

3D Face Recognition by Functional Data Analysis

Dania Porro-Muñoz¹, Francisco José Silva-Mata¹, Anier Revilla-Eng¹,
Isneri Talavera-Bustamante¹, and Stefano Berretti²

¹ Advanced Technologies Application Center, CENATAV, Havana, Cuba

² University of Florence, Florence, Italy

{dporro, fjsilva, arevilla, italavera}@cenatav.co.cu,
{stefano.berretti}@unifi.it

Abstract. This work proposes the use of functional data analysis to represent 3D faces for recognition tasks. This approach allows exploiting and studying characteristics of the continuous nature of this type of data. The basic idea of our proposal is to approximate the 3D face surface through an expansion of a basis functions set. These functions are used for a *global* representation of the entire face, and a *local* representation, where pre-selected face regions are used to construct multiple local representations. In both cases, the functions are fitted to the 3D data by means of the least squares method. Univariate attribute selection is finally applied to reduce the dimensionality of the new representation. The experiments prove the validity of the proposed approach, showing competitive results with respect to the state of the art solutions. Moreover, the dimensionality of the data is considerably reduced with respect to the original size, which is one of the goals of using this approach.

Keywords: 3D face recognition, functional data analysis.

1 Introduction

Biometrics solutions based on face recognition have been studied for a long time using 2D still images and videos. More recently, face recognition based on 3D scans has been experimented as an alternative or complementary approach, which can improve recognition especially in the presence of facial expressions, illumination and pose changes, or spoofing attacks [3,13]. In order to perform recognition by means of scans matching, an attractive solution is represented by functional data analysis [4]. This approach has some potential advantages: the functional representation is capable of capturing the continuous behavior and the dynamic aspects of the original data; the data can be represented as a whole; the most significant features of the function, such as monotony, differentiability and smoothness can be analyzed. An additional interesting property of this representation is that a function not only provides the coordinates of all the points in the domain, but also expresses the relationship between them, an aspect that is commonly lost in vector based representations. However, finding the functions that better approximate the discrete data is a key challenge.

The use of functional basis expansion is a common way to approach the problem, but the selection of the basis set and the fitting, including the selection of the coefficients must be performed in order to obtain a discriminative representation with a small dimension. On the other hand, determining the position of the origin and the coordinate axes with respect to which the representation should be constructed, has a great influence on the accuracy of the functional representation.

Based on these premises, in this paper we propose the use of functional data analysis for representing 3D faces for recognition purposes. To this end, 3D face scans are first aligned to a common reference model, so as to normalize their pose. Then, the cloud of points of a scan is regarded as providing samples of a surface modeled by a function over a spatial domain. Based on the characteristics of the data and the results of state of the art studies [1], Bivariate Splines [9] on a rectangular domain, and Zernike polynomials [12] on a circular domain, together with their local counterparts on multi-rectangles and annular supports [10], have been considered as basis functions to model this type of data. In all the cases, the functions are fitted to the 3D data by using the least squares method. Attribute selection is finally performed to select the most discriminative coefficients with minimal dimension. The schema in Fig. 1 summarizes the proposed solution.

The paper is organized as follows: Preprocessing of 3D face scans and selection of basis functions are described in Section 2. The least squares fitting method for computing the coefficients of the 3D face representation; an attribute selection for reducing the number of coefficients, and their use for functional matching are proposed in Section 3. The experiments and a discussion of the results are reported in Section 4. Finally, conclusions and future work are mentioned.

