Gaze behaviour during a challenging visual search task

James Mardell &lt;james.mardell@imperial.ac.uk&gt;, Mark Witkowski and Robert Spence, Imperial College London

Abstract

We would be happy to discuss the surprising results we have obtained in the context of an extremely challenging visual search task, as well as tentative interpretations we have thereby made concerning the remarkably adaptive nature of the human visual system. Eye-gaze performance has been one of our primary investigative tools.

The Task

If someone is lost in the wilderness it is common to over-fly the area judged most likely to contain that lost person with an Unattended Aerial Vehicle (UAV) fitted with a downward pointing video camera, the output of which is transmitted to a ground station. Here, an observer inspects the live video footage to locate the lost individual. But this is only one way in which the video footage can be inspected: the unanswered question is “what is the best method of presenting the video footage to an observer in order to enhance the likelihood that a lost person can be found?”

Experiment

In a revealing experiment, a simulated video feed from a UAV was converted to a sequence of fifteen static views of the overflown terrain, each view being available for inspection for about 4 seconds. Each view was then segmented into 4 smaller images, each available for inspection for about 1 second in order for the imagery conveyed to the observer to remain ‘live’. Further segmentation was explored, in the limit leading to an observer having sight of \( \frac{1}{6} \)th of each of the original fifteen views, but only for 108 ms. The task performed by subjects was to indicate, by space-bar press as soon as possible, when a human being was identified in any of the terrain images.

Surprisingly, the success of target identification was not significantly affected by segmentation: considering that the targets were hard to spot, a recognition success of about 55% is quite remarkable. Eye gaze records suggest that, before segmentation, gaze behaviour was composed of the conventional sequence of fixations and saccades whereas, at the extreme where \( \frac{1}{5} \)th of each of the original fifteen images was visible for only 108 ms, pre-attentive processing was dominant.

One outcome of our general examination of gaze behaviour was the proposal of gaze-enhanced presentation modes in which the occurrence of ‘target following’ gaze behaviour automatically triggered either a temporary magnification or a slowing of the presentation, in each case with a view to allowing more effective inspection to take place momentarily.