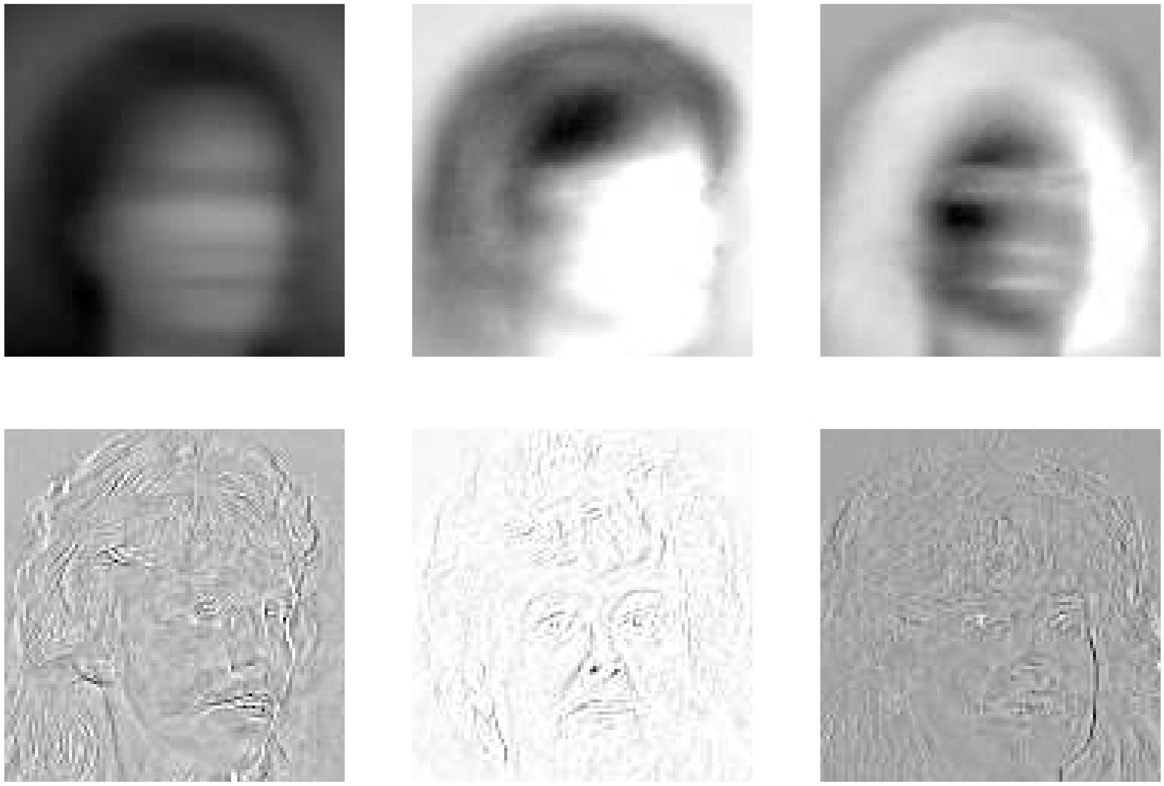

FIGURE 2. First three (top panel) and last three (bottom panel) eigenfeatures of a set of 40 female faces each represented by 10 views sampling the rotation of the head from full-face to profile with 10-degree steps (from Valentin, 1996).

Parallel models of associative memory. Hillsdale: Erlbaum. pp. 213–236.

[6] Hancock, P. J., Burton, A. M., and Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition, 24*, 26–40.

[7] O'Toole, A., Abdi, H., Deffenbacher, K., and Valentin, D. (1993). A low dimensional represention of faces in the higher dimensions of the space. *Journal of the Optical Society of America A, 10*, 405–411.

[8] O'Toole, A., Deffenbacher, K., Abdi, H., and Bartlett, J. (1991). Simulating the other race effect as a problem in perceptual learning. *Connection Sciences, 3*, 163–178.

[9] O'Toole, A., Vetter, T., Troje, N. F., and Bülthoff, H. H. (1997). Sex classification is better with three-dimensional structure than with image intensity information. *Perception, 26* 75–84.

[10] Sirovich, L. and Kirby, M. (1987). Low dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A, 4*, 519–524.

[11] Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neurosciences, 3*, 71–86.

