Perisaccadic mislocalization as optimal percept

Tobias Teichert
Philips University, Department of Physics, NeuroPhysics Group, Marburg, Germany

Steffen Klingenhoefer
Ludwig-Maximilians-Universität München, Department of Biology, Martinsried, Germany, & Philips University, Department of Physics, NeuroPhysics Group, Marburg, Germany

Thomas Wachtler
Philips University, Department of Physics, NeuroPhysics Group, Marburg, Germany

Frank Bremmer
Philips University, Department of Physics, NeuroPhysics Group, Marburg, Germany

The spatially uniform mislocalization of stimuli flashed around the onset of fast eye-movements (perisaccadic shift) has previously been explained by an inaccurate internal representation of current eye position. However, this hypothesis does not account for the observation that continuously presented stimuli are correctly localized during saccades. Here we show that the two findings are not mutually exclusive. The novelty of our approach lies in our interpretation of the extraretinal signal which, in contrast to other models, is not considered an (erroneous) estimate of current eye-position. Based on the reafference principle, our model assumes that the extraretinal signal is optimal in that it accurately predicts the neural representation of the retinal position of a continuously present stimulus. Perisaccadic shift arises as a consequence of maintaining stable perisaccadic position estimates for continuously present stimuli under the physiologically plausible assumption of temporal low-pass filtering in the afferent visual pathway. Consequently, our model reconciles the reafference principle with the finding of perisaccadic shift.

Keywords: perisaccadic shift, reafference principle, visual stability, damped eye-position model, visual persistence


Introduction

Eye movements challenge visual perception (Bremmer & Krekelberg, 2003). Changes in the direction of gaze dramatically change the afferent flow to early visual areas. Nevertheless, we perceive the visual environment as being stable during eye movements. This is generally taken to be the result of an internally generated signal, the so-called extra-retinal signal (exR). The exR is assumed to account for changes in afferent flow which are caused by the movement of the eyes (reafference principle, e.g., von Holst & Mittelstaedt, 1950).

In contrast to this subjective observation of perceptual stability during eye movements in general and saccades in particular, several studies have reported systematic mislocalization of stimuli flashed around the onset of saccades (perisaccadic shift: Cai, Pouget, Schlag-Rey, & Schlag, 1997; Honda, 1989; Lappe, Awater, & Krekelberg, 2000). In the dark, the spatial perception of perisaccadically flashed stimuli shows a biphasic error pattern. Stimuli flashed prior to saccade onset or during the first half of the saccade are mislocalized in the direction of the saccade. Stimuli flashed during the second half of the saccade or shortly after saccade offset are mislocalized in the direction opposite to the saccade.

Perisaccadic shift is generally taken as evidence that the reafference principle is violated perisaccadically. Consequently, it is assumed that alternative mechanisms, such as saccadic suppression, visual masking or the remapping of receptive fields, mediate perceptual stability during saccades (reviewed e.g., by Wurtz, 2008). While these mechanisms certainly play an important role, we show here that the original argument, namely that perisaccadic shift implies a breakdown of the reafference principle, is not compelling. We present a model that, based on the reafference principle, explains both, perisaccadic stability for continuously present stimuli and perisaccadic mislocalization for flashed stimuli.
Review of current models of perisaccadic shift

Our model is the logical next step in a series of models which have so far been put forward to explain perisaccadic shift. We will briefly outline these models to motivate the current work. The first and still widely accepted model of perisaccadic shift explains the mislocalization of flashed stimuli by an erroneous internal estimate of the current eye-position (Matin, 1976; Ross, Morrone, Goldberg, & Burr, 2001; Schlag & Schlag-Rey, 2002). This idea is quantified by Equation 1:

\[ E_f(t) = exR(t) - h(t). \] (1)

Here, \( E_f(t) \) corresponds to the observed localization error of a stimulus flashed at time \( t \) after saccade onset. Further, \( h(t) \) corresponds to the direction of the stimulus at the time of the flash and \( exR(t) \) denotes the extraretinal signal at the time of the flash. In this context it is assumed that in order to guarantee perceptual stability \( exR \) needs to be equal to eye-position: \( exR(t) = h(t) \). Hence, it is assumed that \( exR(t) \) is the visual system’s estimated direction of gaze. The observed pattern of mislocalization can be predicted under the assumption that \( exR(t) \) is a damped version of the actual eye-position (see Honda, 1989, Figure 4, or Pola, 2004, Figure 1). Hence, this model is often referred to as the damped eye position model.

While this elegant approach provides a mechanism of perisaccadic shift, it leaves several issues open. First and foremost, it does not provide a compelling reason as to why the visual system would choose \( exR \) in a way that solves Equation 1. One tentative explanation is that a damped version of eye-position is the visual system’s best guess of actual eye-position under the assumption that it is not able to produce signals that change in time as rapidly as eye position. However, it seems unlikely that the same system which controls eye-position by sending out highly variant neuronal signals to a sluggish eye-plant should not be able to produce signals that describe these very same changes in eye position.

