



## Influence of initial fixation position in scene viewing



Lars O.M. Rothkegel<sup>a,\*</sup>, Hans A. Trukenbrod<sup>a</sup>, Heiko H. Schütt<sup>a,b</sup>, Felix A. Wichmann<sup>b,c,d</sup>, Ralf Engbert<sup>a</sup>

<sup>a</sup> Department of Psychology & Cognitive Science Program, University of Potsdam, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany

<sup>b</sup> University of Tübingen, Faculty of Science, Department of Computer Science, Neural Information Processing Group, Sand 6, 72076 Tübingen, Germany

<sup>c</sup> Bernstein Center for Computational Neurosciences, Tübingen, Germany

<sup>d</sup> Max Planck Institute for Intelligent Systems, Tübingen, Germany

### ARTICLE INFO

#### Article history:

Received 5 January 2016

Received in revised form 30 July 2016

Accepted 16 September 2016

Available online 1 November 2016

#### Keywords:

Visual scanpath

Visual attention

Inhibition of return

Eye movements

Saliency

### ABSTRACT

During scene perception our eyes generate complex sequences of fixations. Predictors of fixation locations are bottom-up factors such as luminance contrast, top-down factors like viewing instruction, and systematic biases, e.g., the tendency to place fixations near the center of an image. However, comparatively little is known about the dynamics of scanpaths after experimental manipulation of specific fixation locations. Here we investigate the influence of initial fixation position on subsequent eye-movement behavior on an image. We presented 64 colored photographs to participants who started their scanpaths from one of two experimentally controlled positions in the right or left part of an image. Additionally, we used computational models to predict the images' fixation locations and classified them as balanced images or images with high conspicuity on either the left or right side of a picture. The manipulation of the starting position influenced viewing behavior for several seconds and produced a tendency to overshoot to the image side opposite to the starting position. Possible mechanisms for the generation of this overshoot were investigated using numerical simulations of statistical and dynamical models. Our model comparisons show that inhibitory tagging is a viable mechanism for dynamical planning of scanpaths.

© 2016 Elsevier Ltd. All rights reserved.

### 1. Introduction

An important problem for research on human vision is to predict where people look in visual scenes (Tatler & Vincent, 2008). Recording of eye movements is among the most important tools to investigate how attention is distributed over a given scene (Findlay & Gilchrist, 2003). In addition to scene content (Henderson, 2003), image-independent viewing strategies exist, e.g., the central fixation tendency (Tatler, 2007) as the most important effect in this category. To obtain a deeper understanding about dynamical aspects of the attention distribution over a scene and possible dependencies between successive fixations, we investigate the influence of the eye's starting position on subsequent viewing behavior based on statistical and dynamical assumptions about eye guidance.

Processes that influence the selection of upcoming saccade targets can be divided into three different categories of theoretical principles. *Bottom-up processes* derive from properties of the viewed stimulus (Itti, Koch, & Niebur, 1998; Mannan, Ruddock, & Wooding, 1996; Parkhurst, Law, & Niebur, 2002). *Top-down*

*processes* depend on the mental state of an observer, e.g., the observers' visual memory (Henderson & Hollingworth, 2003) or the instruction given to the observer before inspection of a scene (Castelhano, Mack, & Henderson, 2009; Yarbus, Haigh, & Riggs, 1967). Finally, *systematic tendencies* describe eye movement behavior found in many experiments independent of stimulus and observer. The initial selection of the center of an image (Bindemann, 2010; Tatler, 2007), the tendency to make initial movements in the leftward direction (Dickinson & Intraub, 2009; Foulsham, Gray, Nasiopoulos, & Kingstone, 2013; Ossandón, Onat, & König, 2014) or the preference for horizontal and vertical over oblique saccades relative to the image (Foulsham & Kingstone, 2010) belong to this category.

Research on bottom-up processes has been particularly popular to predict fixation locations from low-level image features such as contrast, orientation and color (Itti et al., 1998; Kienzle, Wichmann, Franz, & Schölkopf, 2006; Torralba, 2003). For a given scene, computational models generate a *saliency map*, a 2D probability distribution that indicates the probability of receiving a fixation in an eye tracking experiment with human participants (Borji & Itti, 2013; Itti et al., 1998; Itti & Koch, 2000; Judd, Durand, & Torralba, 2012). Thus, a saliency map is a stationary model that computes probabilities for all locations simultaneously.

\* Corresponding author.

E-mail address: [lrothkeg@uni-potsdam.de](mailto:lrothkeg@uni-potsdam.de) (L.O.M. Rothkegel).

However, current computational models for the prediction of fixation locations are not exclusively based on bottom-up features. Recent models incorporate top-down processes like task demands (Navalpakkam & Itti, 2005) and other higher-level image features like face processing (Cerf, Harel, Einhäuser, & Koch, 2008). Moreover, systematic tendencies such as the central fixation bias (Tatler, 2007) are included in the computation of fixation density models. As a result, current models integrate multiple features from all three categories of processes into a coherent computational framework (Cerf et al., 2008; Judd, Ehinger, Durand, & Torralba, 2009; Kümmerer, Wallis, & Bethge, 2015). Although the original meaning of *saliency* refers to the bottom-up features of an image, newer computational models that include other features are also termed saliency models by their authors (Judd et al., 2009; Bylinskii et al., 2015). Because of this unclear terminology we will refer to all stationary models that aim at the prediction of fixation locations as fixation density models. A location that a model tags as likely to receive a fixation will be referred to as conspicuous rather than salient.