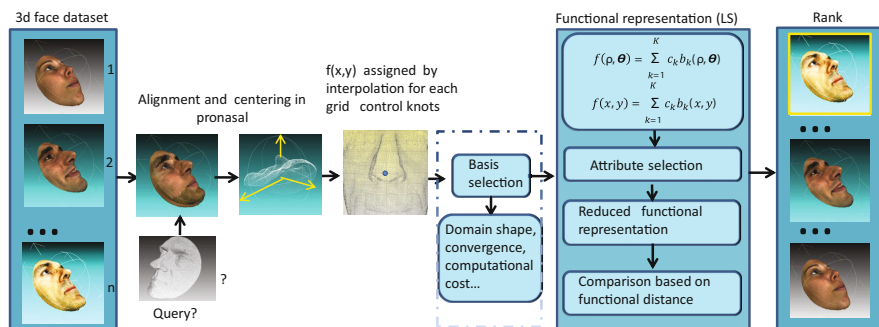


Fig. 1. The proposed solution for 3D face recognition based on functional data analysis

2 Preprocessing and Basis Selection

Preprocessing, including position and orientation normalization is an important step for any 3D face analysis algorithm [13]. In our case, PCA is first used in order to obtain a rough alignment of the principal axis of the face to a common reference system; then, the nose tip is detected and the points of the mesh that fall within a sphere of radius $100mm$ centered on it are retained. Finally, the

cropped region of the face is used to perform fine alignment to a reference model by using the ICP algorithm. A further step, including hole filling and smoothing is performed. The resulting face surface is regarded as a function f directed in the positive z axis, with the positive direction pointing backwards (see Fig. 1), and defined on a 2D support spanned by either Cartesian or Polar coordinates.

In order to derive a functional representation of the 3D surface, a 2D knot control grid for all models is defined. In this work, we consider: a sampling grid in Cartesian coordinates (see Fig. 2(a)), which allows the function $f(x, y)$ be represented as a linear combination of basis functions on a (multi-)rectangular domain; a circular and an annular grid (see Fig. 2(b)), which require a functional representation $f(\rho, \theta)$ in polar coordinates. According to this, for a surface defined on a 2D support, the general equations for the functional representation in Cartesian and polar coordinates are, respectively:

$$f(x, y) = \sum_{k=1}^K c_k b_k(x, y), \quad f(\rho, \theta) = \sum_{k=1}^K c_k b_k(\rho, \theta), \quad (1)$$

where $\{b_k(x, y)\}$ and $\{b_k(\rho, \theta)\}$ denote the sets of the K basis functions, and $\{c_k\}$ represent the coefficients of the expansion. In our work, three sets of basis functions were selected: the *Bivariate Splines* for the rectangular domain (global and local grid); the *circular Zernike polynomials* for the global representation on the circular domain; and the *annular Zernike polynomials* for the local representation on the annular domain.



Fig. 2. (a) Global and local rectangular control grid in Cartesian coordinates; (b) Global circular and local annular control grid in polar coordinates

The Bivariate Splines – The B-spline surfaces are a generalization of the univariate B-spline curves, with the simplest generalization obtained by a tensor product of univariate B-splines [15]. The main advantage of B-splines in representing surfaces of high local variability is their intrinsic zonal character [1]. The B-splines are polynomial curves of degree $k - 1$ along each interval $[t_j, t_{j+1}]$; products of such functions are polynomial surfaces of degree $(k_1 - 1) + (k_2 - 1)$ over each rectangular domain $[s_i, s_{i+1}][t_j, t_{j+1}]$ in \mathbb{R}^2 . The surface is given by:

$$f(x, y) = \sum_{i=1}^n \sum_{j=1}^m c_{ij} \phi_{i,k}(x) \lambda_{j,l}(y), \quad (2)$$

where $c_{i,j}$ are the control points, acting as sub-area control parameters, and $\phi_i(x)$ and $\lambda_j(y)$ are the i -th and j -th B-Spline of degree $k - 1$ and $l - 1$ in the x and y direction, respectively.

The number of breakpoints NBP_x (NBP_y) in the interval, the degree of the B-spline $k - 1$ ($l - 1$), and the level of smoothness at the limits of the subintervals given by the continuity of the function and its derivatives $\mu_s x$ ($\mu_s y$) are parameters that must be set. They have a decisive impact on the final representation by splines. The following equation provides a relationship between these parameters and the number of control points $n + 1$, defined in Eq. 2 [1]:

$$n + 1 = \left[\sum_{s=1}^{[NBP_x]-2} k - \mu_s x \right] + k. \tag{3}$$

In our research, a value of $k = l = 4$ (cubic spline), not multiple control points, and the uniform type of knot vector were used for both x and y directions.

