Computers in Biology and Medicine 42 (2012) 235–244

Contents lists available at SciVerse ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Evaluating gaze control on a multi-target sequencing task: The distribution of fixations is evidence of exploration optimisation

Giacomo Veneri [a,b,c,*], Francesca Rosini [b], Pamela Federighi [a,b], Antonio Federico [b], Alessandra Rufa [a,b,*]

[a] Eye tracking & Visual Applications Lab, University of Siena, Italy
[b] Department of Neurological Neurosurgical and Behavioral Science, University of Siena, Italy
[c] Etruria Innovazione Spa, Italy

ARTICLE INFO

Article history:
Received 19 November 2010
Accepted 28 November 2011

Keywords:
Eye tracking
Trial Making Test
Fixations
Saccades
Attention

ABSTRACT

Many high cognitive applications, such as vision processing and representation and understanding of images, often need to analyse in detail how an ongoing visual search was performed in a representative subset of the image, which may be arranged into sequences of loci, called regions of interest (ROIs). We used the Trial Making Test (TMT) in which subjects are asked to fixate a sequence of letters and numbers in a logical alphanumeric order. The main characteristic of TMT is to force the subject to perform a default and well-known path. The comparison of the expected scan-path with the observed scan-path provides a valuable method to investigate how a task force the subject to maintain a top-down internal representation of execution and how bottom-up influences the performance. We developed a mechanism that analyses the scan path using different algorithms, and we compared it with other methods: we found that fixations outside the ROI are direct influence of exploration strategy. The paper discusses the method in healthy subjects.

1. Introduction

Recent studies are focused on methods and models that evaluate how humans explore real scenes in a naturalistic approach and how machine vision should emulate human searching. In the real world, visual search is a common task that enables humans to explore the environment and direct attention towards regions of interest. In experimental settings in which neurophysiological and cognitive functions are investigated, quantification methods for visual search are used to evaluate the allocation of attention during scene viewing [1, for a short review].

In order to understand how mechanisms drive attention during visual exploration, many studies in which image characteristics are manipulated have been conducted. The main concept is that particular regions of interest in the scene are selected assuming their cognitive relevance or local image's saliency. When image salience is thought to guide visual search, the mechanism is called bottom-up. Conversely, when mechanisms driving visual search depend more on human intention, they are called top-down. These two exemplifications describe the main theories of visual search, although endogenous and exogenous components presumably work together in normal circumstances. By considering this dichotomy, a variety of formal models have been proposed on the last decade, in order to describe the attentional selection mechanism: Feature Integration Theories [2, FIT], Guided Search [3], Theory of Visual Attention [4,5, TVA] and a new purely bottom-up model Winner Take All [6, WTA]. It has been suggested that early selection stages are purely driven by the image saliency factors in a bottom-up prevalence, and later selection is due to the combination of top-down and bottom-up factors. A key debate in this literature is whether bottom-up can override top-down and vice versa. For instance Theeuwes [7] and Daniel Schreij [8] found that the appearance of distracters reduced search efficiency, presumably due to the involuntary capture of attention. Conversely, Chen [9] found that visual search on the real word is dominated by the top-down mechanism.

The focus of recent researches, however, has shifted to how these processes should be combined and their relative contributions to search guidance, how top-down and bottom-up work together to perform an efficient visual exploration and how they interfere with each other (see [10], for a review); therefore, the research is directed towards unified methods of analysis that may better reflect real conditions.

In our research, we aimed to investigate how the mechanism of bottom-up and top-down works together on neuro-psychological context during the ongoing visual search. We aimed to use a task which encourages free exploration, avoiding too high saliency

* Corresponding authors at: Department of Neurological Neurosurgical and Behavioral Science, University of Siena, Viale Bracci 2, 53100 Siena, Italy. Tel.: +39 0577 233136; fax: +39 0577 40327.
E-mail addresses: g.veneri@unisi.it (G. Veneri), visionlab@unisi.it (F. Rosini), federighi@unisi.it (P. Federighi), federico@unisi.it (A. Federico), rufa@unisi.it (A. Rufa).

0010-4825/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2011.11.013