Title
On the Bottom-Up and Top-Down Influences of Eye Movements

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
de Boer, Bart
Koostra, Gert
Nederveen, Arco

Publication Date
2006

Peer reviewed

On The Bottom-Up and Top-Down Influences of Eye Movements

Gert Kootstra (g.kootstra@ai.rug.nl)
Arco Nederveen (a.nederveen@ai.rug.nl)
Bart de Boer (b.de.boer@ai.rug.nl)
Artificial Intelligence,University of Groningen
Grote Kruisstraat 2/1
9712 TS, Groningen, The Netherlands

Keywords: Eye movements; visual attention; saliency.

Introduction

To cope with the enormous amount of visual information in our everyday environment, the human visual system uses a mechanism of visual attention and saccadic eye movements to filter and process only the relevant information. In this study, we try to analyze and model the control of these eye movements.

Eye movements are controlled by bottom-up and top-down mechanisms. The role of these two mechanisms in the control of the eye movements is thought to depend on the amount of semantic content in the image, as suggested by Parkhurst and Ernst (2003). Images with high semantic content are likely to be more top-down controlled, since an internal model of the content exists (Parkhurst *et al.*, 2002). The inspection of only a small part of the image in cooperation with the internal model provides the observer with information about what to expect in the rest of the image. We call these images *high-expectation* images, in contrast to *low-expectation* images, which are unfamiliar and have little meaning to the observer. In low-expectation images, stimulus-driven or bottom-up mechanisms are supposed to play a more prominent role.

In this study, we investigate the differences in eye fixation patterns on low- and high-expectation images. We expect that fixations on low-expectation images are more salient and are better predicted by bottom-up models.

The goal of our study is twofold. On the one hand, we hope to get more insight in the process of visual attention and the role of top-down and bottom-up mechanisms. On the other hand, models of filtering relevant visual information can be an important contribution to object recognition in computer vision and robotics.

Experiment

We performed an experiment where participants were shown images of both types, while their gaze was recorded with an Eyelink eye tracker. We tested 43 participants. The set of low-expectation images consisted of unfamiliar images with natural content, while the set of high-expectation images consisted of highly familiar everyday objects. Both sets contained 10 different images and we included a mirrored version of each image. The elements of both image sets were shown in random order and the participants viewed each image twice for a period of 5 seconds. The participants were given no specific task, that is, it was a free-viewing experiment.

Method

We propose to analyze the rich amount of obtained data in two ways. First, we will analyze the content of the human eye fixations with statistical descriptors, e.g., local contrast, spatial correlation and entropy, similar to the study of Parkhurst and Ernst (2003).

Second, we will compare bottom-up prediction models with the human data. Among others, we will use the saliency model of Itti, Koch and Niebur (1998), the SIFT keypoint detector (Lowe, 2004), which is a popular model in computer vision, and a number of elementary interest-point detectors used by Privitera and Stark (2000). We will follow Le Meur *et al.* (2006) by comparing the fixation density maps using statistical correlation and the Kullback-Leibler divergence. We will investigate whether the models better predict eye fixations on the low-expectation images.

Discussion

A first comparison with the SIFT keypoint detector (Lowe, 2004), shows a significantly better than random prediction of fixation patterns. Although it is too premature to draw conclusions, our proposed method shows promising results for shedding more light upon the role of bottom-up and top-down mechanisms in the control of eye movements.

References

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*(11), 1254-1259.

Le Meur, O., Le Callet, P., Barba, D., & Thoreau, D. (2006). A Coherent Computational Approach to Model Bottom-Up Visual Attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(5), 802-817.

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision, 60*(2), 91-110.

Parkhurst, D. J., & Ernst, N. (2003). Scene Content Selected by Active Vision. *Spatial Vision, 16*(2), 125-154.

Parkhurst, D. J., Law, K., & Niebur, E. (2002). Modeling the Role of Salience in the Allocation of Overt Visual Attention. *Vision Research, 42*, 107-123.

Privitera, C. M., & Stark, L. W. (2000). Algorithms for Defining Visual Regions-of-Interest: Comparison woth Eye Fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(9), 970-982.