

Face Recognition: Would Going Back to Functional Nature Be a Good Idea?

Noslen Hernández, Yoanna Martínez-Díaz, Dania Porro-Muñoz,
and Heydi Méndez-Vázquez

Advanced Technologies Application Center, Havana, Cuba
{nhernandez,ymartinez,dpmunoz,hmendez}@cenatav.co.cu

Abstract. Traditional appearance-based methods for face recognition represent raw face images of size $u \times v$ as vectors in a $u \times v$ -dimensional space. However in practice, this space can be too large to perform classification. For that reason, dimensionality reduction techniques are usually employed. Most of those traditional approaches do not take advantage of the spatial correlation of pixels in the image, considering them as independent. In this paper, we proposed a new representation of face images that takes into account the smoothness and continuity of the face image and at the same time deals with the dimensionality of the problem. This representation is based on Functional Data Analysis so, each face image is represented by a function and a recognition algorithm for functional spaces is formulated. The experiments on the AT&T and Yale B facial databases show the effectiveness of the proposed method.

Keywords: Face recognition, functional data analysis, biometrics.

1 Introduction

Face recognition has received significant attention due to its potential for a wide variety of applications [1]. Within the last several years, numerous face recognition algorithms have been developed [1,2]. Traditional face recognition methods are based on the fact that a class of patterns of interest, i.e. the face, resides in a subspace of the input image space [1]. The high dimensionality of face images is among the challenges that these algorithms have faced. In the context of face recognition, a $u \times v$ gray-scale image is usually identified with the vector $x \in R^{u \times v}$ given by stacking its columns. Performing face recognition in such a high dimensional space might result in several difficulties related to the curse of dimensionality. This matter has been addressed by transforming images to low-dimensional vectors in the face space. Different approaches have been introduced to achieve dimension reduction. They range from traditional approaches such as Principal Component Analysis (PCA) [4], Linear Discriminant Analysis (LDA) [5], Independent Component Analysis (ICA) [6] to more unconventional features such as downsampled images and random projections [7].

Many real-world data, as face images, have this high-dimensional nature. Some of these high-dimensional data usually have a functional nature. What does

this mean? Although in practice, this data is observed and record discretely, the nature of the true process is not a finite vector of measurements, but an underlying continuous function that it is only measured at discrete points. The term functional refers to the intrinsic structure of the data rather than to their explicit form. When this is the case, Functional Data analysis (FDA) is a common way to overcome the effects of the curse of dimensionality [8,9].

Functional Data Analysis is a relatively new and growing research field that deals with the statistical description and modeling of such kind of (functional) data [8,9]. The basic rationale behind functional data analysis is that we should think of observed data functions as single entities described by a continuous real-valued function rather than merely a sequence of individuals observations. In this way, functional data are then supposed to have values in an infinite-dimensional space, often particularized as a Hilbert space.

In addition to handle the problem of dimensionality, why FDA can be considered a natural approach for representing and analyzing face images? A face image (and any image, in general) is a mathematical representation of a physical observation as a function over a spatial domain. Therefore we would like to treat a face image as defined over a continuous spatial domain, not as a mere collection of pixel values. All the aforementioned approaches used a classical discretization of the data as a sequence of numbers and loose some functional characteristics of the data like smoothness and continuity. On the contrary, FDA aggregates consecutive discrete measurements and views them as sampled values of a random variable, keeping track of order or smoothness.

The purpose of this paper is twofold. Firstly, to propose a new representation of face images using the functional data analysis framework. The central idea of this representation is just to describe a gray-scale face image as an observation of a functional random variable. And second, to formulate an algorithm for face identification using that functional representation.

The introduced face recognition algorithm can be categorized as a Nearest Subspace approach. It consists in a generalization to functional spaces of the ideas proposed by [10]. In this way, the proposed formulation lies on the key assumption that patterns from the same class lie on a linear subspace [5,11]. It is shown how our intuition for orthogonality and projections in R^m works fine in general Hilbert spaces which makes the formulation straightforward.

