M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 371
Appears in: *IEEE Conference on Computer Vision & Pattern Recognition, San Francisco, CA, June 1996.*

Bayesian Face Recognition using Deformable Intensity Surfaces

Baback Moghaddam, Chahab Nastar and Alex Pentland
Perceptual Computing Section, The Media Laboratory,
Massachusetts Institute of Technology
20 Ames Street, Cambridge MA 02139, U.S.A.

Abstract

We describe a novel technique for face recognition based on deformable intensity surfaces which incorporates both the shape and texture components of the 2D image. The intensity surface of the facial image is modeled as a deformable 3D mesh in $(x, y, I(x, y))$ space. Using an efficient technique for matching two surfaces (in terms of the analytic modes of vibration), we obtain a dense correspondence field (or *3D warp*) between two images. The probability distributions of two classes of warps are then estimated from training data: interpersonal and extrapersonal variations. These densities are then used in a Bayesian framework for image matching and recognition. Experimental results with facial data from the US Army FERET database demonstrate an increased recognition rate over the previous best methods.

1 Introduction

Current work in the area of image-based object modeling and visual recognition treats the shape and texture components of an object in a separate and often independent manner. The technique of extracting shape and forming a shape-normalized or "shape-free" grayscale component was suggested by Craw & Cameron [3], which used an eigenface technique on shape-free faces for matching and recognition. Recently Craw *et al.* [4] have done a study which combines these two independently derived components (a manually-extracted shape component plus a shape-free texture) for enhanced recognition performance. Similarly, Lanitis *et al.* [7] have developed an automatic face-processing system which is capable of combining the shape and texture components for recognition, albeit independently. Their system detects canonical points on the face and uses these landmarks to warp faces to a shape-free representation prior to implementing an eigenface technique for characterizing grayscale variations (face texture).

Similarly, the face vectorizer system of Beymer & Poggio [2] uses optical flow to obtain a shape representation decoupled from that of texture (in the form of a 2D correspondence field between a given face and a canonical model). However, one of the difficulties with using optical flow for correspondance between two different individuals is that the technique is inherently failure-prone when there are large grayscale variations between the images (*e.g.,* presence/absence of facial hair). A pixel correspondence technique must be able to deal with intensity variations as well as spatial deformations, preferably in a unified framework.

In this paper, we propose a novel representation which combines both the spatial (XY) and grayscale (I) components of the image into a 3D surface (or manifold) and then efficiently solves for a dense correspondence field in the XYI space. These image manifolds are modeled as physically-based deformable surfaces which undergo deformations in accordance with a specified force field. The physical dynamics of the system are efficiently solved for using a formulation in terms of the *analytic* modes of vibration [10]. This manifold matching technique can be viewed as a more general formulation for image correspondence which, unlike optical flow, does *not* require a constant brightness assumption. In fact, by simply disabling the I component of our deformations we can obtain a standard 2D deformable mesh which yields correspondences similar to an optical flow technique with thin-plate regularizers.

This novel image correspondence method is used to match two facial images by deforming the XYI surface of one image into the another (under "physical forces" exerted by nearby mesh nodes). The resulting vector of displacements yields a pixel-dense set of correspondences which can be used for image warping. In addition the vector of modal amplitudes is then used to classify the deformation into one of two categories: *interpersonal* vs. *extrapersonal.* This final classification is performed using the *a posteriori* probabilities computed from the two class-conditional likelihoods which are themselves estimated from training data using an efficient subspace method for density estimation of high-dimensional Gaussian data [9].

2 Deformable Intensity Surfaces

In previous work [13, 12], we formulated a novel image matching technique based on a 3D surface representation of an image $I(x, y)$ — *i.e.,* as the surface $(x, y, I(x, y))$ as shown, for example, in Figure 1 — and developed an efficient method to *warp* one image onto another using a physically-based deformation model. In this section we briefly review the mathematics of this approach (for further details the reader is referred to [11, 13, 12]).

