Spatial updating of attention across eye movements: A neuro-computational approach

Julia Bergelt
Artificial Intelligence, Department of Computer Science, Chemnitz University of Technology, Chemnitz, Germany

Fred H. Hamker
Artificial Intelligence, Department of Computer Science, Chemnitz University of Technology, Chemnitz, Germany

While we are scanning our environment, the retinal image changes with every saccade. Nevertheless, the visual system anticipates where an attended target will be next and attention is updated to the new location. Recently, two different types of perisaccadic attentional updates were discovered: predictive remapping of attention before saccade onset (Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011) and lingering of attention after saccade (Golomb, Chun, & Mazer, 2008; Golomb, Pulido, Albrecht, Chun, & Mazer, 2010). We here propose a neuro-computational model located in lateral intraparietal cortex based on a previous model of perisaccadic space perception (Ziesche & Hamker, 2011, 2014). Our model can account for both types of updating of attention at a neural-systems level. The lingering effect originates from the late updating of the proprioceptive eye-position signal and the remapping from the early corollary-discharge signal. We put these results in relationship to predictive remapping of receptive fields and show that both phenomena arise from the same simple, recurrent neural circuit. Thus, together with the previously published results, the model provides a comprehensive framework for discussing multiple experimental observations that occur around saccades.

Introduction

During natural vision, scene perception depends on exploratory scanning using accurate targeting of attention, saccadic eye movements, anticipation of the physical consequences of motor actions, and continuous integration of visual inputs with stored representations of previously viewed portions of the scene. For example, when there is an impending eye movement, the visual system can anticipate where the target will appear on the retina after the eye movement, and in preparation for this, spatial attention updates and moves to a new location. In subjects instructed to monitor a particular location in a scene while moving the eyes, two different types of spatial-attention shifts have recently been discovered. In one type, spatial attention lingers after a saccade at the (irrelevant) retinotopic position—that is, the focus of attention appears to shift with the eyes but updates to its original world-centered position only after the eyes land at the saccade target location (Golomb, Chun, & Mazer, 2008; Golomb, Pulido, Albrecht, Chun, & Mazer, 2010). Another study by Rolfs, Jonikaitis, Deubel, & Cavanagh (2011) shows that shortly before saccade onset, a locus of attention appears at a position opposite to the direction of the saccade, which suggests an anticipatory correction of the effects of eye movements. While these results initially appear contradictory, Jonikaitis, Szinte, Rolfs, and Cavanagh (2013) have shown that both updating mechanisms occur simultaneously. Around the time of an eye movement, they detected attentional effects at two different locations at the same time, although only one location was cued. This suggests that there must be at least two attention pointers in addition to the saccade target location that are active around saccades. While the anticipatory shift of attention opposite to saccade direction may be potentially useful, although a bit too early, the shift of attention in the direction of the saccade may appear to be an error or at least a delay in the spatial updating. Furthermore, it is not clear whether the recently observed phenomena of spatial updating of attention relate to the observation of predictive remapping of receptive fields. In the seminal study of Duhamel, Colby, and Goldberg (1992), a flashed stimulus in the future receptive field—that is, the location of a neuron’s receptive field after saccade—evoked a neural response prior to saccade. While this can be interpreted as an anticipatory shift of the receptive field, Cavanagh, Hunt, Afraz, and Rolfs (2010) suggested that it may also be explained by...
learned horizontal or lateral connections. Furthermore, they proposed that such a transfer of activation may also be responsible for spatial updating of attention.

Recently, Ziesche and Hamker (2011) proposed a model of perisaccadic perception to describe the underlying mechanisms of perceiving a stable world during eye movements. It uses gain fields and basis-function networks to perform coordinate transformations between different frames of references—more precisely, between eye- and head-centered reference frames—which may possibly take place in the parietal cortex, specifically the lateral intraparietal cortex (LIP). The model accounts for predictive remapping using two eye-position-related signals, a discrete eye-position signal, and a corollary-discharge signal to compute the perceived position of stimuli across saccades. Ziesche and colleagues demonstrated that the model is also able to explain the perisaccadic mislocalization of briefly flashed stimuli in complete darkness (Ziesche & Hamker, 2011) and the observation of saccadic suppression of displacement (Ziesche, Bergelt, Deubel, & Hamker, 2017; Ziesche & Hamker, 2014).

We here explore by means of the neuro-computational model the relationship between predictive remapping of receptive fields (Duhamel et al., 1992) and predictive remapping of attention. How do these phenomena occur at the neural-systems level? Might both recruit the same neural mechanisms? If yes, why does the attention pointer update opposite to saccade direction, while the receptive fields update with the saccade vector?

### Neuro-computational model

Due to lack of data and sometimes even conflicting data, a pure bottom-up approach at the systems level of brain networks is presently not feasible for neuro-computational modeling. Thus, in general a mixture has to be made between theoretical guidance of the computational framework and evidence given by data. Ziesche and Hamker (2011, 2014) made three assumptions when designing the neuro-computational model: First, existing perisaccadic effects may be explainable on basis of two eye-related signals that have been shown to exist in the brain, a (proprioceptive) eye-position signal and a corollary discharge. Second, gain fields are assumed to be Gaussian shaped, although measurements of gain fields show more variability. Third, as the (proprioceptive) eye-position signal is head centered but the corollary discharge is eye displacement—thus retinocentric—and both eye signals jointly contribute to the final decision, they are required to communicate with each other and operate in the same reference frame. Unlike previous models, this model accounts for the full temporal dynamics around saccades. However, it has so far simulated actual experiments using a simplified one-dimensional design of visual space. To allow for realistic simulations of more complex experiments, we started by extending the original model to operate in two image dimensions. This two-dimensional model uses the same input signals, interactions, and concepts as the one-dimensional model proposed by Ziesche and Hamker. In addition to the existing inputs, we expanded the model by including a fourth input signal to introduce top-down attention to the system.

### Core components

As the neuro-computational model is based on gain fields and basis functions, we will first give a brief overview of these two core components.

In gain fields, the response of a neuron is defined by the neuron’s receptive-field location and a modulatory factor. This modulatory factor can be, for example, an eye-position, an attention-position, or a limb-position signal. Importantly, gain fields do not change the position or width of a receptive field, but rather scale the response to a stimulus. Thus, the activity $r$ of a neuron can be written as

$$r(x, y) = f(x)g(y), \quad (1)$$

with $f(x)$ defining the receptive field and $g(y)$ a gain factor (Salinas & Sejnowski, 2001). Gain fields have been found in several brain areas, including the frontal eye field (Cassanello & Ferrera, 2007) and LIP (Andersen, Bracewell, Barash, Gnadt, & Fogassi, 1990). However, the gain field defined in Equation 1 requires nonzero gain signals, which works well for eye or limb position but may lead to difficulties in attention or in all sorts of phasic signals. For such phasic signals an alternative gain field can be used, which follows typical definitions of attentive gain modulation, such as

$$r(x, y) = f(x)(1 + g(y)) = f(x) + f(x)g(y). \quad (2)$$

Although eye-position gain fields often thought to be planar or sigmoidal, Lehky, Sereno, and Sereno (2016) have discussed the fact that the evidence for these particular functions is not overtly strong, and shown that differently shaped gain fields—elliptical, hyperbolic, or even a combination of different shapes—are adequate to represent an eye position. According to that study, the variation within a specific shape is more important than the general shape of gain fields. Following Pouget, Deneve, and Duhamel (2002), we use a Gaussian-shaped gain field; from a computational perspective, it is a special case of an elliptical gain field. Besides, a Gaussian function can be approximated by a linear combination of several.
sigmoid functions, as illustrated in Figure 1. Thus, a Gaussian gain-field response can be interpreted as a weighted sum of multiple sigmoidal gain fields. To be clear, we do not think that this particular gain field is an exact description of the gain fields in the brain; we think of it rather as one possible implementation of this concept.

