Parameterising Images for Recognition and Reconstruction

Ian Craw and Peter Cameron*
Department of Mathematical Sciences,
University of Aberdeen,
Aberdeen AB9 2TY

Introduction

We describe a method based on Principal Component Analysis for extracting a small number of parameters from the whole of an image. These parameters can then be used for characterisation, recognition and reconstruction. The method itself is by no means new, and has a number of obvious flaws. In this paper we suggest improvements, based on purely theoretical considerations, in which the image is preprocessed using prior knowledge of the content. The subsequent Principal Component Analysis (PCA) is both theoretically more attractive, and more effective in practice. We present the work in the context of face recognition, but the method has much wider applicability.

One test of the utility of components extracted by PCA is to see how well they represent data not available to the initial analysis. Figure 1 shows our method representing, with only 50 bytes, a face not present in the original ensemble of faces.

Figure 1: Unknown face and its reconstruction

Superficially, our work has much in common with a neural net approach. Since each method may reduce to PCA in the linear case, the analogy may be much deeper; however the process we present here is one in which each step is comprehensible and justifiable.

Theoretical Objections

The model underlying PCA is crucially a linear space model. It assumes our objects lie in some lower dimensional subspace. Consequently the objects themselves (in our case faces) should form a linear space, and in particular, the sum (or average) of two faces is itself a face. This is patently false; at best the average face is fuzzy, while the failure is marked if the faces combined are of noticeably different sizes or positions.

This failure means that in published work, faces have been very carefully standardised before being subjected to PCA. Thus Kirby and Sirovich[KS90] ensure that their ensemble was chosen to be homogeneous initially, that all the faces have their axes of symmetry and a horizontal axis through the eyes aligned etc.

Modelling Face Space

To explain this *ad hoc* normalisation, a more sophisticated model for "face space" is needed. In order to preserve the linearity locally, we assume face space is some smooth manifold modelled on an underlying Euclidean space of dimension perhaps 50, assumed to be embedded in our high dimensional linear space. As usual, we can map a linear structure locally on the manifold using a chart, or equivalently, perform linear operations in the tangent space. The local linearisation is given as follows: given x_0 in the face manifold M there is a neighbourhood U of x_0 in M and a diffeomorphism

$$\varphi : U \to V \subset \mathbf{R}^{50} \quad \text{(say)}.$$

Formally we have no way of constructing φ, but our discussion on linearity provides a test for the suitability of such a function. Given x_1 and x_2 in U, and hence faces fairly close to x_0, we expect to recognise

$$\varphi^{-1}\left(\frac{\varphi(x_1) + \varphi(x_2)}{2}\right)$$

as a face.

Of course PCA makes perfect sense when applied in the linear space on which the manifold is modelled. Thus we expect to transform our images before performing PCA, and recognise the careful normalisations described above as implementing the map φ.

Implementation

We now describe an implementation of these ideas in which the "linearity" or averaging test has been given precedence. Rather than averaging greyscale values at each pixel, we first choose a set of (typically 76) control points, representing easily located features such as the edge of the mouth. These are arranged in the configuration of the "standard" face — obtained using measurements of the control points on 1000 faces in the Aberdeen face database. A triangulation is built on these points, and induces a triangulation on the original image; the greyscale image in standard position is then obtained by bilinear interpolation

between corresponding triangles. Averaging between standard images is now trivial; to average between real faces, we map back, using as control points the average of the original control-point vectors.

At present the control points are located manually; software to do this automatically is being developed[TCL90].

We can now describe the chart $\varphi(x_0)$; the map takes a face to the corresponding face image in standard position, together with the control-point vector. Our argument above on averaging suggests this maps onto a full neighbourhood in the underlying space. We then use PCA to reduce to the low - dimensional Euclidean space on which the manifold is modelled. At present the results we present come from PCA performed on the normalised images alone. A more rigorous approach would include the control-point vector with the greyscale image in the initial PCA.