The intensity surface is modeled as a deformable mesh and is governed by Lagrangian dynamics [1] :

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{C}\dot{\mathbf{U}} + \mathbf{K}\mathbf{U} = \mathbf{F}(t) \quad (1)$$

where $\mathbf{U} = [\ldots, \Delta x_i, \Delta y_i, \Delta z_i, \ldots]^T$ is a vector storing nodal displacements, $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are respectively the mass, damping and stiffness matrices of the system, and $\mathbf{F}$ is the external force. In warping one image onto a second (reference) image, the external force at each node M_i of

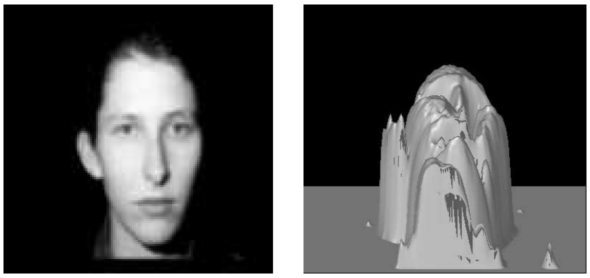


Figure 1: An image and its XYI surface representation

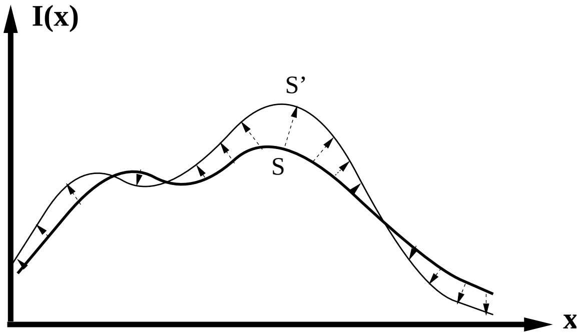


Figure 2: A cross-section of the intensity surface S being pulled towards S' by image forces

the mesh points to the closest 3D point P_i in the reference surface:

$$\mathbf{F}(t) = [\ldots, \overrightarrow{M_i P_i}(t), \ldots]^T \tag{2}$$

The final correspondence (and consequently the resultant XYI-warp) between two images is obtained by solving the governing equation above. Figure 2 shows a schematic representation of the deformation process. Note that the external forces (dashed arrows) do *not* necessarily correspond to the final displacement field of the surface. The elasticity of the surface provides an intrinsic smoothness constraint for computing the final displacement field.

We note that this formulation provides an interesting alternative to optical flow methods for obtaining correspondence, without the classical *brightness constraint* [5]. Indeed, the brightness constraint corresponds to a specific case of our formulation where the closest point P_i has to have the same intensity as M_i — *i.e.,* $\overrightarrow{M_i P_i}$ is parallel to the XY plane. We do not make that assumption here.

Solutions of the governing equation are typically obtained using an eigenvector-based *modal* decomposition [14, 11, 10]. In particular, the vibration modes $\phi(i)$ of the previous deformable surface are the vector solutions of the eigenproblem :

$$\mathbf{K}\phi = \omega^2 \mathbf{M}\phi \tag{3}$$

where $\omega(i)$ is the i-th eigenfrequency of the system. Solving the governing equations in the modal basis leads to scalar equations where the unknown $\tilde{u}(i)$ is the amplitude of mode i [1]

$$\ddot{\tilde{u}}(i) + \tilde{c}_i \dot{\tilde{u}}(i) + \omega(i)^2 \tilde{u}(i) = \tilde{f}_i(t) \qquad i = 1, \ldots, 3N. \tag{4}$$

The closed-form expression of the displacement field is then given by

$$\mathbf{U} \approx \sum_{i=1}^{P} \tilde{u}(i)\phi(i) \tag{5}$$

with $P \ll 3N$, which means that only P scalar equations of the type of (4) need to be solved. The modal superposition equation (5) can be seen as a Fourier expansion with high-frequencies neglected [10]. In our formulation, however, we make use of the *analytic modes* [10, 13], which are known sine and cosine functions for specific surface topologies

$$\phi(p, p') = [\ldots, \cos\frac{p\pi(2i-1)}{2n} \cos\frac{p'\pi(2j-1)}{2n'}, \ldots]^T \tag{6}$$

These analytic expressions avoid costly eigenvector decompositions and furthermore allow the total number of modes to be easily adjusted for the application.

The above modal analysis technique represents a coordinate transform from the nodal displacement space to the modal amplitude subspace:

$$\tilde{\mathbf{U}} = \mathbf{\Phi}^T \mathbf{U} \tag{7}$$

where $\mathbf{\Phi}$ is the matrix of analytic modes $\phi(p, p')$ and $\tilde{\mathbf{U}}$ is the resultant vector of modal amplitudes which encodes the type of deformations which characterize the difference between the two images. In addition, once we have solved for the resultant 3D displacement field we can then warp the original image onto the second in the XYI space and then render a resultant 2D image using simple computer graphics techniques. Figure 3 shows an example illustrating this warping process. We note that the warped image $I_{1\to 2}$ is only an incidental by-product of our correspondence method. Since our main goal is image matching and recognition we are primarily interested in the modal amplitude spectrum represented by $\tilde{\mathbf{U}}$.