According to Salinas and Sejnowski (2001), gain fields as defined in Equation 1 can be used for coordinate transformations between multiple frames of references. As such transformations are usually nonlinear, the output cannot be received through a weighted sum of inputs, and instead a network with at least one intermediate nonlinear map is required (Pouget et al., 2002). Pouget et al. introduced a model where the intermediate map consists of basis-function units. Known from linear algebra, basis functions are used to build more complex functions by a weighted sum of these basis functions. Thus, Pouget et al. combined two input maps (a retinocentric stimulus response and an eye-position signal) multiplicatively to form an intermediate basis-function map and compute via a linear combination an output response, which is the stimulus position in a head-centered reference frame. Although they demonstrated their model only for this kind of transformation, the same principles can be used for other frames of reference, such as limb- or body-centered coordinate systems. Figure 2 shows two examples where a retinotopic response is transformed into a head-centered response using eye position. In the examples, the eye fixates at 30° in head-centered coordinates. The activity of the eye-position signal can be either tonic (Figure 2, left) or phasic (Figure 2, right) and is called $xe_{\text{tonic}}$ or $xe_{\text{phasic}}$, respectively. It is modeled as a Gaussian function centered on the current eye position. Additionally, a stimulus response centered at $-10^\circ$ in retinotopic coordinates is given. It is likewise modeled as Gaussian function, and labeled $xr$. Both signals are multiplicatively combined in the basis-function map $xb$ according to Equation 1 or 2, depending on the temporal dynamics of the eye-position signal. In the tonic case, the activity of $xb$ is formed by $xb = xr \times xe_{\text{tonic}}$; in the phasic case it is calculated by $xb = xr \times (1 + xe_{\text{phasic}})$. This leads in both cases to a two-dimensional Gaussian blob at $(30^\circ, -10^\circ)$. In addition to this blob, there is a band of activity in the phasic case. This band is stretched over the whole eye-position dimension of $xb$ and is centered, like $xr$, at the retinotopic stimulus position at $-10^\circ$. A head-centered response of the stimulus position can be computed from the basis-function map, if each neuron $k$ of $xh$ is connected to the same neuron in $xb$. As a result, the activity of $xh$ is centered at $20^\circ$, the stimulus response in head-centered coordinates. The only difference between the tonic and phasic cases is the baseline activity of $xh$, which is zero in the tonic case and greater than zero in the phasic case due to the band of activity in $xb$.

The example is simplified to a one-dimensional space for the purpose of better visualization. However, our model implements two dimensions of visual space; thus, rather than 2-D basis functions, 4-D basis functions are required. Considering how the 2-D basis-function map is formed, it is easy to imagine that a 4-D basis-function map consists of four dimensions, where the first two belong to horizontal and vertical
information of the first input (e.g., a two-dimensional visual space of stimulus position $x_r$) and the last two belong to horizontal and vertical information of the second input (e.g., a two-dimensional space of eye position $x_e$). Reading out a four-dimensional diagonal can then be done by first projecting the 4-D basis-function map to two separate 2-D maps containing only horizontal and vertical information, respectively, by maximizing over the two dimensions containing the disregarded information. Then we read out the two two-dimensional diagonals separately and finally re-compose them to form again a 2-D space with horizontal and vertical dimensions. For more details of how this diagonal readout is calculated mathematically, see the Appendix.

In perisaccadic visual perception, we have to define how the movement of the eyes should be considered in the model: We can use a continuously updating eye-position signal over time or instead a discretely updating pre- and a post-saccadic eye position. Inspired by physiological data, we decided to use two eye-position-related signals, a tonic signal for the pre- and postsaccadic eye position and a phasic signal for the corollary discharge.

The first signal, called the proprioceptive (PC) signal, encodes the current eye position in head-centered coordinates and may originate in the primary somatosensory cortex (Wang, Zhang, Cohen, & Goldberg, 2007) or in the central thalamus (Tanaka, 2007). In monkeys, the eye-position signal has been shown to update around 60 ms after saccade completion (Y. Xu, Wang, Peck, & Goldberg, 2011). In humans, updating latencies have not been measured, but there is evidence that updating latencies may be species specific, depending on the behavioral task. For example, in the task on saccadic suppression of displacement, data for monkeys and humans appear to differ (Joiner, Cavannaugh, FitzGibbon, & Wurtz, 2013). Summarizing, the PC signal is a tonic, late-updating signal encoded in a head-centered reference frame. The yellow dashed lines symbolize the interactions for the largest activity—that is, how the inputs $x_e*$ and $x_r$ are fed into $x_b$ and how the activity of $x_b$ is read out and projected to output $x_h$.
activity at the old position and a rise of activity at the new one. Due to the late updating, the PC signal is initially a presaccadic eye-position signal.

The second eye-position-related signal, called a corollary-discharge (CD) signal, encodes the eye displacement retinotopically, as it is a copy of the oculomotor command and is only active around saccade onset. It is likely routed from the superior colliculus via the mediodorsal thalamus to the frontal eye field (FEF; Sommer & Wurtz, 2004), where it may be transferred into a gain-field representation. There is no direct evidence for this assumption, but gain-field representations have also been observed in the FEF (Cassanello & Ferrera, 2007). From such a gain field, the expected post-saccadic eye position in head-centered coordinates can be extracted. Thus, the transferred CD signal can be interpreted as the expected post-saccadic eye position. The primary CD signal is a retinotopic, phasic signal rising before and decaying shortly after saccade onset. It is modeled as a two-dimensional spatial signal representing again horizontal and vertical dimensions of visual space.

Next to these two eye-position-related signals, the model receives input from visual stimuli. The stimulus response is encoded in a two-dimensional retinocentric reference frame representing horizontal and vertical dimensions of a visual scene, and models early extrastriate visual areas like MT or V4. The response to a stimulus is modeled by a two-dimensional Gaussian function centered at the retinotopic stimulus position and including decaying strength over time. This means the longer the stimulus is presented, the weaker the response is. The width of the Gaussian is directly proportional to the distance between fixation point and stimulus position, to account for effects of larger receptive fields in the periphery. Additionally, we added a delay in the response to a stimulus to account for the latency in the visual pathway.

We further expanded the model by a fourth input signal to introduce top-down attention to the system. The top-down attention signal is encoded in a two-dimensional head-centered reference frame modeling, for instance, a head-centered position kept in memory. The emergence of this signal is not explicitly modeled here, but it might arise from memory recall in areas such as the medial temporal lobe (Byrne, Becker, & Burgess, 2007). The two dimensions of the reference frame represent, again, horizontal and vertical dimensions of visual space.

Figure 3 shows the temporal dynamics of the four inputs for a simple example. After fixation at the fixation point (FP) for 200 ms, a saccade is executed to the saccade target (ST) lasting 41 ms. The PC signal updates after saccade offset from FP to ST. Thus, the firing rate at FP decays after saccade offset, while the PC signal at ST increases. The CD signal at ST-FP, as it encodes the eye displacement, rises before saccade onset, peaks 10 ms after saccade onset, and decays. At 150 ms before saccade onset, a stimulus is presented at the head-centered stimulus position (SP) for 80 ms. The retinal signal has a delay of 50 ms, and thus it starts 50 ms after stimulus onset at the retinotopic stimulus position—that is, SP-FP. Afterward it decays, first with a Gaussian decay and then, with stimulus offset, linearly. In experiments with top-down attention at the attention position (AP), the attention signal is active all the time.

Model structure and dynamics

Figure 4 shows the structure of the new 2-D model with its maps and their interactions.

