

# Control of fixation duration during scene viewing by interaction of foveal and peripheral processing

Jochen Laubrock\*

Department of Psychology, University of Potsdam,  
Potsdam, Germany



Anke Cajar\*

Department of Psychology, University of Potsdam,  
Potsdam, Germany



Ralf Engbert

Department of Psychology, University of Potsdam,  
Potsdam, Germany



Processing in our visual system is functionally segregated, with the fovea specialized in processing fine detail (high spatial frequencies) for object identification, and the periphery in processing coarse information (low frequencies) for spatial orienting and saccade target selection. Here we investigate the consequences of this functional segregation for the control of fixation durations during scene viewing. Using gaze-contingent displays, we applied high-pass or low-pass filters to either the central or the peripheral visual field and compared eye-movement patterns with an unfiltered control condition. In contrast with predictions from functional segregation, fixation durations were unaffected when the critical information for vision was strongly attenuated (foveal low-pass and peripheral high-pass filtering); fixation durations increased, however, when useful information was left mostly intact by the filter (foveal high-pass and peripheral low-pass filtering). These patterns of results are difficult to explain under the assumption that fixation durations are controlled by foveal processing difficulty. As an alternative explanation, we developed the hypothesis that the interaction of foveal and peripheral processing controls fixation duration. To investigate the viability of this explanation, we implemented a computational model with two compartments, approximating spatial aspects of processing by foveal and peripheral activations that change according to a small set of dynamical rules. The model reproduced distributions of fixation durations from all experimental conditions by variation of few parameters that were affected by specific filtering conditions.

## Introduction

Our conscious perception of a continuous, fully detailed visual environment is illusive. Due to the rapid decrease of visual acuity and spatial resolution with increasing distance from the point of gaze, high-acuity vision is limited to the foveal visual field (Jones & Higgins, 1947; Wertheim, 1894), whereas the periphery looks rather blurry and coarse grained. We therefore move our eyes about three times each second via fast and jerky movements called saccades to bring the regions of interest into high-acuity foveal vision. In this way, we actively sample our visual environment (Findlay & Gilchrist, 2003).

A large field of scene perception research has been investigating the spatial characteristics of eye movements, motivating a number of computational models to predict fixation locations in scene viewing (Hwang, Higgins, & Pomplun, 2009; Itti & Koch, 2000; Itti, Koch, & Niebur, 1998; Kienzle, Franz, Schölkopf, & Wichmann, 2009; Parkhurst, Law, & Niebur, 2002; Torralba, Oliva, Castelhano, & Henderson, 2006; Tsotsos, Culhane, Wai, Lai, Davis, & Nuflo, 1995; Wischniewski, Belardinelli, Schneider, & Steil, 2010; for a recent review, see Borji & Itti, 2013). The temporal aspects of eye-movement control have largely been neglected; so far only one computational model exists for fixation durations during scene viewing (Nuthmann, Smith, Engbert, & Henderson, 2010). The present study therefore focuses on the temporal control of eye movements during scene perception using gaze-contingent spatial frequency filtering.

Citation: Laubrock, J., Cajar, A., & Engbert, R. (2013). Control of fixation duration during scene viewing by interaction of foveal and peripheral processing. *Journal of Vision*, 13(12):11, 1–20, <http://www.journalofvision.org/content/13/12/11>, doi:10.1167/13.12.11.

Since the foveal visual field is most sensitive to high spatial frequencies (Banks, Sekuler, & Anderson, 1991; Hilz & Cavonius, 1974; Robson & Graham, 1981), it is specialized in object identification and the analysis of fine detail. The peripheral visual field, on the other hand, is most sensitive to low spatial frequencies (Banks et al., 1991; Hilz & Cavonius, 1974; Robson & Graham, 1981), and specialized in detecting transients and coarse blobs for the rapid reorienting of overt attention and the selection of new saccade targets (Findlay & Gilchrist, 2003). How is eye-movement control affected by the selective filtering of high or low spatial frequencies in the foveal<sup>1</sup> or peripheral visual field? The moving-window (McConkie & Rayner, 1975) and moving-mask techniques (Rayner & Bertera, 1979) permit gaze-contingent manipulation of the peripheral and the foveal visual fields, respectively, and are therefore ideally suited to investigate this question.

Previous studies that applied spatial frequency filters to peripheral scene regions found consistent effects on saccade amplitudes, but mixed effects on fixation durations. Peripheral low-pass filters, which attenuate high spatial frequencies, decreased the mean saccade amplitude compared with an unfiltered control condition (Foulsham, Teszka, & Kingstone, 2011; Loschky & McConkie, 2002; Loschky, McConkie, Yang, & Miller, 2005; Nuthmann, in press; Shioiri & Ikeda, 1989; van Diepen & Wampers, 1998). Most studies found longer inspection times and fixation durations with peripheral low-pass filters (Loschky & McConkie, 2002; Loschky et al., 2005; Nuthmann, in press; van Diepen & Wampers, 1998), as well as with peripheral high-pass filters (which attenuate low-spatial frequencies; van Diepen & Wampers, 1998). Foulsham et al. (2011), however, reported that both the number of fixations and fixation durations were unaffected by peripheral low-pass filtering. Thus, it is not yet evident how peripheral spatial frequency filtering modulates fixation durations. The studies clearly point out, though, that saccades are preferably programmed to unfiltered scene regions.

Van Diepen, De Graef, and d'Ydewalle (1995) and van Diepen (2001) also confirmed this for foveal degradation in black-and-white line drawings of scenes: Reduced foveal image contrast led to longer saccade amplitudes owing to a decreased amount of short intra-object saccades in the degraded foveal region. Foveal degradation also yielded longer inspection times and fixation durations (Nuthmann, in press; van Diepen, 2001; van Diepen, De Graef, & d'Ydewalle, 1995). This effect is consistent with computational models of eye-movement control that predict longer fixation durations with increased foveal

processing difficulty (Engbert, Nuthmann, Richter, & Kliegl, 2005; Nuthmann et al., 2010; Reichle, Pollatsek, Fisher, & Rayner, 1998).

Taken together, previous studies suggest that image degradation leads to increased fixation durations and the spatial selection of undegraded scene regions. The effects of high-pass filtering, i.e., of selectively attenuating information matching the demands of peripheral vision, have rarely been investigated, though, as have the effects of foveal filtering in natural scenes. Also, to our knowledge, there is no study that filtered foveal and peripheral spatial frequencies independently within a single experiment. In the present study, we therefore filtered high or low spatial frequencies in the foveal or peripheral visual field, while subjects explored natural scenes in preparation for a memory test. We compared four filter conditions with an unfiltered control condition: foveal high-pass filtering, foveal low-pass filtering, peripheral high-pass filtering, and peripheral low-pass filtering.

While the theoretical focus of our study was on the control of fixation durations, effects of filtering on saccade amplitudes are also analyzed and reported. First, we expected peripheral filtering to elicit shorter saccades and foveal filtering to elicit longer saccades compared with the control condition. The attenuation of low spatial frequencies in the periphery should impede saccade target selection more than the attenuation of high spatial frequencies; for foveal filtering, we expected effects of filter type to be reversed. Second, part of the information was missing in all conditions with filtered scenes, and fixation durations should lengthen in all experimental conditions to the extent that the missing information is used in the control of fixation duration. Given the proposed importance of foveal information for the control of fixation durations, foveal filters should particularly lengthen fixation durations. Moreover, we expected the effects to increase with increasing loss of critical spatial frequency content, i.e., with foveal low-pass filters and peripheral high-pass filters.

In the second part of the manuscript, we present a new computational model for the control of fixation durations that was inspired by the experimental results. Since our data seemed most compatible with a dynamical interaction of foveal and peripheral processing, we developed a general modeling framework that permits foveal and peripheral co-activation to modulate ongoing fixation durations. It turns out that the interactive model is in good agreement with our experimental results and that fixation duration is strongly modulated by the difference between foveal and peripheral activations.

## Experiment

### Method

#### Participants

Participants were 11 students from the University of Potsdam (three male, mean age: 24 years), who received course credit or €7 for participation in the experiment. All participants had normal or corrected-to-normal vision. The experiment conformed to the standards set by the Declaration of Helsinki.

#### Apparatus

Stimuli were presented on a 20-in. Iiyama Vision-Master Pro 514 monitor at a resolution of  $1024 \times 768$  pixels and a refresh rate of 100 Hz. A head-chin rest ensured a viewing distance of 60 cm. While viewing was binocular, gaze position of the right eye was recorded using an EyeLink 1000 tower mount system (SR Research, Ontario, Canada) at a sampling rate of 1000 Hz. Stimulus presentation was controlled using Matlab (2009b, The Mathworks, Natick, MA) and the OpenGL-based Psychophysics Toolbox (Kleiner, Brai-nard, & Pelli, 2007; Pelli, 1997).

#### Stimuli

Stimuli were 120 color photographs of natural indoor and outdoor scenes. Of these, 36 were shot in portrait format and 84 in landscape format. Two filtered versions of each image were prepared in advance, one using a high-pass and one using a low-pass filter. Spatial frequency filtering was realized by folding the images with a quadratic kernel with a side length of 25 pixels. Filters were recursively applied seven times in a row. The low-pass filter was a flat kernel standardized to a sum of one; the high-pass filter was a Laplacian kernel, combining a positive center with a negative surround. Filter levels were chosen heuristically so that filtering was above threshold, but the remaining information looked still usable. The signal was attenuated by more than 10 dB at spatial frequencies smaller than 0.8 cycles/° for high-pass filtering and greater than 1.4 cycles/° for low-pass filtering in a comparison of the radially averaged power spectra of filtered and unfiltered images. These numbers correspond fairly well to the maximal sensitivities of magno- and parvocellular cells in the lateral geniculate nucleus (LGN), which has been estimated at 1 and 10 cycles/°, respectively (Derrington & Lennie, 1984).

For gaze-contingent presentation, foreground and background images were merged in real-time using alpha blending. For example, in the foveal low-pass

condition the blurred version of the scene was used as foreground image and the original scene as background image. The mixing ratio was given by a blending function that approximates the human contrast sensitivity function (Geisler, Perry, & Najemnik, 2006, equation B7). The alpha mask was centered at gaze location and scaled so that only the foreground image was visible at the fixation point. The peripheral image was weighted more strongly with increasing eccentricity, so that at far eccentricities only the peripheral background image was visible. The weight of the foreground image was less than one half at eccentricities greater than 2.8°.

