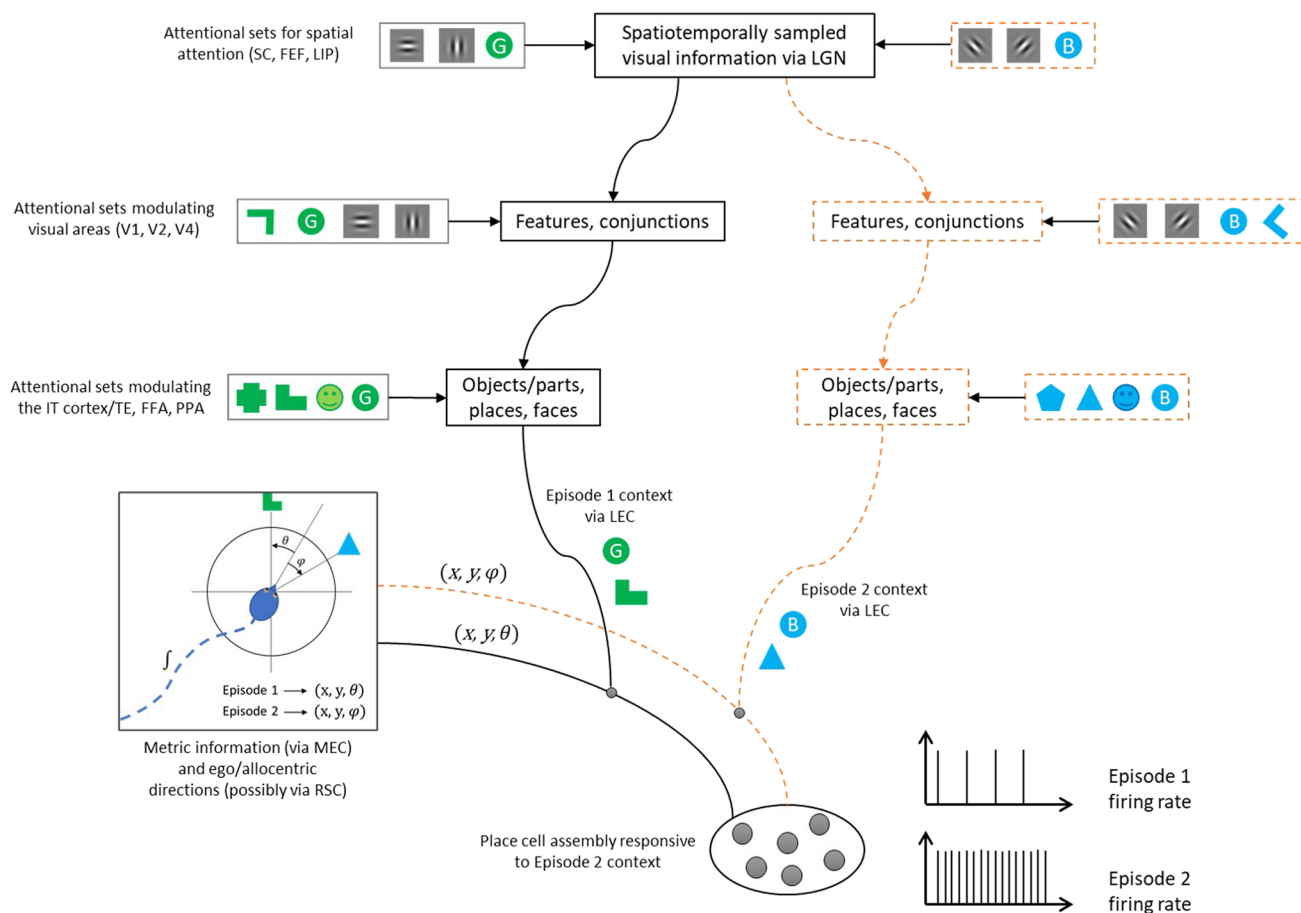
In the primary visual cortex of the mouse, V1 neurons were found to be modulated by the SC, potentially reflecting bottom-up saliency computations in this area which play a role in directing overt attention (Ahmadlou et al. 2018). In primates, electrophysiological studies demonstrated increased responses of V1 neurons when stimuli in their receptive fields were attended (Motter 1993; McAdams and Reid 2005). Through simultaneous stimulation of the FEF and measuring responses from V4 of monkeys, Moore and Armstrong also demonstrated that the responses of V4 neurons to stimuli in their receptive fields changed based on the preparation of a saccade to this area (Moore and Armstrong 2003).

The nature and magnitude of attentional effects change as visual information propagates through the visual areas (Freiwald and Kanwisher 2004). For example, Muller and colleagues found that attentional modulation changed along the visual pathway from purely bottom-up effects in V1 to purely top-down control in V2 and then a combined bottom-up and top-down influence in V4 (Melloni et al. 2012).

Top-down attentional effects are particularly interesting for spatial representations as they encode task- and memory-based strategies. Recordings from macaque V4 neurons showed that the knowledge of the target in a detection task

created an attentional effect which modulated their response to a stimulus array (Chelazzi et al. 2001). A similar experiment showed the same effect for macaque IT neurons, with both findings supporting the biased competition model and the top-down, memory-guided attentional effects in these areas (Chelazzi et al. 1998). Studies using event-related potentials (ERP) showed that IT neurons maintained their firing rates between the cue and search during a cued search task, leading to the hypothesis that the IT cortex is one of the sites that encode top-down attentional sets (Woodman et al. 2013; Conway 2018).