The Circular and Annular Zernike Polynomials – The global representation by Zernike polynomials requires the coordinates to be normalized by the radius, as these polynomials are orthogonal on the unit disk. The original domain of radius R , given by the region of the face to be represented, is transformed so as to obtain the normalized radial coordinate $\rho = r/R$, for each radius r . The basis of circular Zernike polynomials is a tensor product of the Fourier basis in the angular direction, and a special Jacobi polynomial in the radial direction:

$$Z_n^m(\rho, \theta) = \begin{cases} N_n^m R_n^{|m|}(\rho) \cos(m\theta) & \text{for } m \geq 0, 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi \\ -N_n^m R_n^{|m|}(\rho) \sin(m\theta) & \text{for } m < 0, 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi \end{cases}, \tag{4}$$

being θ the azimuthal angle. For a given radial order or polynomial order n , the azimuthal frequency or Fourier order m can only take values of $-n, -n + 2, -n + 4, \dots, n$ [12]. In Eq. (4), N_n^m is the normalization factor:

$$N_n^m = \sqrt{\frac{2(n+1)}{1 + \delta_{m0}}}, \quad \text{with } \delta_{m0} = \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases}, \tag{5}$$

and $R_n^{|m|}(\rho)$ is the representation for the Jacobi polynomial:

$$R_n^{|m|}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-m)! \rho^{(n-2s)} (1/s!)}{[0.5(n+|m|) - s]! [0.5(n-|m|) - s]!}. \tag{6}$$

The annular Zernike polynomials are derived from the circular Zernike polynomials by the Gram-Schmidt orthogonalization process [10]. They maintain their orthogonality, but in an annulus instead of a circle. The parameter $\rho \in [0, 1]$, for the radial coordinate, and $\theta \in [0, 2\pi]$ for the azimuthal component are the same, whereas a new parameter ϵ is used to restrict the inner radius.

One of the main features of these bases is their orthogonality [1], which implies total linear independence among each mode and the others. In the case of a discrete domain, orthogonality is fulfilled only approximately. Due to the fact that Zernike polynomials are orthonormal in the unit circle, then the operations (e.g., inner products and norms) between functions expressed on this basis reduce to operations between their corresponding coefficients.

Determination of the Number of Coefficients – Once the basis set for each domain has been selected, the most suitable number of basis functions for the representation must be determined. A large number of terms may result in a more accurate representation of the surface, but this in turn increases the size of the description. We used the Bootstrap method for determining the number of coefficients [6]. This algorithm randomly reassigns the observations, and re-computes the estimates where a distribution knowledge of the measurement and modeling error is not needed. In the case of local representations, the bootstrapping is performed separately for each sub-domain, so that a different number of coefficients can be obtained for each region.

3 3D Face Functional Representation

Once the number of coefficients is determined, the surface is fitted by the least square method, with SVD as the selected solution method. Although this method has a high computational order ($\sim 2mn^2 + 11n^3$ flops, being n and m the matrix dimensions), it can deal with rank deficiency and matrix singularity problems, providing a unique solution that makes this method more suitable than the QR or the Cholesky decomposition. In the case of gridded data, the problem of approximating a surface is reduced to that of minimizing the square of the difference between the obtained function $f_{i,j}$ and the original function data $(x_i, y_j, g_{i,j})_{i=1, j=1}^{m_1, m_2}$ [9]:

$$\min_{f \in S_1 \otimes S_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} w_i w_j [f(x_i, y_j) - g_{i,j}]^2, \quad (7)$$

where w_i, w_j are positive weights, and S_1, S_2 are univariate spline spaces. The minimization problem using the method of least squares can be decomposed into a sequence of univariate interpolations, simplifying its solution [9]. Then, the total number of weights is $m_1 + m_2$, instead of the number of points $m_1 \times m_2$. The first goal of doing so is to maintain an applicable dissimilarity between the resulting models and the reference scans to allow their identification. Figure 3 shows an example of approximating a 3D face scan using the proposed approach.