#### Design and procedure

Two filter locations were combined with two filter types, resulting in four experimental conditions: foveal low-pass filtering, foveal high-pass filtering, peripheral low-pass filtering, and peripheral high-pass filtering (Figure 1). A control condition presented scenes unfiltered and without a gaze-contingent window. Conditions were varied within subjects and scenes; conditions and scenes were presented in random order.

Data were collected in a single 1-hr session. The eye-tracker was calibrated at the beginning of a session and after every 24 trials. A trial started with a screen showing a central fixation trigger. The stimulus was revealed after the trigger had been fixated for at least 50 ms within 1 s from trial start; otherwise, a recalibration was performed. Each scene was presented for 15 s. Participants were instructed to inspect the scenes carefully and answered a three-alternative question concerning the scene content with the computer mouse after each trial. For instance, the control question for the scene illustrated in Figure 1 was “Which color was the passenger car?” and the three alternative answers were “blue”, “gray”, or “black.”

#### Data preparation

Saccades were detected in the raw time series of gaze positions using a velocity-based adaptive algorithm (Engbert & Kliegl, 2003; Engbert & Mergenthaler, 2006). A total of 71 trials were removed owing to poor recording or too much data loss. Single fixations and saccades were removed if they were neighboring eye blinks or outside of the monitor area. The first and the last event of a trial were excluded from analyses as well, since they were associated with scene onset and offset. Glissades following a saccade were assigned to the saccade; if more than one glissade followed a saccade, the glissades and their adjacent fixation and saccade were removed. In total, 4,961 fixations and 2 654 saccades were removed, leaving a total amount of 51,515 fixations and 52,756 saccades for analyses.





Figure 1. Effects of the four filter conditions. The yellow fixation cross illustrates the current gaze position. (Upper left) Foveal low-pass filter. (Upper right) Peripheral low-pass filter. (Lower left) Foveal high-pass filter. (Lower right) Peripheral high-pass filter.

## Results

The questions about the scene content were answered correctly in about 90% of the cases (see Table 1). The somewhat lower fraction of correct answers for foveal high-pass filtering was probably due to control questions that referred to color information in the scene, which was difficult to extract with foveal high-pass filtering.

Data were analyzed with repeated-measures analysis of variance (ANOVA), including planned comparisons with orthogonal contrasts. We specifically tested (a) the effect of filtering, averaged across all filters and compared with the control condition, (b) the effect of filter location, (c) the effect of filter type, and (d) the interaction of filter location and filter type. We also did post-hoc tests using Bonferroni correction, with alpha set to  $p = 0.0125$  ( $0.05/4$ ), to examine if individual filter conditions differed significantly from the control condition. Fixation durations and saccade amplitudes were log-transformed for the analyses to achieve normal distributions.

### Fixation durations

The mean fixation duration over all participants was 304 ms ( $SD = 174$  ms). The distributions of fixation durations (Figure 2) show that peripheral low-pass filtering and especially foveal high-pass filtering led to a reduced number of short fixations, while the number of long fixations was increased. The distributions for foveal low-pass and peripheral high-pass filtering, however, did not differ markedly from the distribution for the control condition.

The ANOVA confirms the distribution patterns. Fixation duration averaged across the four experimental conditions was significantly longer than the

Condition	Correct answer (%)
Control	90.7
Foveal low-pass filtering	88.8
Foveal high-pass filtering	86.6
Peripheral low-pass filtering	92.4
Peripheral high-pass filtering	89.4

Table 1. Percentage of correct answers to control questions for each condition over all participants.

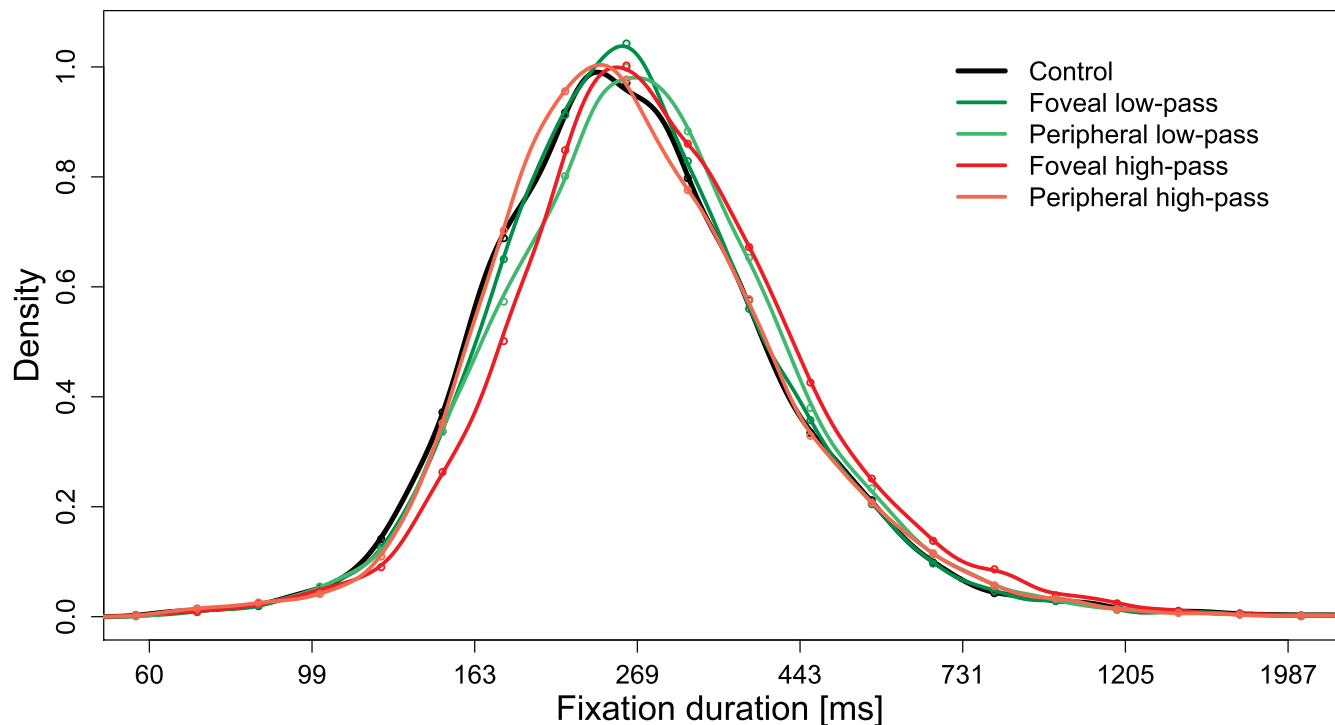


Figure 2. Distributions of fixation durations for the five conditions over all participants. The logarithmic scaling of the abscissa emphasizes the effects for short fixations. Lines represent kernel density estimates using a Gaussian kernel, as implemented in the R function `density` (bandwidth chosen according to Silverman's, 1986, equation 3.31, rule of thumb, with a weight of 1.2). The area under each curve adds up to one.

mean fixation duration of the control condition,  $F(1, 10) = 12.90$ ,  $MSE = 0.001$ ;  $p = 0.005$ . There was no main effect of filter location,  $F(1, 10) = 1.70$ ;  $MSE = 0.002$ ;  $p = 0.221$ , or filter type,  $F(1, 10) = 3.51$ ;  $MSE = 0.002$ ;  $p = 0.090$ , but the interaction of filter location and filter type was pronounced,  $F(1, 10) = 32.94$ ;  $MSE = 0.001$ ;  $p < 0.001$ . As post-hoc tests confirm, this interaction indicates that fixation durations did not differ significantly from the control condition with peripheral high-pass filtering,  $t(10) = -0.48$ ;  $p = 0.641$ , or foveal low-pass filtering,  $t(10) = -0.18$ ;  $p = 0.860$ , but were increased with foveal high-pass filtering,  $t(1,10) = -5.41$ ;  $p < 0.001$ , and peripheral low-pass filtering,  $t(1,10) = -4.23$ ;  $p = 0.002$  (see Figure 3).

### Saccade amplitudes

The mean saccade amplitude was  $5.9^\circ$  ( $SD = 4.5^\circ$ ). The distributions of saccade amplitudes (Figure 4) reveal that short saccades were selectively absent in both foveal filter conditions. Longer saccades occurred more frequently with foveal low-pass filtering than they did with foveal high-pass filtering or the control condition. In the peripheral filter conditions, more short and fewer long saccades occurred than in the control condition, with the pattern being more pronounced for peripheral low-pass filtering.

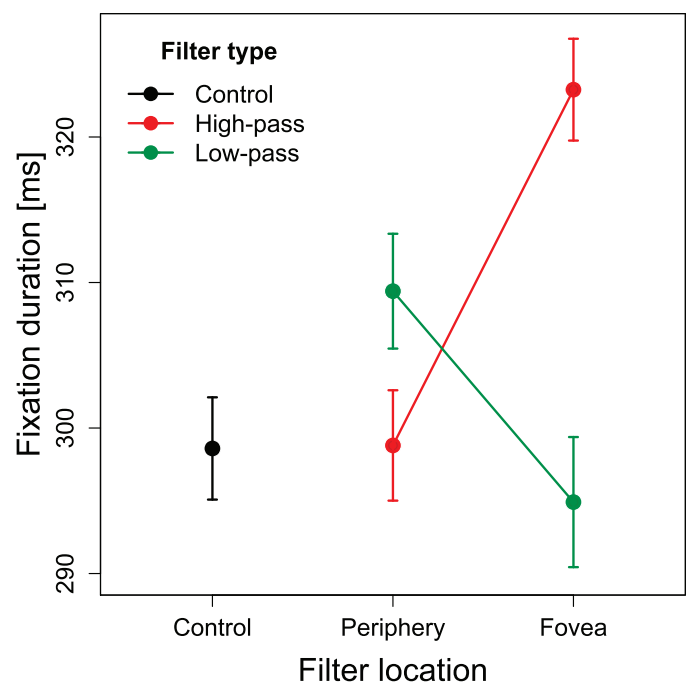


Figure 3. Mean fixation durations for the five conditions. Error bars represent within-subjects standard errors of the mean.