[12] Valentin, D. and Abdi, H. (1996). Can a linear autoassociator recognize faces from new orientations? *Journal of the Optical Society of America A, 13*, 717–724.

[13] Valentin, D., Abdi, H., and Edelman, B. (1996). Connectionist "face-off": Different algorithms for different tasks. *Psychologica Belgica, 36*, 65–92.

[14] Valentin, D., Abdi, H., Edelman, B., and O'Toole, A.J. (in press). Principal component and neural network analyses of face images: How stable is gender information? *Journal of Mathematical Psychology.*

[15] Valentin, D. (1996) *How come when you turn your head I still know who you are: Evidence from computional simulations and human behavior.* Ph.D. Dissertation, The University of Texas at Dallas.

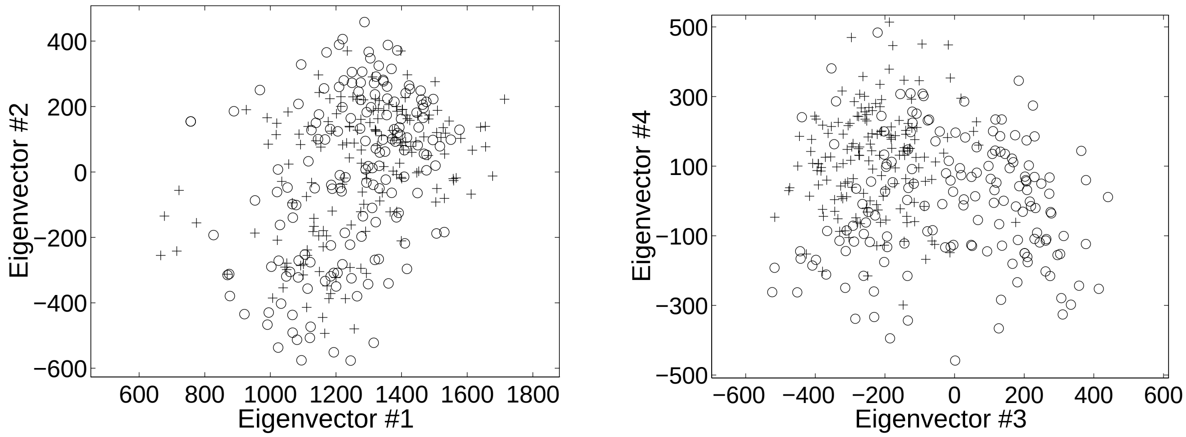

FIGURE 1. An illustration of the other-race effect. New Caucasian faces, (i.e., not learned by the model), denoted by "o", and new Japanese faces, denoted by "+", are projected on eigenfaces obtained with Caucasian faces only. The between faces similarity is larger for other-race faces (i.e. Japanese) than for own-race faces.

Valentin and Abdi (1996). They showed that different eigenfeatures capture different kind of information. As illustrated by Figure 2, eigenfeatures with large eigenvalues contain information relative to the orientation (*e.g.,* full-face, profile) in addition to the categorical assignment (*e.g.,* gender, race) of the faces. These eigenfeatures are robust and can be estimated from a small set of faces (Valentin, Abdi, Edelman & O'Toole, in press). In contrast, eigenfeatures with small eigenvalues contain face identity information.

A potential problem for the PCA approach, as noted by Schyns *et al.*, is that feature extraction operates independently of higher level cognitive processes. Using top-down information to constrain the eigenvector representation may offer a solution to this problem. For example, Abdi, Valentin, Edelman, and O'Toole (1996) derived, recently, a generalization of the PCA model which incorporates *a priori* constraints both at the level of pixel representations and of faces. Using this generalized PCA model, Abdi, Valentin, and O'Toole (1997) showed that constraints on the pixels improves the gender categorization performance of PCA models. The pixels were weighted accordingly to their information content, an idea in tune with Schyns *et al.*'s proposal.

To conclude, eigenfeatures may play the rôle of the "flexible features" hypothesized by Schyns *et al.* even though some other mechanisms certainly coexist. A future line of development for PCA models is to incorporate entry code perceptually or cognitively more realistic than pixels such as the output of Gabor filters, or wavelets.