As the last visual area in the ventral visual pathway, the IT cortex plays a critical role in object recognition and memory. Due to their large receptive fields, always including the fovea, IT neurons can respond to stimuli as complex as complete objects or faces (Miller et al. 1991; Desimone 1996). This leads to higher-level and more meaningful attention criteria being observed in this area. For example, similar to Chelazzi's experiments with macaque monkeys, Peters et al. studied attentional effects in the fusiform face area (FFA) of the IT cortex and the parahippocampal place area (PPA) of the parahippocampal cortex. Human subjects viewed the same image with two different tasks to search for a house and a face. fMRI images showed the increase in activation of the PPA and the FFA, respectively, during the two trials. As the actual stimuli remained unchanged between the



**Fig. 2** Simplified illustration of the proposed attention-based model of metric-contextual integration. Place cell assemblies respond to the same location and visual context (consisting of one green and one blue object) differently based on the attentional tuning in upstream areas

two experiments (i.e., either a house or a face), the study concluded that the activity of the PPA and FFA neurons is modulated by the corresponding attention criteria (Peters et al. 2012).

The projections between the IT cortex and MTL in primates indicate that visual information flows via the dorsal and ventral TE areas of the IT cortex and through the perirhinal cortex and entorhinal cortices, before reaching the hippocampus (Saleem et al. 2000; Buffalo et al. 2006; Ungerleider et al. 1998). Here, perirhinal cortex acts as the interface between the ventral visual pathway and the MTL, receiving strong connections from both the IT cortex and the primate parahippocampal cortex, and relaying these inputs to the hippocampus via the entorhinal cortex (Buffalo et al. 2006).

Based on the demonstrated effects of top-down attention throughout the ventral visual pathway, up to the IT cortex, it can be expected that the encoding of information in the hippocampal formation is similarly modulated by attention. In a recent fMRI study, Aly and Turk-Browne found evidence of such attentional modulation in the hippocampus. In a series

of experiments, they first identified the attentional states of the hippocampus during a cued search task for paintings and rooms. Then, they correlated the hippocampal activity to these templates during another task requiring top-down attention to similar targets. Finally, they tested the subject's memory performance on task-relevant items from the second search task and found that memory performance was better when the attentional state of the hippocampus during encoding matched the information that was being encoded, suggesting that the hippocampus maintained a stronger representation of attended information (Aly and Turk-Browne 2016). Similarly, Scaplen et al. showed that LEC inactivation affects subsequent hippocampus activity in different ways depending on the relevance of visual cues for navigation (Scaplen et al. 2017), suggesting that LEC selectively (rather than uniformly) modulates contextual inputs to the hippocampal place cells.

Combining the above evidence in the current model, Fig. 2 shows a simplified illustration of attention-based integration of visual context into spatial representations. Here, the 'visual context' is assumed to be simply consisting of



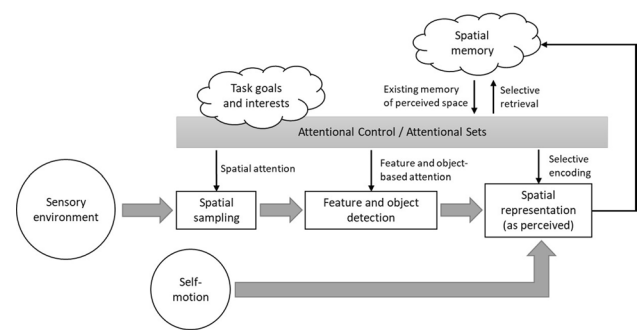
one green and one blue object represented by their geometric and color features and having similar bottom-up saliency. The subject is at a specific location  $(x,y)$  with bearings  $\theta$  and  $\varphi$  to the two objects. A place cell assembly is assumed to be tuned to a specific combination of metric and contextual cues. In the two episodes shown, the subject arrives at  $(x,y)$  with two different top-down attentional sets (e.g., searching for green rectangular objects or blue triangular objects). The scene is spatiotemporally sampled based on these attention criteria and the resulting retinal information is relayed via the LGN to the visual cortex (areas V1, V2, V4), the IT cortex, and the parahippocampal cortex, where feature-based attention further modulates visual processing differently for the two episodes. This highly modulated contextual information is then combined with metric inputs from the grid cell network within the medial entorhinal cortex (MEC) and with ego/allocentric directions to attended objects possibly computed within the retrosplenial cortex (RSC). Consequently, place cell assemblies respond to the same location and visual context differently based on the subject's interests and task goals and the associated attentional tuning in upstream areas.

Note that the above illustration is simplified in a number of ways. First, potential attentional tuning in the hippocampal formation (e.g., in MEC, LEC, or place cells themselves) is not discussed. Secondly, only an instantaneous response is shown and both the encoding phase and the temporal accumulation of visual information across multiple fixations or multiple visits to the same environment are ignored. These topics are discussed below in the conclusion.