Second, it was pointed out that the damped eye position model might be overly simplistic as it relies on physiologically implausible assumptions (Pola, 2004). In particular, the damped eye-position model assumes that the neuronal representation of the flashed stimulus may be modeled as having virtually no extension in time, i.e., as a Dirac delta function. However, while this is a reasonable simplification for the retinal projection of the stimulus, it is certainly not for its neuronal representation (for a similar argument see Schlag & Schlag-Rey, 1995). The temporal low-pass properties of neurons in the afferent visual pathway, also referred to as persistence, complicate Equation 1. A mathematical formulation of this view is given by Equation 2 (for more details see Pola, 2004):

\[ E_f(t) = \left[ R_{f,i}(\tau) - exR(t) \right] \xi_0(t - \tau) d\tau. \] (2)

In this model, the localization error arises as the weighted average of the instantaneous mislocalization error over a period of time \( \tau \). The duration and weight of this epoch is described by \( \xi_0 \) which represents the persistence of sensory preprocessing in afferent neurons. \( \xi_0 \) is normalized such that \( \int \xi_0(\tau) d\tau = 1 \). \( R_{f,i}(\tau) \) corresponds to the time-resolved retinal signal of a stimulus flashed at time \( t \) after saccade onset (see Equation 5 for details). Pola has argued convincingly that \( R_{f,i}(\tau) \) is constant and corresponds to the inverse of direction of gaze at the time of the flash: \( R_{f,i}(\tau) = -h(t) \). For example, a flash at 0 deg in craniocentric coordinates presented while gaze is directed 5 degrees towards the right, will drive neurons with receptive fields 5 deg to the left of the fovea. We can rewrite \( R_{f,i}(\tau) \) as \( R_f(t) \) and remove it from the integral. For more details on the use of functions, their arguments and subscripts, refer to Appendix A. Thus, we can reformulate Equation 2:

\[ E_f(t) = R_f(t) - exR * \xi_0(t). \] (2')

Here * corresponds to the convolution operator (see Appendix A for a definition). We see that Equation 1 is a special case of Equation 2 by setting \( \xi_0 \) equal to a Dirac delta function (see Appendix A for details). For a wide range of choices of \( \xi_0 \) we can find an entire family of functions \( exR \) which solve the equation. In other words, Equation 2 describes an infinite number of models of perisaccadic shift, including the damped eye-position model. In order to come up with a unique solution, Equation 2 needs to be restricted. One way to do so is to measure the persistence of the afferent neurons in question and hence determine \( \xi_0 \) explicitly. However, it is not obvious which neurons at what level of the visual hierarchy should be considered. Alternatively, \( \xi_0 \) may be estimated from psychophysical data as has been done by Pola (2004). Following this approach, it is possible to identify a family of models of perisaccadic shift all of which use physiologically plausible temporal dynamics.

A final criticism of the damped eye position model is that it predicts perisaccadic mislocalization also for continuously present stimuli (Pola, 2004). Hence, it fails to account for the finding that flashed stimuli are mislocalized relative to continuously present stimuli (Cai et al., 1997). It is not immediately obvious whether the family of models described by Equation 2 predict mislocalization for continuously present stimuli or not. Based on the assumption of linearity (see below) we show...
that in general, these models do predict mislocalization for continuously present stimuli. However, we show that some of the models of this kind come very close to predicting visual stability for continuously present stimuli.

**Reafference principle: A model of perisaccadic shift?**

In the current manuscript we take a novel approach which at the same time resolves the issues of current models of perisaccadic shift which were raised above and imposes restrictions on the family of models described by Equation 2'. The novelty and power of our approach lies in the idea of deriving restrictions for \( \xi_R \) from a setting that is independent from the one used to measure perisaccadic shift. In particular, we start with the following assumption: visual processing is based on and optimized for continuously present stimuli. Thus, assuming that the reafference principle is valid during saccades, we choose \( \xi_R \) such that continuously present stimuli are not mislocalized. This restriction is described by Equation 3:

\[
E_c(\tau) = 0.
\]

In Equation 3, \( E_c \) corresponds to the localization error of a continuously present stimulus at time \( \tau \) after saccade onset. Whenever Equation 3 holds, continuously present stimuli are not mislocalized perisaccadically and the visual world appears stable during saccades. It is essential to point out the difference to previous models which choose \( \xi_R \) in order to predict the previously measured perisaccadic shift. Such an approach will usually lead to a violation of the reafference principle, i.e., these models will not be able to explain perisaccadic stability for continuously present stimuli. Here we invert the problem in order to determine the properties of the sensory preprocessing, i.e., \( \xi_0 \), which are necessary to produce perisaccadic shift. In contrast to other models, our approach does **not** necessarily guarantee a solution.

In the current paper, we show that it is possible to find preprocessing parameters \( \xi_0 \) which reconcile the reafference principle with perisaccadic shift, i.e., simultaneously solve Equations 2' and 3. Hence, perceptual stability for continuously present stimuli, i.e., Equation 3, on the one hand and perisaccadic shift, i.e., Equation 2', on the other hand are not mutually exclusive. Further, we show that the preprocessing parameters \( \xi_0 \) which simultaneously solve Equations 2' and 3 have physiologically plausible low-pass filter properties. In contrast, we show that for a wide range of un-physiological choices of \( \xi_0 \), our model does **not** predict perisaccadic shift of realistic amplitude. In particular, we can rule out the case that \( \xi_0 \) equals a Dirac delta function which underlies the damped eye position model. In summary, we argue that the visual system chooses \( \xi_R \) in order to guarantee perisaccadic stability, and in doing so causes perisaccadic shift under the assumption of physiologically plausible temporal low-pass filtering in the afferent visual pathway. In addition, our analysis reduces the dimensionality of the family of potential models of perisaccadic shift described by Equation 2'.

**Model**

We simulate neuronal activity in a retinocentric visual area around the time of a saccade. From the resulting population activity we derive a scalar estimate of retinocentric position (the so-called retinal signal, \( R \)). \( R \) is converted to a craniocentric position estimate \( W \) by subtracting the extraretinal signal \( \xi_R \).

