
Sex classification is better with three-dimensional head structure than with image intensity information

Alice J O'Toole

School of Human Development, GR4.1 The University of Texas at Dallas, Richardson, TX 75083-0688, USA; e-mail: otoole@utdallas.edu.

Thomas Vetter, Nikolaus F Troje, Heinrich H Bülthoff

Max-Planck-Institut für biologische Kybernetik, Spemannstrasse 38, D 72076 Tübingen, Germany; e-mail: vetter@mpik-tueb.mpg.de

Received 17 April 1996, in revised form 22 November 1996

Abstract. The sex of a face is perhaps its most salient feature. A principal components analysis (PCA) was applied separately to the three-dimensional (3-D) structure and graylevel image (GLI) data from laser-scanned human heads. Individual components from both analyses captured information related to the sex of the face. Notably, single projection coefficients characterized complex differences between the 3-D structure of male and female heads and between male and female GLI maps. In a series of simulations, the quality of the information available in the 3-D head versus GLI data for predicting the sex of the face has been compared. The results indicated that the 3-D head data supported more accurate sex classification than the GLI data, across a range of PCA-compressed (dimensionality-reduced) representations of the heads. This kind of dual face representation can give insight into the nature of the information available to humans for categorizing and remembering faces.

1 Introduction

In recent computational work, principal components analysis (PCA)⁽¹⁾ has been applied widely to analyzing the information in images of human faces. The information captured by PCA has been shown to be reliable, in purely computational terms, for recognizing faces (O'Toole et al 1993; Turk and Pentland 1991) and for classifying them by race and sex (O'Toole et al 1991). Additionally, different components of PCA have been shown to relate reliably to human performance on some of these same tasks (O'Toole et al 1994).

The human head, however, is a complex three-dimensional (3-D) object with a characteristic shape and an associated surface image (ie surface marking information) from which individual faces, and categories of faces, vary. In the present study, we applied PCA to a more complete physical model of the human face than is available from a facial image. This model included both a 3-D structure and a component based on graylevel image (GLI). Recently, Hancock et al (1996) highlighted the importance of considering separately image and 'shape-based' components in modeling human performance on a face recognition task. They used images of faces to create a 'shape-free' face by hand-selecting facial landmarks and morphing the faces to an average shape (Craw and Cameron 1991). They then applied PCA separately to the shape-free image of the face and to the deviation of the individual faces from the average shape. Their results indicated that different components of human recognition performance related to the structure and GLI-based information.

A more explicit representation of 3-D head information than that available from morphing between 2-D images can be obtained from laser scan technology, which provides both the 3-D coordinates of the head (structure) and a wrap-around image

⁽¹⁾ For related analyses, see also Pearson (1901), Hotelling (1933), Karhunen (1946), and Loève (1955).

(GLI) that maps point for point onto the head surface.⁽²⁾ While this representation lacks the advantageous feature of established correspondences between a reduced set of facial features [eg 34 facial landmarks such as the corners of the mouth and eyes (Hancock et al 1996)], it has the important advantage of including much more detailed information about subtle variations in 3-D structure than that retained with a relatively small set of interfacial landmark distances. The purpose of the present study was to compare the overall utility of the structure versus GLI information for predicting the sex of a face. We did this in two steps. First, we applied PCA separately to the head structure and GLI data taken from a large number of male and female faces. Second, we carried out a series of simulations to compare the quality of the information available in the head structure versus GLI data for categorizing the faces by sex.

2 Principal component analysis

2.1 Stimuli

Laser scans (Cyberware™) of 130 heads of young adults (65 male and 65 female) were used as stimuli. The subjects were scanned wearing bathing caps, which were removed digitally prior to the PCA, consequently eliminating most of the hair region of the head. Additionally, further preprocessing of the head scans was done by making a vertical cut behind the ears, and a horizontal cut to remove the shoulders.

The laser scans provide two kinds of information. First, the head structure data consisted of the lengths of 512×512 radii from a vertical axis in the middle of the subject's head to 'sample' points on the surface of the head. This is a cylindrical representation of the head surface, with surface points sampled at 512 equally-spaced angles around the circular slices of the cylinder, and at 512 equally spaced vertical distances along the long axis of the cylinder. The radii in our data set varied from a minimum of 4 cm (neck area) to a maximum of 12 cm (nose area). The resolution of the scanner was $16 \mu\text{m}$.

The graylevel data consisted of 512×512 256-graylevel image intensities that mapped point for point onto the 3-D sample points of the head. The 3-D head models and GLI maps were aligned so that regions around the nose and eyes roughly coincided spatially. In both cases, the mean of the data was subtracted prior to the application of the PCA.

