Vision Research 99 (2014) 12–18

Contents lists available at ScienceDirect

Vision Research

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

Implicit learning of geometric eigenfaces

Xiaoqing Gao *, Hugh R. Wilson

Centre for Vision Research, York University, Canada

ARTICLE INFO

Article history:
Received 1 May 2013
Received in revised form 15 July 2013
19 July 2013
Available online 30 July 2013

Keywords:
Face learning
Implicit learning
Summary statistics
Prototype
Principal components

ABSTRACT

The human visual system can implicitly extract a prototype of encountered visual objects (Posner & Keele, 1968). While learning a prototype provides an efficient way of encoding objects at the category level, discrimination among individual objects requires encoding of variations among them as well. Here we show that in addition to the prototype, human adults also implicitly learn the feature correlations that capture the most significant geometric variations among faces. After studying a group of synthetic faces, observers mistook as seen previously unseen faces representing the first two principal components (eigenfaces, Turk & Pentland, 1991) of the studied faces at significantly higher rates than the correct recognition of the faces actually studied. Implicit learning of the most significant eigenfaces provides an optimal way for encoding variations among faces. The data thus extend the types of summary statistics that can be implicitly extracted by the visual system to include several principal components.

1. Introduction

The human visual system has the ability to extract statistical regularities from the environment. It has been demonstrated that humans can automatically extract the central tendency (prototype) of a series of exemplars. Human observers tend to regard the unseen prototype as more familiar than the experienced exemplars. This phenomenon is known as the prototype effect (Posner & Keele, 1968). The prototype effect has been observed with a wide range of visual stimuli, from simple geometric shapes (e.g., dot patterns, Posner & Keele, 1968; circles, Chong & Treisman, 2003) to complex visual objects (e.g., faces, Baudouin & Brochard, 2011; Cabeza et al., 1999; de Fockert & Wolfenstein, 2009; de Haan et al., 2001; Haberman & Whitney, 2009; Or & Wilson, 2013; Solso & McCarthy, 1981; Wallis et al., 2008).

The prototype effect indicates an efficient mechanism for encoding objects at the category level, as the prototype permits easy classification of new exemplars. However, for many object categories, it is also crucial to recognize individual exemplars, and human faces are one clear example (Tanaka, 2001). Although it has been demonstrated that humans can implicitly learn the prototype of encountered faces (Baudouin & Brochard, 2011; Cabeza et al., 1999; de Fockert & Wolfenstein, 2009; de Haan et al., 2001; Haberman & Whitney, 2009; Or & Wilson, 2013; Solso & McCarthy, 1981; Wallis et al., 2008), learning the prototype is not sufficient for encoding individual faces. On the other hand, remembering all the exemplars is not an efficient way of encoding. We know, however, little about what kind of statistical regularity is learned in addition to the prototype. Principal components (PC) have proved effective in capturing the major variations among faces for computer recognition (Sirovich & Kirby, 1987; Turk & Pentland, 1991) and for modeling human perception (Calder et al., 2001; Hancock, Burton, & Bruce, 1996; O'Toole et al., 1991, 1993; Said & Todorov, 2011). However, it is not clear whether the human brain utilizes a mechanism that is similar to Principal Component Analysis (PCA) in encoding faces, nor has the ability of learning PC been demonstrated with any other visual objects. In the current study, we investigated whether human observers can learn PC from geometric information of faces, given that learning PC from a set of exemplars is biologically plausible as demonstrated by neural network architectures based on Hebbian learning mechanisms (Diamantaras & Kung, 1996; Rubner & Schulten, 1990).

We calculated summary statistics from a set of synthetic faces. Each synthetic face was derived from a frontal face photograph and specified by 37 parameters capturing the major geometric information in the face (Wilson, Loffler, & Wilkinson, 2002). Although the synthetic faces are simplified representations of real faces, they are sufficiently complex to capture salient shape information of real faces as evidenced by high accuracy in matching the synthetic faces to grayscale photographs from which the synthetic faces were derived (Wilson, Loffler, & Wilkinson, 2002). The synthetic faces can be precisely manipulated as with Cartoon faces (e.g., Brunswik & Reiter, 1937; Freiwald, Tsao, & Livingstone, 2009; Sigala & Logothetis, 2002), while having an advantage over the Cartoon faces as they were derived from the geometric measures

* Corresponding author. Address: Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. Fax: +1 416 736 5857.

E-mail address: xgao@cvr.yorku.ca (X. Gao).

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.07.015

of real faces. The synthetic faces also have an advantage over pixel-based representation of faces (i.e., photographs). The synthetic faces provide a precise representation of feature locations while pixel-based representation can only provide an approximate representation of the feature location. As the result of the approximate nature of the representation of the feature location, eigenfaces derived from face photographs are far from realistic looking. These eigenfaces cannot be used in combination to create new facial identities that are realistic looking. With the synthetic faces, we are able to derive eigenfaces that have the same quality of representation as the original synthetic faces. We are also able to create new facial identities by combining several eigenfaces. Most importantly, studies of PCs derived from face photographs show that the first several PCs contain only low spatial frequency information that is related to shadows and shading but not to individual identity. Synthetic faces, on the other hand, are bandpass filtered in the optimal band for identity processing (Gao & Maurer, 2011;Gold, Bennett, & Sekuler, 1999a; Näsänen, 1999) and are comprised exclusively of geometric information indicative of individual identity. These characteristics make synthetic faces optimal for our investigation of the learning of PCs from faces.