3 Recognition: Bayesian Analysis of Deformations

We now consider the problem of characterizing the type of deformations which occur when matching two images in a face recognition task. We define two distinct and mutually exclusive classes: Ω_I representing *intrapersonal* variations between multiple images of the same individual (*e.g.,* with different expressions and lighting conditions), and Ω_E representing *extrapersonal* variations which result when matching two different individuals. We will assume that both classes are Gaussian-distributed and seek to obtain estimates of the likelihood functions $P(\tilde{\mathbf{U}}|\Omega_I)$ and $P(\tilde{\mathbf{U}}|\Omega_E)$ for a given deformation's modal amplitude vector $\tilde{\mathbf{U}}$.

Given these likelihoods we can define the "similarity score" between a pair of images directly in terms of the intrapersonal *a posteriori* probability as given by Bayes rule:

$$P(\Omega_I|\tilde{\mathbf{U}}) = \frac{P(\tilde{\mathbf{U}}|\Omega_I)P(\Omega_I)}{P(\tilde{\mathbf{U}}|\Omega_I)P(\Omega_I) + P(\tilde{\mathbf{U}}|\Omega_E)P(\Omega_E)} \tag{8}$$

where the priors $P(\Omega)$ can be set to reflect specific operating conditions (*e.g.,* number of test images *vs.* the size of the database) or other sources of *a priori* knowledge regarding the two images being matched.

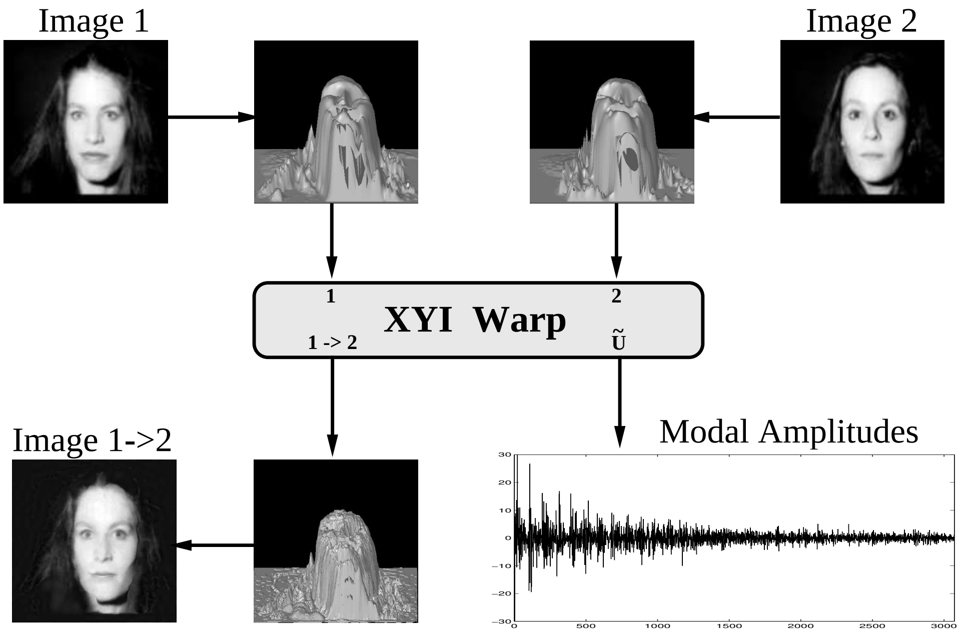


Figure 3: Example of XYI warping two images.

Additionally, this Bayesian formulation casts a face recognition task into a classical binary pattern classification problem which can then be solved using the maximum *a posteriori* (MAP) rule — *i.e.,* two facial images are determined to belong to the same individual if $P(\Omega_I|\tilde{\mathbf{U}}) > P(\Omega_E|\tilde{\mathbf{U}})$.

3.1 Statistical Modeling of Modal Amplitudes

One difficulty with this approach is that the modal amplitude vectors are very high-dimensional, with $\tilde{\mathbf{U}} \in \mathcal{R}^N$ with $N = O(10^3)$. Therefore we typically lack sufficient independent training observations to compute reliable 2nd-order statistics for the likelihood densities (*i.e.,* singular covariance matrices will result). Even if we were able to estimate these statistics, the computational cost of evaluating the likelihoods is formidable. Furthermore, this computation would be highly inefficient since the *intrinsic* dimensionality or major degrees-of-freedom of $\tilde{\mathbf{U}}$ for each class is likely to be significantly smaller than N.