In summary, the head structure data provide an object-centered representation of the 3-D coordinates of a head. This representation is, therefore, inherently independent of viewpoint. Likewise, the 'wrap-around' GLI maps point for point onto the 3-D head structure and so can be considered, also, as a viewpoint independent representation.

2.2 Analysis procedure and results

Principal components analysis was applied separately to the structure and GLI data. Figure 1 shows the average head structure in the top row, and the two eigenvectors (the first and the sixth), that related most reliably to face sex for the head data in the second and third rows, respectively. The eigenvectors are displayed by adding (subtracting) them to (from) the mean head. In both cases, distinctive shape differences between male and female heads can be captured simply by adding versus subtracting these single components to/from the average head. This demonstration replicates a similar finding for images (O'Toole et al 1993) and indicates that, in this representation

⁽²⁾ We use the term 'graylevel image' (GLI) to refer to this complete wrap-around image. This wrap-around image is distinct from a view-based image of the head (ie an image taken from a particular view), since it contains image intensity information about the entire head surface. In computer graphics, this graylevel wrap-around image is referred to as a 'texture map' in contrast to the pure geometry of an object. We do not use the term 'texture' here, because in the computer vision literature it generally connotes a repetitive pattern.

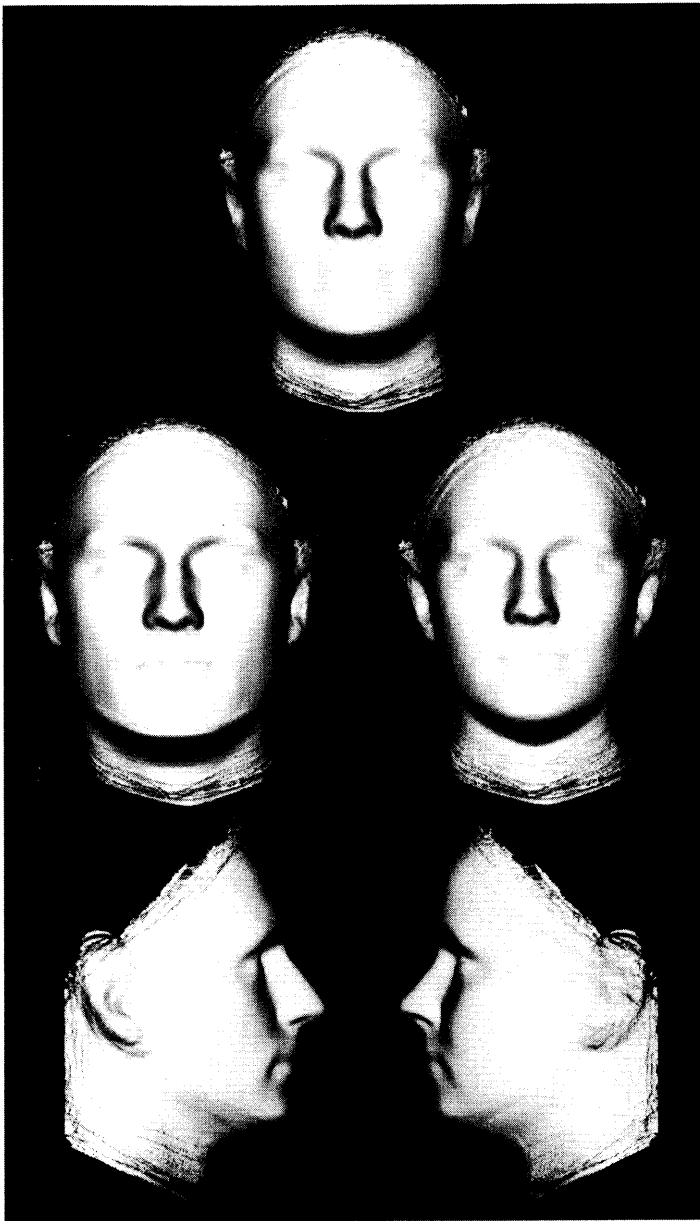


Figure 1. Row 1: average of 130 heads (65 male and 65 female).

Row 2: left head constructed by adding the shape defined by the first eigenvector to the average head; right head constructed by subtracting the shape defined by the first eigenvector from the average head. More precisely, the display shows the average head combined linearly with $\pm 2\sigma$ times the first eigenvector, where σ is the standard deviation of the projection coefficients on this eigenvector computed across all faces. It is interesting to note that a single coefficient describes a compelling and complex structural transformation between male and female heads, including aspects of brow protrusion and jaw shape.

Row 3: the same as row 2 but with the sixth eigenvector, viewed from the side where it is easiest to see that it captures differences in nose size and shape between male and female heads.