### Reflective attention as a mechanism to stabilize spatial representations

So far in the above model, I suggest that attentional control of sensory information may cause different spatial representations to be associated with the same physical space. However, in real-world situations, small changes in the environment (such as motion, occlusions, lighting and visibility) can dramatically change the perceived image. These changes can affect feature detection and spatial attention mechanisms, causing significantly different looking behavior. Additionally, the subject's interests and goals in a place can change from one episode to another, resulting in different top-down attention criteria. Under such variations, and given the influence of sensory information, how is it possible that an animal can maintain a stable representation of a particular environment during multiple visits and recognize this as the same place?

One possible mechanism (e.g., in contextual gating) is the weighting of contextual and metric cues based on their reliability. Such a mechanism can be enhanced by the above attentional effects, where certain features of the environment



**Fig. 3** Initial perception of a place may trigger selective retrieval of previous experiences, setting expectations and attentional sets for new sensory exploration, thereby gradually stabilizing perception in a feedback loop

can be selected to drive spatial representations, while others are ignored or suppressed. This idea is consistent with recent observations that spatially relevant objects are processed differently than non-relevant objects (Scaplen et al. 2014), as well as reports of attentional modulation of hippocampal representations discussed earlier (Fenton et al. 2010; Kentros et al. 2004; Muzzio et al. 2009). However, attention-based encoding of place memories does not completely solve the stability problem. Variations in top-down attention criteria may still result in significant changes in attentional bias, thereby causing contextual input to be significantly different between two or more episodes.

In their model of scene perception and memory, Hollingworth and Henderson reported a similar problem from a visual perception perspective. They considered that subsequent episodes with the same scene could involve the retrieval of previous LTM representations consisting of a scene map and object codes, but how the correct scene map was retrieved from the previous episode remained an open question (Hollingworth and Henderson 2002). Such a mechanism is clearly required to prevent spatial representations from diverging between old and new episodes due to lack of correspondence of sensory information.

To answer this question, I speculate that reflective attention might play a stabilizing role, where spatial memories (including sensory information) are not entirely retrieved, but attentively recalled based on an initial perception of where the animal thinks it is (e.g., based on initial self-motion and visual/olfactory cues). These memories of an initially perceived location then set attention criteria and guide spatial attention in a way to collect further contextual information to verify or update the animal's initial perception. In other words, the initial perception of the environment triggers selective retrieval of previous spatial memories, which then influence sensory perception through attention, thereby gradually consolidating old memories and new sensory experiences related to the same place in a feedback loop.

Figure 3 illustrates this idea. Here, sensory information is first sampled via spatial attention. This information is processed through the visual pathway under attentional control to eventually form episodic memories, which are also encoded and emphasized by attention. Later, in a new episode, these memories are recalled partly and selectively based on initial sensory experiences and current task goals. This in turn sets an expectation about the sensory context of the space, which contributes, together with task goals, to attentional tuning that guides further sensory exploration. Therefore, a feedback loop is created to resolve conflicts between where we *think* we are and where we *actually* are, and to gracefully merge earlier experiences and metric cues with new sensory information.

Note that, although visual context is attentively sampled and encoded in spatial memories, such sampling is based on the attentional sets and saliency maps of past episodes. With experience (multiple episodes or visits to the same place), such memories can be expected to approach a complete representation (possibly still with significant and surprising gaps that we sometimes experience in everyday life). Conversely, the retrieval process employs attention to memory with attentional sets reflecting current goals and using bottom-up saliency from the current state of the environment. Since the same attentional sets and saliency maps are used for perception in the current episode, the proposed mechanism helps to reduce conflict between what is perceived and what is remembered, while also updating memories when there is a real conflict not caused by sampling (i.e., a real change in the environment between the old and new episodes).

The idea of past memories being attentively retrieved and used to set expectations for sensory exploration is supported by a number of studies. First, the ability of spatial cells to generate memory-based responses was demonstrated by Miller et al., who recorded place-responsive cells in human patients during virtual navigation and later during retrieval of navigation-related memories without actual navigation. They found that the same cells were activated during episodic memory retrieval, and the firing pattern was similar to the activity observed during the initial encoding of the locations (Miller, et al. 2013). Secondly, Ciaramelli and Cabeza showed that such retrieval of episodic memories is indeed under attentional control. In their attention to memory (AtoM) model, they proposed that the dorsal and ventral parietal cortices (DPC and VPC) are associated with attention-based retrieval of memories, respectively, for top-down (goal-driven) and bottom-up (memory-driven) attention to memory (Ciaramelli et al. 2008; Cabeza et al. 2008). Combined with Aly and Turk-Browne's results (Aly and Turk-Browne 2016) showing attentional modulation in the hippocampus during encoding of episodic memories, it can be argued that spatial representations are both attentively

encoded and retrieved. Finally, to complete the proposed feedback mechanism, it was recently shown that V1 neurons in the mouse visual cortex shifted their responses according to the animal's perception of distance traveled (Fournier, et al. 2020). The same study found that both hippocampal CA1 and visual V1 neurons could both be used to decode the animal's position and the associated errors were highly correlated. The authors provided two possible explanations, one based on the same self-motion cues feeding both areas and the other based on a feedback signal from the hippocampus to V1, possibly via the RSC.

