Vision Research 49 (2009) 2026–2036

Contents lists available at ScienceDirect

Vision Research

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

The role of view in human face detection

A. Mike Burton *, Markus Bindemann

Department of Psychology, University of Glasgow, G12 8QQ, UK

ARTICLE INFO

Article history:
Received 19 February 2009
Received in revised form 8 May 2009

Keywords:
Face detection
View
Frontal
Profile
3/4 View
Eyes
Scenes

ABSTRACT

The ability to detect faces in visual scenes is little understood. Across three experiments we examined whether particular facial views (for example those revealing a pair of eyes) facilitate detection while observers are searching for faces in complex visual scenes. Viewers' performance was equivalent for faces shown in frontal and mid-profile pose, but declined in profile (Experiment 1). These differences persisted when only half the face was shown, so that one eye was visible in frontal and profile view but both eyes were preserved in mid-frontal faces (Experiment 2). The same pattern was found when only the upper region of a face appeared in visual scenes, but the presentation of lower half faces eliminated all differences (Experiment 3). These findings demonstrate that the upper face mediates detection across different views, but 'a pair of eyes' cannot explain differences in detectability.

1. Introduction

Few stimuli, if any, can match the social and biological importance of the human face. Even a fleeting look at a person's face can provide information about their identity, gender, emotional and attentive state, attractiveness, approximate age and so forth. The failure to notice the presence of a face within our visual environment would inevitably lead to a loss of this information. Face *detection* therefore not only represents one of the most fundamental but also one of the most important aspects of face processing, and more generally, of human social cognition. In spite of this, the ability to locate faces in our visual environment is little studied and remains poorly understood (for reviews, see Lewis & Edmonds, 2005; Lewis & Ellis, 2003; Tsao & Livingstone, 2008). In this study, we begin to address this disparity by exploring one of the most elementary questions in this field, namely whether detection depends on the particular view in which a face is seen.

Changes in view can induce substantial variation in a face's visual appearance. In full-face (or frontal) view, for example, faces contain a contiguous pair of eyes, which are located either side of a centrally positioned nose. By comparison, only a single eye is visible in a profile view of the head, and this eye is located much more peripherally than both eyes in a full face. The appearance of other facial landmarks, such as the nose and mouth and more global visual characteristics, such as the head outline and hair, also vary across different face views and can change the overall appearance of a face substantially. This variation is such that observers often fail to match two different views of the same face (e.g., Burke, Taubert, & Higman, 2007; Hancock, Bruce, & Burton, 2000; Hill, Schyns, & Akamatsu, 1997; Lee, Matsumiya, & Wilson, 2006; Liu & Chaudhuri, 2002; Newell, Chiroro, & Valentine, 1999). Moreover, matching accuracy falls continuously as the distance between two to-be-matched views increases, pointing at independent perceptual coding of different face views (see, e.g., Longmore, Liu, & Young, 2008; O'Toole, Edelman, & Bulthoff, 1998).

These observations converge with studies of cell recordings in non-human primates (see, e.g., Perrett, Oram, & Ashbridge, 1998; Perrett et al., 1985, 1991), brain imaging studies of human observers (e.g., Ewbank & Andrews, 2008; Ewbank, Smith, Hancock, & Andrews, 2008; Grill-Spector et al., 1999; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005), and behavioural visual adaptation studies (e.g., Benton, Jennings, & Chatting, 2006; Fang & He, 2005; Jeffery, Rhodes, & Busey, 2006, 2007; Jiang, Blanz, & O'Toole, 2007), which have consistently found evidence for view-specific face coding. Single cell recordings, for example, have revealed separate assemblies of cells for processing characteristic face views, such as full-face and profile views (Perrett et al., 1985, 1991). These findings have been extended to human observers by studies of neural adaptation, which show a reduced response (adaptation) in face-sensitive brain areas when successive images of faces are shown in the same view. This is contrasted by a release of adaptation when faces are presented at different viewing angles, indicating the operation of view-dependent face coding mechanisms (e.g., Ewbank & Andrews, 2008; Grill-Spector et al., 1999). Similar approaches have been employed in behavioural adaptation studies, which show, for example, that prolonged viewing of one face identity (adaptation) leads to a reduced

* Corresponding author. Fax: +44 (0)141 330 4606.
E-mail address: mike@psy.gla.ac.uk (A.M. Burton).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.05.012

perception of that identity in an immediately succeeding test face (Benton et al., 2006). This effect is maximal when adaptation and test face are shown in the same view, and decreases as the angle of the test face moves further away from the adaptation view.

There is, then, considerable evidence for view-specific visual encoding of different face views. However, in previous studies faces were always presented in isolation, on a plain background and in the centre of the visual field. As a consequence, these studies cannot address whether view affects our ability to locate a face in the visual field in first place, prior to any of the other face processing tasks (e.g., face identification, matching, adaptation, etc.) that have been studied in this domain, and whether some canonical face views exist that are detected more proficiently than others. The aim of this study, therefore, is to investigate how variations in view affect our face detection ability, with a series of three experiments.

2. Experiment 1

The first experiment explored how view generally affects the ability to locate faces in our visual environment. For this purpose, observers were presented with natural visual scenes in which a face was either present or absent. Faces were embedded in target-present scenes in a frontal, mid-profile, or profile view of the head, and detection performance was measured as a function of face view.