Figure 1. The model simulates neuronal activity in a one-dimensional retino-centric layer around the time of saccade onset while presenting a stimulus at 0 deg cranio-centric position (see Methods for details). Neuronal activity \( N(x, \tau) \) (panel C) is simulated as the convolution of the retinal projection of the visual stimulus \( S(x, \tau) \) (panel A) with the spatio-temporal receptive field \( \xi \) (panel B). The time-resolved retinal signal \( R(\tau) \) is calculated as the center of gravity of neuronal activity at any given time (green line in panel D). Due to the properties of the temporal receptive field the retinal signal only begins to reflect the eye-movement well after saccade onset.
Preprocessing in early visual areas was simulated in a layer of 201 rate-coding neurons in a one-dimensional retinocentric area spanning roughly 10° of visual space. Neuronal activity was simulated around the onset of a simulated 6° leftward saccade (from \( x = +3° \) to \( x = -3° \), craniocentric coordinates) while presenting stimuli of variable duration at \( x = 0° \) in craniocentric coordinates. The saccade was simulated with a Gaussian velocity profile over time with a standard deviation of 8 ms. This setting corresponded to a saccade duration of 40 ms (using a speed criterion of 15 deg/sec) and a peak velocity of 300 deg/sec (Carpenter, 1988).

The neurons were modeled as linear filters, yielding a spatio-temporal neural activity profile \( N(x, \tau) \) that corresponded to the linear convolution of the retinal stimulus \( S \) with a spatio-temporal receptive field \( \xi \). The activity of the neurons is described by Equation 4 (see Figure 1D):

\[
N(x, \tau) = S \ast \xi(x, \tau).
\] (4)

Here \( x \) corresponds to the receptive field position in retinal coordinates and \( \tau \) indicates time relative to saccade onset. The input \( S \) corresponds to the retinal projection of the visual scene coded as zeros and ones, depending on the presence of a stimulus. Note that the temporal kernel \( \xi_0 \) in Equation 2 is identical to the temporal aspect of the spatio-temporal kernel \( \xi \) in Equation 4, i.e., \( \xi_0(\tau) = \xi(0, \tau) \). The spatial receptive field was modeled as a Gaussian with a standard deviation of 0.15°, normalized to a maximum amplitude of 1 (see Figure 1B, upper panel). The temporal impulse response was described as a Gamma distribution with various scale and shape parameters (see Figure 1B, lower panel). A gamma distribution can be described as the convolution of \( n \) exponential distributions. The shape parameter of the gamma distribution corresponds to the number \( n \); the scale parameter corresponds to the time constant \( \lambda \) of the underlying exponential distributions. Hence, the entire temporal receptive field is determined by the shape parameter \( n \) and the scale parameter \( \lambda \). Unless mentioned otherwise, the \( n \) was set to 5. The resulting kernel is a low-pass filter of order \( n \) and cutoff wavelength \( 2\pi\lambda \). The extent of the kernel in time can be quantified by its standard deviation which is given as \( \sqrt{n\lambda^2} \). Kernels with large standard deviation give rise to a system with slow temporal dynamics. Vice versa, small standard deviation is an indicator of fast temporal dynamics. In addition, a delay of 15 ms was added to the kernel to account for conduction delays caused by preprocessing in the afferent visual pathway.

Whenever the maximum firing rate was above zero, an object was deemed to be present in the scene and its retinal signal \( R(\tau) \) was derived as the center of gravity of the neuronal activity at time \( \tau \) in question (see Figure 1D):

\[
R(\tau) = \int xN(x, \tau)dx.
\] (5)

This instantaneous retinal position estimate was converted to instantaneous craniocentric coordinates \( W(\tau) \), by subtracting \( exR(\tau) \):

\[
W(\tau) = R(\tau) - exR(\tau).
\] (6)

Note, that \( R(\tau) \) is only defined when an object is present in the scene. Hence, the same restrictions apply to \( W(\tau) \). Finally, the time-resolved craniocentric position estimate was converted into a global position estimate \( \bar{W} \) as the weighted mean of the instantaneous craniocentric position estimate:

\[
\bar{W} = \int W(\tau)g(\tau)d\tau.
\] (7)

Here the weights \( g(\tau) \) were calculated as the maximum of the neuronal activity at time \( \tau \). In addition, the weights were normalized to a sum of one: \( g(\tau) = \max[N(x, \tau)] / \max[N(x, \tau)]d\tau \). For stimuli flashed at time \( t \) after saccade onset, the weights \( g(\tau) \) are defined by the temporal impulse response function: \( g(\tau) = \xi_0(\tau - t) \).

The actual position of the stimulus was always fixed at 0°. Hence, \( \bar{W} \) also corresponds to the localization error \( \bar{E} \). By substituting \( \bar{W} \) for \( \bar{E} \), \( exR(\tau) = R_c(\tau) \) for \( W(\tau) \) and \( g(\tau) \) for \( \xi_0(\tau - t) \) we see that for flashed stimuli, Equation 7 corresponds to Equation 2.

**Optimization of the extraretinal signal**

In the context of our model, the reafference principle holds if continuously present stimuli are not mislocalized during saccades as described in Equation 3. Thus, for stationary stimuli presented continuously at 0° in craniocentric coordinates, the extraretinal signal has to be identical to the retinal signal:

\[
exR(\tau) = R_c(\tau).
\] (8)

This is achieved by a simple calibration procedure. The model is run while presenting a stimulus continuously at 0° in craniocentric coordinates (see Figure 1 or Supplementary movie). \( R_c \) is determined and for all subsequent runs \( exR \) is set to \( R_c \), regardless of whether presenting flashed or continuously present stimuli. In a
real system that performs saccades of variable direction and amplitude, such a simple mapping is not feasible. Instead, exR actually has to be trained with a number of different saccade vectors.

Results

Veridical localization of continuously present stimuli

Figures 2A through 2D show the response of the model for a continuously present stimulus for four different choices of $\xi$ with increasingly slower temporal dynamics, i.e., longer time constants $\lambda$ (see Methods). Prior to saccade onset the eye is fixating 3° to the right of the stimulus which consequently drives neurons with receptive fields 3° to the left of the fovea. At time 0 the eye starts to move to the left and finally reaches a position 3° to the left of the stimulus which consequently drives neurons with receptive fields 3° to the right of the fovea. Between 25 and 50 ms after saccade onset, the neuronal representation of the stimulus begins to reflect this change in eye-position as the center of gravity of the neuronal activity, i.e., $R_c$, moves from $-3^\circ$ to $+3^\circ$ (green lines in Figure 2, see also Supplementary movie). Note that for slower temporal dynamics, described by increases in the scale parameter $\lambda$ (panels A through D), $R_c$ begins to move later and at lower speeds. Accordingly, each of the four models will use different extraretinal signals which, according to Equation 8, are set to $R_c$.