Finally, the anterior cingulate cortex (ACC), implicated in the control of attention, decision making and conflict resolution, might provide further top-down attentional control signals. Strong connections between the ACC and the entorhinal cortex and the error detection function of the ACC are thought to support error-driven learning (Ku et al. 2021; Calderazzo et al. 2021). It is possible that ACC might play a similar role during the encoding of spatial memories to support the above comparison of memories to new sensory experiences.

## Conclusion and future work

This paper presented an account of spatial representation based on attentional control of associated contextual and sensory inputs. I started with the widely accepted notion of spatial representations being both metric and sensory in nature. Then, I reviewed the evidence of attentional modulation in several visual and MTL areas and incorporated these into a conceptual model where sensory information is integrated into long-term spatial representations under attentional control, thereby creating task-oriented episodic memories of space. I further speculated that attention could play a role in stabilizing such representations by attentively retrieving past experiences and using these memories to guide new sensory exploration.

Further development of the model can provide more detailed information on the interaction of specific brain areas involved in controlling attention and in encoding and retrieving spatial memories. The nature and loci of attentional sets or memory representations guiding attention require further investigation. These are so far considered to be embedded in connection weights and in responsiveness, sensitivity and receptive fields of individual cells. Recent accounts of time- and experience-dependent responses in the hippocampal formation (e.g., in the LEC) provide new insights into the representation of time (Tsao et al. 2013, 2018; Eichenbaum 2017). These results can help to further develop the temporal aspects of the model to include gradual improvement and consolidation of old and new experiences beyond the initial ideas discussed above.

Finally, experimental verification of the proposed model requires simultaneous control of the sensory environment and the subject's attention while carrying out a spatial task. Visual attention mechanisms in humans and other primates are extensively studied in the literature; however, similar studies with other mammals are scarce. Conversely, spatial representations are generally studied in rodents due to the possibility of electrophysiological measurements on freely moving animals. Species-specific aspects make it difficult to translate and combine findings in these two areas, as attempted in this paper. Although rodents do not possess a retinal fovea like primates and they are considered to have limited visual abilities, a recent study found that mice do have improved visual resolution in the front and slightly upper part of their visual field (termed *fovea* to indicate a cortical specialization), mainly due to lower scatter of single-cell receptive fields corresponding to this region in their visual cortex (Van Beest, et al. 2021). They have also been shown to use visual selective attention and eye movements to explore their environment similar to primates (Wang and Krauzlis 2018; Meyer et al. 2020). These studies show that there may be opportunities in future to study attentional effects at neural level in rodent navigation experiments. In humans and non-human primates, advances in virtual environments and fMRI technology can facilitate experiments on visual attention and spatial representations.

**Author contributions** The author would like to thank the anonymous reviewers for their valuable contributions.

**Funding** No funding was received for this study.

**Availability of data and material** Not applicable.

**Code availability** Not applicable.

## Declarations

**Conflict of interests** The author has no conflicts of interest or conflict of interests.

**Ethics approval** The study does not involve any work with humans or animals, their data or biological material.

## References

- Aagten-Murphy, D., Bays, P. (2018). Functions of memory across saccadic eye movements. In: T. Hogsten (ed) Processes of visuospatial attention and working memory. Current topics in behavioral neurosciences, Springer, Berlin. vol 41, 155–183. [https://doi.org/10.1007/7854\\_2018\\_66](https://doi.org/10.1007/7854_2018_66)
- Ahmadlou M, Zweifel LS, Heimel JA (2018) Functional modulation of primary visual cortex by the superior colliculus in the mouse. *Nat Commun* 9:3895. <https://doi.org/10.1038/s41467-018-06389-6>
- Aly M, Turk-Browne NB (2016) Attention promotes episodic encoding by stabilizing hippocampal representations. *PNAS* 113(4):420–429. <https://doi.org/10.1073/pnas.1518931113>
- Awh E, Armstrong K, Moore T (2006) Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends Cogn Sci* 10(3):124–130. <https://doi.org/10.1016/j.tics.2006.01.001>
- Awh E, Belopolsky AV, Theeuwes J (2012) Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends Cogn Sci* 16(8):437–443
- Battaglia FP, Sutherland GR, McNaughton BL (2004) Local sensory cues and place cell directionality: additional evidence of prospective coding in the hippocampus. *J Neurosci* 24(19):4541–4550. <https://doi.org/10.1523/JNEUROSCI.4896-03.2004>
- Bicanski A, Burgess N (2018) A neural-level model of spatial memory and imagery. *eLife*. <https://doi.org/10.7554/eLife.33752>
- Binda PM (2018) Vision during saccadic eye movements. *Ann Rev Vis Sci* 4:193–213
- Boccignone G, Cuculo V, D'Amelio A (2019) Problems with saliency maps. In: Ricci E, Rota BS, Snoek C, Lanz O, Messelodi S, Sebe N (ed) Image analysis and processing – ICIAP 2019. Lecture notes in computer science. 11752. Springer, Cham. [https://doi.org/10.1007/978-3-030-30645-8\\_4](https://doi.org/10.1007/978-3-030-30645-8_4)
- Broadbent D (1958) Perception and communication. Pergamon Press, London
- Buffalo EA, Bellgowan PS, Martin A (2006) Distinct roles for medial temporal lobe structures in memory for objects and their locations. *Learn Mem* 13(5):638–643
- Burgess N (2006) Spatial memory: how egocentric and allocentric combine. *Trends Cogn Sci* 10(12):551–557. <https://doi.org/10.1016/j.tics.2006.10.005>
- Burr DC, Ross J (1982) Contrast sensitivity at high velocities. *Vis Res* 22:479–484
- Byrne P, Becker S, Burgess N (2007) Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychol Rev* 114(2):340–375. <https://doi.org/10.1037/0033-295X.114.2.340>
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 9:613–625. <https://doi.org/10.1038/nrn2459>
- Calderazzo SM, Busch S, Moore T, Rosene D, Medalla M (2021) Distribution and overlap of entorhinal, premotor, and amygdalar connections in the monkey anterior cingulate cortex. *J Comp Neurol* 529:885–904. <https://doi.org/10.1002/cne.24986>
- Carrasco M (2011) Visual attention: the past 25 years. *Vision Res* 51(13):1484–1525. <https://doi.org/10.1016/j.visres.2011.04.012>
- Castet E, Jeanjean S, Masson GS (2002) Motion perception of saccade-induced retinal translation. *Proc Natl Acad Sci* 99(23):15159–15163
- Castet E, Jeanjean S, Montagnini A, Laugier D, Masson G (2006) Dynamics of attentional deployment during saccadic programming. *J vis* 6:196–212
- Cavanagh PH (2010) Visual stability based on remapping of attention pointers. *Trends Cogn Sci* 14(4):147–153
- Chelazzi L, Duncan J, Miller EK, Desimone R (1998) Responses of neurons in inferior temporal cortex during memory-guided search. *J Neurophysiol* 80(6):2918–2940
- Chelazzi L, Miller EK, Duncan J, Desimone R (2001) Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb Cortex* 11(8):761–762. <https://doi.org/10.1093/cercor/11.8.761>
- Chun MM, Wolfe JM (2005) Visual Attention. In: Goldstein BE (ed) Blackwell handbook of sensation and perception. Blackwell Publishing, pp 272–310
- Ciaramelli E, Grady CL, Moscovitch M (2008) Top-down and bottom-up attention to memory: a hypothesis (AtoM) on the role of the