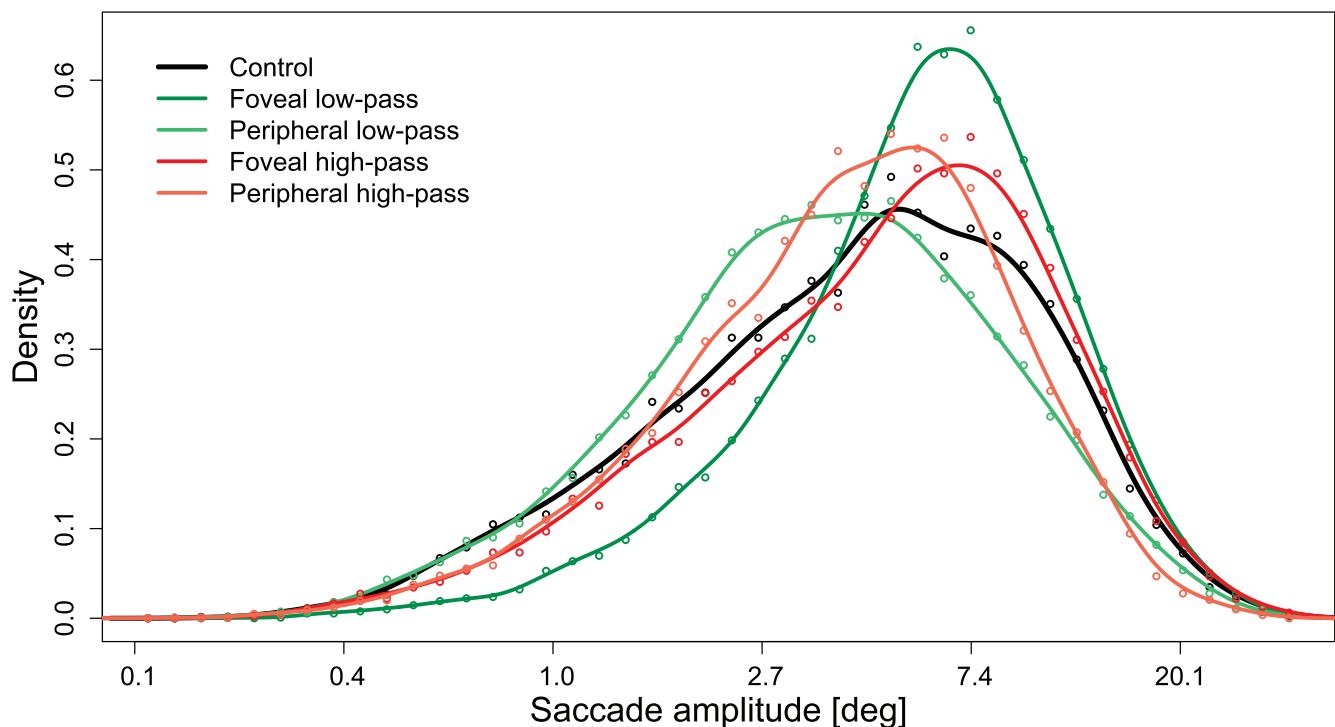


Figure 4. Distributions of saccade amplitudes for the five conditions over all participants. The logarithmic scaling of the abscissa emphasizes the effects for short saccades. Lines represent kernel density estimates using a Gaussian kernel, as implemented in the R function `density` (bandwidth chosen according to Silverman's, 1986, equation 3.31, rule of thumb, with a weight of 1.1). The area under each curve adds up to one.

The ANOVA confirms these patterns, with no main effect of filter type,  $F(1, 10) = 3.48$ ;  $MSE = 0.004$ ;  $p = 0.092$ , but a pronounced effect of filter location,  $F(1, 10) = 211.57$ ;  $MSE = 0.005$ ;  $p < 0.001$ . Peripheral filtering triggered shorter and foveal filtering longer saccades than the unfiltered control condition. Although the sign of these effects was different between filter locations, their magnitude was of similar size; hence, amplitudes averaged across all filter conditions did not differ from the control condition,  $F(1, 10) = 4.88$ ;  $MSE = 0.006$ ;  $p = 0.052$ . Mean saccade amplitude differed from the control condition in three of the four filter conditions, though: peripheral low-pass filtering,  $t(10) = 6.44$ ;  $p < 0.001$ , foveal low-pass filtering,  $t(10) = -8.47$ ;  $p < 0.001$ , and foveal high-pass filtering,  $t(10) = -3.25$ ;  $p = 0.009$ . Peripheral high-pass filtering did not significantly decrease saccade amplitude,  $t(10) = 1.85$ ;  $p = 0.094$ . The interaction of filter type and filter location was also significant,  $F(1, 10) = 49.45$ ;  $MSE = 0.004$ ;  $p < 0.001$ —foveal filtering caused longer saccades with a low-pass filter, and peripheral filtering caused slightly longer saccades with a high-pass filter (see Figure 5).

## Discussion

The effects of fixation durations, which are the focus of the present study, turned out to be strongly in conflict

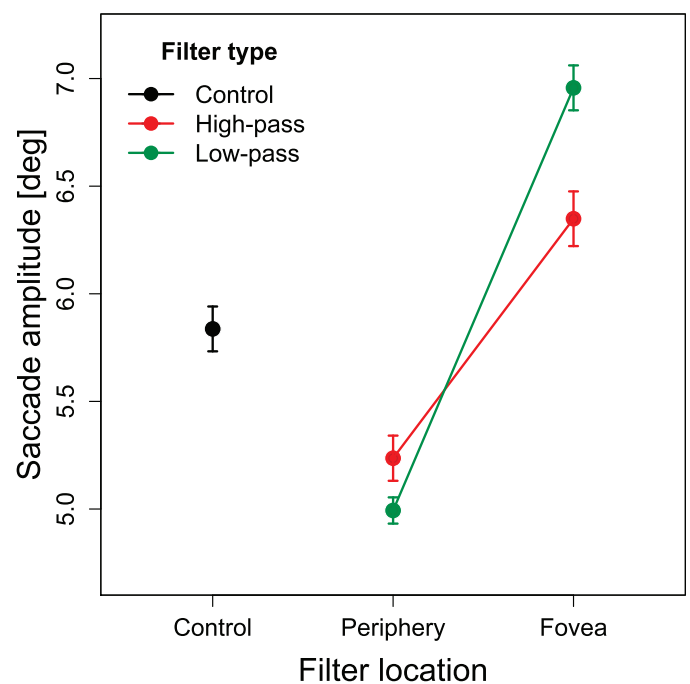


Figure 5. Mean saccade amplitudes for the five conditions. Error bars represent within-subjects standard errors of the mean.

with hypotheses derived from the existence of a functionally segregated visual system. Compared with the control condition, fixation durations were only increased with peripheral low-pass filtering (replicating Loschky & McConkie, 2002; Loschky et al., 2005; van Diepen & Wampers, 1998) and with foveal high-pass filtering, but were unaffected with peripheral high-pass and foveal low-pass filtering. Thus, the two filter conditions with the most serious loss of information did not affect fixation durations, but the relatively more informative filter conditions did.<sup>2</sup> Here, the spatial frequencies that the respective region of the visual field is specialized for were still largely available if attenuated, and participants fixated longer to extract as much useful information as possible. Foveal low-pass and peripheral high-pass filtering, on the other hand, left little useful information to analyze that would make it worthwhile to prolong fixation. Similar to the inverted-optimal viewing position (IOVP) effect in reading (Vitu, Lancelin, & Marrier d'Unienville, 2007), our results show that fixation durations increase when greater amounts of information are anticipated in a specific region.

As expected from earlier findings, the selective attenuation of spatial frequencies in the present experiment affected saccade amplitudes as well. Results were consistent with previous studies. Foveal filtering involved longer saccades with both filter types; apparently, participants preferred to explore scene regions outside the central mask, which caused a decreased number of short inspection saccades (replicating van Diepen, 2001; van Diepen et al., 1995). The amplitude effect was rather strong with foveal low-pass filtering, but less pronounced with foveal high-pass filtering. Thus, when details within the filtered region were still open for exploration with a high-pass filter, a larger proportion of short saccades remained. With peripheral filtering, on the other hand, mean saccade amplitude was shorter than it was in the control condition. This tunnel-vision effect confirms the results of previous studies (Foulsham et al., 2011; Loschky & McConkie, 2002; Shioiri & Ikeda, 1989; van Diepen & Wampers, 1998). Surprisingly, amplitudes were somewhat longer with peripheral high-pass than low-pass filtering, suggesting that high spatial frequencies were more important for the selection of peripheral saccade targets than previously assumed. This is consistent with the finding by Baddeley and Tatler (2006) that high spatial frequency content might be a reliable predictor for fixation locations.

Our findings on the influence of spatial frequency filtering on fixation durations are challenging for theoretical models of eye-movement control. One prominent concept for the control of fixation duration has been termed direct control (for a theoretical overview on direct and indirect control theories, see Henderson & Smith, 2009; Reingold, Reichle, Glaholt, & Sheridan, 2012; see also Trukenbrod & Engbert,

2013), which indicates that processing difficulty of the currently fixated stimulus immediately transfers into observed fixation durations. Under this framework, our findings conflict with the assumption that higher processing difficulty involves longer fixation durations than lower processing difficulty does, since we found (a) longer fixation durations when critical information was partially available and (b) unaffected fixation durations when critical information was strongly attenuated.

One solution to this problem might be to relate our findings to a model that implements the interaction between foveal and peripheral processing for the control of fixation duration. Such a generalization involves a very different control mechanism termed indirect control (see Trukenbrod & Engbert, 2013). While the Controlled Random-walk with Inhibition for Saccade Planning (CRISP) model (Nuthmann et al., 2010) for the control of fixation duration in scene viewing is based on indirect control, it does not address spatial aspects of processing, so that simulation studies on differences in foveal and peripheral processing are precluded. Therefore, we developed a new computational model of eye-movement control with two compartments (fovea, periphery) that interact via a small set of activation-based dynamical rules. The model is inspired by the model framework of Findlay and Walker (1999), who proposed that saccade timing and saccade target selection proceed largely independently and only interact at the lower levels of the oculomotor system.