As a consequence of this calibration, all models will meet the requirements of Equation 3, i.e., the cranio-centric representation of the stimulus $W_c(\tau)$ (blue line in Figure 2) accurately remains at 0° during the entire perisaccadic time period. Hence, continuously present stimuli are not mislocalized, regardless of the receptive field properties defined by $\xi$. In the following we will explore the predictions of the models for flashed stimuli.

Shift-like mislocalization of flashed stimuli

Based on $\xi$ and exR we can derive the mislocalization of flashed stimuli by solving Equation 2. Alternatively, we can run the model multiple times and simulate flashes at various times relative to saccade onset. Figures 3A–3D show the responses of the model from Figure 2B to stimuli flashed at four different times relative to saccade onset. Stimuli flashed before saccade onset or in the first half of the saccade are mislocalized in the direction of the saccade (Figures 3A and 3B). Stimuli flashed in the second half of the saccade and briefly after saccade completion are mislocalized in direction opposite to the saccade (Figure 3D). Figure 3E depicts a summary of mislocalization as a function of flash onset relative to saccade onset. It verifies that the model indeed predicts the biphasic pattern of mislocalization that is typically referred to as perisaccadic shift. It is important to note that the mislocalization starts well before saccade onset and that the maximal amplitude is observed for stimuli flashed around saccade onset. At the same time it is important to note again that the model does not predict mislocalization for continuously present stimuli. Figure 4 shows the...
predicted perisaccadic shift for the four choices of $\lambda$ which were used in Figures 2A through 2D. Note that it is only in the conditions with slower temporal dynamics, i.e., conditions with longer time constants, that the predicted mislocalization error reaches an amplitude that is comparable to experimental findings (see below for a quantitative analysis).

**Read-out of population activity**

The retinal signal can be derived from the population activity $N(x, \tau)$ in a number of different ways. In the simulations above, we assumed that $R$ corresponds to the center of gravity (CG) of the neuronal activity (see Methods). Alternatively, $R$ may, for example, be defined as the retinal location with the strongest activity (MAX) or be derived by a maximum likelihood method (ML). Several recent studies have provided neuronal mechanisms for ML calculations (Deneve, Latham, & Pouget, 2001) as well as evidence in favor of the brain using ML-like methods (e.g., Knill & Pouget, 2004). Hence, it is important to test how our model is affected by using the MAX or the ML method instead of the CG method to estimate the retinal signal.

Figure 5 shows the results of the MAX method (results for the ML method are virtually identical). There are two remarkable points. First, the retinal signals derived with

![Figure 3](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933477/)
the MAX method differ substantially from the ones calculated with the CG method. This difference is most obvious for the two slowest conditions. For the CG method, the retinal signal changes smoothly as a function of time. In contrast, the retinal signal derived with the MAX method is essentially a step function. From one time point to the next, the retinal signal jumps from the pre- to the post-saccadic value. Second, despite these differences, the main effect, i.e., the biphasic mislocalization of the flashed stimuli can still be observed. Finally, we find the same dependence of the amplitude of the perisaccadic shift on the time constant as previously observed for the CG method.

Slow dynamics are necessary for perisaccadic shift

In the previous simulations we held the shape of the temporal impulse response constant. In the following simulations we systematically vary the shape of this temporal kernel. As it is impossible to explore all possible shapes we settled for the subspace of kernels described by the two-dimensional Gamma distribution family (see Methods). Our simulations show that the amplitude of the predicted mislocalization increases both with the shape parameter \( n \), as well as the scale parameter \( \lambda \) (see Figures 6A and 6D for results of the CG and MAX method, respectively). Closer investigation revealed that the mislocalization amplitude could be described reasonably well as a function of the standard deviation of the kernel which is given by \( \sqrt{n \lambda^2} \) (see inset Figures 6A and 6D).

To quantify the fit of the predictions to actual data we compared it to the mislocalization error reported by Honda (1989). As we did not have access to the actual data we emulated these previous results by estimating the parameters of the damped eye-position model which best fit the mislocalization amplitude reported in his paper. For saccades of 8° amplitude mislocalization ranging from \(-2.4^\circ\) to \(+2.4^\circ\) was reported. The range of this mislocalization corresponds to 0.6 times the saccade amplitude.

A number of models accurately predicted the emulated data reasonably well (see Figures 6B and 6E). The family of models that fits Honda’s data best can be described as having temporal receptive fields with a standard deviation around 24 ms and 35 ms for the CG and MAX method, respectively. For example, using a shape parameter \( n \) of 5, we can find a good model for the CG method by setting \( \lambda \) equal to 10.6 ms, as the standard deviation of the temporal receptive field, i.e., a gamma-distribution with shape parameter 5 and scale parameter 10.6 ms, will be close to 24 ms: \( \sqrt{5 \times (10.7 \text{ ms})^2} = 23.9 \text{ ms} \). Similarly, by setting \( \lambda \) to 15.7 ms we will find a good model for the MAX method. Despite having different temporal receptive fields, and consequently different extraretinal signals, these models predict almost identical mislocalization profiles (three examples each are plotted in Figures 6C and 6F).