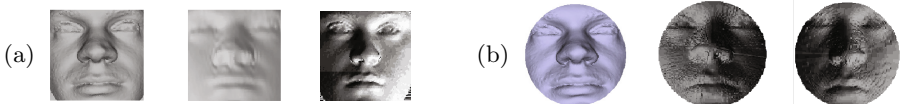


Fig. 3. 3D scan of a subject and its reconstructed models on the rectangular and circular domains (both centered on the nose tip): (a) global/local bivariate splines (a different area of the face is covered); (b) circular/annular Zernike polynomials

Attribute Selection and Functional Matching – In order to reduce the number of coefficients obtained in the representations, while maintaining their discriminative power, an *attribute selection* approach has been used. In our case, where attributes coincide with the coefficients of the functional representation, this method has the advantage that it does not transform the original attributes, so that the original values of the coefficients of the function are maintained unchanged. This is particularly relevant for the functional case, since the coefficients may provide a direct interpretation of the features of the face. In this work, we use the strategy of the univariate selection and evaluation for each feature (coefficient) as mentioned and formalized in [5]. The most successful method used in our case for the selection of attributes was the Chi-Square statistics. It evaluates the features individually by measuring their chi-squared statistic with respect to the classes. Then, they are ranked according to the result of the evaluation. After the feature selection, the matching step is performed. Given a probe face P and a gallery face G , their distance can be computed as in Eq. 8, where p and g are their corresponding functions, defined on a common domain $[a, b] \times [c, d]$. For the norm L_n , the following distance is used [7]:

$$d(p, g) = \left(\int_a^b \int_c^d |p(x, y) - g(x, y)|^n dx dy \right)^{1/n} \quad (8)$$

4 Experimental Results

The proposed 3D face recognition approach has been evaluated on the 2D/3D Florence dataset [2], and the Gavab database [14]. These datasets include scans acquired with different devices and show different challenges (i.e., non-frontal pose, presence of hair, neck, shoulders). For each dataset, four different representations were constructed based on: *global circular Zernike polynomials* (GCZP) on a circular domain covering the face; *local annular Zernike polynomials* (LAZP) on two circular crowns, the innermost containing the nose region, the outermost the lips, cheeks and eyes; *global and local bivariate splines* (GBVS and LBVS) defined, respectively, on a rectangular domain covering the face, and on three disjoint rectangular regions, the upper one containing the eyes and eyelashes, the central the nose and cheekbones, the lower the lips and jaw. The Bootstrap algorithm and attribute selection are used to determine the number of terms in each case. A summary of these representations is given in Table 1 (refer also to Fig. 2 and Fig. 3 to visualize the different domains).

Florence Dataset – The 3D part of the 2D/3D Florence face dataset [2] (UF-3D), consists of high-resolution scans of 54 subjects (14 females and 40 males) of Caucasian ethnicity, with age ranging from 20 to 60 years. Two 3D frontal scans with neutral expression (N) have been acquired for each subject in the same session, though some scan exhibits moderate facial expressions.

Table 1. Summary of the representations used in the experiments. The *pronasale* has been used as origin of the domain in all the cases.

function	domain	control-grid (<i>mm</i>)	step
GCZP	circular	radial 60	1° angular / 1 <i>mm</i> radial
LAZP	annular	0 ÷ 30 / 30 ÷ 60	1° angular / 1 <i>mm</i> radial
GBVS	rectangular	64 × 64	1 <i>mm</i>
LBVS	multi-rectangular	(three) 24 × 64	1 <i>mm</i>

Gavab Dataset – The Gavab database [14] comprises 3D facial scans with large pose and expression variations, and noisy acquisitions. It includes scans of 61 adult Caucasian individuals (45 males and 16 females). For each individual, we consider the two frontal scans with neutral expression (N), and the frontal scans in which the person smiles (S), laughs (L), or shows a random gesture (G).