Recently, an efficient density estimation method was proposed by Moghaddam & Pentland [9] which divides the vector space $\mathcal{R}^N$ into two complementary subspaces using an eigenspace decomposition. This method relies on a Principal Components Analysis (PCA) [6] to form a low-dimensional estimate of the complete likelihood which can be evaluated using only the first M principal components, where $M << N$. This decomposition is illustrated in Figure 4 which shows an orthogonal decomposition of the vector space $\mathcal{R}^N$ into two mutually exclusive subspaces: the principal subspace F containing the first M principal components and its orthogonal complement $\bar{F}$, which contains the residual of the expansion. The component in the orthogonal subspace $\bar{F}$ is the so-called "distance-from-feature-space" (DFFS), a Euclidean distance equivalent to the PCA residual error. The component of $\tilde{\mathbf{U}}$ which lies *in* the feature space F is referred to as the "distance-in-feature-space" (DIFS) and is a *Mahalanobis* distance for Gaussian densities.

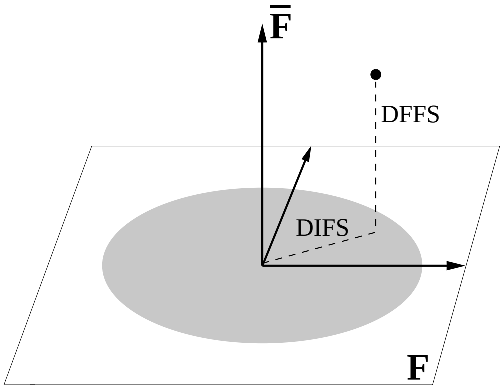
As derived in [9], the complete likelihood estimate can be written as the product of two independent marginal Gaussian densities (or equivalently as an appropriately weighted sum of the DIFS and DFFS)

$$\begin{aligned}\hat{P}(\tilde{\mathbf{U}}|\Omega) &= \left[\frac{\exp\left(-\frac{1}{2}\sum_{i=1}^{M}\frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{M/2}\prod_{i=1}^{M}\lambda_i^{1/2}}\right]\cdot\left[\frac{\exp\left(-\frac{\epsilon^2(\tilde{\mathbf{U}})}{2\rho}\right)}{(2\pi\rho)^{(N-M)/2}}\right] \\ &= P_F(\tilde{\mathbf{U}}|\Omega)\ \hat{P}_{\bar{F}}(\tilde{\mathbf{U}}|\Omega)\end{aligned} \tag{9}$$

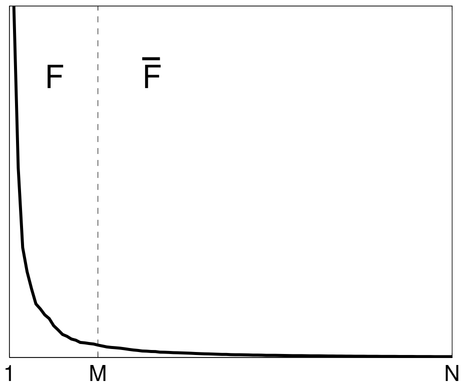
where $P_F(\tilde{\mathbf{U}}|\Omega)$ is the true marginal density in F, $\hat{P}_{\bar{F}}(\tilde{\mathbf{U}}|\Omega)$ is the estimated marginal density in the orthogonal complement $\bar{F}$, y_i are the principal components and $\epsilon^2(\tilde{\mathbf{U}})$ is the residual (or DFFS). The optimal value for the weighting parameter ρ is simply the average of the $\bar{F}$ eigenvalues

$$\rho = \frac{1}{N-M}\sum_{i=M+1}^{N}\lambda_i \tag{10}$$

We note that in actual practice, the majority of the $\bar{F}$ eigenvalues are unknown but *can* be estimated, for example, by fitting a nonlinear function to the available portion of the eigenvalue spectrum and estimating the average of the eigenvalues beyond the principal subspace.



(a)



(b)

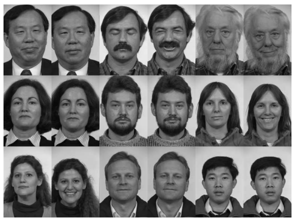
Figure 4: (a) Decomposition of $\mathcal{R}^N$ into the principal subspace F and its orthogonal complement $\bar{F}$ for a Gaussian density, (b) a typical eigenvalue spectrum and its division into the two orthogonal subspaces.

4 Experiments

To test our recognition strategy we used a collection of images from the US Army's FERET face database. This collection of images consists of hard recognition cases that have proven difficult for all face recognition algorithms previously tested on the FERET database. The difficulty posed by this dataset appears to stem from the fact that the images were taken at different times, at different locations, and under different imaging conditions.