The prototype effect in face recognition shows that the unseen face prototype is more likely to be recognized than the actually studied exemplar faces (de Fockert & Wolfenstein, 2009; Haberman & Whitney, 2009; Or & Wilson, 2013; Solso & McCarthy, 1981). We hypothesize that if the prototype face and the most significant eigenfaces of the studied synthetic faces are implicitly learned, the observers would identify the unseen prototype face and the unseen eigenfaces as having been seen during a subsequent face memory test.

2. Experiment 1

2.1. Method

2.1.1. Participants

Ten adults (27.6 ± 5.1 years, five males) participated in Experiment 1. All the participants had normal or corrected-to-normal vision. We obtained informed written consent from all participants. The procedures were approved by the York University research ethics board.

2.1.2. Stimulus

A detailed description of the design of the synthetic faces has been reported elsewhere (Wilson, Loffler, & Wilkinson, 2002). Briefly, each synthetic face is defined by 37 parameters. Among the 37 parameters, 23 of them define the head shape and hairline, while the remaining 14 parameters define the locations and sizes of the facial features. All the 37 measures were normalized with the unit change on each measure representing a percentage relative to the mean head radius of 41 synthetic faces. The reconstructed synthetic faces were grayscale and were filtered with a band pass difference of Gaussians filter centered on 10 cycles per face with a bandwidth of two octaves to keep the most important information for facial identity. The Face stimuli were presented on a 20-in. LCD monitor with a mean luminance of 74 cd/m^2. From a viewing distance of 127 cm, each face subtended an angle of 6.9° (height) by 4.6° (width).

We submitted 41 synthetic faces of Caucasian males to PCA. The resulting 37 PCs were used to define a multidimensional face space. The distance between any two synthetic faces in this face space is defined as the Euclidean distance between the two faces in the 37-dimensional face space as a fraction of the mean radius of the 41 faces. New faces created using a single eigenvector will be referred to as eigenfaces (Turk & Pentland, 1991). We constructed eigenfaces using both positive and negative values of the PC, and we refer to these as PC+ and PC−. The first eigenface incorporates the maximum amount of variance among facial features as defined by the covariance matrix, and subsequent eigenfaces incorporate the maximum of the remaining variance.

As shown in Fig. 1A, we created 16 faces for the study phase by combining an eigenvector on one direction of PC1 (PC1+ or PC1−) with an eigenvector on one direction of a higher PC (PC2, PC4, PC6, or PC8; + or −). We also created 16 faces as the new faces in the testing phase by combining an eigenvector on one direction of PC3 (PC3+ or PC3−) with an eigenvector on one direction of a higher PC (PC5, PC7, PC9, or PC10; + or −), so that the new faces would be in a non-overlapping and orthogonal volume of the face space from the studied faces. The distance of each eigenvector was set to 0.15 from the average face.

2.1.3. Procedures

In the study phase, participants studied 16 faces each for a total of 40 s. There were four blocks in the study phase. Within each block, each face appeared once for 10 s in a random order. Before the study phase started, the participants were informed that they would be tested on their memory of the studied faces following the study phase. Immediately after the study phase, the participants performed a studied/novel recognition task. In this task, each trial started with a central fixation cross for 500 ms, followed by a

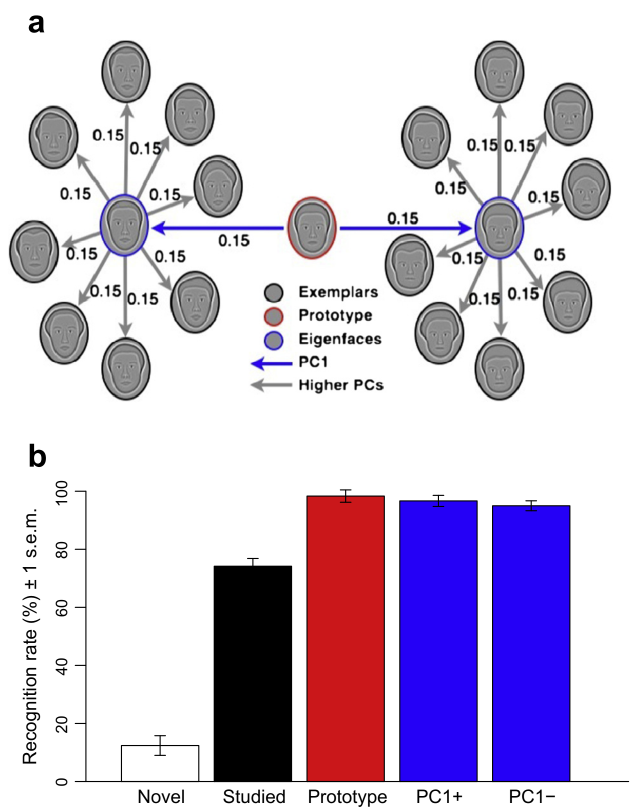

Fig. 1. Stimulus composition and recognition rates in Experiment 1. (A) Each studied exemplar face (black oval) is a linear combination of two eigenvectors with a length of 0.15. The length or distance (shown as a number on each arrow) is defined as the Euclidean distance between two faces in the 37-dimensional face space as a fraction of the mean head radius of the faces. The blue arrows represent the first PC. The gray arrows represent four mutually orthogonal higher PC. The eigenfaces (blue oval) and the prototype face (red oval) were never studied and only presented in the testing phase. (B) Mean recognition rates (±1 s.e.m) for novel faces (white bar), studied faces (black bar), prototype face (red bar), and eigenfaces of PC1 (blue bars). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)