The two eye-related input signals, eye position and corollary discharge, feed into two two-dimensional maps $X_{PC}$ and $X_{CD}$. Note that according to the input signal, $X_{PC}$ is organized in a head-centered reference frame, whereas $X_{CD}$ is organized in an eye-centered reference frame. As the two signals should be progressed in the same manner, they have to operate in the same frame of reference. A simple solution to transferring the eye-centered CD signal into a head-centered signal is to use the PC signal with the concept of coordinate transformation described previously. Following this concept, the basis-function map, called $X_{CEF}$, that multiplicatively combines the CD and PC signal is required to be—from a

![Figure 3. Example of typical temporal dynamics of input signals. At 150 ms before saccade onset, a stimulus is presented for 80 ms. The retinal signal starts 50 ms after stimulus onset. The signal decays with Gaussian distribution for the duration of stimulus presentation and linearly afterwards. The corollary-discharge (CD) signal rises before saccade onset, peaks 10 ms after saccade onset, and decays. The proprioceptive (PC) signal updates after saccade offset by decaying at the presaccadic position (fixation point; FP) and increasing at the post-saccadic position (saccade target; ST). If we use top-down attention, the signal at the attention position (AP) is active throughout the whole trial.](image-url)
computational point of view—a four-dimensional map, as each input is two-dimensional. The structural equation for the activity in map $X_{e_{\text{FEF}}}$ can be written as

$$X_{e_{\text{FEF}}} = X_{e_{\text{PC}}} \times X_{e_{\text{CD}}}.$$ 

The retinal signal feeds into two maps assumed to be located in LIP, where it is gain modulated by either the eye-position signal or the corollary-discharge signal to obtain a joint representation of the stimulus position and the eye position or eye displacement, respectively. The two LIP maps, called $X_{b_{\text{PC}}}$ and $X_{b_{\text{CD}}}$ according to the modulating signal, combine two two-dimensional signals multiplicatively and work like two separate basis-function maps. They again—from a computational point of view—have to be designed as four-dimensional maps. As the modulating signal is either tonic (PC) or phasic (CD), the feedforward activity in these maps is calculated by either Equation 1 or 2. Both LIP maps interact with each other via neurons organized in a two-dimensional head-centered reference frame $X_{h}$ using feedback projections. This feedback is modulated by the corresponding eye-related signal as well, and is crucial for predictive remapping, as shown later. The structural equations for the activity of both LIP maps are then

$$X_{b_{\text{PC}}} = X_r \times X_{e_{\text{PC}}} + X_h \times X_{e_{\text{PC}}}$$  
$$X_{b_{\text{CD}}} = X_r \times (1 + X_{e_{\text{FEF}}}) + X_{e_{\text{FEF}}} \times X_{e_{\text{FEF}}}.$$ 

Due to the feedforward and feedback components, there will be at least two distinct activity blobs in each LIP map simultaneously in the perisaccadic case. In addition, as the update of the PC signal is defined by a decreasing activity at the presaccadic position and an increasing activity at the postsaccadic position, the activity in $X_{b_{\text{PC}}}$ updates jump-like, but smooth shifts of activity are possible to a certain degree.

The neurons of $X_{h}$ combine the activities from both LIP maps in an additive manner as follows:

$$X_{h} = X_{b_{\text{PC}}} + X_{b_{\text{CD}}}.$$ 

In $X_{h}$ is encoded the perceived spatial position of a stimulus in a head-centered reference frame. Following the idea of LIP as a priority map (Bisley, 2011; Goldberg, Bisley, Powell, & Gottlieb, 2006), any activation in LIP is interpreted as an attentional signal, and we refer to this as an attention pointer. In order to simulate the dynamics of spatial attention...
around eye movements, we can apply the model in two operational modes: a bottom-up mode where we cue attention by a stimulus and observe how the initially cued activity evolves during saccade, and a top-down mode where we set a head-centered top-down pointer into the model by providing a static signal to $X_h$ and observe the dynamics of attention while the top-down signal is being transformed by the two eye-position signals traveling through the feedback pathways toward an eye-centered representation of spatial attention.

Results

Predictive remapping

Predictive remapping is a phenomenon where the retinotopic position of a spatial-restricted neuronal receptive field (RF) shifts in anticipation of a saccade such that the neuron responds to stimuli in the future receptive field (FRF) before the eye-movement onset. Predictive remapping was first observed in LIP by Duhamel et al. (1992), and has been proposed to play a key role in establishing the percept of a visually stable world (Mathôt & Theeuwes, 2010; Sommer & Wurtz, 2006; Wurtz, 2008). Predictive remapping has been observed in many other areas: extrastriate visual cortex (Nakamura & Colby, 2002), superior colliculus (Walker, Fitzgibbon, & Goldberg, 1995), and FEF (Sommer & Wurtz, 2006; Umeno & Goldberg, 1997). However, in these early studies predictive remapping was not clearly delineated from other RF shifts, as discussed by Zirnsak, Lappe, and Hamker (2010)—which triggered new experimental observations (Hartmann, Zirnsak, Marquis, Hamker, & Moore, 2017; Neupane, Guittion, & Pack, 2016a, 2016b; Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014).

This report focuses on modeling the neural substrates of the original predictive-remapping phenomenon, as described by Duhamel et al. (1992). In our model, predictive remapping arises from the feedback connection from $X_h$ to $X_{bCD}$. To demonstrate this, we conducted two simple experiments, similar to those of Duhamel et al. In the first experiment, called the fixation task, a stimulus is presented in the RF for 100 ms while the eyes remain fixed at a particular location—FP at $(0^\circ, 0^\circ)$. The second experiment is called the saccade task, as a saccade is executed from the FP to the ST at $(10^\circ, 8^\circ)$. At 150 ms before saccade onset, a stimulus is presented for 100 ms in the FRF, which is at the position of the RF after the eyes have landed at the saccade target. The eye movement is simulated with the model of Van Wetter and Van Opstal (2008), and for this saccade length, it was 65 ms in duration. The setup as well as the positions and temporal dynamics of the three input signals—eye position, corollary discharge, and retinal signal—are shown in Figure 5.

The results of the simulations are shown in Figure 6. To visualize the 4-D LIP maps, we projected each map to two two-dimensional planes representing either the horizontal or the vertical information by maximizing over the two disregarded dimensions. Further, we show the neural activities of both LIP maps projected into the retinotopic space together with the spatial setup. In the fixation task (Figure 6A), a stimulus is presented in the RF, no saccade is planned, and the eyes fixate at the FP. The LIP map $X_{bPC}$ combines the eye-position signal at $(0^\circ, 0^\circ)$ and the retinal signal at $(-5^\circ, -2^\circ)$ multiplicatively, thus resulting in a single locus of activity at the crossing of both signals, at $(0^\circ, -5^\circ)$ for horizontal and $(0^\circ, -2^\circ)$ for vertical, respectively. As the retinal signal feeds continuously into the LIP map $X_{bCD}$, we get an activity line along one axis at the position of the stimulus (at $-5^\circ$ for horizontal and $-2^\circ$ for vertical, respectively). In the saccade task (see Figure 6B), the CD signal rises shortly before saccade onset and modulates the retinal signal of the stimulus presented in the FRF in such a way that it increases the neural gain at a single blob where the CD signal position crosses the retinal signal, at $(10^\circ, 5^\circ)$ for horizontal and $(8^\circ, 6^\circ)$ for vertical, respectively. The activity in LIP PC is the result of the multiplicative interaction of the PC signal and the retinal signal, at $(0^\circ, 5^\circ)$ for horizontal and $(0^\circ, 6^\circ)$ for vertical, respectively.

Activity along each diagonal of $X_{bPC}$—for example, along the yellow dashed line—is read out and fed to one neuron in $X_h$. Likewise, the activity in the LIP CD map is read out along each diagonal and projected to $X_h$. At the same time, the activity of $X_h$ is fed back to both LIP maps along each diagonal, which allows an interaction between $X_{bPC}$ and $X_{bCD}$ via $X_h$. In LIP CD, the feedback from $X_h$ is combined multiplicatively with the CD signal. This leads to a second locus of activity in LIP CD, at $(10^\circ, -5^\circ)$ for horizontal and $(8^\circ, -2^\circ)$ for vertical, respectively. Note that the position of this second locus of activity has the same “visual” position as if we presented the stimulus in the RF (compare Figure 6A). That means the neurons in $X_{bCD}$ anticipate the updating of the stimulus position from FRF to RF with the help of the CD signal and the feedback from $X_h$. Thus, the $X_{bCD}$ neurons show predictive remapping. In contrast, the LIP map for the PC signal has no such predictive component, and contains only one single activation blob that represents the stimulus position with respect to the current eye position. Supplementary Movie S1 shows the development of the activities.
in both LIP maps over the course of a single simulated trial.