Relative position judgments

Intuitively, relative position judgments should be unaffected by changes in direction of gaze. Nevertheless, errors of relative position judgments around saccades have been reported both between a flashed and a continuously present stimulus (e.g., Cai et al., 1997) and between two flashed stimuli (e.g. Schlag & Schlag-Rey, 1995). Hence, it is of special importance to elaborate how our model deals with relative position judgments.

In the context of our model, relative position judgments could be defined either as 1) the comparison of the two global craniocentric position estimates, 2) the time-averaged comparison of the two instantaneous craniocentric position estimates or 3) the time-averaged comparison of the two retinal signals.
First, we will elaborate the predictions of the three methods for the comparison of a flashed and a continuously present stimulus. The mathematical description of the three relative position methods is given by the Equations D1, D2 and D3 below:

\[
\Delta_{f,c}^W(t) = W_c - \bar{W}_f(t), \quad \text{(D1)}
\]

\[
\overline{\Delta}_{f,c}^W(t) = \int [W_c(\tau) - W_{f,\tau} \tau] g(\tau) d\tau, \quad \text{(D2)}
\]

\[
\overline{\Delta}_{f,c}^R(t) = \int [R_c(\tau) - R_{f,\tau} \tau] g(\tau) d\tau. \quad \text{(D3)}
\]

Using Equation 8, it becomes obvious that all three estimates of relative position \(\Delta_{f,c}(t)\) are identical and equal to \(E_f(t)\). Hence, the relative error between the flashed and the continuously present stimulus is identical to the absolute error of the flashed stimulus. Further, it is clear from Equation D3 that the relative mislocalization does not depend on the execution of an eye-movement: as the extraretinal signal does not figure into Equation D3, the same pattern of relative mislocalization will be observed if the continuously present stimulus is actively moved across the retina by a saccade or is moved across the retina due to stimulus motion. Consequently, our model makes identical predictions for relative position judgments during real and simulated saccades.

Now we will turn to the relative localization of two stimuli flashed at times \(u\) and \(v\), respectively. In analogy to Equations D1, D2 and D3, we define the Equations D1',

Figure 5. Perisaccadic shift with a fast retinal signal. (A) Neuronal responses for the model with four different temporal dynamics, as in Figure 2. In contrast to Figure 2, the retinal signal is defined as the location with the strongest activity (MAX method, see text for details). Note that the retinal signals differ substantially from the ones observed with the center of gravity method (CG) in Figure 2. Instead of becoming slower with increasing \(\lambda\) they are fastest for the slowest dynamics where they resemble a step function. B) Predicted mislocalization for flashed stimuli with the MAX method. As observed for the CG method in Figure 4, mislocalization amplitude increases with slower temporal dynamics \(\lambda\).
Within the framework of our model, these formulas provide a mathematical description of the relative localization for two stimuli flashed at times \( u \) and \( v \), respectively:

\[
\overline{\Delta_f^W}(u, v) = \overline{W_f(u)} - \overline{W_f(v)},
\]

\[
\overline{\Delta_f^R}(u, v) = \overline{R_f(u)} - \overline{R_f(v)}.
\]

\( D'_2 \) and \( D'_3 \). Within the framework of our model, these formulas allow us to calculate the mislocalization as a function of the temporal receptive fields which were modeled as gamma distributions with different scale and shape parameters (\( \lambda \) and \( n \) plotted on the x and y axis, respectively). Upper and lower row illustrate results for the CG method and the MAX method, respectively. (A),(D) Magnitude of the mislocalization defined as the range of mislocalization values (max–min) normalized to saccade amplitude as a function of shape and scale parameter. Across all models, mislocalization amplitude can be predicted reasonably well by the standard deviation of the temporal receptive fields alone (see methods and inset). (B),(E) Similarity between model predictions and data from Honda (1989) was quantified as the maximum difference. The black lines are iso-standard deviation curves of the temporal receptive fields. Across all models, the error can be predicted reasonably well by the standard deviation of the temporal receptive fields alone (see inset). Models that provide a good fit (i.e., error close to zero) have a standard deviation around 24 and 35 ms for the CG and MAX method, respectively. (C),(F) Extragential signal (solid lines) and predicted mislocalization (dashed lines) for the four models indicated by the red, green, blue and black dots in (B) and (E), respectively. The red, green and blue models were selected randomly from the family of models which provide a good fit to the data (thick gray line, see text for details). Note that despite the variability between the extraretinal signals, predicted mislocalization is very similar for all models. The black model uses the temporal receptive field selected by Pola (2004) on the basis of flicker fusion thresholds. Despite the fact that it was not selected based on its fit to the data it matches it quite well.
In contrast to the relative position judgments between a flashed and a continuously present stimulus, here the three equations yield different results. From Equation D3 the relative position difference between the two flashed stimuli is equal to the difference in the direction of gaze at the time of the flash: \( \Delta_f^R(u, v) = R_f(u) - R_f(v) \). In particular, if both stimuli are flashed before saccade onset, the two stimuli will not be mislocalized relative to each other. This is not in keeping with experimental findings. If we use the fact that \( W_f(\tau) = R_f(\tau) - eR(\tau) \) we see that Equation D2 yields the same results as Equation D3. In contrast, Equation D1 describes the relative mislocalization as the difference in the global cranio-centric position estimates. Hence, in line with experimental findings, two stimuli flashed before saccade onset will be mislocalized relative to each other.