## A computational model for the control of fixation durations

For a range of visuomotor tasks, computational models of eye-movement control have been proposed within the framework of indirect control; in particular, such models successfully explained saccade timing in reading (Engbert et al., 2005), scene viewing (Nuthmann et al., 2010), and scanning tasks (Trukenbrod & Engbert, 2013). These models postulate a random saccade timer, which determines fixation durations and is modulated by ongoing visual and cognitive processing. With this approach, a variety of effects concerning mean values and distributions of fixation durations can be reproduced. The models do not distinguish between processing in the foveal and peripheral visual field, though, which is a crucial aspect in the present experiment. Therefore, we were motivated to develop a new model of eye-movement control in scene viewing with explicit assumptions on (a) foveal and peripheral processing and (b) their interaction and modulation of the random saccade timer. Such a model seems to be well constrained by our experimental data on gaze-



contingent manipulations of the foveal or peripheral visual field and the resulting changes in eye-movement behavior.

## Core principles of the model

Our model is based on four fundamental principles. First, we assume that a random timer accumulates activation toward a threshold to generate stochastic intervals between saccades, with some preferred mean value. This form of timing has been termed indirect control, since saccades are triggered autonomously, i.e., without cognitive trigger signals. There is experimental as well as theoretical support for autonomous saccade timing across a range of visuomotor tasks (e.g., Engbert & Kliegl, 2001; Engbert et al., 2005; Hooge & Erkelens, 1996, 1998; Nuthmann & Engbert, 2009; for an overview, see Trukenbrod & Engbert, 2013). Second, we assume that the autonomous random timer can be inhibited by ongoing visual-cognitive processing. Note that these two assumptions are also implemented in the CRISP model (Nuthmann et al., 2010), in which a scene-onset delay induces prolonged fixation durations by inhibition of the saccade timer. Third, we account for spatial visual processing by introducing two compartments. Both a foveal and a peripheral compartment are described by temporal activations representing the unfolding of foveal and peripheral processing over time. This choice is motivated by the functional segregation of the visual stream as outlined above. Fourth, we explicitly model the interaction of foveal and peripheral processing for the inhibition of the saccade timer. Principles 3 and 4 are qualitatively different from the CRISP model, which does not account for spatially distributed processing.

The temporal activations of the three model components, i.e., the random saccade timer,  $a_T(t)$  the foveal compartment,  $a_F(t)$ , and the peripheral compartment,  $a_P(t)$ , are implemented as stochastic processes in the form of parallel, discrete-state continuous-time random walks with exponentially distributed waiting times between elementary transitions (Gillespie, 1976). Models based on random walk timing have already been very successful in explaining reaction times in simple saccadic decision tasks (see Smith & Ratcliff, 2004, for an overview).

## A random walk for saccade timing

The random walk of the saccade timer controls the generation of the next saccade. The start of the random walk at time  $t = 0$  signals the beginning of a new

fixation; this state is related to an activation  $a_T(t) = 0$ . The random walk of activation then accumulates by increments of one toward a predefined threshold  $N_T$  with a certain rate. The time when activation reaches threshold corresponds to the fixation duration. Once the threshold is reached, a saccade program is triggered and a new fixation begins by resetting activation to a value of zero (note that for simplicity we do not model the actual saccade programming in the current version of the model; see Engbert et al., 2005, for an explicit model of saccade programming). We define the transition rate  $w_T$  for increments of the timer's random walk as

$$w_T = \frac{N_T}{t_{sac}}, \quad (1)$$

where  $N_T$  represents the number of states the process can adopt, and  $t_{sac}$  represents the mean duration of the timing signal.

## Parallel processing of foveal and peripheral information

In parallel to the timer's random walk, both foveal and peripheral activations,  $a_F(t)$  and  $a_P(t)$  respectively, evolve over time. Here, we follow the basic framework underlying the Saccade-generation with Inhibition by Foveal Targets (SWIFT) model for reading (Engbert et al., 2005). Both foveal and peripheral activations are oculomotor variables in our model, since they directly influence saccade timing. However, they might also be interpreted psychologically in terms of visual processing, which can modulate oculomotor activation: Before processing, the foveal or peripheral stimulus is unknown; after processing, the stimulus is considered completely processed. Both cases are related to an activation of zero. During stimulus processing, activations first accumulate toward predefined thresholds,  $N_F$  or  $N_P$  for foveal and peripheral activation, respectively. During these random walk processes, activations can either increment by one unit with probability  $p > 0.5$  or decrement by one unit with probability  $q = 1 - p$ . Each threshold can be interpreted as the maximum processing difficulty of the respective stimulus. After the threshold is reached, activation gradually declines to a value of zero, with activation either decrementing by one unit with probability  $p$  or incrementing by one unit with probability  $q$ . The transition rates for the increments of the foveal and peripheral random walks,  $w_F$  and  $w_P$ , are multiples of the timer rate  $w_T$ , with

$$w_F = j \cdot w_T \quad \text{and} \quad w_P = k \cdot w_T. \quad (2)$$



## Inhibition of saccade timing by interaction of foveal and peripheral processing

Activations in the two compartments of fovea and periphery can influence the decision to start the next saccade program. This key hypothesis is motivated psychologically and neurophysiologically by Findlay and Walker's (1999) model of saccade generation by parallel processing and competitive inhibition. In this model, a hierarchy of levels of parallel processing in spatial and temporal pathways generates neural activations in fixate and move centers for the control of saccadic eye movements (Level 2 in the model by Findlay & Walker, 1999, tentatively located in the superior colliculus). Foveal and peripheral activations in our model relate to these oculomotor control centers. Therefore, the activations in the two compartments represent oculomotor tendencies to maintain fixations or to move to the next target location (i.e., the periphery). As a consequence, activations are related to cognitive processing, since higher levels of processing clearly influence neural activations in the fixate and move centers (see Findlay & Walker, 1999). The saccade timer can be considered the trigger signal for motor commands to the oculomotor muscles that generate a saccadic movement (Level 1 in the “when” pathway of the Findlay and Walker model); it can be inhibited by upstream processing in the fixate and move centers.

Mathematically, we assume for moment-by-moment control by visual stimulus processing that the transition rate  $w_T$  of the downstream saccade timer can be modulated by the dynamical interaction between foveal and peripheral processing at any point in time during a fixation. We specifically propose that the timer rate decreases when foveal processing demands are higher than peripheral processing demands. We refer to this process as foveal inhibition (see also Engbert et al., 2005; Nuthmann et al., 2010) and define the modulation of the timer by

$$w'_T = \frac{w_T}{h} \quad \text{with} \quad h = 1 + \rho \frac{[a_F(t) - a_P(t)]}{N_F}, \quad (3)$$

where  $\rho$  represents the strength of foveal inhibition, and  $[\cdot]$  indicates that the positive part is taken (i.e., negative values of the argument are set to zero). Thus, any kind of foveal activation leads to inhibition, as long as it is not disinhibited by a higher amount of peripheral activation at a given point in time. Peripheral activation cannot increase the transition rate of the timer, but decreases the proportion of time that inhibition is active. Numerical simulations indicated that a generalization by introducing weighting factors in the activation difference, Equation 3, did not improve the results for the current version of the model. The time course of foveal and peripheral activations

plays an important role for the impact on saccade timing. Foveal inhibition has the greatest effect on the saccade timer when foveal activation accumulates faster than peripheral activation during a fixation and thus leads to a positive activation difference to effectively inhibit the timer early on (large value of  $h$ ). When peripheral activation rises earlier and foveal activation accumulates more slowly during a fixation, inhibition is less effective, since the timer's random walk is a stochastic process and, consequently, can reach its threshold by chance when it is close to the threshold.

Figure 6 schematically illustrates the dynamics of the three model components. Since foveal and peripheral activations of our model can be interpreted as neural activations in fixate and move centers of the model by Findlay and Walker (1999), high foveal activation is in favor of a decision to prolong fixation: High foveal activation induces foveal inhibition that slows down the saccade timer and maintains the current fixation. The case of an early rise of foveal inhibition is illustrated in the upper panel of Figure 6. In addition to this inhibitory influence on the saccade timer from the fovea, our model proposes that high peripheral activation reduces the likelihood for foveal inhibition, effectively disinhibiting the saccade timer in favor of a move response (see Figure 6, lower panel). As a result, there is a dynamical interaction of foveal inhibition and peripheral disinhibition of the saccade timer; it is this interplay that determines the tendency to prolong the current fixation or to move to another fixation location by generating a saccade. Note that the interplay between foveal and peripheral activations is a specific version of Findlay and Walker's push-pull interaction between the (foveal) fixate and the (peripheral) move centers.

## Numerical simulation procedures

To illustrate how the random walks for the timer, the foveal, and peripheral activation work in parallel, consider a system in state  $S(t) = (a_T, a_F, a_P)$  at time  $t$ , which changes to an adjoined state  $S(t + \tau)$  at time  $t + \tau$ . With each time step, only one of the three random walks changes its state, while the other two random walks remain unchanged. Table 2 summarizes all possible state transitions in the model. The three transition rates each represent the probability for a specific state transition, so that the random walk with the highest rate has the highest probability of changing its state (for further details, see Gillespie, 1976). The total transition rate  $W$  is defined as the sum of the three individual rates,