To illustrate the necessity of the feedback from $X_h$ to LIP for predictive remapping, we simulated the same experiments without the connections from $X_h$ to $X_b$ PC and $X_b$ CD, respectively. As can be seen in Figure 7, the main effect of the missing feedback besides reduced activity in both LIP maps is the nonexistent activity blob in LIP CD in the saccade task—that is, no predictive remapping.

To summarize, our model suggests that predictive remapping in LIP is generated with the help of the CD signal. When a stimulus is presented, the eye-centered retinal signal feeds into both LIP maps. Thus, neurons whose RFs match with the stimulus position become active. Long before a saccade is planned, the eye-centered stimulus position (from $X_r$) is transformed into a head-centered position through $X_b$ PC with the help of the eye-position signal and encoded in $X_h$. Shortly before saccade onset, the CD signal rises, and with the feedback from $X_h$ it activates neurons in $X_b$ CD which encode the stimulus with respect to the future eye position. Thus, if the stimulus presented in the FRF is encoded with reference to ST, this stimulus—in a retinocentric reference frame—leads to a neural activation at RF, as the eyes are still at FP, and consequently predictive remapping emerges. Predictive remapping does not require all-to-all connections, only a diagonal connectivity scheme within a small recurrently connected network of neurons.

Spatial updating of attention

We particularly compare our simulation results to the data of Jonikaitis et al. (2013), who demonstrated both types of attentional updating in their data: predictive remapping of attention reported by Rolfs et al. (2011) and lingering of attention reported by Golomb and colleagues (Golomb et al., 2008; Golomb et al., 2010). Predictive remapping of attention occurs shortly before saccade onset, and spatial attention is remapped to a position opposite to the direction of the saccade. In contrast, Golomb and colleagues have reported that after a saccade, spatial attention lingers at the (irrelevant) retinotopic position—that is, the focus of attention shifts with the eyes but updates to the original fixed world-centered position only after the eyes have landed. Using our model we replicated the observation of Jonikaitis et al. and shed light onto the
possible mechanisms that might account for the experimental findings.

Jonikaitis et al. (2013) conducted two slightly different variants of their experiment. In one the attentional cue was turned off during saccade (task with transient cue), and in the other the cue was shown until the end of the experiment (task with sustained cue). They determined the locus of attention by measuring performance in a discrimination task before and after the saccade at three different positions: the attention position (AP), where the attentional cue was presented; the lingering attention position (LAP)—that is, the AP shifted by the saccade vector; and the remapped attention position (RAP)—that is, AP shifted by the reverse saccade vector. For both tasks, they found an attentional benefit at AP and LAP after the saccade. They also observed a small but significant attentional benefit at RAP after saccade in the sustained-cue task, which did not appear
in the transient-cue task. Since both transient and sustained cues led to similar attentional effects, it is possible that attention is largely driven by cue onset. Further, previous recordings in V4 have shown a strong decay in activation of a permanently present stimulus (Fischer & Boch, 1985). Thus, we simulated the study using a brief appearance of the attention cue. We used the same spatial setup as Jonikaitis et al.: The model executes an 8° saccade to the right (modeled as described previously), lasting 53 ms. Before saccade onset, attention is directed at the AP located 6° above the fixation point due to a cuing stimulus at this position shown 180 ms before saccade onset for 10 ms. To account for the latency in the visual pathway, the activation of Xr in the model starts 50 ms after stimulus onset. The spatial layout and temporal dynamics of the input signals—eye position, corollary discharge, and retinal signal—are shown in Figure 8.

While Jonikaitis et al. (2013) used a behavioral paradigm to estimate the amount of attention distributed across the saccade, we directly plot the LIP activity as a measure of attention. Figure 9 shows the activity of both maps at different times. Again, since the LIP maps are four-dimensional, for visualization we projected them to two two-dimensional planes representing either the horizontal or the vertical information of this map. In our model, each LIP map triggers an attention pointer in a retinotopic reference frame. At the beginning, before the saccade, the only inputs to the model are an eye-position (PC) signal encoding the current eye position at FP and a retinal signal to Xr encoding the cued attention position at AP. The retinal signal is combined multiplicatively with the PC signal in LIP PC, which leads to a single activation blob, at (0°, 0°) for horizontal and (0°, 6°) for vertical, respectively. Projecting the activity of the LIP map back to visual space shows the attention pointer (red blob) at the desired attention position, as shown in Figure 9A. Additionally, the cue generates an activity line in the second LIP map XrCD, which leads to an attention pointer at AP (covered by the red blob).

Meanwhile, both LIP maps interact with each other via
the neurons of $X_h$. All activity in $X_{bPC}$ and $X_{bCD}$ is summed up along each diagonal, fed into $X_h$, and projected back along the same diagonal. Thus, the activity of $X_h$ is projected back into $X_{bPC}$ along the diagonal, where it is combined multiplicatively with the PC signal. This sustains the same position as initially triggered by the cue. Importantly, as the saccade is being planned the CD signal rises. It feeds into $X_{bCD}$ and is multiplicatively combined with the reentrant signal from $X_h$ feeding in along the diagonal. This leads to an activity blob in LIP CD, at $(8^\circ, 0^\circ), (0^\circ, 6^\circ)$ for horizontal and $(0^\circ, 6^\circ)$ for vertical, respectively. Projected to visual space, this activity triggers a second attention pointer at RAP $8^\circ$ to the left of AP (see Figure 9B, blue blob). This means that shortly before saccade onset, there are two spatial locations that exhibit attentional facilitation: the remapped attention position and the attention position itself. During saccade, the triggered attention pointers are shifted along with the eye movement as they are encoded in a retinotopic reference frame. Thus, after the saccade, the attention pointers are shifted by $8^\circ$ to the right. Therefore, the attention pointer induced by $X_{bPC}$ is now at LAP and the attention pointer induced by $X_{bCD}$ is at AP (Figure 9C). As the CD signal decays after saccade onset, the activity in $X_{bCD}$ also decays, and the attention pointer induced by this map is gradually removed until it is completely extinguished. After the eyes have landed at the saccade target (ST), the PC signal updates to the new eye position, and thus the activity in $X_{bPC}$ updates as well—to $(8^\circ, 0^\circ)$ for horizontal and $(0^\circ, 6^\circ)$ for vertical, respectively—and the attention pointer triggered by this map is remapped back to AP (see Figure 9D). Supplementary Movie S2 shows the development of the activities in both LIP maps over the course of a single simulated trial.

Figure 10 shows the time course of attention at AP, RAP, and LAP over time. The three panels show the effect separated for the two LIP maps (red and blue lines, respectively) as well as the sum of the activities of both LIP maps (purple dashed lines) for each position. To obtain the activity of a LIP map at a certain spatiotopic position, we first project the LIP maps to visual space. Then the firing rate of the neuron which corresponds to the position is read out. As the projections of the LIP maps are retinotopic, the neurons which correspond to one location (AP, RAP, or LAP) differ before and after saccade. Thus, we plotted the activity of the two corresponding neurons in the pre- and postsaccadic condition separated by the time (in milliseconds) is aligned to saccade onset. The still image shows the inputs marginalized over time.
Figure 9. Simulation results of spatial updating of attention with cued attention for different time steps. The activities of both maps of lateral intraparietal cortex (LIP; $X_{PC}, X_{CD}$) are plotted, as well as the setup and the triggered attention pointers. Both LIP maps interact with each other by summing up all activations along each diagonal toward each neuron in $X_H$ and projecting back along the same diagonal. The yellow dashed line shows the diagonal which has the largest activation. The symbols are identical to Figure 8. The
saccade period (gray bar). RAP is attended only before saccade onset generated by LIP CD cells. In contrast, LAP is attended only after saccade offset through an attention pointer maintained by LIP PC cells. AP receives attention over the whole period, as attention at AP before saccade is triggered by LIP PC and LIP CD, and after saccade the origin of the attention pointer switches from LIP CD to LIP PC.