The remainder of the paper is organized as follows. Section 2 introduces the functional face recognition algorithm and describes how face images can be represented as functions. Section 3 analyzes the performance of the introduced approach on two standard public available face databases. Finally, some conclusions and future works are drawn in section 4.

2 Face Recognition in Functional Space

Let (X, Y) be a pair of random variables taking values in $\mathcal{X} \times \{1, 2, \dots, N\}$, where $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ is the space of square integrable functions from $[a, b] \times [c, d]$ to \mathbb{R} (i.e., $L_2([a, b] \times [c, d])$) and $\{1, 2, \dots, N\}$ represents the class labels for X , with

$N \in \mathbb{N}$ the number of distinguished classes. Suppose also that n independent and identically distributed (i.i.d) realizations of (X, Y) are given (training data), denoted by $(x_i, y_i), i = 1, \dots, n$. The problem of interest is to build, from $(x_i, y_i)_i$, a predictor of the value of Y corresponding to a future observed value of X .

For convenience, the training data $(x_i, y_i)_i$ will be organized according to the class membership. Suppose then that, for each class $j \in \{1, \dots, N\}$, n_j realizations of X are given, denoted by $(x_i^j), i = 1, \dots, n_j$. This means that in the training data, each person j constitutes a class and we will have n_j images of this person.

For each class j , $(x_i^j)_i$ spans a (closed) finite-dimensional linear subspace M_j of the Hilbert space \mathcal{X} . Based on the concept that patterns from the same class lie on a linear subspace [5,11], any future observation x_0 belonging to class j must lie on M_j . Therefore, an estimate \hat{y}_0 of the class label y_0 corresponding to a new observation x_0 , can be obtained by looking for, on each subspace M_j , the element $x_0^j \in M_j$ closest to x_0 (in the sense that it minimizes $d_j = \|x_0 - x_0^j\|$) and then, choosing the j for which d_j is minimized. Formally, this can be written as:

$$\hat{y}_0 = \min_{j \in \{1, \dots, N\}} \|x_0 - x_0^j\|, \quad (1)$$

where $\|\cdot\|$ is the norm in \mathcal{X} induced by the inner product.

We will see that such best approximations $x_0^j \in M_j$ exist, are unique and can be characterized in a way that they can be computed. It is known that if $(x_i^j)_i$ are elements of the Hilbert space \mathcal{X} , and $M_j = \text{span}\{x_i^j\}$ is a subspace of \mathcal{X} , for an arbitrary $x_0 \in \mathcal{X}$, there exist a unique $x_0^j \in M_j$ such that $\|x_0 - x_0^j\| = \inf_{x \in M_j} \|x_0 - x\|$. This unique minimizing vector is the orthogonal projection $x_0^j = P_{M_j}(x_0)$ of x_0 onto M_j .

The projection x_0^j can be written as $x_0^j = \sum_{k=1}^{n_j} \alpha_k^j x_k^j, \alpha_k^j \in \mathbb{R}$, and must satisfy $\langle x_0 - x_0^j, x_k^j \rangle = 0$, for $k = 1, \dots, n_j$, hence the $(\alpha_k^j)_k$ can be estimated by solving the following system of linear equations, called normal equations:

$$G^j(x_1^j, x_2^j, \dots, x_{n_j}^j)\alpha^j = \beta^j, \quad (2)$$

where

$$G^j(x_1^j, x_2^j, \dots, x_{n_j}^j) = \begin{pmatrix} \langle x_1^j, x_1^j \rangle & \langle x_2^j, x_1^j \rangle & \dots & \langle x_{n_j}^j, x_1^j \rangle \\ \langle x_1^j, x_2^j \rangle & \langle x_2^j, x_2^j \rangle & \dots & \langle x_{n_j}^j, x_2^j \rangle \\ \dots & \dots & \dots & \dots \\ \langle x_1^j, x_{n_j}^j \rangle & \langle x_2^j, x_{n_j}^j \rangle & \dots & \langle x_{n_j}^j, x_{n_j}^j \rangle \end{pmatrix}$$

is called the Gram matrix of $\{x_1^j, x_2^j, \dots, x_{n_j}^j\}$, and α^j and β^j are the column vectors $\alpha^j = (\alpha_1^j, \alpha_2^j, \dots, \alpha_{n_j}^j)^T$ and $\beta^j = (\langle x_0, x_1^j \rangle, \langle x_0, x_2^j \rangle, \dots, \langle x_0, x_{n_j}^j \rangle)^T$. Note that, although the normal equations do not possess a unique solution if the x_i^j 's are linearly dependent, there is always at least one solution. Thus the degeneracy that could arise as a result of $\det(G^j) = 0$ always results in a multiplicity of solutions rather than an inconsistent set of equations.

