Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 46 (2006) 1857–1862
Vision Research
www.elsevier.com/locate/visres

Brief communication

The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task

Benjamin W. Tatler [a,*], Roland J. Baddeley [b], Benjamin T. Vincent [b]

[a] *Department of Psychology, University of Dundee, Dundee DD1 4HN, Scotland, UK*
[b] *Department of Experimental Psychology, University of Bristol 8, Woodland Road, Bristol BS8 1TN, UK*

Received 13 June 2005; received in revised form 15 November 2005

Abstract

We recorded over 90,000 saccades while observers viewed a diverse collection of natural images and measured low level visual features at fixation. The features that discriminated between where observers fixated and where they did not varied considerably with task, and the length of the preceding saccade. Short saccades (<8°) are image feature dependent, long are less so. For free viewing, short saccades target high frequency information, long saccades are scale-invariant. When searching for luminance targets, saccades of all lengths are scale-invariant. We argue that models of saccade behaviour must account not only for task but also for saccade length and that long and short saccades are targeted differently.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Eye movement; Salience; Image features; Natural images; Saccade amplitude; Task

1. Introduction

When viewing complex scenes, we are highly selective in the locations that we choose to fixate. While it is clear that the task at hand is important in determining the locations we choose to fixate (Buswell, 1935; Nelson, Cottrell, Movellan, & Sereno, 2004; Yarbus, 1967), low level visual features can also influence eye movements (Findlay, 1981, 1997; Zelinsky, Rao, Hayhoe, & Ballard, 1997), and are likely to play a role in selection even if fixation location choice is dominated by high level factors (Tatler, Baddeley, & Gilchrist, 2005).

A recent framework for investigating the factors involved in saccade targeting is that of the *salience map* (Itti & Koch, 2000; Kadir & Brady, 2001; Koch & Ullman, 1985; Parkhurst & Niebur, 2003; Renninger, Coughlan, & Vergheese, 2005). These authors suggest that a spatial map of the *salience* of potential fixation locations is constructed by combining multiple low level visual features at many spatial scales. One natural question in this framework is whether the visual salience (in terms of a given set of low level features) differs between fixated and non-fixated locations (Parkhurst, Law, & Niebur, 2002; Tatler et al., 2005). Such studies have shown that the visual features at locations selected for foveation differ statistically from those at randomly selected locations. It appears that low level visual features are on average more extreme at fixated than non-fixated locations, and these differences tend to be larger for edges and contrast than luminance and colour. Furthermore, the largest differences between fixated and non-fixated locations is for high frequency information, and this can be interpreted as reflecting a dominance of high frequency information in saccade target selection (Tatler et al., 2005). It is this final point, and its interaction with task, that we will consider in more detail in the present report.

A system in which saccade target selection is dominated by high frequency information in scenes does encounter an obvious problem. Selection of the target to fixate must have occurred prior to the initiation of the saccade that brought the fovea to bear on this location and therefore was selected

* Corresponding author. Fax: +44 0 1382 229993.
E-mail address: b.w.tatler@dundee.ac.uk (B.W. Tatler).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.12.005

using peripheral vision. Visual acuity declines with increasing eccentricity in the retina (Østerberg, 1935). Thus we are faced with the question of how selection can be driven by high frequency information when it is performed by low resolution peripheral vision.

The decline in visual acuity with retinal eccentricity suggests that it may be important to consider the eccentricity of a location when it is selected for fixation. In their recent study, Tatler et al. (2005) found that the largest difference between fixated and non-fixated locations was for spatial scales of information as high as 10.8 cycles per degree (cpd) when viewing natural images. Given estimates of how the human modulation transfer function changes with eccentricity (Rovamo & Virsu, 1979) and assuming a central visual acuity of 60 cpd (Curcio, Sloan, Kalina, & Hendrickson, 1990) we can calculate that from approximately 14° from the centre of fixation, information at scales of 10.8 cpd or higher should not be resolved. However, a significant proportion of saccades are made to locations in excess of 14° when viewing complex scenes or during real world tasks (see, Land & Hayhoe, 2001). Logically, the high frequency information should not dominate this subset of saccades.

In the present study, we analysed a total of 68,983 eye movements as observers viewed images of complex real world scenes. We then constructed salience maps of the same visual features explored by Tatler et al. (2005) at six spatial scales. The same approach for comparing salience at fixation and at randomly selected locations was employed to speculate upon the relative contribution of different spatial scales of information in target selection. However, in this study we also considered the amplitude of the saccade that had brought the fovea to the fixated location. In this way, we were able to assess the interaction between saccade amplitude and the scale of information present at fixation and thus make inferences about their possible involvement in saccade target selection.

