Saliency predicts change detection in pictures of natural scenes

MICHAEL J. WRIGHT*
Centre for Cognition and Neuroimaging, Brunel University, Uxbridge, Middlesex UB8 3PH, UK

Received 26 June 2004; accepted 15 November 2004

Abstract—It has been proposed that the visual system encodes the salience of objects in the visual field in an explicit two-dimensional map that guides visual selective attention. Experiments were conducted to determine whether salience measurements applied to regions of pictures of outdoor scenes could predict the detection of changes in those regions. To obtain a quantitative measure of change detection, observers located changes in pairs of colour pictures presented across an inter-stimulus interval (ISI). Salience measurements were then obtained from different observers for image change regions using three independent methods, and all were positively correlated with change detection. Factor analysis extracted a single saliency factor that accounted for 62% of the variance contained in the four measures. Finally, estimates of the magnitude of the image change in each picture pair were obtained, using nine separate visual filters representing low-level vision features (luminance, colour, spatial frequency, orientation, edge density). None of the feature outputs was significantly associated with change detection or saliency. On the other hand it was shown that high-level (structural) properties of the changed region were related to saliency and to change detection: objects were more salient than shadows and more detectable when changed.

Keywords: Change blindness; attention; saliency; natural scenes.

1. GENERAL INTRODUCTION

‘Change blindness’ experiments (Simons and Levin, 1998; Simons, 2000) offer a method of investigating the availability of an observer’s representation of a scene. Observers may fail to notice changes in a scene when these occur across a blank ISI or other visual transient (Rensink et al., 1997). If a change is detectable across an ISI, this implies that it has been represented in memory. Certain objects in a scene may have priority of encoding in order to guide attention and behaviour. The following experiments address two related questions about change blindness and change detection. First, does the salience of an image region determine change

*E-mail: Michael.Wright@brunel.ac.uk
detection, and second, what determines saliency: low-level or high-level properties? Low-level properties are those that reflect early visual processing, for example, spatiotemporal luminance contrast, and chromatic or textural changes. High-level properties include the both structural properties of objects and scenes, and semantic properties (Yarbus, 1967). An unresolved issue is the extent to which the detection of changes is predictable from the magnitude of low-level changes in the image. In a psychophysical analogue of change blindness experiments, Wright et al. (2000) and Wright et al. (2002) have shown that for arrays of Gabor targets, the probability of detection of a change depended on the spatial frequency difference, and this in turn is related to the spatiotemporal Fourier energy in the image pair. The slope of the psychometric function depended on the number of Gabor targets. Moreover, Olds and Engel (1998) have demonstrated linearity in object recognition in images of common objects suggesting a dependence on spatiotemporal contrast sensitivity. The recognition of spatially filtered images and their combinations showed that different spatial frequency bands combine linearly in predicting the recognition of the compound image. The responses of sets of spatial frequency filters to the difference in a pair of images can represent early visual response to the image change, but focal attention and comparison in memory may be necessary to overcome the spatiotemporal masking effect of transients in change-blindness stimuli (Lakha and Wright, 2004).

2. DECISION PROCESSES FOR ONE-SHOT CHANGE LOCALISATION

Change detection may be studied using single (one-shot) or repeated (flicker) presentation of image pairs. In the ‘flicker paradigm’ (Rensink, 2002) a pair of images that contain a large difference are cycled repeatedly with a blank interval between each presentation. Typically, observers search for the change but completely fail to see it until it is suddenly noticed. Thereafter, it becomes almost impossible not to see the change. This seems to imply that there is a threshold for seeing the change that may be crossed only when attention is given to the appropriate location and image property. The unattended detection threshold must be very much higher than the attended sensory detection threshold for the change. This would predict that changes are not reported unless the change is explicitly ‘seen’. Thus, false positive change responses should be relatively rare.

Rensink (2002) however argued that change detection can operate on an explicit, semi-explicit or implicit level. Explicit responses occur when the change is consciously seen, and are most commonly tested in a ‘yes/no’ task. Semi-explicit responses are triggered by the awareness that a change has occurred, but the change is not consciously seen, so that identification of the changed object is not possible. Instead of regarding explicit and semi-explicit responses as representative of different processes, it may be possible to explain both by a signal detection theory (SDT) approach. It may then not be necessary to invoke different levels of conscious and unconscious processing. For example, we could apply a standard SDT model with