The set of images consists of pairs of frontal-views and are divided into two subsets: the "gallery" (training set) and the "probes" (testing set). The gallery images consisted of 74 pairs of images (2 per individual) and the probe set consisted of 38 pairs of images, corresponding to a subset of the gallery members. These images are shown in Figure 5.

These images were subsequently aligned with an automatic face-processing system which extracts faces from the input image and normalizes for translation, scale as well as slight rotations (both in-plane and out-of-plane). This system is described in detail in Moghaddam & Pentland [9] and uses the same PCA-based density estimation technique described earlier to obtain maximum-likelihood (ML) estimates of object location (in this case the position and scale



(a)

(b)

Figure 5: Examples of FERET frontal-view image pairs used for (a) the Gallery set (training) and (b) the Probe set (testing).

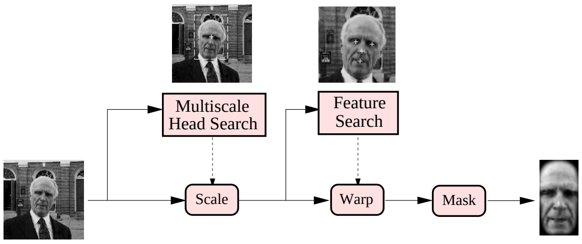


Figure 6: The face alignment system

of a face and the location of individual facial features).

The block diagram of this system is shown in Figure 6. The first step in this process is obtaining the ML estimate of the position and scale of the face (indicated by the cross-hairs and bounding box in the figure). Once these regions have been identified, the estimated scale and position are used to normalize for translation and scale, yielding a standard "head-in-the-box" image. Next a second feature detection stage operates at this fixed scale to estimate the position of four facial features: the left and right eyes, the tip of the nose and the center of the mouth. Once these facial features have been detected, the face image is transformed to align these features with those of a canonical model. Then the facial region is extracted (by applying a fixed mask) and subsequently normalized for contrast. The geometrically aligned and normalized image is then ready for comparison with other (similarly processed) faces.

4.1 Method 1: Eigenfaces

As a baseline comparison we first used a standard eigenface technique for recognition. The normalized images from the gallery and the probe sets were projected onto a 100-dimensional eigenspace and a nearest-neighbor rule based on a Euclidean distance metric was used to match each probe image to a gallery image. A few of the lower-order eigenfaces used for this projection are shown in Figure 7. We note that these eigenfaces represent the principal components of an entirely different set of images — *i.e.,* none of the individuals in the gallery or probe sets were used in obtaining these eigenvectors. In other words, neither the gallery nor the probe sets were part of the "training set."

The rank-1 recognition rate obtained with this method was found to be 84% (64 correct matches out of 76), and the correct match was always in the top 10 nearest neighbors. Note that this performance is better than or similar to recognition rates obtained by any algorithm tested on this database, and that it is lower (by about 10%) than