2.1. Method

2.1.1. Participants

Thirty undergraduate students from the University of Glasgow participated in this experiment for a small fee. All had normal or corrected-to-normal vision.

2.1.2. Stimuli

The stimuli consisted of 24-bit RGB photographs of 120 indoor scenes, which were taken from inside houses, apartments and office buildings, and measured 1000 (H) $\times$ 750 (W) pixels at a resolution of 72 pixels/inch (sustaining a visual angle of $30.5^\circ \times 24.8^\circ$ at a viewing distance of 60 cm). For each scene, six versions existed that were identical in all aspects, except for the following differences. Five of the scene versions contained a face (for face-present trials) and one version did not (face absent condition). In face-present scenes, faces were either depicted in a frontal view, a mid-profile view, or a profile view. In mid-profile faces both eyes were always clearly visible, and each face was shown in a mid-profile left view (with the face pointing towards the left side of the screen, viewed from the observer's perspective) or a mid-profile right view. Similarly, profile faces were shown in a profile left and profile right view (see Fig. 1 for an illustration of these face conditions). Applying these manipulations to each of the scenes therefore resulted in a total of 720 different displays, comprising 120 face-absent displays and 600 face-present displays, in which a face was present in either a frontal view (120 images), mid-profile left or mid-profile right view (120 images each), and profile left or profile right view (120 images each).

The faces depicted in these scenes were of twenty unfamiliar models (10 male). Faces can attract visual attention (see, e.g., Bindemann & Burton, 2008; Bindemann, Burton, Langton, Schweinberger, & Doherty, 2007; Langton, Law, Burton, & Schweinberger, 2008), but so do human bodies (Downing, Bray, Rogers, & Childs, 2004; Ro, Friggel, & Lavie, 2007). To avoid any potential influence on face detection from body parts, the faces were therefore embedded in the scenes as a photograph, devoid of any body cues (see Fig. 2 for an example scene). To ensure that the face locations were unpredictable

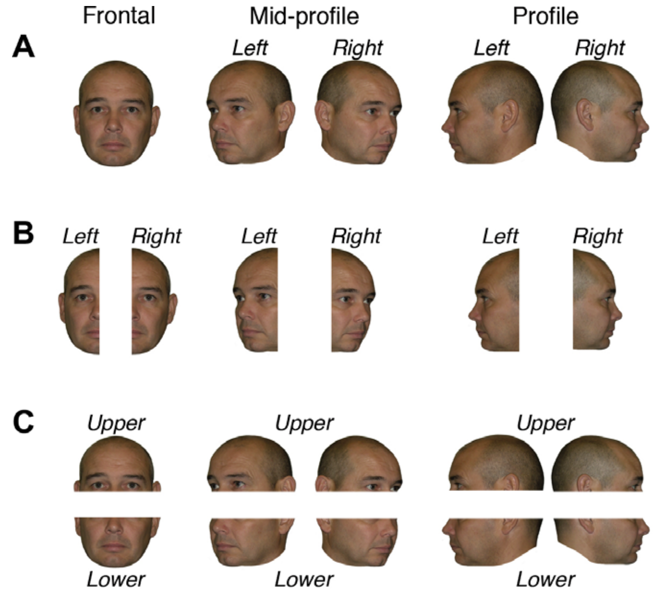

Fig. 1. An illustration of the face conditions for Experiment 1 (A), 2 (B) and 3 (C).

throughout the experiment, the scene images were divided into an invisible 3 $\times$ 2 grid of six equally sized rectangles. Across the set of scene images, faces were rotated around these areas, so that they were equally likely to appear in each of the six regions. In addition, the size of the faces was varied across the scenes, ranging from 0.08% of the total scene area for the smallest head (comprising the face, hair and external features) to 1.73% for the largest head, to ensure that participants could not adopt a simple search strategy based on target size (mean size and std., frontal: 0.28% (0.20%), mid-profile: 0.33% (0.24%), profile, 0.35% (0.25%)).

2.1.3. Procedure

Each participant was shown 360 randomly intermixed trials, consisting of 240 face-absent trials and 120 face-present trials. Face-present trials consisted of 40 scene stimuli for each of the three conditions (frontal, mid-profile and profile view). For mid-profile and profile face views, these comprised 20 trials in a left view and 20 trials in a right view. The scene stimuli were rotated around these conditions so that each face-present scene was only shown once to each participant. Overall, however, the presentation of the scenes was counterbalanced across participants, so that each scene appeared in each condition an equal number of times.

A trial began with a central fixation cross for 1 s, followed by a scene stimulus, which was displayed until response. Participants were briefed about the different experimental conditions prior to the experiment and were asked to make speeded key-press responses concerning whether a face was present in a scene or not. In addition, participants were made aware that only a proportion of scenes contained a face, and were encouraged to guess when they were uncertain regarding the presence of a face.

2.2. Results

In a first step, the time taken to detect a face was plotted against the surface area of a face. By item analyses of this data shows that large faces were more likely to be detected faster than smaller faces in all conditions; for frontal faces, $r = -.24$, mid-profile faces, $r = -.21$ and profile faces, $r = -.17$, all $ps < 0.001$. This shows that the process by which faces are detected in these scenes preserves a simple physical property, namely face size. However, these correlations produced the same outcome in all of the experiments