$$W = w'_T + w_F + w_P. \quad (4)$$

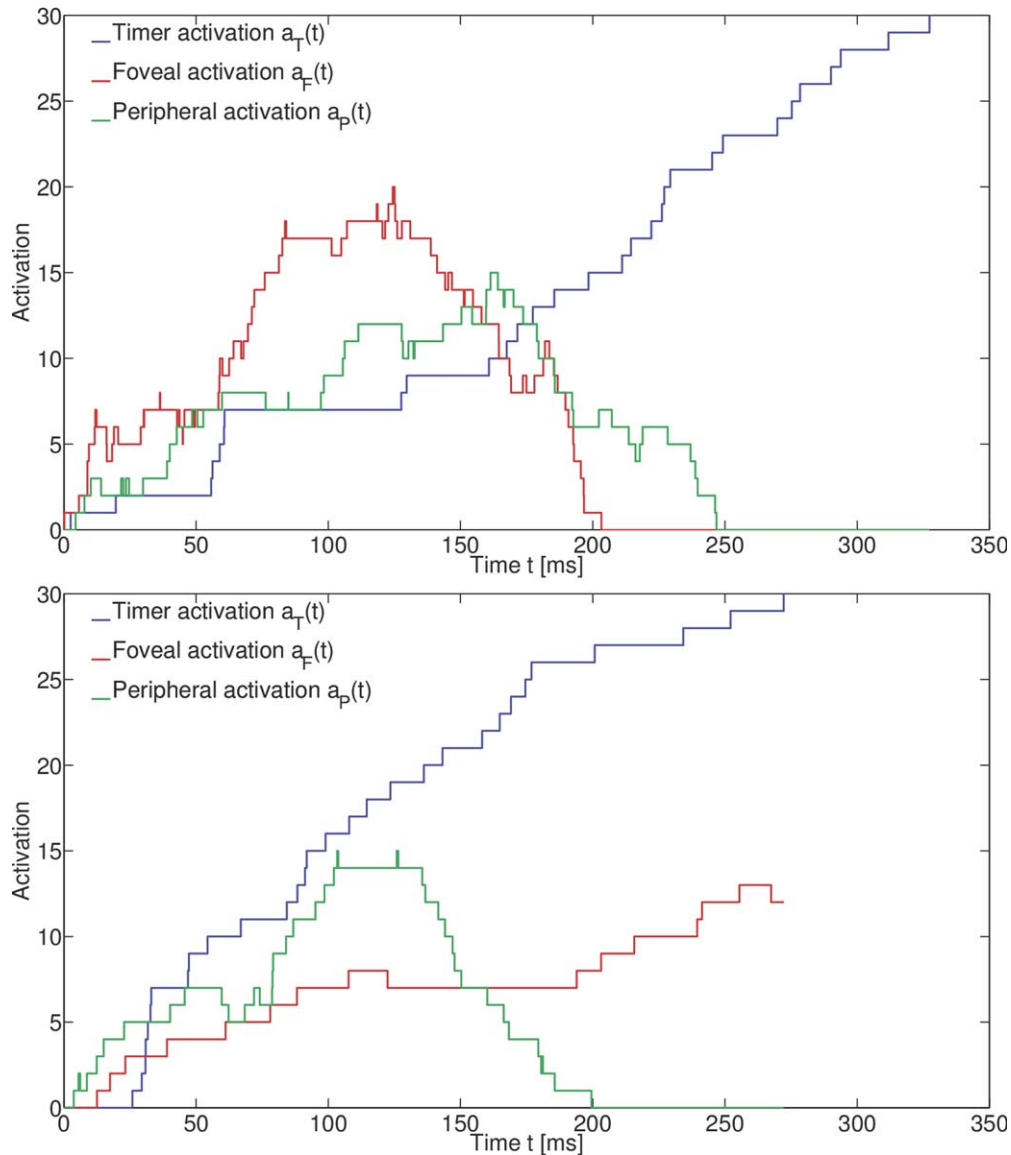


Figure 6. Illustration of the dynamical behavior of saccade timer (blue line), foveal compartment (red line), and peripheral compartment (green line) obtained from numerical simulation of the model. The timer accumulates activation toward a threshold that terminates the current fixation. Foveal and peripheral activations simultaneously and independently accumulate toward a maximum and then decline toward zero. The saccade timer is inhibited when foveal activation is higher than peripheral activation. (Upper panel) Foveal activation builds up early during the fixation, and foveal inhibition of the saccade timer is effective during the initial part of the fixation (duration 327 ms). (Lower panel) Peripheral activation builds up early, leading to disinhibition during the initial part of the fixation (duration 272 ms).

The algorithm consists of two steps. First, a time step is chosen. For each random walk, the transition probability from the current state to the next state depends on the past only through the current state; this is characteristic for Markov processes (e.g., Gardiner, 2004; van Kampen, 1981). Consequently, the waiting time  $\tau$  between different transitions follows an exponential distribution (Gillespie, 1976), and thus can be transformed from a uniformly distributed random number by

$$\tau = -\frac{1}{W} \log(1 - r), \quad (5)$$

where the inverse of the total transition rate  $W$  represents the mean waiting time in a given state  $S(t)$ , and  $r$  is a random number with equal probability in  $0 \leq r < 1$ . Second, a transition is selected in proportion to the transition rate of the walks. The probabilities for selecting a transition in the saccade timer, foveal, and peripheral compartments are given by  $w'_T/W$ ,  $w_F/W$ , and  $w_P/W$ , respectively.

## Numerical simulation study

A simulation study was conducted to investigate if the model was able to (a) reproduce the mean values and distributions of fixation durations from the present experiment and (b) provide a viable mechanism for the experimental findings on fixation duration effects of foveal and peripheral filtering. We thus simulated fixation durations from 11 subjects for foveal low-pass filtering, foveal high-pass filtering, peripheral low-pass filtering, peripheral high-pass filtering, and for the unfiltered control condition.

### Hypotheses

We assume that selective spatial frequency filtering in the foveal or peripheral visual field modulates saccade timing by affecting the random walks for foveal or peripheral processing. Specifically, foveal filtering should affect foveal processing, and peripheral filtering should affect peripheral processing. The time course of foveal and peripheral processing can therefore be modulated in three ways. First, filtering could change the transition rates  $w_F$  and  $w_P$  for foveal and peripheral activations by changing the factors  $j$  and  $k$  (see Equation 2). Second, filtering could change the probability  $p$  of the random walks to increment (or decrement, after the threshold has been reached) by one unit. Third, filtering could change the values  $N_F$  and  $N_P$  for the maximum difficulty of the foveal or peripheral stimulus. We assumed that these modulations could co-occur, of course, but kept a number of model parameters fixed across conditions for psychological plausibility of the model. Since there is neurophysiological evidence for fixed thresholds but variable growth rates obtained from experiments on the generation of voluntary eye movements (Hanes & Schall, 1996), we assumed that the thresholds for foveal and peripheral activations in our model,  $N_F$  and  $N_P$ , are constant across all five filter conditions. The growth of activation, on the other hand, was supposed to change with the different filter conditions—either by changing the transition rates, or by changing the probability of the random walk to increment/decrement by one unit.

### Modeling results

Numerical simulations were run to find the best-fitting model on fixation durations. We aimed at a minimal model with as few free parameters as possible, which captures all the main qualitative effects of fixation durations observed in the experimental data. A genetic algorithm (Goldberg, 1989; Holland, 1975) was used for parameter estimation. Predefined parameter ranges (see Table 3) were chosen for mathematical

Random walk	Transition from $S(t) = (a_T, a_F, a_P)$ to $S(t + \tau) = \dots$			Transition rate
Saccade timer	$a_T + 1$	$a_F$	$a_P$	$w_T'$
Foveal activation	$a_T$	$a_F \pm 1^*$	$a_P$	$w_F$
Peripheral activation	$a_T$	$a_F$	$a_P \pm 1^*$	$w_P$

Table 2. Transition events and transition rate. *Notes:* \*, Foveal and peripheral activations are incremented by +1 before the respective threshold  $N_F$  or  $N_P$  is reached and are decremented by  $-1$  thereafter with probability  $p$ . Activations decrease by  $-1$  after the threshold is reached and increase by +1 thereafter with probability  $q = 1 - p$ .

reasons or for neurophysiological or psychological plausibility. A quantile maximum likelihood approach (Heathcote, Brown, & Mewhort, 2002) served as a goodness-of-fit measure and quantified how much the simulated fixation duration distributions deviated from the experimental distributions (see Appendix for details on the parameter estimation and fitting procedure).

Explorative numerical simulations of two model variants—Model A with variable transition rates for the five filter conditions, and Model B assuming variable probabilities of incrementing/decrementing—indicated that Model A provided a qualitatively better fit to the experimental distributions. This model has 12 free parameters. Six parameters were not allowed to change across the five filter conditions—the three thresholds of the random walks,  $N_T$ ,  $N_F$ , and  $N_P$ , the mean duration of the timing signal,  $t_{sac}$ , the probability  $p$  for the foveal and peripheral random walk to increment (or decrement) by one unit, and the strength of foveal inhibition,  $\rho$ . The transition rates  $w_F$  and  $w_P$  for the foveal and peripheral random walk, however, were allowed to change for the different filter conditions by changing the factor  $j$  with foveal filtering, and the factor  $k$  with peripheral filtering. Six parameters are necessary to describe this modulation, since for foveal as well as for peripheral filtering, there is one transition rate for the control condition, one for high-pass filtering, and one for low-pass filtering.

We fitted this model to the experimental data of each subject separately and obtained 11 sets of model parameters. Each parameter set was used to simulate as many observations (fixation durations) as the respective subject contributed to the experimental data set. Figure 7 illustrates for 1 of the 11 subjects how the simulated fixation duration distributions fit the respective experimental distributions. The model captured the shape for each experimental distribution well, with the characteristic positive skew including a longer tail and the mode below the mean. Despite individual differences between the subjects, the model captured the distributions of all subjects equally well. Although only the distributions were fitted, Figure 8 illustrates

Parameter	Symbol	<i>M</i>	<i>SE</i>	Predefined range
Default random walk transition rate for the saccade timer	$w_T = N_T / t_{sac}$	$N_T = 9$ $t_{sac} = 253 \text{ ms}$	0.28 5.39 ms	3–50 150–300 ms
Random walk transition rate for foveal processing	$w_F = j \cdot w_T$	$j_C = 2.61$ $j_{fLP} = 2.58$ $j_{fHP} = 3.44$	0.15 0.14 0.12	1–5 1–5 1–5
Random walk transition rate for peripheral processing	$w_P = k \cdot w_T$	$k_C = 3.15$ $k_{pLP} = 2.53$ $k_{pHP} = 3.06$	0.23 0.17 0.13	0.01–5 0.01–5 0.01–5
Maximum foveal processing difficulty	$N_F$	33	0.39	3–50
Maximum peripheral processing difficulty	$N_P$	22	1.49	3–50
Probability of foveal and peripheral random walk to increment by one	$\rho$	0.66	0.01	0.55–1.0
Strength of foveal inhibition	$\rho$	4.37	0.48	0.1–10

Table 3. Average parameter values for the best model.

that the simulated data set combined from all 11 subjects recovered the experimental pattern of mean fixation durations remarkably well; only the mean value for foveal high-pass filtering was slightly underestimated by the model. Additionally, the simulation of fixation durations with the average values of the model parameters across subjects yielded very similar results.