- posterior parietal cortex in memory retrieval. *Neuropsychologia* 47(7):1828–1851. <https://doi.org/10.1016/j.neuropsychologia.2008.03.022>
- Conway BR (2018) The organization and operation of inferior temporal cortex. *Ann Rev Vis Sci* 4:381–402. <https://doi.org/10.1146/annurev-vision-091517-034202>
- Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215. <https://doi.org/10.1038/nrn755>
- Curcio CA, Sloan KR, Kalina RE, Hendrickson AE (1990) Human photoreceptor topography. *J Comput Neurol* 292(4):497–523
- Desimone R (1996) Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci* 93:13494–13499
- Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222
- Deubel H, Schneider WX, Bridgeman B (2002) Transsaccadic memory of position and form. In: Hyona J, Munoz DP, Heide W, Radach R (eds) *Progress in brain research*, vol 140, pp 165–180. Elsevier Science.
- Deubel H, Schneider WX (1996) Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vis Res* 36(12):1827–1837
- Duncan J, Humphreys GW (1989) Visual search and stimulus similarity. *Psychol Rev* 96(3):433–458
- Egeth H, Yantis S (1997) Visual attention: control, representation, and time course. *Annu Rev Psychol* 48:269–297. <https://doi.org/10.1146/annurev.psych.48.1.269>
- Eichenbaum H (2017) On the integration of space, time, and memory. *Neuron* 95:1007–1018. <https://doi.org/10.1016/j.neuron.2017.06.036>
- Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H (1999) The hippocampus, memory, and place cells: is it spatial memory or memory space? *Neuron* 23:209–226
- Epstein R (2008) Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci* 12(10):388–396. <https://doi.org/10.1016/j.tics.2008.07.004>
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392(6676):598–601
- Etienne AS (1992) Navigation of a small mammal by dead reckoning and local cues. *Curr Dir Psychol Sci* 1(2):48–52. <https://doi.org/10.1111/1467-8721.ep11509737>
- Etienne AS, Jeffery KJ (2004) Path integration in mammals. *Hippocampus* 14:180–192
- Etienne AS, Maurer R, Seguinot V (1996) Path integration in mammals and its interaction with visual landmarks. *J Exp Biol* 199(1):201–209
- Fehlmann B, Coyne D, Schickanz N, Milnik A, Gschwind L, Hofmann P, de Quervain DJ (2020) Visual exploration at higher fixation frequency increases subsequent memory recall. *Cerebral Cortex Commun* 1:1–14. <https://doi.org/10.1093/txcom/tgaa032>
- Fenton AA (1998) Place cell discharge is extremely variable during individual passes of the rat through the firing field. *Proc Natl Acad Sci* 95:3182–3187
- Fenton AA, Lytton WW, Barry JM, Lenck-Santini PP, Zinyuk LE, Kubík Š, Olypher AV (2010) Attention-like modulation of hippocampus place cell discharge. *J Neurosci* 30(13):4613–4625
- Fiete IR, Burak Y, Brookings T (2008) What grid cells convey about rat location. *J Neurosci* 28(27):6858–6871
- Filimon F (2015) Are all spatial reference frames egocentric? Reinterpreting evidence for allocentric, object-centered, or world-centered reference frames. *Front Hum Neurosci* 9(648).
- Fournier J, Saleem AB, Diamanti EM, Wells MJ, Harris KD, Carandini M (2020) Mouse visual cortex is modulated by distance travelled and by theta oscillations. *Curr Biol* 30:3811–3817. <https://doi.org/10.1016/j.cub.2020.07.006>
- Freiwald WA, Kanwisher NG (2004) Visual selective attention: Insights from brain imaging and neurophysiology. In: Gazzaniga MS (ed) *The cognitive neurosciences*. Boston Review, pp 575–588