In top-down mode we cue the desired location of attention to the model by an endogenous signal via $X_h$ modeling top-down attention. To demonstrate that this version of the model equally well explains the observations of Jonikaitis et al. (2013), we simulated the same experiment as before, but instead of a flashed cuing stimulus we introduced top-down attention directed to the attention position as a constant input to $X_h$ (see Figure 4). The spatial layout and temporal dynamics of the input signals—here eye position, corollary discharge, and top-down attention—are shown in Figure 11.

Figure 12 shows the activity of both maps at different times. At the beginning, before the saccade, the only inputs to the model are an eye-position (PC) signal encoding the current eye position at FP and a top-down attention signal to $X_h$ encoding the attention position at AP. The attention signal is projected backward into the LIP map $X_{hPC}$ along the diagonal and is combined multiplicatively with the PC signal, which leads to a single activation blob, at $(0^\circ, 0^\circ)$ for horizontal and $(0^\circ, 6^\circ)$ for vertical, respectively. This activity maintains an attention pointer (red blob) at the desired attention position, as shown in Figure 12A. Shortly before saccade onset, the CD signal rises and feeds into the second LIP map, $X_{hCD}$, where it is multiplicatively combined with the attention signal from $X_h$ feeding in along the diagonal. This leads to an activity blob in LIP CD—at $(8^\circ, C_0 8^\circ)$ for horizontal and $(0^\circ, 6^\circ)$ for vertical, respectively—and a second attention pointer at RAP $8^\circ$ to the left of the attention position (see Figure 12B, blue blob). Thus, like in the previous simulation, the two attention pointers triggered by the LIP maps direct attention to RAP and AP itself shortly before saccade onset. Likewise, shortly after saccade the attention pointers are shifted to AP and LAP, respectively (see Figure 12C), and long after the saccade, when the CD signal has decayed and the PC signal is updated to the saccade target, there is again only one attention pointer triggered by LIP PC.

Figure 10. Attentional effect at different spatial positions over time for cued attention. The three plots show the attentional effect at the three spatial positions RAP (remapped attention position, left), AP (attention position, middle), and LAP (lingering attention position, right) over time. The red lines mark attention triggered in the lateral intraparietal cortex (LIP) map modulated by the proprioceptive (PC) eye position; the blue lines mark attention triggered by the LIP corollary discharge (CD) map. Note that before and after saccade, we have to read out the firing rate of different neurons to obtain the LIP activity corresponding to the same spatial position. The purple dashed lines mark the overall attention at this position. The gray bar covers the period of the eye movement and separates the activity from the two different neurons. The time (in milliseconds) is aligned to saccade onset.
directing attention to AP (see Figure 12D). Supplementary Movie S3 shows the development of the activities in both LIP maps over the course of a single simulated trial.

Discussion

Our simulation results suggest that predictive remapping of receptive fields and spatial updating of attention are two sides of the same coin. Predictive remapping in our model arises from a relatively small, and simple, recurrent neural circuit. Although the model likely simplifies from the biophysical implementation in the brain, much can be explained at this abstract level. The remapping results from the interaction of the phasic corollary-discharge signal and the tonic feedback from LIP, which means it is synchronized with the CD signal—that is, saccade onset. The CD signal has recently been well explored in electrophysiological studies (Sommer & Wurtz, 2004, 2008). The model suggests that this signal is integrated into LIP gain fields. As the CD signal contains information about the future position of the eyes, it contributes to a new reference to which the flashed stimulus is anchored. Thus, predictive remapping of receptive fields can be well understood by the integration of a second spatial reference system within a gain-field-like representation. As the gain fields operate in a retinocentric coordinate frame and the internal eye position updates after a saccade, the flashed stimulus is transferred against saccade direction due to the future eye-position reference, leading to the activation of LIP neurons at the RF location if the stimulus has been flashed at FRF. From the viewpoint of these LIP neurons, it looks like their RFs have shifted forward.

Regardless of the particular experimental paradigm, our model suggests the presence of two pointers around saccades which are linked to different eye-related signals, one to a proprioceptive eye position and the other to the corollary discharge. In the spatial-attention task, the corollary-discharge signal leads to a retinotopic attention pointer at the remapped attention position shortly before saccade onset, which is shifted with the saccade to the original position of attention and is finally removed shortly after the eyes have landed at the saccade target, consistent with Rolfs et al. (2011). Furthermore, the model explains the lingering of attention (Golomb et al., 2008;
Figure 12. Simulation results of spatial updating of attention with top-down attention for different time steps. The activities of both maps of lateral intraparietal cortex (LIP; $X_{\text{APC}}, X_{\text{CD}}$) are plotted, as well as the setup and the triggered attention pointers. Additionally, the diagonal on which the attention signal is fed into the LIP maps is plotted with a yellow dashed line. The symbols and definitions are identical to Figure 9. The time (in milliseconds) is aligned to saccade onset. (A) Long before saccade, the attention
pointer at the desired attention position (AP) is encoded by the LIP map for the proprioceptive (PC) signal (red blob). (B) Shortly before saccade, the corollary-discharge (CD) signal rises and activates neurons in LIP CD, which trigger a second attention pointer at the remapped attention position (RAP, blue blob). (C) Shortly after saccade, both attention pointers are shifted according to the eye movement, as they are retinotopic. This leads to an attention pointer at the lingering attention position (LAP, red blob) and another one at AP (blue blob). (D) Long after saccade, the PC signal updates to the new eye position (saccade target; ST) and the CD signal decays, so there is again only one attention pointer triggered by LIP PC at AP (red blob).

Golomb et al., 2010) by a late-updating, proprioceptive eye-position signal which triggers a retinotopic attention pointer first at the desired position. This pointer is then shifted by the saccade and finally, with the updated signal, remapped to the initial position without covering intermediate positions (Golomb, Marins, Chun, & Mazer, 2011). Hence, we have two attention pointers generated by LIP that are retinotopic. These pointers are shifted during the eye movement according to the saccade vector. Thus, during saccades there is one pointer which moves from the remapped attention position toward the actual attention position and one pointer moving away from the attention position. Consequently, while the eyes are moving, there is no attention directed to the original attention position itself, as shown by Yao, Ketkar, Treue, and Krishna (2016). Furthermore, Yao et al. concluded from their experimental results that the attention position must be attended again within 30 ms after the eyes have landed. In our model, attention is available at the attention position immediately after the saccade, as the remapped attention pointer is then shifted to this position. Taking into account that the decision process in their experiment itself needs time (processing the visual input and making the decision), this fits the data. In addition to attentional cuing, we show that attention may be incorporated by an endogenous attention signal.

The functional roles of the two eye-related signals used in our model have been recently discussed in a review (Sun & Goldberg, 2016). As the (proprioceptive) eye-position signal is inaccurate after saccade, LIP gain fields may not be suitable to solve the spatial-accuracy problem. Therefore, Sun and Goldberg conclude that there are “two different representations of space: a rapid retinotopic one and a slower craniotopic one” (p. 80). In the model presented here, these two representations interact to trigger predictive remapping and spatial updating of attention. Both phenomena can be well explained by a lateral or reentrant network at the level of LIP, such that the presaccadic activation is recomputed in the future reference frame by means of the corollary discharge. For simplicity, we assumed two different neuron types, CD and PC eye-related cells. However, it is likely that the separation is not so complete, and there may be a continuum of cells. The joint effect of both eye-position signals may also explain the quite accurate eye-position decoding observed by Morris, Kubischik, Hoffmann, Krekelberg, and Bremmer (2012).