The average values across subjects for the 12 model parameters are listed in Table 3. The mean threshold  $N_T$  for the saccade timer is much lower than the values obtained for the thresholds  $N_F$  and  $N_P$  for foveal and peripheral activation, respectively. With a low timer threshold, fixation duration intervals are more variable, so that the corresponding distributions are consider-

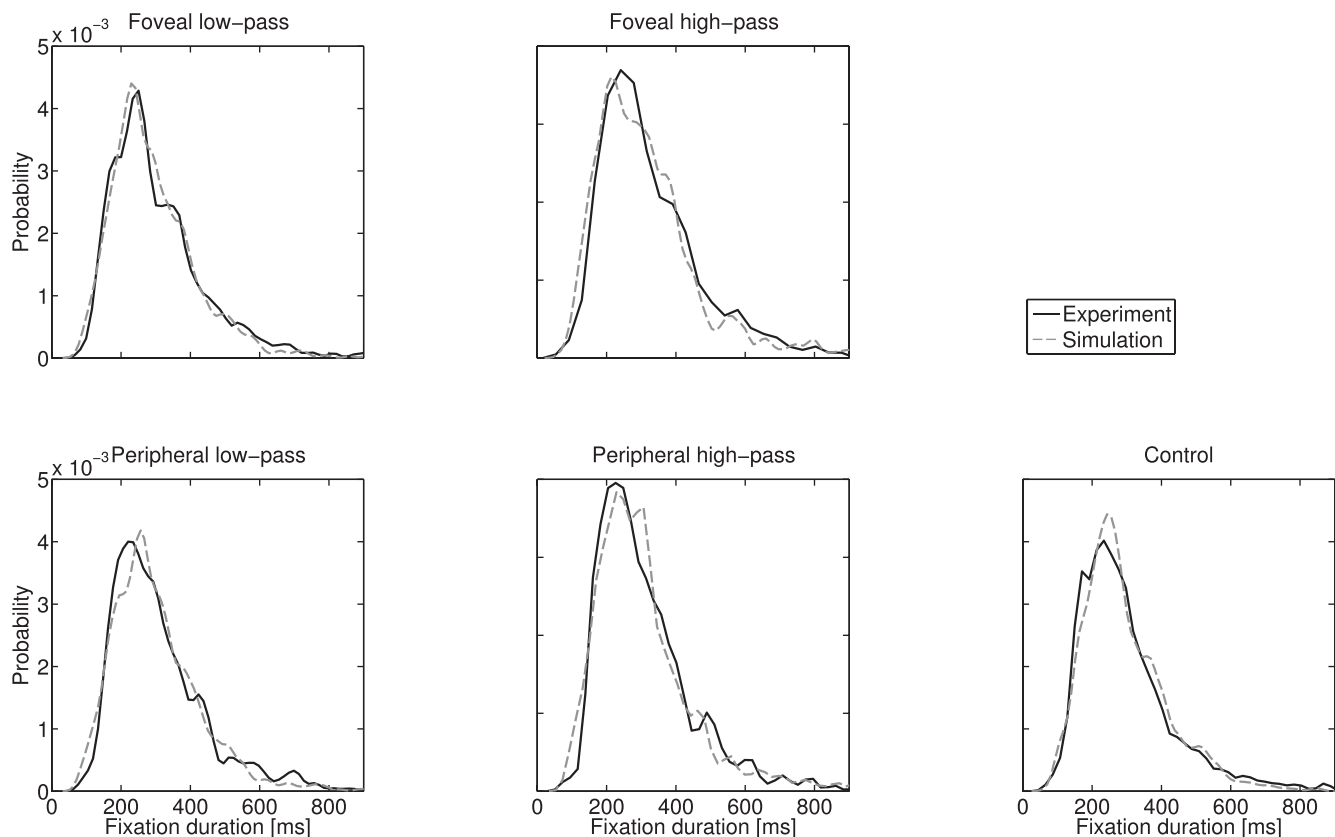


Figure 7. Distributions of fixation durations for the experimental data (solid black lines) and the simulated data (dashed gray lines) for one subject. Simulated fixation durations were generated with the set of individual model parameters that resulted from fitting the model to the experimental distributions of this particular subject. Quantile maximum likelihood served as a quantitative goodness-of-fit measure for fitting the distributions.



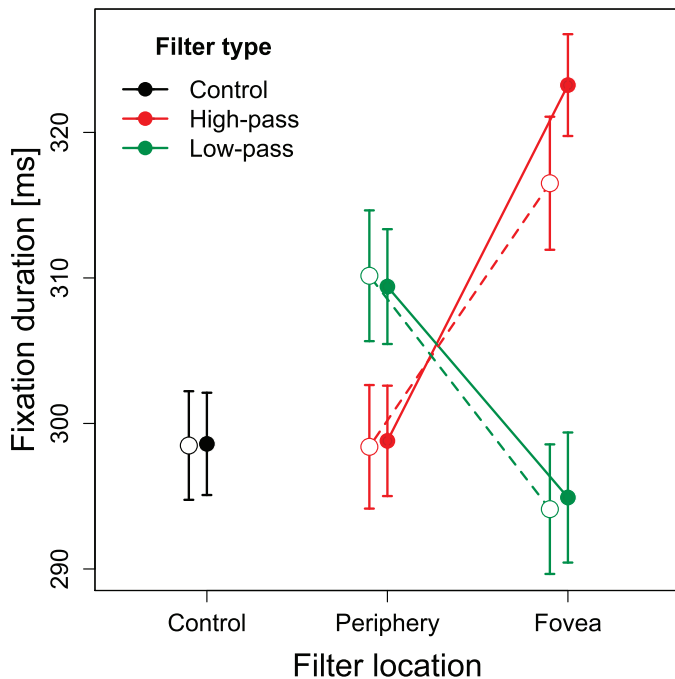


Figure 8. Mean fixation durations for the experimental data (filled symbols with solid lines) and the simulated data (open symbols with dashed lines). Error bars represent within-subjects standard errors of the mean. The simulated data represent the fixation durations for all subjects that were simulated with the 11 sets of individual model parameters.

ably skewed (for a numerical simulation on this issue, see Trukenbrod & Engbert, 2013, figure 1b). Furthermore, the peripheral and foveal activations mostly remain in their incrementing phase (see Figure 9, upper row). The mean duration of the saccade timer ( $t_{sac} = 253$  ms) is slightly larger than the mode for distributions of fixation durations in scene viewing (Henderson & Hollingworth, 1998). The estimated mean value  $\rho = 4.37$  for the strength of foveal inhibition differs significantly from zero, indicating that foveal inhibition is an important model mechanism for explaining the experimental effects. As the numerical values of the parameters  $j$  and  $k$  indicate, the transition rates for the foveal and peripheral random walks are always higher than the rate for the saccade timer, suggesting that foveal and peripheral activations build up faster than the timer activation. This effect is partly counteracted by the higher thresholds for the foveal and peripheral random walks, so that inhibition can potentially contribute during a substantial part of the fixation duration.

The differing fixation duration effects for the five filter conditions emerge from a modulation of the foveal and peripheral transition rates by variation of the parameters  $j$  and  $k$ . Table 3 indicates that the rates for foveal low-pass and peripheral high-pass filtering,  $j_{fLP}$  and  $k_{pHP}$ , are similar to the corresponding rates  $j_C$

and  $k_C$  for the control condition. For foveal high-pass and peripheral low-pass filtering, however, transition rates  $j_{fHP}$  and  $k_{pLP}$  change drastically compared with the control condition. The foveal transition rate for foveal high-pass filtering increases compared with the foveal rate in the control condition ( $j_{fHP} = 3.44$  vs.  $j_C = 2.61$ ), indicating that foveal activation accumulates faster, leading to stronger inhibition. With peripheral low-pass filtering, the peripheral transition rate slows down compared with the corresponding rate for the control condition ( $k_{pLP} = 2.53$  vs.  $k_C = 3.15$ ), indicating that peripheral activation accumulates slower, leading to less disinhibition. Both mechanisms increase the activation difference and thus effectively prolong fixation durations by foveal inhibition. Since the increase of fixation durations is not as pronounced for peripheral low-pass filtering as it is for foveal high-pass filtering, the difference between the transition rates (control vs. experimental) is smaller for peripheral low-pass filtering.

Figure 9 shows the result of 500 model runs per condition, using the fitted parameter values. It is evident that inhibition is on average active earlier and for a longer time period in the foveal-high pass and the peripheral low-pass than in the other three conditions. However, inhibition plays a significant role in all conditions, especially in generating long fixations. The fraction of time during which inhibition was on is estimated at 43%, 43%, 57%, 53%, and 45% for the control, foveal low-pass, foveal high-pass, peripheral low-pass, and peripheral high-pass conditions, respectively.

## Discussion

We developed a model for the control of fixation durations in scene viewing based on the interaction between foveal and peripheral information processing. Numerical simulations of the model reproduced empirical means and distributions of fixation durations well. The model simulations demonstrate that the concept of a dynamical interaction of foveal and peripheral processing is a promising framework for the control of fixation durations.

Specifically, for unimpaired scene viewing (control condition), the model parameters revealed a lower threshold and a higher transition rate of activations in the peripheral compartment compared to the corresponding values in the foveal compartment (see Table 3). Therefore, spatially distributed processing in the model is compatible with the current view that, due to the physiological properties of the visual system, processing of information from the peripheral visual field is faster than processing of information from the foveal visual field. In the model, faster rise and decay of