- Gottlieb J (2007) From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron* 53:9–16. <https://doi.org/10.1016/j.neuron.2006.12.009>
- DeGraef P, Verfaillie K (2002) Transsaccadic memory of visual object detail. In: Hyona J, Munoz DP, Heide W, Radach R (eds) *Progress in brain research*, vol 140, pp 181–196. Elsevier Science.
- Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436(7052):801–806
- Hayman RM, Jeffery KJ (2008) How heterogeneous place cell responding arises from homogeneous grids - a contextual gating hypothesis. *Hippocampus* 18:1301–1313. <https://doi.org/10.1002/hipo.20513>
- Henderson JM, Castelano MS (2005) Eye movements and visual memory for scenes. In: Underwood G (ed) *Cognitive processes in eye guidance*. Oxford University Press, pp 213–235
- Henderson JM, Hollingworth A (2003) Eye movements, visual memory, and scene representation. In: Peterson MA, Rhodes G (eds) *Perception of faces, objects, and scenes: analytic and holistic processes*. Oxford University Press, pp 356–377
- Hollingworth A (2006) Visual memory for natural scenes: Evidence from change detection and visual search. *Vis Cogn* 14:781–807. <https://doi.org/10.1080/13506280500193818>
- Hollingworth A, Henderson JM (2002) Accurate visual memory for previously attended objects in natural scenes. *J Exp Psychol* 28(1):113–136. <https://doi.org/10.1037/0096-1523.28.1.113>
- Irwin DE, Andrews RV (1996) Integration and accumulation of information across saccadic eye movements. In: Inui T, McClelland JL (eds) *Attention and performance 16: information integration in perception and communication*. The MIT Press, pp 125–155
- Jeffery KJ (2011) Place cells, grid cells, attractors, and remapping. *Neural Plact*. <https://doi.org/10.1155/2011/182602>
- Jeffery KJ, Anderson MI, Hayman R, Chakraborty S (2004) A proposed architecture for the neural representation of spatial context. *Neurosci Biobehav Rev* 28:201–218. <https://doi.org/10.1016/j.neubiorev.2003.12.002>
- Jeffery KJ (2007) Integration of the sensory inputs to place cells: What, where, why, and how? *Hippocampus* 17:775–785
- Johansson R, Johansson M (2014) Look here, eye movements play a functional role in memory retrieval. *Psychol Sci* 25(1):236–242
- Johnson A, Proctor RW (2004) *Attention: theory and practice*. Sage Publications, New York
- Kahneman D, Treisman A, Gibbs BJ (1992) The reviewing of object files: object-specific integration of information. *Cogn Psychol* 24(2):175–219. [https://doi.org/10.1016/0010-0285\(92\)90007-o](https://doi.org/10.1016/0010-0285(92)90007-o)
- Kandel E, Mason C (1995) Perception of form and motion. In: Kandel E, Schwartz JH, Jessell TM (eds) *Essentials of neural science and behavior*, pp 425–451. Appleton and Lange
- Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER (2004) Increased attention to spatial context increases both place field stability and spatial memory. *Neuron* 42:283–295
- Kim S, Jung D, Royer S (2020) Place cell maps slowly develop via competitive learning and conjunctive coding in the dentate gyrus. *Nat Commun* 11:4550. <https://doi.org/10.1038/s41467-020-18351-6>
- Klatzky RL (1998) Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. *Spatial Cognition*
- Knierim J, Neunuebel JP, Deshmukh SS (2014) Functional correlates of the lateral and medial entorhinal cortex: objects path integration and local-global reference frames. *Philos Trans R Soc B*. <https://doi.org/10.1098/rstb.2013.0369>
- Koch C, Ullman S (1985) Shifts in selective visual attention: towards the underlying neural circuitry. *Hum Neurobiol* 4(4):219–227