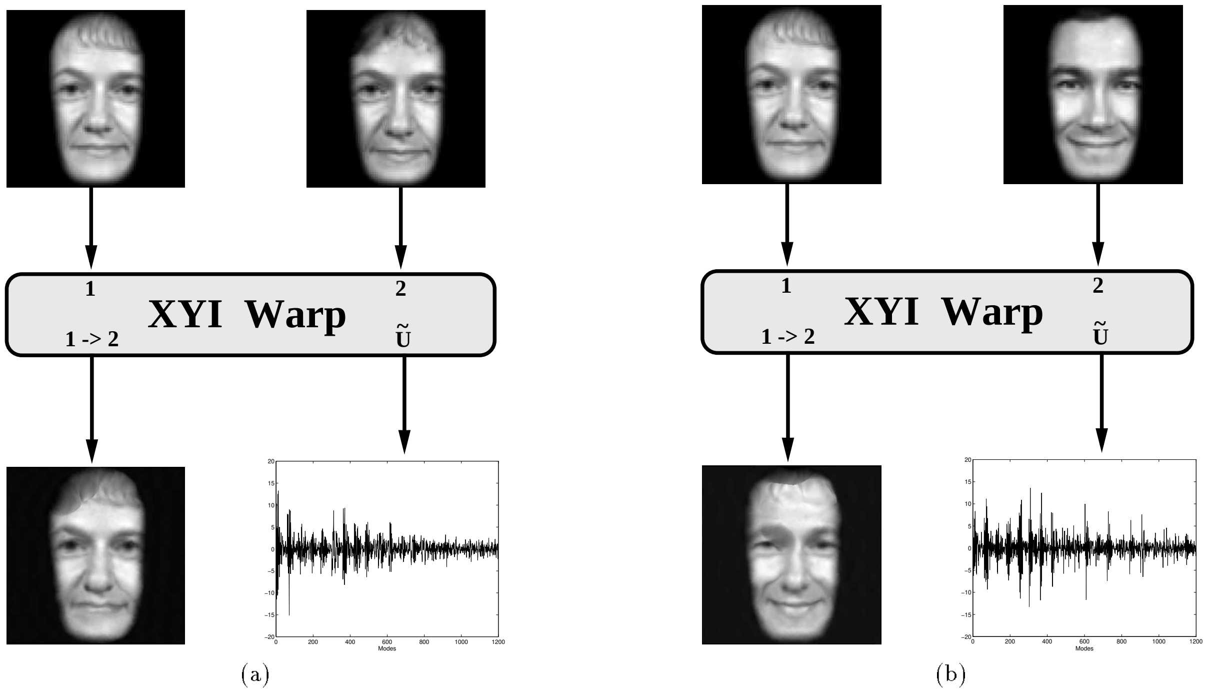


Figure 8: Examples of (a) intrapersonal and (b) extrapersonal facial warps.

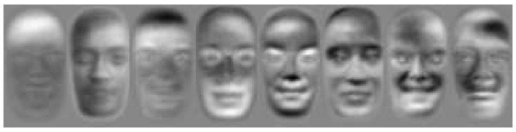


Figure 7: The first 8 eigenfaces.

the typical rates that we have obtained with the FERET database [8]. We attribute this lower performance to the fact that these images were selected to be particularly challenging. In fact, using an eigenface method to match the first views of the 76 individuals in the gallery to their second views, we obtain a higher recognition rate of 89% (68 out of 76), suggesting that the gallery images represent a less challenging data set since these images were taken at the same time and under identical lighting conditions.

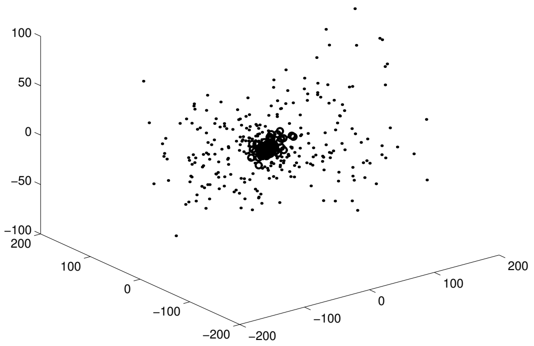
4.2 Method 2: XYI Warps

For our Bayesian method, we first gathered training data by computing the modal amplitude spectra for a training subset of 74 intrapersonal warps (by matching the two views of every individual in the gallery) and a random subset of 296 extrapersonal warps (by matching images of *different* individuals in the gallery), corresponding to the classes Ω_I and Ω_E, respectively. An example of each of these two types of warps is shown in Figure 8.

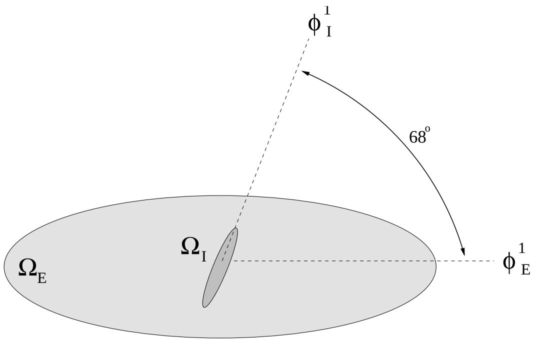
It is interesting to consider how these two classes are distributed; for example, are they linearly separable or embedded distributions? One simple method of visualizing this is to plot their mutual principal components — *i.e.,* perform PCA on the *combined* dataset and project each vector onto the principal eigenvectors. Such a visualization is shown in Figure 9(a) which is a 3D scatter plot of the first 3 principal components. This plot shows what appears to be two completely enmeshed distributions, both having near-zero means and differing primarily in the amount of scatter, with Ω_I displaying smaller modal amplitudes as expected. It therefore appears that one can not reliably distinguish low-amplitude extrapersonal warps (of which there are many) from intrapersonal ones.

However, direct visual interpretation of Figure 9(a) is very misleading since we are essentially dealing with low-dimensional (or "flattened") hyper-ellipsoids which are intersecting near the origin of a very high-dimensional space. The key distinguishing factor between the two distributions is their relative orientation. Fortunately, we can easily determine this relative orientation by performing a separate PCA on each class and computing the dot product of their respective first eigenvectors. This analysis yields the cosine of the angle between the major axes of the two hyper-ellipsoids, which was found to be 68°, implying that the principal orientations of the two hyper-ellipsoids are indeed quite different. Figure 9(b) is a schematic illustration of the geometry of this configuration, where the hyper-ellipsoids have been drawn to approximate scale using the corresponding eigenvalues.