Once estimated α^j for each class j , the projections x_0^j can be obtained and consequently the class \hat{y}_0 , by solving (1).

2.1 Functional Approximation of Face Images

The previous section presented the theoretical formulation of the introduced face recognition algorithm for an arbitrary Hilbert space. It is based on the fact that the face image is represented as a function $x \in L_2([a, b] \times [c, d])$. However, as stated in the introduction, what we observed in practice is a discretization $\tilde{x} \in \mathbb{R}^{u \times v}$ of the continuum data x . So, we need to construct an approximation of the functions x . This is usually done by a projection approach. Each image is approximated (smoothed) by a weighted sum (a linear combination) of basis functions, and the coefficients of the expansion are determined by fitting data by least squares. Each image is therefore, completely determined by the coefficients on this basis and each function is computable for any desired argument value (h, w) .

Since we are interested in approximating functions of two variables, we need bivariate basis functions for the expansion. One way for obtaining such bivariate basis functions is through tensor product of univariate basis. In this way we can approximate the function $x(h, w)$ by

$$x(h, w) = \sum_i^p \sum_j^q c_{ij} g_{ij}(h, w) \quad (3)$$

where $g_{ij}(h, w) = e_i(h)f_j(w)$ is a basis on $L_2([a, b] \times [c, d])$, and $(e_i)_i, (f_j)_j$ are basis on $L_2[a, b]$ and $L_2[c, d]$, respectively.

We need not only that (3) approximates well the original function but also that the operations performed on the reconstructed functions approximate as exactly as possible the corresponding operations on the original functions. Let see how computing inner products and norms (integrals) between functions, which are mainly the involved operations in the algorithm, can be easily calculated.

Let E and F be the matrices defined by $E_{ij} = \langle e_i, e_j \rangle, i, j = 1, \dots, p$ and $F_{kl} = \langle f_k, f_l \rangle, k, l = 1, \dots, q$, and $E = R'_e R_e, F = R'_f R_f$ their corresponding Cholesky decompositions. It can be proved (we will not show this due to space limitations) that the inner product between two functions $x(h, w)$ (as expressed in (3)) and $z(h, w) = \sum_i^p \sum_j^q a_{ij} g_{ij}(h, w)$, with coefficients matrices $C = (c_{ij})$ and $A = (a_{ij})$ respectively, is equal to:

$$\langle x, z \rangle = \sum_i \sum_j \tilde{c}_{ij} \tilde{a}_{ij}, \quad (4)$$

where $\tilde{C} = (\tilde{c}_{ij}) = R'_e C R_f$ and $\tilde{A} = (\tilde{a}_{ij}) = R'_e A R_f$. With this, we have found an effective way to perform the operations between the functional approximations of face images involved in the algorithm.

3 Experimental Results

To validate the proposal, experiments were conducted on two standard public available face databases: AT&T (also known as ORL) [12] and Extended Yale B

[13]. The introduced Functional Near Subspace Classification (FNSC) algorithm is compared with existing methods that have been tested on these databases. The results reported in all of our experiments are given by the recognition rate, which measures the percent of face images well classified.

For the experiments on both databases, we shall use a tensor product of univariate B-spline basis of order 4 (which gives a suitable smooth approximation of the faces) over the rectangular domain of interest. An advantage of B-splines is their local character. B-splines capture the information locally, which is suitable for the problem at hand.