Previous simulations with the 1-D-model version indicate that mislocalization in total darkness around saccade and saccadic suppression of displacement can be accounted for by the same neural circuits used here (Ziesche et al., 2017; Ziesche & Hamker, 2011, 2014). Thus, as the 1-D version of this model was designed prior to the observation of updating of attention, from the model’s point of view this observation can be considered an inherent prediction. This model prediction might be tested by suppressing or inactivating specific brain areas related to the two eye-position signals, similar to previous studies (Sommer & Wurtz, 2008). For example, if either the PC or CD pathway is deactivated, one could investigate whether lingering and remapping of attention are still present. Additionally, the explicit role of LIP in updating of attention can be tested with such a method. Furthermore, other models of perisaccadic space perception can be investigated in how far they can account for updating of attention as well. For example, with models using only one extraretinal signal—like those of Pola (2004) or Binda, Cicchini, Burr, and Morrone (2009)—one can test whether two extraretinal signals, as proposed with our model, are actually necessary.

Although the newly presented model and the previous 1-D model do a good job of accounting for several outstanding issues, and accurately replicate the behavioral findings of Jonikaitis et al. (2013), there are nevertheless a number of issues that have to be addressed in future work. First, the exact timing of the proprioceptive eye-position signal needs to be explored in more detail. While a recent study in LIP suggests that gain fields may update even after 150 ms (B. Y. Xu, Karachi, & Goldberg, 2012), which is about 80–100 ms later than in our model, Y. Xu et al. (2011) report an update after 60 ms in somatosensory cortex. No such information exists in humans. Our parameters are based on previous versions of the model (Ziesche et al., 2017; Ziesche & Hamker, 2011, 2014) and are mainly motivated to account for human behavioral data. Second, there is some variation in experimental studies regarding the dominance of the
effects in the data which requires future clarifications. For instance, two recent studies investigating attentional effects in V4 (Marino & Mazer, 2018) and MT (Yao, Treue, & Krishna, 2018) seem to reveal contradictory results. Marino and Mazer (2018) recorded the neural response of V4 cells in monkeys performing a saccade while attention was cued at different positions, either in the receptive field, in the future receptive field, or at a control position. For tasks in which attention was cued at the receptive field, the neuronal response decreased after saccade onset to control level, with no indication of a lingering effect. When attention was cued at the future receptive field, the response became greater than control level about 40 ms after saccade onset. Taking into account the neuronal latency, this has been taken as evidence for predictive remapping. Yao et al. (2018) conducted a quite similar experiment but recorded neural responses in MT. In their study, the response became higher than control level after saccade offset (either right after saccade offset or 50 ms later for two different monkeys), which is later than observed by Marino and Mazer in V4. Thus, Yao et al. (2018) concluded that MT shows no predictive remapping of attention. However, they observed evidence for the lingering of attention. In another study, Yao et al. (2016) found neither remapping nor lingering of attention in a psychophysical study with humans, though these results are maybe biased by their experimental design, where remapped and lingering attention position were not only irrelevant but importantly were never target locations that tested the amount of attention. Lisi, Cavanagh, and Zorzi (2015) suggest that spatial updating of attention depends on the setup of the experiment, namely whether visual objects, which can serve as spatial landmarks, are presented or not. In their experiments, they found that the lingering effect vanishes if a placeholder is shown at the attended location for the whole trial. However, the ability to maintain attention at a spatiotopic location during saccades increases with the presence of placeholders. These findings may help to classify the location during saccades increases with the presence of placeholders. These findings may help to classify the various contradictory results and might be used to further elaborate our model of perisaccadic space perception.

**Keywords:** eye movement, neuro-computational model, space perception, predictive remapping, spatial updating of attention

**Acknowledgments**

We thank J. Mazer for his comments and discussions. This work was supported by the US-German collaboration on computational neuroscience (BMBF 01GQ1409) and partly supported by the European Union’s Seventh Framework Programme (FET, Neuro-Bio-Inspired Systems: Spatial Cognition) under grant agreement No. 600785. The publication costs of this article were funded by the German Research Foundation/DFG-392676956 and the Technische Universität Chemnitz in the funding program Open Access Publishing. The source code of the model implemented in Python is available at https://www.tu-chemnitz.de/informatik/KI/supplement/BergeltHamker2019/. The simulations were done with the neural simulator ANNarchy (Vitay, Dinkelbach, & Hamker, 2015).

Commercial relationships: none.
Corresponding author: Fred H. Hamker.
Email: fred.hamker@informatik.tu-chemnitz.de.
Address: Artificial Intelligence, Department of Computer Science, Chemnitz University of Technology, Chemnitz, Germany.

**References**


Ferraina, S., Paré, M., & Wurtz, R. H. (2002). Comparison of cortico-cortical and cortico-collic-
ular signals for the generation of saccadic eye movements. *Journal of Neurophysiology*, 87(2), 845–858.


position signals in the primate central thalamus. *Cerebral Cortex*, 17(7), 1504–1515.


---

Appendix: Neuro-computational model

The neurons in each map follow different ordinary differential equations (ODEs). In our extension of the original 1-D model to the two-dimensional one, we use the same ODEs as stated in Ziesche and Hamker (2011), but with some simplifications and extensions. For all ODEs that compute firing rates $r$, we set negative values to zero.

- Firing rates of neurons in map $X_r$ representing the stimulus position in an eye-centered reference frame are given by

$$
\tau \frac{d}{dt} r_{Xr} = r_{Xr, in} \left( 1 + \left[ A_{Xr} - r_{Xr} \right]^+ \right) \times \sum_{FB} w_{X_{FB} \rightarrow Xr, X_{FB}} - r_{Xr},
$$

with $[A_{Xr} - r_{Xr}]^+$ a scalar saturation factor controlling the feedback from $X_{FB}$, and $r_{Xr, in}$ the sensory bottom-up input created by a given stimulus:

$$
r_{Xr, in} = S_{Xr} K_{Xr} \exp \left( -\frac{\|p_{Xr} - c_{Xr}\|^2}{2(\sigma_{Xr})^2} \right).
$$

Here, $S_{Xr}$ is the short-term synaptic depression simulated as in Hamker (2005), modeling the decaying response strength over time while the stimulus is presented; $K_{Xr}$ defines the strength of the stimulus; and $\|p_{Xr} - c_{Xr}\|^2$ is the distance between stimulus position $p_{Xr}$ and receptive-field center $c_{Xr}$ for each neuron of the map. As we now have a two-dimensional visual scene, the stimulus position and the receptive-field center are two-dimensional.
Firing rates of neurons in maps \( X_{ePC} \) and \( X_{eCD} \) representing eye position in head and retinotopic eye displacement, respectively, are given by

\[
\tau \frac{d}{dt} r_{X_{ePC}} = r_{X_{ePCin}} - r_{X_{ePC}},
\]

\[
\tau \frac{d}{dt} r_{X_{eCD}} = r_{X_{eCDin}} - r_{X_{eCD}},
\]

where \( r_{X_{ePCin}} \) and \( r_{X_{eCDin}} \) are Gaussian input signals modeling the proprioceptive eye-position signal and the corollary-discharge signal, respectively:

\[
r_{X_{ePCin}} = K_{X_{ePC}} \exp \frac{-\|p_{X_{ePC}} - c_{X_{ePC}}\|^2}{2(\sigma_{X_{ePC}})^2},
\]

\[
r_{X_{eCDin}} = TC^{X_{eCD}}(t) K^{X_{eCD}} \exp \frac{-\|p_{X_{eCD}} - c_{X_{eCD}}\|^2}{2(\sigma_{X_{eCD}})^2}.\]