- Kowler E, Anderson E, Doshier B, Blaser E (1995) The role of attention in the programming of saccades. *Vis Res* 35(13):1897–1916
- Ku S, Hargreaves EL, Wirth S, Suzuki WA (2021) The contributions of entorhinal cortex and hippocampus to error driven learning. *Commun Biol* 4:628. <https://doi.org/10.1038/s42003-021-02096-z>
- Land MF, Hayhoe M (2001) In what ways do eye movements contribute to everyday activities? *Vision Res* 41:3559–3565
- Lever C, Burton S, Jeewajee A, O’Keefe J, Burgess N (2009) Boundary vector cells in the subiculum of the hippocampal formation. *J Neurosci* 29(31):9771–9777. <https://doi.org/10.1523/JNEUROSCI.1319-09.2009>
- Long NM, Kuhl BA, Chun MM (2018). Memory and attention. In: Wixted JT (ed), Steven’s handbook of experimental psychology and cognitive neuroscience. John Wiley & Sons. <https://doi.org/10.1002/9781119170174.epcn109>
- Marchette SA, Vass LK, Ryan J, Epstein RA (2014) Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nat Neurosci* 17(11):1598–1606. <https://doi.org/10.1038/nn.3834>
- Matsumoto M, Inoue K, Takada M (2018) Causal role of neural signals transmitted from the frontal eye field to the superior colliculus in saccade generation. *Front Neural Circuits*. <https://doi.org/10.3389/fncir.2018.00069>
- McAdams CJ, Reid RC (2005) Attention modulates the responses of simple cells in monkey primary visual cortex. *J Neurosci* 25(47):11023–11033. <https://doi.org/10.1523/JNEUROSCI.2904-05.2005>
- Melcher D (2011) Visual stability. *Philos Trans R Soc Lond Ser B Biol Sci*. 366(1564):468–475
- Melloni L, van Leeuwen S, Alink A, Müller NG (2012) Interaction between bottom-up saliency and top-down control: how saliency maps are created in the human brain. *Cereb Cortex* 22(12):2943–2952. <https://doi.org/10.1093/cercor/bhr384>
- Meyer AF, O’Keefe J, Poort J (2020) Two distinct types of eye-head coupling in freely moving mice. *Curr Biol* 30(11):2116–2130. <https://doi.org/10.1016/j.cub.2020.04.042>
- Miller EK, Li L, Desimone R (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 254(5036):1377–1379. <https://doi.org/10.1126/science.1962197>
- Miller JF, Neufang M, Solway A, Brandt A, Trippel M, Mader I, Schulze-Bonhage A (2013) Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science* 342(6162):1111–1114. <https://doi.org/10.1126/science.1244056>
- Moore T (2006) The neurobiology of visual attention: finding sources. *Curr Opin Neurobiol* 16:159–165
- Moore T, Armstrong KM (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421(6921):370–373. <https://doi.org/10.1038/nature01341>
- Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229:782–784
- Moser M, Rowland DC, Moser EI (2015) Place cells, grid cells, and memory. *Cold Spring Harb Perspect Biol*. <https://doi.org/10.1101/cshperspect.a021808>
- Motter BC (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol* 70(3):909–919
- Muller R (1996) A quarter of a century of place cells. *Neuron* 17:979–990
- Muller RU, Kubie JL (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J Neurosci* 7(7):1951–1968
- Muzzio IA, Kentros C, Kandel E (2009) What is remembered? Role of attention on the encoding and retrieval of hippocampal representations. *J Physiol* 587(12):2837–2854
- O’Connor D, Fukui M, Pinsk M, Kastner S (2002) Attention modulates responses in the human lateral geniculate nucleus. *Nat Neurosci* 5:1203–1209. <https://doi.org/10.1038/nn957>
- O’Keefe J (1976) Place units in the hippocampus of the freely moving rat. *Exp Neurol* 51(1):78–109. [https://doi.org/10.1016/0014-4886\(76\)90055-8](https://doi.org/10.1016/0014-4886(76)90055-8)
- O’Keefe J, Burgess N (1996) Geometric determinants of the place fields of hippocampal neurons. *Nature* 381:425–428. <https://doi.org/10.1038/381425a0>
- O’Keefe J, Conway DH (1978) Hippocampal place units in the freely moving rat: why they fire where they fire. *Exp Brain Res* 31:573–590. <https://doi.org/10.1007/BF00239813>
- O’Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res* 34(1):171–175. [https://doi.org/10.1016/0006-8993\(71\)90358-1](https://doi.org/10.1016/0006-8993(71)90358-1)
- O’Keefe J, Nadel L (1978) *The hippocampus as a cognitive map*. Oxford University Press
- Peters JC, Roelfsema PR, Goebel R (2012) Task-relevant and accessory items in working memory have opposite effects on activity in extrastriate cortex. *J Neurosci* 32(47):17003–17011. <https://doi.org/10.1523/JNEUROSCI.0591-12.2012>
- Posner MI (1980) Orienting of attention. *J Exp Psychol* 32(1):3–25. <https://doi.org/10.1080/00335558008248231>
- Rensink RA, O’Regan JK, Clark JJ (1997) To see or not to see: the need for attention to perceive changes in scenes. *Psychol Sci* 8:368–373
- Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647. <https://doi.org/10.1146/annurev.neuro.26.041002.131039>
- Rowland DC, Roudi Y, Moser M-B, Moser EI (2016) Ten years of grid cells. *Annu Rev Neurosci* 39:19–40. <https://doi.org/10.1146/annurev-neuro-070815-013824>
- Saleem KS, Suzuki W, Tanaka K, Hashikawa T (2000) Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J Neurosci* 20(13):5083–5101
- Savelli F, Yoganarasimha D, Knierim J (2008) (2008) Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus* 18:1270–1282. <https://doi.org/10.1002/hipo.20511>
- Scaplen KM, Gulati AA, Heimer-McGinn VL, Burwell RD (2014) Objects and landmarks: hippocampal place cells respond differently to manipulations of visual cues depending on size, perspective, and experience. *Hippocampus* 24(11):1287–1299
- Scaplen KM, Ramesh RN, Nadvar N, Ahmed OJ, Burwell RD (2017) Inactivation of the lateral entorhinal area increases the influence of visual cues on hippocampal place cell activity. *Front Syst Neurosci*. <https://doi.org/10.3389/fnsys.2017.00040>
- Schall JD (2004) On the role of frontal eye field in guiding attention and saccades. *Vision Res* 44:1453–1467. <https://doi.org/10.1016/j.visres.2003.10.025>
- Schall JD (2013) Production, control, and visual guidance of saccadic eye movements. *ISRN Neurol*. <https://doi.org/10.1155/2013/752384>
- Schütz AC, Braun DI, Gegenfurtner KR (2011) Eye movements and perception: a selective review. *J vis* 11(5):1–30. <https://doi.org/10.1167/11.5.9>
- Skaggs WE, McNaughton BL (1998) Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. *J Neurosci* 18(20):8455–8466
- Stewart EE, Valsecchi M, Schutz AC (2020) A review of interactions between peripheral and foveal vision. *J vis* 20(12):2
- Tatler BW, Hayhoe MM, Land MF, Ballard DH (2011) Eye guidance in natural vision: reinterpreting salience. *J Vis* 11(5)