3.1 Results on the AT&T Database

The AT&T database consists of 400 face images from 40 subjects. It has 10 images per person with different variations in expressions and deviations from the frontal pose up to 20 degrees. Face images from some of the subjects have additional changes such as the use of glasses. We adopt here two evaluation protocols [10]. The first protocol (EP1) uses the first five images of each person for training, leaving the remaining five for testing. The second protocol (EP2) consists on the “leave-one-out” cross-validation strategy. In both experiments we have used the images of size 92×112 pixels as provided, without any further processing, i.e., no geometric or photometric normalization is applied.

As described in Section 2, we first have to represent face images through B-spline basis expansion. For this, we need to select the number of basis functions ($p \times q$ in Section 2.1) to be used. The choice of the number of basis is very important because it can be regarded as a smoothing parameter. Statistically, keeping a few coefficients in the expansion is equivalent to conducting heavy amount of smoothing for the original data. We have selected here the number of basis functions by expert knowledge (comparing row data images with its reconstructions) and by trial and error. A more reasoned approach like cross-validation could have been used in this step.

The number of basis functions used in this database is 49, i.e., $p = 7$ and $q = 7$. This implies that the image is characterized only by 49 coefficients in the basis expansion, which considerably reduce the dimension of the problem. We will not refer to this vector of coefficients as feature vector, because they are not going to be new features for us. There is some subtle but important difference that should be accounted for. Our data, is still a function and we operate on it like this: we have formulated an algorithm for that, we are performing norms and inner product in a functional sense. But in practice, as the calculations are reduced to operate on these coefficient matrices, we can argue that the dimension of the problem has been reduced to that number of coefficients.

Table 1 shows the results obtained on both evaluation protocols. For EP1, the introduced approach does not achieve the best results, but its recognition accuracy of 95.5% is comparable with the other methods based on subspace analysis. Only the 2D-PCA and the Eigenfeature Regularization and Extraction (ERE) algorithms outperform the proposal by a 0.5 % and 1.5 %, respectively. For the second protocol, FNSC reports the best results.

Table 1. Recognition rates on AT&T database

Method	EP1	EP2
	Recognition Rate	Recognition Rate
PCA [14]	93.50 %	97.50 %
LDA [14]	94.50 %	98.50 %
ICA [14]	85.00 %	93.50 %
Kernel-PCA [14]	94.00 %	98.00 %
2D-PCA [14]	96.00 %	98.30 %
ERE [15]	97.00 %	99.25 %
LRC [10]	93.50 %	98.75 %
FNCS	95.50 %	99.50 %

3.2 Results on the Extended Yale B Database

The Extended Yale B database contains images of 38 subjects seen under 64 different illumination conditions, in which the angle between the light source direction and the camera axis was changed each time, in a way that the larger the angle, the more unfavorable the lighting conditions are. This database is usually divided into 5 subsets according to the angle of the incident illumination. Subset S1 is composed by 7 face images per subject with frontal or almost frontal incident lighting and is used as gallery. Subset S2 is composed by 456 images (12 per person) with incident lighting angle between $13^{\circ} - 25^{\circ}$, S3 have also 456 images with angles between $26^{\circ} - 50^{\circ}$, 532 images (14 per subject) with angles between $51^{\circ} - 70^{\circ}$ are in S4 and S5 contains 722 images (19 per person) with angles greater than 70° . Also in this case, the provided cropped version of the original images of size 168×192 are used without any pre-processing.

Figure 1 shows (for an image of this database) the original image and its functional representation. Also for this database, the number of basis functions was selected by expert knowledge to be $p = 31$ and $q = 35$, which gives a total of 1085 coefficients.

Also in this case, the dimensionality of the problem is considerably reduced. As was explained before, the number of coefficients chosen has the role of a smoothing parameter. This can be observed in this figure, where the functional