\( K_{X_{ePC}} \) and \( K^{X_{eCD}} \) are the strengths of the corresponding signal, \( \|p_{X_{ePC}} - c_{X_{ePC}}\|^2 \) is the distance between eye position \( p_{X_{ePC}} \) and center of eye-position tuning \( c_{X_{ePC}} \) for each neuron of map \( X_{ePC} \), and likewise \( \|p_{X_{eCD}} - c_{X_{eCD}}\|^2 \) is the distance between eye displacement \( p_{X_{eCD}} \) and center of eye-displacement tuning \( c_{X_{eCD}} \) for each neuron of map \( X_{eCD} \). Again, the positions are now two-dimensional. \( TC^{X_{eCD}}(t) \) models the time course of the phasic corollary-discharge signal, namely rise and decay around saccade onset:

\[
TC^{X_{eCD}}(t) = \begin{cases} 
\exp \frac{-|t - t^{CD}|^2}{2(\sigma_{CD,in})^2}, & \text{if } t \leq t^{CD} \\
\exp \frac{-|t - t^{CD}|^2}{2(\sigma_{CD,dec})^2}, & \text{if } t > t^{CD}
\end{cases},
\]

with \( t^{CD} \) the time where the CD signal reaches its maximum. In our model, this maximum is at 10 ms after saccade onset, consistent with data of Ferraina, Paré, and Wurtz (2002).

Firing rates of neurons in map \( X_{eFF} \) representing the eye displacement in a head-centered reference frame are given by

\[
\tau \frac{d}{dt} r_{X_{eFF}} = \sum_{FF} W_{X_{eFF} \rightarrow X_{eFF}} r_{X_{eFF}} \\
\sum_{FF} W_{X_{ePC} \rightarrow X_{eFF}} r_{X_{ePC}} \\
-r_{X_{eFFin}} \sum_{inh} W_{X_{eFF} \rightarrow X_{eFFin}} r_{X_{eFFin}}.
\]

In contrast to the original ODE, we simplified the firing rate for \( X_{eFF} \) by removing the saturation term as well as the gain-modulation term. Thus, this map is now a classical basis-function map instead of a gain-modulation map to combine PC and CD signals.

Firing rates of neurons in map \( X_{hPC} \) representing the joint representation of stimulus position and eye position are given by

\[
\tau \frac{d}{dt} r_{X_{hPC}} = \sum_{FF} W_{X_{hPC} \rightarrow X_{hPC}} r_{X_{hPC}} \\
(1 + [A_{X_{hPC}} - r_{X_{hPC}}])^{-1} \\
\sum_{FF} W_{X_{hPC} \rightarrow X_{hPC}} r_{X_{hPC}} \\
+r_{X_{hPCin}} \sum_{inh} W_{X_{hPC} \rightarrow X_{hPCin}} r_{X_{hPCin}} \\
-r_{X_{hPCin}}.
\]

with \( [A_{X_{hPC}} - r_{X_{hPC}}]^{-1} \) a scalar saturation factor controlling the feed-forward input from \( X_{eFF} \), and \( (r_{X_{hPC}} + D_{X_{hPC}}) \) a scalar regulating the global inhibition.

\[
\tau \frac{d}{dt} r_{X_{hCD}} = \sum_{FF} W_{X_{hCD} \rightarrow X_{hCD}} r_{X_{hCD}} \\
(1 + [A_{X_{hCD}} - r_{X_{hCD}}])^{-1} \\
\sum_{FF} W_{X_{hCD} \rightarrow X_{hCD}} r_{X_{hCD}} \\
+r_{X_{hCDin}} \sum_{inh} W_{X_{hCD} \rightarrow X_{hCDin}} r_{X_{hCDin}} \\
-r_{X_{hCDin}}.
\]

with \( [A_{X_{hCD}} - r_{X_{hCD}}]^{-1} \) a scalar saturation factor controlling the feed-forward input from \( X_{eFF} \), and \( (r_{X_{hCD}} + D_{X_{hCD}}) \) a scalar regulating the global inhibition.
• Firing rates of neurons in map $X_h$ are given by

\[ \tau \frac{d}{dt} X_h = S^{X_h} I^{X_h} + \sum w_{\text{exc}}^{X_h} X_h \]

\[ - \left( r^{X_h} + D^{X_h} \right) w_{\text{inh}}^{X_h} \sum X_h - r^{X_h}, \]

with $(r^{X_h} + D^{X_h})$ a scalar regulating the global inhibition, and $\tau^{X_h}$ the input consisting of the feedforward input from both LIP maps and a newly introduced, attentional top-down signal $r^{X_h, \text{in}}$:

\[ I^{X_h} = \sum w_{X_b}^{X_h} \cdots X_h + X_{h}^{X_h} \cdots X_h \]

\[ r^{X_h, \text{in}} = K^{X_h} \exp \left( - \frac{\| p^{X_h} - c^{X_h} \|^2}{2(\sigma^{X_h})^2} \right), \]

where $K^{X_h}$ denotes the strength of the attention and $\| p^{X_h} - c^{X_h} \|^2$ the distance between attention position $p^{X_h}$ and center of attention-position tuning $c^{X_h}$ for each neuron of map $X_h$, each with two dimensions. $S^{X_h}$ is the synaptic suppression simulated as in the study by Hamker (2005):

\[ S^{X_h} = 1 - a_s^{X_h} s \]

\[ \tau_s^{X_h} \frac{d}{dt} s = I^{X_h} - s. \]

As the dimension of each map has to be doubled compared to the definition of the maps in Ziesche and Hamker (2011), we reduced the number of neurons in each dimension to compensate for the higher computational effort. More precisely, we have 21 neurons for the horizontal and 16 neurons for the vertical, covering a rectangular visual field of $40^\circ \times 30^\circ$. Thus, $X_r$, $X_{EP}$, $X_{CE}$, and $X_h$ are now two-dimensional, containing $21 \times 16$ neurons, and $X_{FEF}$, $X_{PC}$, and $X_{BC}$ are four-dimensional, with $21 \times 21 \times 21 \times 16$ neurons.

The connections between the different maps are defined through the different weights used in the ODEs. The connection weights follow Gaussian functions dependent on the distance between the position of the neuron in the presynaptic and postsynaptic maps:

\[ w_{\text{pre} \rightarrow \text{post}}^{X_h} = K_{\text{pre} \rightarrow \text{post}}^{X_h} \exp \left( - \frac{\text{distance}}{(\sigma_{\text{pre} \rightarrow \text{post}})^2} \right), \]

with $w_{\text{pre} \rightarrow \text{post}}$ set to 0 if the value is lower than 0.001. The measurement of the distance differs among the connections due to the different dimensions of the maps and the different ways of connecting maps. There are three types of connections: horizontal, vertical, and diagonal. The horizontal and vertical connections are used to connect two-dimensional maps with four-dimensional maps where two of the four dimensions are disregarded (either the first two or the last two). For example, we connect $X_r$ with $X_{BC}$ horizontally independent of the last two dimensions of $X_{BC}$. That means that the distance used for the Gaussian function depends only on the position of the neuron in $X_r$ and the first two position parameters of the neuron in $X_{BC}$. More precisely, suppose we want to connect neuron $(i, j)$ of $X_r$ with neuron $(k, l, m, n)$ of $X_{BC}$. The distance between these neurons is then calculated by

\[ \text{distance} = ||i - k||^2 + ||j - l||^2. \]

The weight between neuron $(i, j)$ and neuron $(k, l, m, n)$ is then

\[ w_{X_r \rightarrow X_{BC}}^{(i,j),(k,l,m,n)} = K_{X_r \rightarrow X_{BC}}^{(i,j),(k,l,m,n)} \exp \left( - \frac{||i - k||^2 + ||j - l||^2}{(\sigma_{X_r \rightarrow X_{BC}})^2} \right). \]

This connection pattern is also used to connect $X_r$ with $X_{PC}$, $X_{EP}$ with $X_{FEF}$, and $X_{BC}$ with $X_{PC}$.