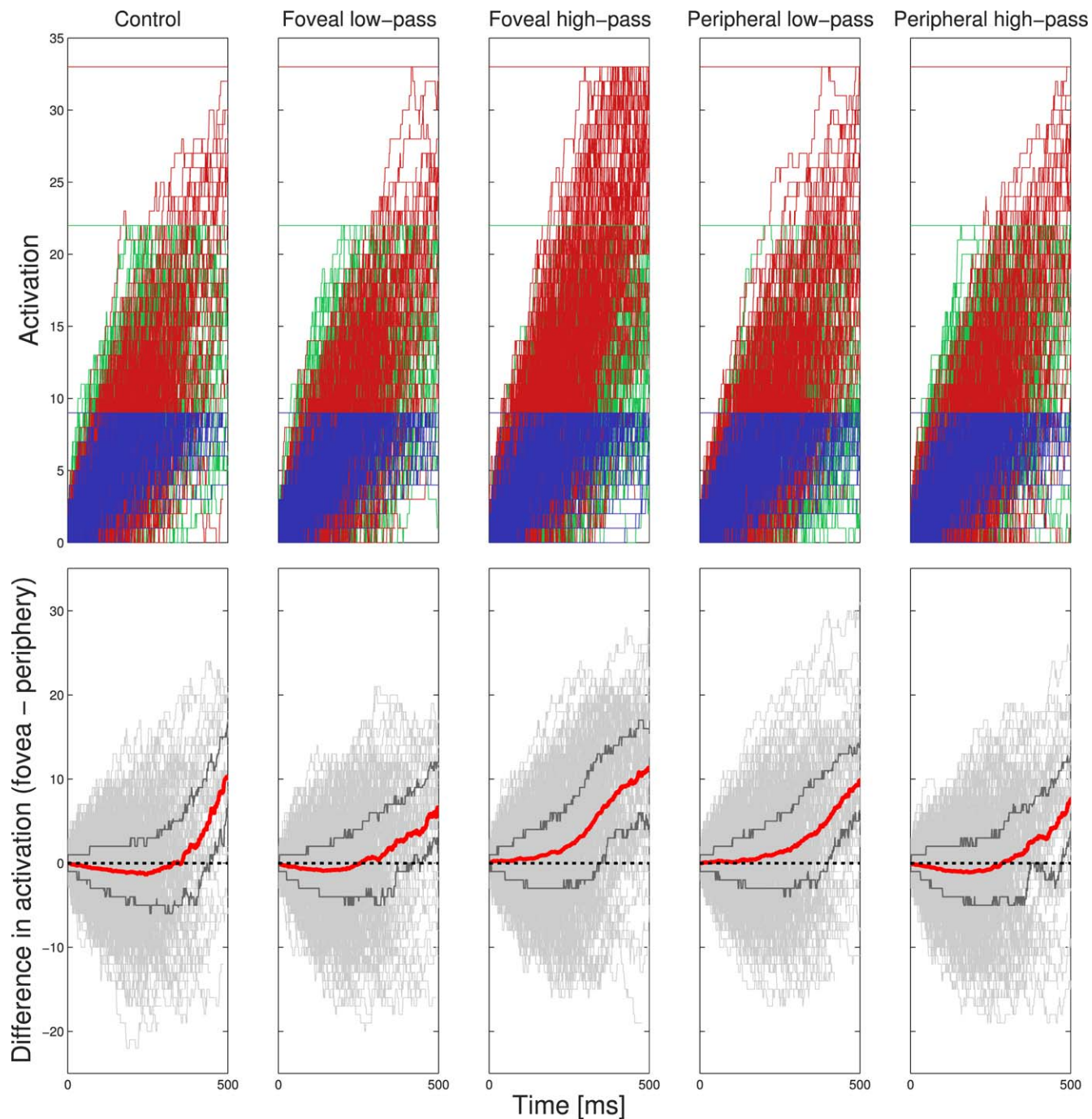


Figure 9. Results of 500 model runs per condition, using the fitted parameter values. (Upper row). Evolutions of peripheral (green), foveal (red), and timer (blue) activations over time for control, foveal low-pass, foveal high-pass, peripheral low-pass, and peripheral high-pass conditions (left to right), with the time axis limited to the first 500 ms. Respective thresholds are marked by horizontal lines. (Lower row). Differences of foveal-peripheral activations per condition over time. The thick red line is the average difference, dark gray lines mark the first and third quartile, and the black horizontal dashed line marks a difference of zero. Positive values indicate periods during which inhibition is active. In the control condition and also in foveal low pass and peripheral high pass, inhibition is initially suppressed by fast rise of peripheral activations. In contrast, in the peripheral low pass and especially the foveal high-pass conditions, inhibition on average starts earlier and lasts longer. For short fixations, inhibition does not play a major role except in the latter two conditions. Long fixations tend to be such because of the contribution of inhibition.

peripheral processing leads to a longer time period during which foveal inhibition is active. Therefore, the parameter estimates suggest that fixation duration during normal scene viewing is often controlled by foveal vision.

Model simulations revealed that spatial frequency filtering affected the transition rates for the foveal and peripheral random walks. More precisely, foveal high-pass filtering affected the transition rate for the foveal compartment in the model, and peripheral low-pass filtering affected the transition rate of the peripheral compartment. Consequently, when the model dynamics change with differing transition rates between the five conditions, the amount of foveal inhibition in these conditions changes as well. Since foveal inhibition was active for a substantial proportion of time in each condition, it is an important mechanism for delaying ongoing fixations in each condition.

Although inhibition is important in all conditions, it makes a stronger contribution in conditions that evoked a larger number of long fixation durations. The model simulations reflect this process in the change of the transition rates. For filter conditions that did not increase mean fixation durations in comparison to the control condition, i.e., foveal low-pass and peripheral high-pass filtering, the model's transition rates were similar to the corresponding rates of the control condition. Thus, the model reproduced the three distributions with unchanged parameter values across conditions and adopted a similar mechanism for saccade timing. Consequently, the mechanism of foveal inhibition was equally important with foveal low-pass filtering, peripheral high-pass filtering, and the control condition. This indicates that the attenuation of critical spatial frequency information did not induce a stronger activation in the fixate center (i.e., higher foveal activation) to take more time for processing the filtered regions, but that the remaining information was not useful enough to make the investment of additional processing time worthwhile.

With foveal high-pass and peripheral low-pass filtering, however, where fixation durations increased significantly compared to the other conditions, the transition rates of the random walks changed markedly. Compared with the rate of the control condition, the foveal transition rate for foveal high-pass filtering increased and was even higher than the peripheral transition rate; for peripheral low-pass filtering, the peripheral transition rate decreased compared to the control condition. Thus, the model adopts two different mechanisms to prolong fixation durations by means of foveal inhibition. With a high foveal transition rate, activation in the foveal compartment builds up faster than activation in the peripheral compartment; with a low peripheral transition rate, peripheral activation builds up slower than foveal activation. Both processes

increase the difference between activations in the foveal and peripheral compartments (see Equation 3), thus strengthening the foveal inhibition process that slows down the growth rate of the saccade timer and delays the end of the ongoing fixation. The model dynamics for foveal high-pass and peripheral low-pass filtering are essentially different from the model dynamics in the other three conditions. Since foveal high-pass filtering still preserves detailed information that the fovea is specialized in processing, although attenuating other information, fixation durations are prolonged, i.e., more processing time is taken. The model thus induces a stronger activation in the fixate center (i.e., higher foveal activation). Since peripheral low-pass filtering, on the other hand, still preserves useful coarse information for saccade target selection in the periphery (although strongly attenuating edge information that might be useful for separating objects from background), additional processing time is taken in this condition as well. The model reflects this process by reducing activation in the move center.

From these results, we can derive a first interpretation of the relation between time course of activation and processing difficulty. In the case of high processing difficulty, the saccade timer needs to be inhibited, i.e., the ongoing fixation is prolonged—at least as long as an investment of additional processing time seems worthwhile. In the model, such a prolongation is achieved by a fast build-up of activation in the foveal compartment, i.e., a high transition rate, or by a slow build-up of activation in the peripheral compartment, i.e., a low transition rate. Because of the push-pull type interaction between peripheral and foveal compartments, a fast build up of activation in the peripheral compartment can compensate the foveal activation and cancel the inhibitory influence from the fovea (disinhibition).

## General discussion

In the present study we investigated the temporal control of eye movements during scene viewing using an experimental manipulation of the spatial frequency content in foveal and peripheral vision and computational modeling. This approach was motivated by the functional segregation of foveal and peripheral vision for high and low spatial frequencies (Banks et al., 1991; Hilz & Cavonius, 1974; Robson & Graham, 1981), and by the lack of knowledge about the foveal and peripheral contributions to the control of fixation durations in scene viewing (but see Trukenbrod & Engbert, 2012).

In the experiment, human observers viewed natural scenes in preparation for a memory test, while high or



low spatial frequencies were gaze-contingently filtered in the foveal or peripheral visual field. Following the common notion that fixation durations lengthen with increased processing difficulty (Henderson, 2003; Rayner, 2009), we expected fixation durations to increase with spatial filtering in any case compared with the unfiltered control condition. Fixations were also expected to increase more strongly when filters attenuated the spatial frequency bands that the region of the visual field is best suited to process (fovea: high spatial frequencies, periphery: low spatial frequencies).

Our experimental findings concerning fixation durations turned out to be in partial opposition to the hypotheses. Foveal high-pass filtering and peripheral low-pass filtering caused longer fixation durations compared with the control condition, but foveal low-pass filtering and peripheral high-pass filtering did not affect fixation durations. Thus, with spatial filters that were assumed to impair information processing the most, the temporal control of eye movements was similar to the control condition. Filter conditions where processing should have been relatively easier because the more useful spatial frequency information was still mostly available did increase fixation durations significantly, however. For saccade amplitudes, we replicated effects obtained in previous studies. Peripheral filtering caused a tunnel-vision effect with a preference for saccade targets inside the tunnel (replicating Foulsham et al., 2011; Loschky et al., 2005; Shioiri & Ikeda, 1989), and foveal filtering caused a preference for saccade targets outside the central scotoma (replicating van Diepen, 2001).

The experimental effects of spatial frequency filtering on fixation durations can neither be explained by the different sensitivity of the foveal and peripheral visual field to certain spatial frequency bands, nor by the predicted increase of fixation durations with higher (foveal) processing difficulty. As a solution, we developed a new computational model based on more complicated, dynamical interactions of foveal and peripheral processing for the control of fixation durations. The model was implemented via temporal activations for (a) a random saccade timer, which generates saccadic commands, and (b) foveal and peripheral compartments, which represent the unfolding of foveal and peripheral processing over time. The interaction of foveal and peripheral activations can modulate the saccade timer by foveal inhibition: Higher foveal compared to peripheral activations reflect a bias for inhibiting the saccade timer and prolonging the current fixation; higher peripheral activations, on the other hand, reflect a bias for disinhibiting the saccade timer. This interaction of activations in the foveal and peripheral compartment resembles the balance between the fixate center and the move center in the model framework by Findlay and Walker (1999).

The simulations demonstrated that the interactive model reproduced both the experimental distributions and the mean values for the fixation durations by changing the foveal and peripheral transition rates between the five filter conditions. The transition rates for foveal low-pass filtering and peripheral high-pass filtering did not differ markedly from the rates of the control condition, indicating a similar mechanism for saccade timing. Thus, foveal inhibition was equally strong in these three conditions, suggesting that critical information was so heavily impaired with foveal low-pass and peripheral high-pass filtering that an investment of more processing time in terms of stronger inhibition of the saccade timer did not seem useful. With foveal high-pass filtering and peripheral low-pass filtering, however, the transition rates differed significantly from the rates of the control condition. With foveal high-pass filtering, foveal processing tended to evolve faster and early after the beginning of a fixation, causing a strong inhibition of the saccade timer. High spatial frequencies that the fovea is specialized in processing were still largely preserved, so that the prolongation of fixations by increasing activation in the fixate center was useful to extract as much information as possible. With peripheral low-pass filtering, peripheral processing tended to evolve slower, which also allows a stronger inhibition of the timer. Since filtering impaired processing of peripheral information, but still preserved useful low spatial frequencies, additional processing time was taken by decreasing activation in the move center. Thus, the model generated increased mean fixation durations with both foveal high-pass and peripheral low-pass filtering by a pronounced inhibition of the saccade timer, but the dynamics that produced this behavior were completely different.