While it has long been understood that task has a strong influence upon the distribution of fixations on complex scenes (e.g., Buswell, 1935; Yarbus, 1967), the influence of task upon any salience map framework that might underlie saccade targeting still remains unclear (see, Tatler et al., 2005). It may be that there are common salience-based targeting mechanisms underlying different tasks (this is possible even in the face of different spatial *locations* being chosen for different tasks), or it may be that the task constraints alter salience-based criteria for saccade target selection. We are by no means the first to consider the limitations of salience-based models of fixation behaviour under varying task conditions. Using an information theoretic approach to model scene statistics at the centre of gaze Krieger, Rentschler, Hauske, Schill, and Zetzsche (2000) highlighted the need for a unitary model of fixation behaviour that integrated high and low level factors; a goal toward which they have since been working (e.g., Schill, Umkehrer, Beinlich, Krieger, & Zetzsche, 2001). Raj, Geisler, Frazor, and Bovik (2005) proposed a model based upon minimising contrast entropy, which they suggested was adequate for certain tasks (in which the observer must gain as much information about the structure of a scene as possible) but would not generalise to all tasks. More explicit models of the role of task in salience-based approaches have been proposed recently. Torralba and colleagues (e.g. Torralba, 2001, 2003; Torralba, Murphy, & Freeman, 2005) have suggested a specific Bayesian framework in which low level salience maps are spatially weighted depending on the most probable location of target objects. Navalpakkam and Itti (2005) have suggested that the high level component is manifest as a bias toward particular features (or feature conjunctions) in the salience map framework, effectively weighting particular feature channels over others.

Given the body of evidence for top down effects in the control of fixation behaviour, we decided to include a manipulation of task within our exploration of saccade length effects on fixation selection. We collected eye movement data under two different task situations. In the first, participants were merely asked to look at the images freely (free viewing). In the second, they were given a search task in which they had to search for a small, localised artificial increase in brightness at a random location in the image: specifically a Gaussian luminance bump that had been added to 50% of the images (search task). For the second task only images in which the target was absent were analysed for this study. Thus, the stimuli viewed under the two task constraints in this report were identical. In this way, we were able to assess whether selection criteria for saccades of various amplitudes varied according to the task.

2. Method

In the free viewing task, 22 participants aged 18 to 29 years (mean = 21.7, $SD = 3.2$) viewed 120 photographic images of real world scenes. In the search task, 30 participants aged 18 to 53 years (mean = 22.9, $SD = 6.6$) viewed the same 120 images. In this task, half of the images had a small ($SD = 0.3°$) Gaussian brightness blip added in a random location. The task was to decide whether there was a brightness blip present and to respond using a button box.

Images were recorded using a Nikon D2 digital SLR using the highest resolution (4 megapixels). Images were displayed in 1600 × 1200 pixel format on a 21 in. SVGA colour monitor with a refresh rate of 100 Hz and a maximum luminance of 55 cd m^{-2}. The monitor was positioned at a viewing distance of 60 cm; consequently, the images presented subtended 40° horizontally and 30° vertically. Each trial was preceded by a fixation target positioned randomly within 10° of the centre of the screen before displaying the image for 5 s.

Eye movements were recorded during viewing using the SR Research. EyeLink II eye tracker, which samples eye position data at 500 Hz. Eye position data were collected binocularly and analysed for the eye that produced the better spatial accuracy as determined using the calibration. Nine-point target displays were used for calibration and validation of eye position. Saccade detection required a deflection of greater than 0.1°, with a minimum velocity of 35° s^{-1} and a minimum acceleration of 9500° s^{-2}, maintained for at least 4 ms. We used a minimum fixation duration of 50 ms.

Using this procedure, data were collected for 40,011 saccades in the free view task and for 55,170 saccades in the search task. We only analysed eye movements made when viewing images in which the search target was absent in order to ensure that the stimuli viewed in the two tasks were identical. This resulted in 28,972 being available for analysis for the search task.

Image features were made explicit using the same procedures as detailed in Tatler et al. (2005) for luminance, contrast and edge information. The only departure from the image feature extraction methodology is in the spa-

Download English Version:

https://daneshyari.com/en/article/4036277

Download Persian Version:

https://daneshyari.com/article/4036277

Daneshyari.com