**Discussion**

For almost 200 years, the reafference principle has been one of the most widely accepted theories in sensory-motor research in general, and oculomotor control in particular (e.g., Helmholtz, 1866; Purkinje, 1825; Steinbuch, 1811; von Holst & Mittelstaedt, 1950; for a thorough historical review see Gruesser, 1995). For many years, the finding of perisaccadic shift was interpreted as evidence against the reafference principle. We argue that this interpretation was based on disregarding slow temporal dynamics in afferent visual neurons, which in turn, led to the faulty assumption that the extraretinal signal should represent eye position. Recent modeling work has acknowledged temporal low-pass filtering (Pola, 2004) and considerably changed our interpretation of perisaccadic shift. However, the implications of this low-pass filtering for the retinal signal of continuously present stimuli had not been acknowledged so far. Our model closes this link and hence provides a stringent implementation of the reafference principle. As required by the reafference principle, our model views the extraretinal signal not as an erroneous estimate of eye-position, but rather as an accurate estimation of the retinal signal of continuously present stimulus. By inverting this model we can deduce the preprocessing parameters that predict perisaccadic shift in the framework of the reafference principle. Our results show that for physiologically plausible preprocessing parameters, the reafference model does indeed predict perisaccadic shift. Hence, the reafference principle provides a very simple and elegant account of both, perisaccadic shift and perisaccadic stability.

In addition to the reafference principle, our model is indebted to a number previous approaches such as an informal suggestion by MacKay (1970) to account for mislocalization during simulated saccades, a model by Krekelberg and Lappe (2000) to account for the flash lag effect and the models by Pola (2004, 2007, 2008) to account for mislocalization during real saccades. Our new model provides five main advances to the field.

1. Our model provides a novel and compelling explanation as to why perisaccadic shift occurs. While previous models could accurately model perisaccadic shift, they did not convincingly answer the question why the extraretinal signal would happen to be chosen in a way that is necessary to predict perisaccadic shift. In our framework, perisaccadic shift follows from two simple principles: perisaccadic visual stability and physiologically plausible temporal dynamics in the afferent visual pathway. The former principle can easily be verified by introspection: the world does appear stable during saccades; the latter has been documented extensively in electrophysiological and psychophysical studies.

In the context of our model the extraretinal signal is optimal in that it guarantees perisaccadic visual stability for continuously present stimuli. This interpretation removes perisaccadic shift from the list of mislocalization phenomena such as the Filehne Illusion and the phenomenon of autokinesis which are thought to be caused by an inadequate extraretinal signal. Note, that our explanation of perisaccadic shift by an accurate extraretinal signal does not imply that the extraretinal signal would have to be accurate in other conditions as well which give rise, for example, to the Filehne Illusion.

2. Our model demonstrates that mislocalization of flashed stimuli and correct localization of continuously present stimuli are not mutually exclusive. Hence, to our best knowledge, it provides the first account for erroneous perisaccadic relative position judgments between flashed and continuously present stimuli (Cai et al., 1997; Teichert, Klingenhofer, Wachtler, & Bremmer, 2008).

3. Our model provides a link between two of the three frameworks of visual stability during eye movements: efferent (e.g., Helmholtz, 1866; von Holst & Mittelstaedt, 1950) and reafferent theories (Murakami & Cavanagh, 1998, 2001). The former theory holds that the efference copy provides a means to correct afferent signals for changes of the position of the retinae in space, the latter assumes that changes of direction of gaze can be estimated from the afferent visual signals themselves without knowledge of either motor efference or proprioceptive afference. For the absolute localization of flashed stimuli in the dark our model depends exclusively on the efferent theory of visual stability. In contrast, the relative mislocalization during simulated saccades depends exclusively on the reafferent theory. The link between the two theories can best be appreciated for relative position judgments during real saccades; if the extraretinal signal accurately predicts the retinal signal of the continuously present stimulus, both theories make exactly the same predictions.

4. The model predicts relative mislocalization of a flashed relative to a continuously present stimulus during simulated saccades. This visual stimulus configuration is
very similar to the one in which the flash lag effect is typically observed. Hence, we establish a common ground between models of perisaccadic shift and the flash lag effect (e.g., Krekelberg & Lappe, 2000). The field of perisaccadic mislocalization may benefit from certain ideas which have been brought up in the flash lag literature. For example, the mechanism of motion extrapolation (Nijhawan, 1994) by lateral connections as well as subtle latency differences between flashed and continuously present stimuli (e.g., Purushothaman, Patel, Bedell, & Ogmen, 1998) may play an important role in the fine-tuning of more sophisticated versions of our model (see also below).

(5) Our approach imposes novel restrictions on the family of possible models described by Equation 2'. Realistic patterns of perisaccadic mislocalization can only be observed with slow temporal dynamics. Our simulations suggest that the standard deviation of the temporal kernels needs to be on the order of 24 and 35 ms for the CG and MAX method, respectively. Thus, we can definitely rule out the damped eye-position model which relies on the assumption that the temporal kernel is a Dirac delta function.

Relation to previous models

Our approach supports and complements Pola’s work. First, independent of the flicker-fusion experiments that motivated Pola’s choice of a slow temporal kernel, we find that only such slow temporal kernels may produce perisaccadic shift in the context of our model. This converging evidence should considerably strengthen the case in favor of the slow temporal kernels. Second, the restrictions imposed by Pola on the set of solutions of Equation 2', are orthogonal to the restrictions we can impose. While he identifies a single temporal receptive field with the associated family of extraretinal signals, we identify a family of temporal receptive fields each with its unique extraretinal signal derived via Equation 8. These restrictions may be combined to yield a unique model of perisaccadic shift, i.e., we can use the temporal receptive field identified by Pola in our model. Using the CG and the MAX method, respectively, we identify two unique models of perisaccadic shift. For the CG method the predicted mislocalization amplitude exceeds the one observed by Honda (1989), for the MAX method it matches it pretty closely (see Figures 6C and 6E, black dotted line). As there is considerable variability in reported mislocalization amplitude, we do not consider this a definite argument against the CG method. Further, the estimation of the temporal receptive field properties from flicker fusion data is certainly subject to variability. Hence, we regard both of these models as feasible candidates. Note that the two models have very different extraretinal signals. The CG model features a slow exR, while the MAX/ML model features a very fast one.