REFERENCES

[1] Abdi, H. (1988). Generalized approaches for connectionist auto-associative memories: Interpretation, implication, and illustration for face processing. In J., Demongeot. (Ed.) *Artificial Intelligence and Cognitive Sciences.* Manchester: Manchester University Press. pp. 151–164.

[2] Abdi, H., Valentin, D., Edelman, B., and O'Toole, A. (1995). More about the difference between men and women: Evidence from linear neural networks and the principal component approach. *Perception, 24*, 539–562.

[3] Abdi, H., Valentin, D., Edelman, B., and O'Toole, A. (1996). A Widrow-Hoff learning-rule for a generalization of the linear auto-associator. *Journal of Mathematical Psychology, 40*, 175–182.

[4] Abdi, H., Valentin, D., and O'Toole, A. (1997). A generalized auto-associator model for face semantic processing In D. Levine, & W.R. Elsberry (Eds.) *Optimality in Biological and Artificial Networks?* Hillsdale: Erlbaum. pp. 317–337.

[5] Anderson, J. and Mozer, M. (1981). Categorization and selective neurons. In Hinton, G. and Anderson, J. (Eds.)

Vol.# , No.#, December 1997, pp. ***–***. / Brain and Behavioural Sciences

Eigenfeatures as intermediate level representations: The case for PCA models

Hervé Abdi *†, Dominique Valentin *†, Betty G. Edelman †

Eigenfeatures are created by the principal component approach used on objects described by a low level code (i.e., pixels, Gabor jets). We suggest that eigenfeatures act like the flexible features of Schyns *et al.* They are particularly suited for face processing and give rise to class-specific effects such as the other race effect. The PCA approach can be modified to accommodate top-down constraints.

1. Introduction

How can the gender, the race, or the identity of a face be inferred from a digitized picture? We can imagine that the "flexible features" proposed by Schyns *et al.* to perform this type of tasks correspond to some intermediate level features which are progressively extracted from the exemplars of the relevant categories. The problem, however, is to find out mechanisms responsible for extracting such features. We propose that, for categorization tasks involving high similarity object classes such as faces, the principal component analysis (PCA) model is a good candidate. The applicability of this model to face processing, first suggested in the late eighties (Abdi, 1988; Sirovich & Kirby, 1987; Turk & Pentland, 1991), is still current (Hancock, Burton & Bruce, 1996; O'Toole, Vetter, Troje & Bülthoff, 1997).

The PCA approach represents faces by their projections on a set of orthogonal features (principal components, eigenvectors, "eigenfaces") epitomizing the statistical structure of the set of faces from which they are extracted. These orthogonal features are ordered according to the amount of variance (or eigenvalue) they explain, and are often referred to as "macro-features" (Anderson & Mozer, 1981) or *eigenfeatures* by opposition with the high level features traditionally used to describe a face (e.g., nose, eyes, mouth). Eigenfeatures are flexible in that they evolve with the faces encountered (Valentin, Abdi, & Edelman, 1996) and depend on the set of faces from which they are extracted (O'Toole, Deffenbacher, Abdi & O'Toole, 1991).

Because they are optimal for the set of faces from which they are extracted, eigenfeatures are less efficient for representing faces from a different population and thus generate class-specific effects such as the other race effect. As an illustration, Figure 1 displays the projection of 160 Caucasian faces and 160 Japanese faces on the first 4 eigenvectors derived from Caucasian faces (none of the faces have been used to compute the eigenvectors). The Japanese faces are more similar to each other than the Caucasian faces are. This shows that Caucasian eigenfeatures give rise to the other race effect when used for categorizing Japanese faces.

Eigenfeatures can be used to perform higher level categorization tasks such as face categorization or identification. For example, Abdi, Valentin, Edelman, and O'Toole (1995) showed that a neural network, trained to classify faces according to their gender, generalizes its learning to new faces better when the faces are represented by eigenfeatures than by arrays of pixel intensities. This superiority of the eigenfeatures over pixel representations suggests that eigenfeatures are semantically relevant.

Semantic relevance has also been demonstrated by O'Toole, Abdi, Deffenbacher, and Valentin (1994) and

Send correspondence to: Hervé Abdi, The University of Texas at Dallas, School of Human Development, MS:Gr.4.1, Richardson, TX 75083-0688, USA. email: herve@utdallas.edu http://www.utdallas.edu/~herve