All fixation density models need to predict the density of the eye's fixation locations (so-called first-order statistics). Thus, the evaluation of the models is primarily based on the assumption of statistically independent fixations without reference to previous fixations, i.e., the scanpath (Kümmerer et al., 2015). In contrast to static models, dynamic models try to capture some additional aspects of the scanpath. Dynamical principles for saccade planning are *inhibitory tagging* (Bays & Husain, 2012; Itti et al., 1998; Klein, 1988; Le Meur & Liu, 2015), *saccadic momentum* (Smith & Henderson, 2009, 2011; Wilming, Harst, Schmidt, & König, 2013) and *facilitation of return* (Luke, Schmidt, & Henderson, 2013; Smith & Henderson, 2009; Smith & Henderson, 2011).

Inhibitory tagging is motivated by the effect of inhibition of return, a neural mechanism that inhibits the processing at recently attended locations (Klein, 2000; Posner & Cohen, 1984; Posner, Rafal, Choate, & Vaughan, 1985) and is often interpreted as a foraging facilitator. While this mechanism was first discovered as an effect on a temporal scale, i.e., increased processing time at a previously attended stimulus for a specific time window, inhibition of return might carry over to spatial effects. In the case of spatial inhibition of return, recently fixated positions are inhibited from being re-fixated shortly afterwards (Gilchrist & Harvey, 2000). Several studies were unable to report evidence for inhibition of return during scene viewing; quite the contrary, a facilitation of return saccades to currently fixated locations has been found (Smith & Henderson, 2009; Smith & Henderson, 2011; Wilming et al., 2013).

However, compared to a statistical baseline model without memory based on inhibitory tagging, return saccades occur less often in experiments than expected (Bays & Husain, 2012), when the density map of fixations and the distribution of angles between two subsequent saccades are reproduced. Therefore, there is at least weak support for a memory-producing mechanism during scene exploration. In agreement with this result, we recently published a computational model of saccade generation in scene viewing that implements both inhibitory tagging and dynamical attention mechanisms (Engbert, Trukenbrod, Barthelmé, & Wichmann, 2015). In this model inhibitory tagging is combined with a dynamical activation map representing attention allocation, allowing the model to reproduce second-order statistics that include spatial correlation functions characterizing the clustering of fixations in a scanpath in addition to the first-order density of fixations. Thus, inhibitory tagging seems to be important to reproduce higher-order scanpath statistics (Engbert et al., 2015), despite the current lack of direct experimental support for inhibition of return in scene viewing (Luke et al., 2013; Smith & Henderson, 2009, 2011).

Saccadic momentum, another dynamical principle of saccade planning in scene viewing, describes the tendency to maintain the direction of the previous saccade for the upcoming saccade (Smith & Henderson, 2009, 2011; Wilming et al., 2013). Similar to inhibition of return, saccadic momentum could serve as a foraging facilitator in visual search. Finally, facilitation of return describes the tendency that it is actually more likely to produce return saccades than it would be by chance (Hooge, Over, van Wezel, & Frens, 2005; Smith & Henderson, 2009). On the time scale of one fixation duration (~ 300 ms), such a facilitation seems to be in contradiction to spatial inhibitory tagging. Because of these behaviorally relevant processes, we were interested to find experimental support for the presence of *inhibitory tagging*, *saccadic momentum*, *facilitation of return* or a mixture of these fundamental principles in attentional and oculomotor control.

Smith and Henderson (2009) ruled out inhibitory tagging, since they found an increased number of return saccades in comparison to a probabilistic baseline (Smith & Henderson, 2009). However, it has also been argued that there is a reduced number of return saccades compared to a memoryless system (Bays & Husain, 2012). Given the current mixed evidence on return saccades, we focus on the time window of events. Return saccades are limited to a time window of one fixation duration, i.e., about 300 ms. Since attention moves to the future fixation location before a saccade is executed (Deubel & Schneider, 1996), inhibition of return is at its maximum shortly after the saccade is planned if we assume that the typical time-course transfers to scene viewing (Klein, 2000; Posner & Cohen, 1984). However, first, it would not be surprising to find that more time than a single fixation duration is needed to build-up spatial inhibition. Second, return saccades might be planned before the inhibition of return mechanism is activated, so that saccades to previously inspected image regions could be produced while inhibition is on the rise. Third, it has been reported that the time scale of IOR is dependent on task difficulty (Klein, 2000). Therefore, the current lack of direct evidence for inhibition of return does not rule out inhibitory tagging as a saccade-planning mechanism.

To investigate inhibitory tagging, saccadic momentum, and facilitation of return, we recorded observers' scanpaths on natural scenes starting from one of two predefined starting positions close to either side of the monitor. Participants were forced to maintain fixation at an initial location in an image for one second under gaze-contingent monitoring. Under the hypothesis that spatial inhibitory tagging is active at the starting position, we expected observers (i) to leave their starting positions when fixation markers disappeared, and (ii) not to return immediately to the region of the experimentally controlled starting position. Since we hypothesized that both behaviors depend on the conspicuity of the region of the starting position, we classified natural images into three categories with left-sided and right-sided conspicuity asymmetry as well as images with an approximately symmetrical distribution. First, we expected that initial fixations stay closer to the starting position when the starting position was in interesting side of a scene; second, gaze was expected to move immediately to the opposite side of a scene, when the starting position was opposite to the scenes interesting side. Third, according to the saccadic momentum and facilitation of return hypothesis, we expected a behavior where subsequent eye movements depend on the direction of the first saccade. With the typical center bias we assume that the gaze had to shift to the center and, subsequently, either maintain direction and move to the opposite image side (saccadic momentum) or return close to the starting position (facilitation of return).

Below we report that gaze positions of the participants moved further away from the starting position than predicted by the empirical fixation map or a saccadic momentum mechanism. Next,

Download English Version:

<https://daneshyari.com/en/article/5705956>

Download Persian Version:

<https://daneshyari.com/article/5705956>

[Daneshyari.com](https://daneshyari.com)