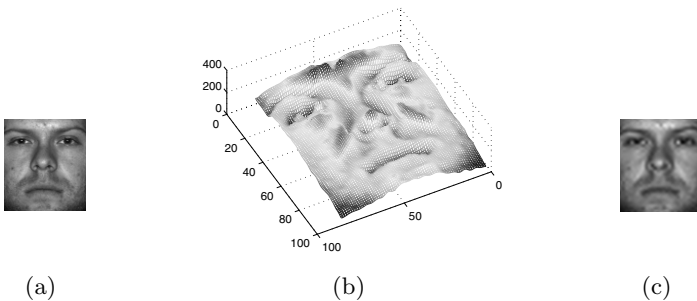


Fig. 1. a) Original image and b-c) its functional approximation

Table 2. Recognition rates on Extended Yale B database

Method	S2	S3	S4	S5
PCA [10]	98.46 %	80.04 %	15.79 %	24.38 %
ICA [10]	98.03 %	80.70 %	15.98 %	22.02 %
LRC [10]	100.0 %	100.0 %	83.27 %	33.61 %
FNCS	100.0 %	100.0 %	87.97 %	38.50 %

approximation shown in Figure 1(c) is smoother than the original image that appears in Figure 1(a).

The Extended Yale B database contains large illumination variations and it is usually used to test face recognition methods in front of this kind of problems. We are not addressing here the illumination problem and we have not applied any photometric normalization, although this could improve the results.

Table 2 shows the obtained results and comparisons with the Linear Regression Classification (LRC) algorithm and the other methods reported in [10]. The FNCS method reached the best results, achieving 100% of recognition rate in subset 2 and subset 3 and outperforming LRC in almost 5% on subsets 4 and 5.

4 Further Issues and Conclusions

This paper presents a new face recognition method based on Functional Data Analysis. First, a new representation of face images is proposed in which images are considered observations of functional random variables taking values in a Hilbert space. Second, a face recognition algorithm based on Near Subspace approach was generalized to functional spaces. Unlike the majority of the appearance-based methods for face recognition, our proposal attempts to take into account the smoothness and continuity of the face image.

Experiments were conducted on two popular face databases, AT&T and Yale B. Comparisons were done with previous approaches, and it was shown that the proposed face recognition method performs very well.

Although the obtained results are satisfactory, there are still many issues to investigate on the functional method proposed in this paper. One of these, is the selection of the basis. What is the more appropriate basis to represent face images? Is there a criteria to choose such a basis? Also, a two-dimensional basis can be obtained not only through tensor products of univariate basis. It is possible to use finite element analysis or thin-plate spline method to define bivariate basis over more complicated domains, such as non uniform grid. This new domain can be for example a Delaunay triangulation that takes better into account the structure of the face.

Another slightly more ambitious aspects would be to incorporate other kind of information in the construction of the functional approximation of the face image, not only the pixels intensity values. This would make the representation more robust to different problems affecting the image, like illumination variations.

Despite all that remains to be done or can be improved in this approach, we think that going back to the functional nature of face images could be very promising in the face recognition tasks, would you?

References

1. Jain, A.K., Li, S.Z.: Handbook of Face Recognition. Springer-Verlag New York, Inc., Secaucus (2005)
2. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. *ACM Comput. Surv.* 35, 399–458 (2003)
3. Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. *Journal of Information Processing Systems* 5(2) (2009)
4. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cognitive Neuroscience* 3(1), 71–86 (1991)
5. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 711–720 (1997)
6. Yuen, P.: Face representation using independent component analysis. *Pattern Recognition* 35(6), 1247–1257 (2002)
7. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 31(2), 210–227 (2009)
8. Ramsay, J., Silverman, B.: Functional data analysis. Springer series in statistics. Springer (2005)
9. Ferraty, F., Vieu, P.: Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics. Springer-Verlag New York, Inc., Secaucus (2006)
10. Naseem, I., Togneri, R., Bennamoun, M.: Linear regression for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 32(11), 2106–2112 (2010)
11. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 218–233 (2003)
12. Samaria, F.S., Harter, A.C.: Parameterisation of a stochastic model for human face identification. In: *Second IEEE Workshop Applications of Computer Vision* (1994)
13. Lee, K.-C., Ho, J., Kriegman, D.J.: Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(5), 684–698 (2005)
14. Yang, J., Zhang, D., Frangi, A.F., Yang, J.-y.: Two-dimensional pca: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 131–137 (2004)
15. Jiang, X., Mandal, B., Kot, A.: Eigenfeature regularization and extraction in face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 30(3), 383–394 (2008)