Similarly, a vertical connection between a two-dimensional map and a four-dimensional map makes use of only the last two position parameters of the neurons in the four-dimensional map to calculate the distance. Thus, the distance between neuron $(i, j)$ of a two-dimensional map and neuron $(k, l, m, n)$ of a four-dimensional map is

\[ \text{distance} = ||i - m||^2 + ||j - n||^2. \]

Such a vertical connection is used to connect $X_{EP}$ with $X_{FEF}$ and with $X_{PC}$.

These definitions of the connection patterns allow the interpretation of the four dimensions of a map as follows: The first two dimensions represent the horizontal and the vertical information of a horizontally connected input, and the last two dimensions represent the horizontal and the vertical information of a vertically connected input.

For a diagonal connection pattern we use both horizontal and vertical information of the four-dimensional map to connect it with a two-dimensional map. We use such a connection pattern to connect $X_h$ with the LIP maps $X_{BC}$ and $X_{CD}$ and vice versa. The distance between neuron $(i, j)$ of $X_h$ and neuron $(k, l, m, n)$ of an LIP map is defined as

\[ \text{distance} = ||i - k - m||^2 + ||j - l - n||^2. \]

For the remaining connection pattern between $X_{FEF}$ and $X_{BC}$ that connects two four-dimensional maps, we read out the presynaptic map diagonally and connect this vertically with the postsynaptic map—that is, we use all four dimensions of $X_{FEF}$ and only the last two dimensions of $X_{BC}$ to calculate the distance.
between the neurons. More precisely, if we want to connect neuron \((i, j, k, l)\) of \(X_{\text{FEF}}\) with neuron \((m, n, o, p)\) of \(X_{\text{BCD}}\), the distance between these neurons is

\[
\text{distance} = \|i + k - o\|^2 + \|j + l - p\|^2.
\]

The lateral excitatory connections used in map \(X_{\text{BCD}}\) and \(Xh\) are defined as Gaussian functions dependent on the distance between the positions of the neurons in the maps considering all dimensions of the map. Thus, the calculation of the weights follows

\[
w_{\text{exc}}^{X_i} = K_{\text{exc}}^{X_i} \exp\left(-\frac{\text{distance}_{\text{exc}}}{\sigma^2}\right),
\]

with

\[
\text{distance}_{\text{exc}} = \begin{cases} \|i - k\|^2 + \|j - l\|^2, & \text{for } (i, j, k, l) \in Xh \\ \|i - m\|^2 + \|j - n\|^2 + \|k - o\|^2 + \|l - p\|^2, & \text{for } (i, j, k, l), (m, n, o, p) \in X_{\text{BCD}} \end{cases}
\]

To reduce the computational effort, the lateral inhibitory connections with fixed weights \(w_{\text{inh}}^{X_i}\) are not created as explicit connections but rather are calculated with the help of the mean over all firing rates multiplied by the total number of neurons:

\[
w_{\text{inh}}^{X_i} \sum r_i = w_{\text{inh}}^{X_i} \text{mean}(r_i) \times \text{number of neurons}.
\]

Table 1 lists all parameters whose values have changed in comparison to Ziesche and Hamker (2011). Mainly, the adaptation of the values is due to the three major changes in the model: reduction of the number of neurons, simplification of \(X_{\text{FEF}}\), and addition of feedback from \(Xh\) to \(X_{\text{BCD}}\). The new parameter values for this feedback connection are listed in the last two rows and are equal to those for the feedback connection from \(Xh\) to \(X_{\text{BCD}}\).

### Supplementary material

**Supplementary Movie S1.** Simulation results of the two predictive-remapping tasks over time. The activity of both maps of lateral intraparietal cortex (LIP) split into two two-dimensional planes is plotted representing horizontal and vertical information. Additionally, the setup with the current eye position (red cross) and stimulus (green star), as well as the activities of both LIP maps projected to visual space, are plotted. The yellow dashed line shows the diagonal which has the largest activation. The symbols and definitions are identical to Figure 6. The time (in milliseconds) is aligned to saccade onset. (Left) In the fixation task, a stimulus is presented in the receptive field (RF), the eyes fixate at the fixation point (FP), and no saccade is planned. Due to the multiplicative interaction of eye-position signal and retinal signal representing the stimulus, there is a single activation blob in the LIP proprioceptive (PC) eye position map. Projected to visual space, this results in activity at RF (red blob). In the LIP coronary discharge (CD) map, there is an activity line indicating the position of the stimulus, which also leads to activity at RF (covered by red blob). (Right) In the saccade task, the stimulus is presented in the future receptive field (FRF) and a saccade is going to be executed from FP to the saccade target (ST). In LIP PC, there is an activity blob at the crossing of the current eye position and the stimulus.
position, similar to the fixation task. Thus, we see activity at FRF triggered by LIP PC (red blob). Shortly before saccade onset, the CD signal rises, which produces an additional peak of activity along the activity line from the stimulus signal at ST in LIP CD. Additionally, there is a second activity blob resulting from the interaction of the CD signal with the feedback signal from $X_h$ along the yellow dashed diagonal. This leads to a second activity blob at RF (blue blob). Both activity blobs are encoded in a retinotopic reference frame, and thus they move according to the eye movement. After the saccade, the CD signal decays and with it the activity in LIP CD (and hence the blue blob, now at FRF). Additionally, the activity blob by LIP PC moves back to FRF as the PC signal updates to the correct postsaccadic eye position. There it decays as the retinal signal decays after turning off the stimulus.

Supplementary Movie S2. Simulation results of spatial updating of attention with cued attention over time. The activities of both maps of lateral intraparietal cortex (LIP; $X_{hPC}$, $X_{hCD}$) are plotted as well as the setup and the triggered attention pointers. The yellow dashed line shows the diagonal which has the largest activation. The symbols and definitions are identical to Figure 9. The time (in milliseconds) is aligned to saccade onset. At the beginning, the activity in the LIP proprioceptive (PC) eye position map resulting from the interaction of the PC signal and the retinal signal triggers an attention pointer at the attention position (AP; red blob). Shortly before saccade onset, a second attention pointer is triggered at the remapped attention position (RAP) by the LIP corollary discharge (CD) map through the interaction of the rising CD signal and the attention signal (blue blob). Both attention pointers are shifted with the eye movement, as they are retinotopic. After the saccade, the CD signal decays and with it the activity in LIP CD, as well as the second attention pointer. Furthermore, the PC signal updates to the correct postsaccadic eye position and thus the attention pointer triggered by LIP PC updates to the correct position (AP).

Supplementary Movie S3. Simulation results of spatial updating of attention with top-down attention over time. The activities of both maps of lateral intraparietal cortex (LIP; $X_{hPC}$, $X_{hCD}$) are plotted as well as the setup and the triggered attention pointers. Additionally, the diagonal on which the attention signal is fed into the LIP maps is plotted with a yellow dashed line. The symbols and definitions are identical to Figure 12. The time (in milliseconds) is aligned to saccade onset. At the beginning, the activity in the LIP proprioceptive (PC) eye position map resulting from the interaction of the PC signal and the attention signal triggers an attention pointer at the attention position (AP; red blob). Shortly before saccade onset, a second attention pointer is triggered at the remapped attention position (RAP) by the LIP corollary discharge (CD) map through the interaction of the rising CD signal and the attention signal (blue blob). Both attention pointers are shifted with the eye movement, as they are retinotopic. After the saccade, the CD signal decays and with it the activity in LIP CD, as well as the second attention pointer. Furthermore, the PC signal updates to the correct postsaccadic eye position and thus the attention pointer triggered by LIP PC updates to the correct position (AP).

Amended August 20, 2019: The movie files were replaced with new files to improve reader experience.