Note that for both of these models which lie at the intersection of Pola’s and our restrictions, the extraretinal signal starts moving only well after saccade onset. Furthermore, within the framework of our simulations there is not a single plausible model (i.e., a model with causal receptive fields) with an exR that starts moving before saccade onset. This provides further evidence against the common assumption that the mislocalization of stimuli flashed before saccade onset can only be explained by an anticipatory extraretinal signal. This argument certainly does not deny the existence of anticipatory signals that may help to prepare visual areas for upcoming saccades. It merely argues against the involvement of such signals in perisaccadic shift.

Effect of stimulus duration

Our model predicts no mislocalization for continuously present stimuli, and shift-like mislocalization for flashed stimuli. Therefore, it seems well suited to make predictions regarding the transition between these two extremes. We simulated the localization of flashes with different durations from 1 to 80 ms (data not shown). As expected, the maximal mislocalization amplitude drops gradually to about a quarter of its original value as stimulus duration is increased from 1 to 80 ms. For longer stimuli, a small effect of saccades on mislocalization in saccade direction is predicted when stimulus offset occurs around saccade onset. Conversely, mislocalization opposite to saccade direction is predicted when stimulus onset occurs around saccade offset.

Assumption of linearity

In the context of the current model we assumed that the visual system processes information linearly and strictly feed-forward. Rational for this assumption was to keep the model both easy to understand and mathematically tractable. However, we acknowledge that the assumption of linearity is a simplification that almost certainly will be violated especially when using very short visual stimuli. Several studies have shown that responses to very brief visual stimuli are stronger and last longer than responses predicted from linear models (Duysens, Orban, Cremieux, & Maers, 1985; Kratz & May, 1990; Levick & Zacks, 1970). Other studies suggest different nonlinearities that may lead to latency differences between responses to the flashed and moving present stimuli (e.g., Orban, Hoffman, & Duysens, 1985). Similarly, there may be nonlinearities present in the responses to the continuously present stimuli. These nonlinearities may be mediated by lateral or feedback projections (Matin, Clymer, & Matin, 1972). Naturally, our idea of calibrating the extraretinal signal to cancel out the retinal signal of a continuously present stimulus can be applied to models that incorporate such
We begin our discussion of the matter by noting that our model by itself does not predict the observed mislocalization of stimuli flashed during simulated saccades: the extraretinal signal stays constant and hence flashed stimuli are not mislocalized in craniocentric coordinates. However, as noted above, our model does predict relative mislocalization between the flashed stimulus and the continuously present background. In other words, the flashed stimulus is mislocalized relative to the coordinate system defined by the visual references in the background.

We speculate that the empirically observed absolute mislocalization of the flashed stimulus may be causally related to this mislocalization relative to the visually defined coordinate system. To that aim we assume that in the presence of visual references, the visual system estimates the craniocentric position of the flashed stimulus indirectly, i.e., relative to the visual references:

\[ W_f = W_c + \Delta. \]  

Here \( W_f \) corresponds to the craniocentric position of the continuous stimulus and \( \Delta \) is given by our model as calculated by Equations D1'/2'/3'. In the context of our model, \( W_c \) does not correspond to the true location of the continuously present stimulus. Consequently, Equation 9 does not predict the empirically observed mislocalization. However, if we assume that the visual system has an independent way to correctly estimate \( W_c \), Equation 9 would predict the observed mislocalization during simulated saccades. We will not speculate in detail about the mechanisms that may give rise to an accurate estimate of \( W_c \). However, we assume that it would involve neurons with inherently faster temporal dynamics, i.e., the magnocellular pathway, in combination with cross-validation by other sensory-motor systems.

It is important to note that for simulated saccades the reported effect of mislocalization in the direction opposite to the simulated saccade tended to be weaker or even absent (Morrone et al., 1997). Assuming that our model provides an accurate prediction of the relative position error, this implies that in the second half of the simulated saccade part of the error is actually carried by an absolute mislocalization of the continuously present stimulus, i.e., an error in \( W_c \). Further, it needs to be mentioned that one recent study (Ostendorf et al., 2006) reported compression of space around the onset of simulated saccades. Our model does not predict such behavior (see next paragraph).

**Relation to perisaccadic compression**

Depending on the presence of visual references and/or background luminance, different patterns of perisaccadic mislocalization have been reported (Lappe et al., 2000).

---

**Active eye-movement versus passive image motion**

Flashed stimuli are not only mislocalized during real but also during simulated saccades. Stimuli flashed around the onset of fast uniform background motion will be mislocalized in direction opposite to the background motion, i.e., in the direction of the simulated saccade. The pattern of mislocalization during simulated saccades is quite similar to the one observed during real saccades (Honda, 1995; MacKay, 1970; Morrone, Ross, & Burr, 1997, see however, Ostendorf, Fischer, Gaymard, & Ploner, 2006). Despite these similarities, no common mechanism has been suggested for mislocalization during real saccades in the dark and simulated saccades in ambient lighting conditions. In the following we want to relate the predictions of our model to the observation of mislocalization during simulated saccades.
Perisaccadic shift is generally observed in the dark and in the absence of (postsaccadic) visual references. With ambient light and in the presence of visual references a different pattern of mislocalization is observed, the so-called perisaccadic compression (e.g., Morrone et al., 1997). So far, no model of perisaccadic mislocalization, including ours, has been able to accurately predict perisaccadic shift and compression as a function of the visual references and/or background luminance.

However, it is important to note that our model accurately predicts relative mislocalization during simulated saccades, which are necessarily conducted with visual references (see above). Such conditions typically lead to perisaccadic compression if measured with real, not simulated saccades. Our tentative explanation of this finding is the following. We assume that perisaccadic compression depends in one way or another on the remapping of receptive fields. Hence we do not expect perisaccadic compression during real saccades only if visual stimuli are present in the scene. In the dark, the remapping of receptive fields which is thought to link the pre- and postsaccadic neuronal representation in retinocentric visual areas seems pointless. Under this assumption we would predict no perisaccadic compression during real saccades in the absence of visual references.