We note that since these classes are not linearly separable, simple linear discriminant techniques (*e.g.,* hyper-planes) can not be used with any degree of reliability. The proper decision surface is inherently nonlinear (quadratic, in fact, under the Gaussian assumption) and is best defined in terms of the *a posteriori* probabilities — *i.e.,* by the equality $P(\Omega_I|\tilde{\mathbf{U}}) = P(\Omega_E|\tilde{\mathbf{U}})$. Fortunately, the optimal discriminant surface is automatically implemented when invoking a MAP classification rule.



(a)



(b)

Figure 9: (a) distribution of the two classes in the first 3 principal components (circles for Ω_I, dots for Ω_E) and (b) schematic representation of the two distributions showing orientation difference between the corresponding principal eigenvectors.

Having analyzed the geometry of the two distributions, we then computed the likelihood estimates $P(\tilde{\mathbf{U}}|\Omega_I)$ and $P(\tilde{\mathbf{U}}|\Omega_E)$ using the PCA-based method outlined in Section 3.1. We selected principal subspace dimensions of $M_I = 10$ and $M_E = 30$ for Ω_I and Ω_E, respectively. These density estimates were then used with a default setting of equal priors, $P(\Omega_I) = P(\Omega_E)$, to evaluate the *a posteriori* intrapersonal probability $P(\Omega_I|\tilde{\mathbf{U}})$ for matching probe images to those in the gallery.

In order to avoid an unnecessarily large number of XYI warps, we only matched a probe image to the top 10 gallery images retrieved by the eigenface method. This significantly reduces the computational cost of our system, since computing eigenface similarity scores is negligible compared to computing XYI warps (the former requires several milliseconds whereas the latter takes approximately 20 seconds on an HP 735 workstation).

Therefore, for each probe image we computed a set of 10 probe-to-gallery warps and re-sorted the matching order, this time using the *a posteriori* probability $P(\Omega_I|\tilde{\mathbf{U}})$ as the similarity metric. This probabilistic ranking yielded an improved rank-1 recognition rate of 92% (70 out of 76). Furthermore, out of the 608 extrapersonal warps performed in this recognition experiment, only 2% (11) were misclassified as being intrapersonal — *i.e.,* with $P(\Omega_I|\tilde{\mathbf{U}}) > P(\Omega_E|\tilde{\mathbf{U}})$.

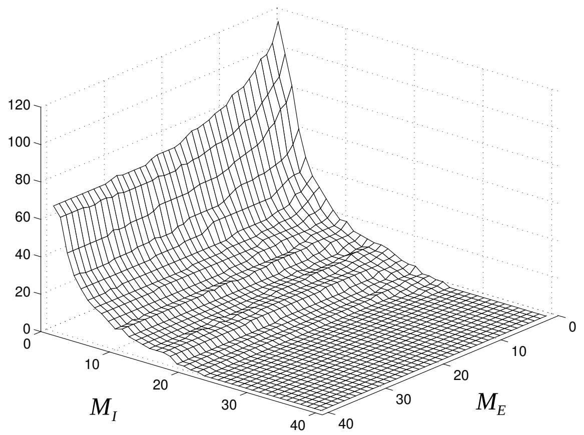


Figure 10: Total number of misclassified extrapersonal matches (with $P(\Omega_I|\tilde{\mathbf{U}}) > 0.5$) as a function of the principal subspace dimensionalities M_I and M_E.

We also analyzed the sensitivity of our Bayesian matching technique with respect to the principal subspace dimensionalities M_I and M_E, which are used in estimating the likelihoods $P(\tilde{\mathbf{U}}|\Omega_I)$ and $P(\tilde{\mathbf{U}}|\Omega_E)$. The higher we set these parameters the more accurate an estimate of the likelihoods we obtain, while also requiring more principal projections. These parameters therefore represent an accuracy *vs.* complexity tradeoff in our Bayesian approach. To quantify this tradeoff, we repeated the probe set recognition experiment while varying both parameters and noted that the recognition rate never dropped below 79%, even when the two subspaces used in estimating the likelihoods were as low as one-dimensional. However, we noted that the total number of extrapersonal matches which were misclassified as being intrapersonal — *i.e.,* $P(\Omega_I|\tilde{\mathbf{U}}) > P(\Omega_E|\tilde{\mathbf{U}})$ — varied in a systematic way with the subspace dimensionalities. This variation is shown in Figure 10 and is clearly the type of behavior one would expect: the total number of misclassified matches decreases with increasing subspace dimensionalities. From the figure, it is apparent that these errors are more sensitively dependent on M_I, the dimensionality of the intrapersonal subspace (possibly because this class has a much lower *intrinsic* dimensionality and its distribution can be modeled using fewer principal eigenvectors).

