Fixation sequences made during visual examination of briefly presented 2D images

S. K. MANNAN,1 K. H. RUDDOCK1,* and D. S. WOODING2,†

1Imperial College, London SW7 2BZ, UK
2Department of Diagnostic Radiology, University of Wales College of Medicine, Heath Park, Cardiff, CF4 4XN Wales, UK

Received 7 November 1996; revised 28 January 1997; accepted 28 January 1997

Abstract—Eye movements made by eighteen observers in response to brief (3 s) presentations of eleven different images, each in three forms (unfiltered, high-pass filtered and low-pass filtered), have been analysed in order to identify both repeated sequences of fixations and image locations which attract re-fixations. It is shown that eye-movement traces made by different observers in response to the same image have few common temporal sequences involving the same fixation locations, even for sequences of only two fixations. There is a greater incidence of such sequences in eye-movement traces made by the same observer in response to two presentations of the same image, but average numbers are still low. Conserved sequences involving more than two identical locations occur at a much lower frequency, and the incidence of repeated sequences is not increased if consideration is restricted to regions of the image which attract large numbers of fixations. It is concluded that the temporal sequence in which fixations are made is not a significant factor in the analysis of the eye-movement data considered in this report. Calculations based on a least squares index of similarity are consistent with this conclusion. The analysis shows a relatively high incidence of re-fixation on certain locations in the images and there is evidence that such re-fixations are a significant factor in the high similarity between fixation locations established by different observers when viewing the same image.

1. INTRODUCTION

Fixations made during extended viewing of images are non-randomly distributed (Buswell, 1935; Yarbus, 1967; Jeannerod et al., 1968), but the mechanisms which control them are largely unknown. Both peripheral visual mechanisms responsive to specific image features and higher level cognitive mechanisms have been implicated by different investigators. A number of studies have been concerned with the relationship between fixation locations and the spatial structure of the visual stimulus, and a variety of significant factors have been identified (Buswell, 1935; Attneave, 1954; Yarbus, 1967; Mackworth and Morandi, 1967; Richards and Kaufman, 1967;
Antes, 1971; Loftus and Mackworth, 1978). These studies have in the main examined local rather than global spatial properties of the stimuli. The influence of cognitive mechanisms on the pattern of eye movements is demonstrated by dependence on the instructions given to the observer (Yarbus, 1967) and on the task undertaken (Land and Lee, 1994; Land and Harwood, 1995).

Previous studies have examined the temporal characteristics of eye-movement responses with reference both to the rate at which information about the image is acquired and to the sequence in which fixations are established. Informative regions of an image appear to be identified within the first few seconds during a sequence of eye movements (Mackworth and Morandi, 1967; Antes, 1974) and knowledge about the principal features of a scene is derived during the first fixation (Potter, 1976; Biederman et al., 1982). It is postulated that the visual system takes a snapshot of the image by utilising low spatial frequency content to achieve global recognition at low spatial resolution (Gould, 1967; Oliva et al., 1993). Noton and Stark (1971a, b) proposed that observers adopt regular spatio-temporal sequences of fixations and saccades whilst familiarising themselves with an image, these ‘scan paths’ forming a memory trace through which subsequent recognition of the image is achieved. Scan paths, which are specific to the image under examination and are idiosyncratic between observers, are attributed to cognitive brain activity associated with recognition of visual images. Quantitative analysis of scan paths has been undertaken by measuring transition probabilities for saccades made between different regions of the images, which can be represented by Markov conditional matrices (Stark and Ellis, 1981). Subsequent studies have investigated the scan path hypothesis, and results suggest that well-defined spatial distributions of fixations are established without specific temporal sequences (Walker-Smith et al., 1977; Locher and Nodine, 1987). Groner et al. (1984) provided a critique of the analytical methods used in the analysis of scan paths, and introduced the concept of a ‘global’ scan path to describe the distribution over a larger time scale than that of the strictly sequential, ‘local’ scan paths discussed by Stark and his colleagues. They argue that ‘local’ scan paths represent moment-to-moment control of eye movements, regulated by the momentary fixation and by peripheral input, whereas ‘global’ scan paths established, for example, between the beginning, the middle and the end of the scanning period, are determined by a central control mechanism (Groner and Groner 1982, 1989; Groner, 1988). Experimental observations on images of faces (Groner et al., 1984) provided evidence of ‘local’ scan paths in four of six subjects and analysis of the data led to the proposal regarding ‘global’ scan paths. Studies with random dot images indicated that both ‘local’ and ‘global’ scan paths are observer specific, and that ‘global’ scan paths are specific to the task undertaken and therefore reflect activity of central processing (Groner and Menz, 1985).

In this paper, we are concerned with the analysis of temporal sequences of fixations established by a group of eighteen observers who made brief (3 s) examination of eleven images presented in three different versions: unfiltered, low-pass filtered and high-pass filtered. In a previous communication (Mannan et al., 1995) we analysed similarities between the spatial distributions of fixations in eye movements made by different observers in response to the same image and in those made by the same