### Conclusions

For almost 200 years, the reafference principle was thought to mediate perceptual stability during eye-movements. Several experiments used brief perisaccadically presented stimuli to test the validity of the reafference principle around the occurrence of fast eye-movements; the finding of perisaccadic shift was interpreted as evidence against the reafference principle. Consequently, perisaccadic visual stability was attributed to the general decrease of visual sensitivity (saccadic suppression), the remapping of visual receptive fields and visual masking. Our modeling efforts show that the reafference principle actually provides a very simple and elegant account of perisaccadic shift, if slow temporal dynamics in the afferent visual pathway are acknowledged: instead of interpreting the extraretinal signal as a faulty estimate of eye position, we interpret it as an accurate estimate of the neuronal representation of a continuously present stimulus. Hence, the model solves the apparent paradox of why mislocalization occurs for flashed, but not for continuously present stimuli. Rather than being mutually exclusive, we argue that mislocalization of flashed stimuli arises as a consequence of providing accurate craniocentric position estimates for continuously present stimuli under the assumption temporal low-pass filtering the afferent visual pathway.

### Appendix A

#### Function arguments and subscripts:

To specify the input to the functions (see below) we use a number of different arguments. The main argument(s) is (are) presented in parenthesis following the function symbol. Additional qualifying arguments are represented as subscripts.

- \( \text{cif} \) qualifies a function as pertaining either to conditions with a flashed or a continuously present stimulus.
- \( \tau \) denotes time within a trial relative to saccade onset.
- \( t \) specifies a condition in which the flash was present at time \( t \) relative to saccade onset.

#### Constants and Operators

- \( n \) — shape parameter of the Gamma distribution.
- \( \lambda \) — scale parameter of the Gamma distribution. Corresponds to the time constant of the underlying exponential distribution.
- \( * \) — convolution operator. The spatio-temporal convolution operator was defined by:

\[
S * \xi(x, \tau) = \int S(u, v)\xi(u - x, v - \tau)du\,dv. \tag{A1}
\]

#### Functions

- \( h(\tau) \) — direction of gaze as a function of time from saccade onset.
- \( S(x, \tau) \) — retinal projection of the visual scene as a function of one-dimensional space \( x \) and time \( \tau \) after saccade onset.
- \( \xi(x, \tau) \) — (spatio-) temporal receptive field. A kernel which describes the neuronal processing of the retinal...
stimulus. \( \xi(0, \tau) \) describes the one-dimensional temporal receptive field (spatial position \( x \) is being held constant at zero). To simplify the notation we use \( \xi_0(\tau) \) or simply \( \xi_0 \) to refer to the same expression, i.e., the temporal receptive field.

\( N(x, \tau) \) — neuronal activity in the retinocentric visual area as a function of position \( x \) and time \( \tau \) after saccade onset. In Equation 4 \( N(x, \tau) \) is described as the convolution of \( S \) with \( \xi \).

\( R_{ef}(\tau) \) — retinal signal defined in Equation 5 as the retinocentric position of a stimulus as estimated from the neuronal activity. For example, \( R_{ef}(\tau) \) denotes the retinal signal of a stimulus flashed at time \( t \) after saccade onset as a function of time \( \tau \) after saccade onset. As \( R_{ef}(\tau) \) does not vary as a function of \( \tau \), we will at times rewrite the same expression as \( R_{ef}(t) \). Note that \( R_{ef}(t) \) is the inverse of direction of gaze at the time of the flash \( R_{ef}(t) = -h(t) \).

\( exR(\tau) \) — extraretinal signal. In the context of the damped eye-position model, \( exR \) is interpreted as the visual system’s (erroneous) estimate of eye-position. In the current manuscript \( exR \) is defined as the retinal signal of a continuously present stimulus (see Equation 8). Hence \( exR \) is not an (erroneous) estimate of eye-position, but an accurate estimate of \( R_{ef} \). \( exR \) can be thought of as the output of the forward model (e.g., Kalveram, 1993) which explicitly predicts the reafference as a function of the efference and the neuronal preprocessing (reafference Principle: von Holst & Mittelstaedt, 1950). If \( \xi_0 \) is a Dirac impulse our definition and the definition used in the damped eye-position model are identical (except for the sign).

\( W_{ef}(\tau) \) — instantaneous craniocentric position. Is derived by subtracting \( exR \) from \( R_{ef} \) (see Equation 6). Additional subscripts may specify the instantaneous craniocentric position of the continuous or the flashed stimulus. If the flashed stimulus is specified, a second subscript \( t \) may indicate the time of the flash relative to saccade onset.

\( W_{fc}(t) \) — global craniocentric position. For clarity, the second subscript indicating the time \( t \) of the flash after saccade onset is now the explicit argument. \( W_{fc}(t) \) represents the craniocentric position estimate of a stimulus flashed at time \( t \) after saccade onset. Notice the difference to the instantaneous craniocentric position which is expressed as a function of \( \tau \), i.e. time after saccade onset.

\( \tilde{E}_{f}(t) \) — global craniocentric position error. Because the actual stimulus position was always 0° in craniocentric coordinates, it is identical to the global craniocentric position: \( \tilde{E}_{f}(t) = W_{fc}(t) \).

Acknowledgments

This research was supported by DFG GRK 885, DFG FOR 560, and EU-MEMORY.

Thanks to Vincent Ferrera and Jordan Pola for very helpful discussions.

Also, we want to thank the two anonymous reviewers for their very constructive criticism of earlier versions of the manuscript.

Commercial relationships: none.

Corresponding author: Tobias Teichert.

Email: tt2288@columbia.edu.

Address: 1042 Riverside Drive, New York, NY 10046, USA.

References