Finally, we note that the set of training warps used in forming an estimate of the intrapersonal density $P(\tilde{\mathbf{U}}|\Omega)$ consisted entirely of gallery images. In retrospect, this is probably an insufficient training set, since as argued in the previous section, a gallery-only set of intrapersonal training samples can not present adequate examples for learning temporal and/or lighting variations in appearance. We plan to further investigate this phenomenon by estimating $P(\tilde{\mathbf{U}}|\Omega)$ from a suitably heterogeneous FERET training set which includes multiple images of the same individual separated in time by days, weeks and months.

5 Conclusions

Template-based approaches to face recognition have been shown to be the most successful techniques to date, however they require multiple templates to span the space of typical variations in expression and pose. One of the distinct advantages of our approach is that by learning the statistics of interpersonal variations we do not require multiple templates for every individual: a single-view prototype supplemented with a probabilistic model of intrapersonal shape/texture deformations seems to be adequate to capture variations due to expression, some pose change, and lighting. This has resulted in higher recognition accuracies.

The disadvantage of our approach is that it is more computationally expensive. To reduce this expense, we have used our standard eigenface approach to quickly retrieve a small number of potential matches from the gallery; in our experiments the correct match is almost always within such a list of top candidate matches. These candidates are then XYI warped to determine the final ranking based on the intrapersonal *a posteriori* probability. By using this hybrid approach we were able to keep the total processing time under three minutes per query on a HP 735 workstation.

Finally, we should point out that this general warping method is applicable to data of higher dimensionality. For instance, the Cyberware scanner produces XYZI data. The techniques described here are directly generalizable to this type of data, and we are now pursuing this avenue of research.

References

[1] K. J. Bathe. *Finite Element Procedures in Engineering Analysis.* Prentice-Hall, 1982.

[2] David Beymer. Vectorizing face images by interleaving shape and texture computations. A.I. Memo No. 1537, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1995.

[3] I. Craw and P. Cameron. Face recognition by computer. In D. Hogg and R. Boyle, editors, *Proc. British Machine Vision Conference*, pages 498–507. Springer-Verlag, 1992.

[4] I. Craw and et al. Automatic face recognition: Combining configuration and texture. In Martin Bichsel, editor, *Proc. Int'l Workshop on Automatic Face and Gesture Recognition*, Zurich, 1995.

[5] B.K.P. Horn and G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[6] I.T. Jolliffe. *Principal Component Analysis.* Springer-Verlag, New York, 1986.

[7] A. Lanitis, C. J. Taylor, and T. F. Cootes. A unified approach to coding and interpreting face images. In *IEEE Proceedings of the Fifth International Conference on Computer Vision (ICCV'95)*, Cambridge, MA, June 1995.

[8] B. Moghaddam and A. Pentland. Face recognition using view-based and modular eigenspaces. *Automatic Systems for the Identification and Inspection of Humans*, 2277, 1994.

[9] B. Moghaddam and A. Pentland. Probabilistic visual learning for object detection. In *IEEE Proceedings of the Fifth International Conference on Computer Vision (ICCV'95)*, Cambridge, USA, June 1995.

[10] C. Nastar. Vibration modes for nonrigid motion analysis in 3D images. In *Proceedings of the Third European Conference on Computer Vision (ECCV '94)*, Stockholm, May 1994.

[11] C. Nastar and N. Ayache. Fast segmentation, tracking, and analysis of deformable objects. In *IEEE Proceedings of the Third International Conference on Computer Vision (ICCV'93)*, Berlin, May 1993.

[12] C. Nastar, B. Moghaddam, and A. Pentland. Generalized image matching: Statistical learning of physically-based deformations. In *Proceedings of the Fourth European Conference on Computer Vision (ECCV'96)*, Cambridge, UK, April 1996.

[13] C. Nastar and A. Pentland. Matching and recognition using deformable intensity surfaces. In *IEEE International Symposium on Computer Vision*, Coral Gables, USA, November 1995.

[14] A. Pentland and S. Sclaroff. Closed-form solutions for physically based shape modelling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(7):715–729, July 1991.