Results In figure 1 we gave an example applying PCA with an ensemble of 50 faces to reproduce a new face not in the ensemble. We present in figure 2 below the intermediate stages of the calculations. Here, the face is first standardised (left) and then the standardised face is reconstructed using PCA (right).
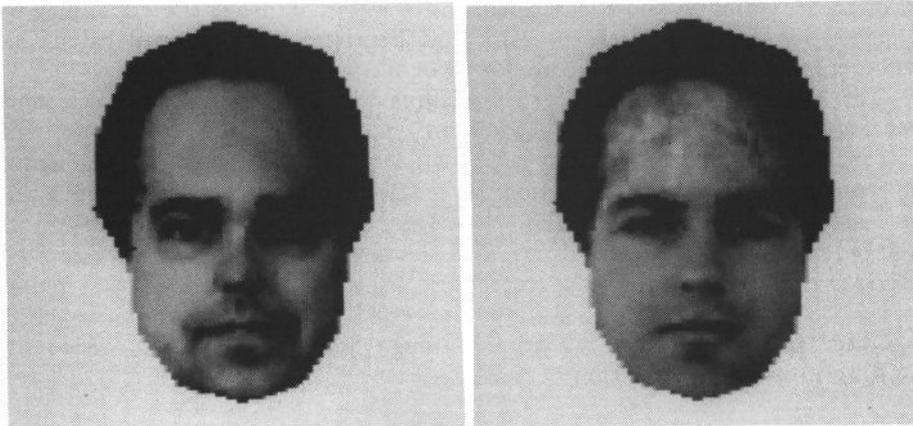

Figure 2: Intermediate stages

To give a feel for the axes used, the first principal component is displayed in figure 3. We also display the conventional approximation of the unknown face from figure 1 alongside. This should be contrasted with the reconstruction shown in figure 1.

We have performed PCA on 12 sets of 50 faces chosen "at random", and used each of the resulting sets of axes to represent an unknown face. The corresponding approximations have been almost identical, suggesting we have succeeded in representing the "faceness" of the image.

Improvements and Implications One way to improve the face descriptions above is to add more detail, perhaps by concentrating on individual face features. This was originally suggested by Barron [Bar81], although he was unable to do this automatically. Working in parallel with us, Shackleton and Welsh [SW91] have implemented model-based PCA for eyes. Adding such a

Figure 3: First principal component (left) and conventional approximation (right)

set of co-ordinates to those described above would improve the capability of a recognition system. Recently Turk and Pentland [TP91] investigated the usefulness of principal components for recognition of ensemble faces.

The system provides a way of modelling the learning of new faces, and even forgetting old ones. While representing new face images, some cases will arise in which the representation is particularly bad. This is detected during the representation, giving the option to repeat the initial PCA either replacing the least descriptive of the initial set of faces with the new one or using an enlarged ensemble. Representations with the new PCA axes may have greater descriptive power; certainly they are more relevant to the set of faces being learned "now". However, faces already encoded using the original axes need re-coding. In the absence of the original image, fidelity will be lost, eventually resulting in lack of recognition of that face.

References

[Bar81]   Robert J. Baron. Mechanisms of human facial recognition. *International Journal of Man–Machine Studies*, 15:137–178, 1981.

[KS90]   M. Kirby and L. Sirovich. Application of the Karhunen-Loève procedure for the characterisation of human faces. *IEEE: Transactions on Pattern Analysis and Machine Intelligence*, 12(1):103–108, 1990.

[SW91]   Mark A Shakleton and W.J. Welsh. Classification of facial features for recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVIP-91)*, 1991.

[TCL90]   David Tock, Ian Craw, and J.Rowland Lishman. A knowledge based system for measuring faces. In *BMVC90 — Proceedings of the British Machine Vision Conference*, pages 401–406, 1990.

[TP91]   Matthew Turk and Alex Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.