- Taube JS, Muller RU, Ranck BJ Jr (1990) Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J Neurosci* 10:420–435
- Theeuwes J (2019) Goal-driven, stimulus-driven, and history-driven selection. *Curr Opin Psychol* 29:97–101
- Theeuwes J, Bogaerts L, van Moorselaar D (2022) What to expect where and when: how statistical learning drives visual selection. *Trends Cogn Sci* 26(10):860–872
- Tolman EC (1948) Cognitive maps in rats and man. *Psychol Rev* 55(4):189–208
- Treisman AM (1964) Selective attention in man. *Br Med Bull* 20:12–16
- Treisman AM, Gelade G (1980) A feature-integration theory of attention. *Cogn Psychol* 12:97–136
- Triesch J, Ballard D, Hayhoe MM, Sullivan BT (2003) What you see is what you need. *J vis* 3:86–94. <https://doi.org/10.1167/3.1.9>
- Tsao A, Moser M-B, Moser E (2013) Traces of experience in the lateral entorhinal cortex. *Curr Biol* 23(5):399–405. <https://doi.org/10.1016/j.cub.2013.01.036>
- Tsao A, Sugar J, Lu L, Wang C, Knierim JJ, Moser M-B, Moser EI (2018) Integrating time from experience in the lateral entorhinal cortex. *Nature* 561:57–62. <https://doi.org/10.1038/s41586-018-0459-6>
- Ungerleider LG, Courtney SM, Haxby JV (1998) A neural system for human visual working memory. *Proc Natl Acad Sci* 95:883–890
- Van Beest EH, Mukherjee S, Kirchberger L, Schnabel UH, van der Togt C, Teeuwen RR, Self MW (2021) Mouse visual cortex contains a region of enhanced spatial resolution. *Nat Commun* 12:4029. <https://doi.org/10.1038/s41467-021-24311-5>
- Van der Stigchel S, Hollingworth A (2018) Visuospatial working memory as a fundamental component of the eye movement system. *Curr Dir Psychol Sci* 27(2):136–143
- Wang L, Krauzlis RJ (2018) Visual selective attention in mice. *Curr Biol* 28:676–685. <https://doi.org/10.1016/j.cub.2018.01.038>
- Ward LM (2008) Attention. *Scholarpedia* 3(10):1538. <https://doi.org/10.4249/scholarpedia.1538>
- Wilber AA, Clark BJ, Demecha AJ, Mesina L, Vos JM (2015) Cortical connectivity maps reveal anatomically distinct areas in the parietal cortex of the rat. *Front Neural Circuits* 8:146. <https://doi.org/10.3389/fncir.2014.00146>
- Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261(5124):1055–1058. <https://doi.org/10.1126/science.8351520>
- Witter MP, Doan TP, Jacobsen B, Nilssen ES, Ohara S (2017) Architecture of the entorhinal cortex. A review of entorhinal anatomy in rodents with some comparative notes. *Front Syst Neurosci* 11:46. <https://doi.org/10.3389/fnsys.2017.00046>
- Wolfe JM (1994) Guided search 2.0 a revised model of visual search. *Psychon Bull Rev* 1(2):202–238
- Wolfe JM (2007) Guided search 4.0: current progress with a model of visual search. In: Gray WD (ed) *Integrated models of cognitive systems*, Oxford University Press. pp. 99–119, <https://doi.org/10.1093/acprof:oso>
- Woodman GF, Carlisle NB, Reinhart RM (2013) Where do we store the memory representations that guide attention? *J Vis* 13(3):1–17. <https://doi.org/10.1167/13.3.1>
- Wurtz R (2008) Neuronal mechanisms of visual stability. *Vis Res* 48:2070–2089
- Yarbus AL (1967) *Eye movements and vision*. Plenum Press, New York
- Zeki SM (1976) The functional organization of projections from striate to prestriate visual cortex in the rhesus monkey. *Cold Spring Harb Symp Quant Biol* 40:591–600

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.