Based on a small set of dynamical rules, the model provided a very good fit to the empirical data. In its current version, however, it also has some limitations. First, although the division into two compartments is compatible with and inspired by the two-streams hypothesis (Goodale & Milner, 1992), we do not yet model how the relative weight of the foveal and peripheral compartment is adjusted. It would be interesting to add a mechanism that controls the transitions rates and thresholds for foveal and peripheral processing. Implementation of such a mechanism, which should also be sensitive to different filter characteristics, is left for future work. Second, the model is limited to temporal aspects of eye-movement control. Although a full model of eye-movement behavior during scene viewing certainly needs to include a mechanism for spatial target selection, we constrained the first version of our computational framework to temporal control in order to obtain a simpler model that might be more easily understood and analyzed. This decision is also justified by the fact



that “where” and “when” decisions in the oculomotor system are largely independent (Findlay & Walker, 1999). There is no reason, however, why our model cannot be integrated with one of the many existing models of spatial selection.

We consider our model a fruitful first step at addressing the contribution of different areas of the visual field to the control of fixation duration during scene viewing. Specifically, both experimental data and computational simulations indicate that not only foveal, but also peripheral vision plays a critical role in regulating fixation duration during scene viewing. While current theoretical models of saccade timing suggest that the large variance of fixation durations is due to the decision of where to move the eyes next (Reddi & Carpenter, 2000), our model proposes that the amount of variance in fixation durations might be partially due to the complicated interaction of foveal and peripheral information processing. The model moves beyond the widespread notion that in the time course of a fixation, peripheral information is usually processed to select the next saccade target after the foveal stimulus has largely been analyzed (Rayner, 2009; van Diepen & d’Ydewalle, 2003). Future experimental and computational studies may delineate the general mechanisms that underlie the interaction between foveal and peripheral processing and the underlying adaptivity for a variety of visual tasks.

*Keywords:* scene perception, spatial frequencies, fixation durations, computational modeling

## Acknowledgments

This work was funded by Deutsche Forschungsgemeinschaft (DFG) by grants LA 2884/1 to J. L. and EN 471/10 to R. E. We thank Petra Schienmann for help with data collection. We are grateful to Casimir Ludwig and an anonymous reviewer for helpful comments.

\* J. L. and A. C. contributed equally to this work. Corresponding author: Jochen Laubrock. Email: laubrock@uni-potsdam.de. Address: Department of Psychology, University of Potsdam, Potsdam, Germany.

## Footnotes

<sup>1</sup>Note that, in the following, the terms “foveal visual field” and “foveal filtering” actually refer to a larger part of the visual field that includes the parafovea (up to about 5° of visual angle).

<sup>2</sup>Note that although unpredicted, this result is not aberrant; we have since replicated it in other experiments with different tasks and filtering parameters (Cajar, Laubrock, & Engbert, 2013).

## References

- Baddeley, R. J., & Tatler, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*, 46, 2824–2833.
- Banks, M. S., Sekuler, A. B., & Anderson, S. J. (1991). Peripheral spatial vision: Limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America A*, 8, 1775–1787.
- Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, 185–207.
- Cajar, A., Laubrock, J., & Engbert, R. (2013). Control of saccade timing by foveal and peripheral spatial frequencies in natural scenes. In K. Holmqvist, F. Mulvey, & R. Johansson (Eds.), *Book of Abstracts of the 17th European Conference on Eye Movements. Journal of Eye Movement Research*, 6(3), 266.
- Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *The Journal of Physiology*, 357, 219–240.
- Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics*, 85, 77–87.
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43, 1035–1045.
- Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of the Sciences, USA*, 103, 7192–7197.
- Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777–813.
- Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. New York, NY: Oxford University Press.
- Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22, 661–721.

- Foulsham, T., Teszka, R., & Kingstone, A. (2011). Saccade control in natural images is shaped by the information visible at fixation: Evidence from asymmetric gaze-contingent windows. *Attention, Perception & Psychophysics*, 73, 266–283.
- Gardiner, C. W. (2004). *Handbook of stochastic methods (3rd Ed.)*. Berlin: Springer.
- Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, 6(9):1, 858–873, <http://www.journalofvision.org/content/13/8/27>, doi:10.1167/6.9.1. [PubMed] [Article]
- Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22, 403–434.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15, 20–25.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.
- Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary eye movement initiation. *Science*, 274, 427–430.
- Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, 9, 394–401.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7, 498–504.
- Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. In G. D. M. Underwood (Ed.), *Eye guidance in reading and scene perception*. (pp. 269–295). Oxford, UK: Elsevier.
- Henderson, J. M., & Smith, T. J. (2009). How are eye fixation controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, 17, 1055–1082.
- Hilz, R., & Cavanaugh, C. R. (1974). Functional organization of the peripheral retina: Sensitivity to periodic stimuli. *Vision Research*, 14, 1333–1337.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.
- Hooze, I. T. C., & Erkelens, C. J. (1996). Control of fixation durations in a simple search task. *Perception & Psychophysics*, 58, 969–976.
- Hooze, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration in visual search. *Vision Research*, 38, 1295–1302.
- Hwang, A. D., Higgins, E. C., & Pomplun, M. (2009). A model of top-down attentional control during visual search in complex scenes. *Journal of Vision*, 9(5):25, 1–18, <http://www.journalofvision.org/content/9/5/25>, doi:10.1167/9.5.25. [PubMed] [Article]
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489–1506.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254–1259.
- Jones, L. A., & Higgins, G. C. (1947). Photographic granularity and graininess. III. Some characteristics of the visual system of importance in the evaluation of graininess and granularity. *Journal of the Optical Society of America*, 37, 217–263.
- Kienzle, W., Franz, M. O., Schölkopf, B., & Wichmann, F. A. (2009). Center-surround patterns emerge as optimal predictors for human saccade targets. *Journal of Vision*, 9(5):7, 1–15, <http://www.journalofvision.org/content/9/5/7>, doi:10.1167/9.5.7. [PubMed] [Article]
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, 36(Suppl.), 14.
- Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied*, 8, 99–117.
- Loschky, L. C., McConkie, G. W., Yang, J., & Miller, M. E. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition*, 12, 1057–1092.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17, 578–586.
- Nuthmann, A. (in press). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*. doi:10.1037/a0033854
- Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research*, 49, 322–336.
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model

- of fixation durations in scene viewing. *Psychological Review*, 117, 382–405.
- Parkhurst, D. J., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual selective attention. *Vision Research*, 42, 107–123.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, 206, 468–469.
- Reddi, B. A. J., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience*, 3, 827–830.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125–157.
- Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology*, 65, 177–206.
- Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21, 409–418.
- Shioiri, S., & Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. *Perception*, 18, 347–361.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27, 161–168.
- Silverman, B. W. (1986). *Density estimation*. London, UK: Chapman and Hall.
- Torralba, A., Oliva, A., Castelhana, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, 113, 766–786.
- Trukenbrod, H. A., & Engbert, R. (2012). Eye movements in a sequential scanning task: Evidence for distributed processing. *Journal of Vision*, 12(1): 5, 1–12, <http://www.journalofvision.org/content/12/1/5>, doi:10.1167/12.1.5. [PubMed] [Article]
- Trukenbrod, H. A., & Engbert, R. (2013). ICAT: A computational model for the adaptive control of fixation durations. Manuscript under review.
- Tsotsos, J. K., Culhane, S. M., Wai, W. Y. K., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual-attention via selective tuning. *Artificial Intelligence*, 78, 507–545.
- van Diepen, P. M. J. (2001). Foveal stimulus degradation during scene perception. In F. H. Columbus (Ed.), *Advances in psychology research* (Vol. 2, pp. 89–115). Huntington, NY: Nova Science Publishers.
- van Diepen, P. M. J., & d'Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition*, 10, 79–100.
- van Diepen, P. M. J., De Graef, P., & d'Ydewalle, G. (1995). Chronometry of foveal information extraction during scene perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 349–362). New York, NY: Elsevier.
- van Diepen, P. M. J., & Wampers, M. (1998). Scene exploration with Fourier-filtered peripheral information. *Perception*, 27, 1141–1151.
- van Kampen, N. G. (1981). *Stochastic processes in physics and chemistry*. Amsterdam, The Netherlands: North-Holland.
- Vitu, F., Lancelin, D., & Marrier d'Unienville, V. (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1220–1249.
- Wertheim, T. (1894). Über die indirekte Sehschärfe [Translation: On indirect visual acuity]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 7, 121–187.
- Wischnewski, M., Belardinelli, A., Schneider, W. X., & Steil, J. J. (2010). Where to look next? Combining static and dynamic proto-objects in a TVA-based model of visual attention. *Cognitive Computation*, 2, 326–343.

## Appendix

### Estimation of the model parameters

A genetic algorithm (Goldberg, 1989; Holland, 1975) was used to estimate the model parameters. We started with 50 combinations (i.e., individuals) of random parameter values selected from a predefined range that was restricted by mathematical constraints or by neurophysiological plausibility (see Table 3). Using selection, mutation, and crossover, individuals were evolved over several thousand generations. After

convergence of the population, we used the best 50 individuals of the population from each iteration.

### **Quantile maximum likelihood**

Quantile maximum likelihood estimation (Heathcote, Brown, & Mewhort, 2002) was used as a fitting procedure for the fixation duration distributions. Every experimental distribution was divided into 20 quantiles that each covered 5% of the data. For each parameter combination, simulated fixations for the five distribu-

tions were tested for their probability to fall in each of the experimental quantile boundaries (a probability of 0.05 was considered to be a perfect fit). The log likelihood of the data given the parameter values was then computed from the logarithmized probabilities for fixations to fall in each of the 20 quantiles. The 50 individuals with the largest log likelihood were chosen from the population. The source code for the computational model is available at the Potsdam Mind Research Repository (PMR2), <http://read.psych.uni-potsdam.de/pmr2>.