Results – In a first set of experiments, we modeled the face recognition problem as a classification task, using a k-NN classifier with Euclidean and Manhattan distance. Results are reported in Table 2. For each variant of the proposed approach, we reported the number of coefficients that resulted in the best performance. As for the UF-3D, it can be observed that the two solutions based on bivariate splines (GBVS and LBVS) using Manhattan distance, and the global Zernike polynomials (GCZP) achieve 100% recognition. The same result cannot be achieved by the local annular Zernike polynomials (LAZP). Considering the Gavab dataset, which also includes expressive scans (we report results for N and S), the approaches based on bivariate splines outperform both the methods based on Zernike polynomials. Interestingly, the proposed method is able of achieving high accuracy with a very small number of coefficients. In this respect, Zernike polynomials achieve a more compact representation than B-splines.

Table 2. Results obtained using the best setting on the UF-3D and Gavab datasets

database	method	#coef.	#sel. coef.	variations	RR (%) / dist.
UF-3D	GCZP	48	24	N	100 / Euc.
	LAZP	64	44	N	93.6 / Euc.
	GBVS	4096	48	N	100 / Manh.
	LBVS	4608	20	N	100 / Manh.
Gavab	GCZP	48	32	N,S	95.6 / Manh.
	LAZP	96	36	N,S	95.6 / Manh.
	GBVS	4096	250	N,S	100 / Manh.
	LBVS	4608	1000	N,S	100 / Manh.

To compare our approach with existing state of the art solutions, we performed a face recognition experiment on the Gavab dataset by including one of the neutral scans per subject in the gallery, and using all the other scans (neutral and expressive) as probes. The obtained rank-1 recognition rates are reported in

Table 3, where the best settings of Table 2 are used for our approach. It can be observed, the two variants based on B-splines (GBVS and LBVS) outperform the state of the art solutions for both the case of expressive and neutral scans. The GCZP solution (i.e., we do not report the LAZP since it shows the lowest accuracy in Table 2) is still competitive, although at a lower extent, but with the clear advantage of using less coefficients and thus a more compact representation.

Table 3. Comparative evaluation on the Gavab dataset. Rank-1 RR is reported (best results for each case are marked in bold).

	[8]	[11]	[3]	Ours		
				GBVS	LBVS	GCZP
Neutral	96.7	95.0	100	100	100	96.7
Expressive	93.3	72.0	94.5	94.9	95.1	81.6
Neutral + Expressive	94.7	77.8	95.9	96.2	96.3	85.4

5 Conclusions

In this paper, an original 3D face recognition approach based on functional data analysis has been proposed. The basic idea of our approach is to consider the 3D face as a surface defined on a 2D support, and to approximate the surface by a basis functions set expansion in both a global and local approach. Results show that the studied basis functions, B-Splines and Zernike polynomials, are suitable options for modeling this type of data. Selecting the most discriminative coefficients of the basis expansion resulted in a compact representation that allows a fast and effective recognition. Experiments conducted on two datasets show promising results, also in comparison with state of the art solutions.

As future work, we will experiment the proposed approach on larger datasets (e.g., FRGC v2 and Bosphorus), and extend it to manage more challenging scenarios (e.g., 3D face scans with pose variations and missing parts).

References

1. Ares, M., Royo, S., Caum, J., Pizarro, C.: Comparison of B-spline and Zernike fitting techniques in complex wavefront surfaces. In: Optical Measurement Systems for Industrial Inspection IV, Munich, Germany (2005)
2. Bagdanov, A.D., Del Bimbo, A., Masi, I.: The Florence 2D/3D hybrid face dataset. In: Proc J-HGBU 2011, Arizona, USA, pp. 79–80 (2011)
3. Drira, H., Ben Amor, B., Srivastava, A., Daoudi, M., Slama, R.: 3D face recognition under expressions, occlusions, and pose variations. *IEEE Trans. Pattern Analysis and Machine Intelligence* 35(9), 2270–2283 (2013)
4. Ferraty, F., Vieu, P.: *Nonparametric Functional Data Analysis Theory and Practice*. Springer (2006)
5. Harol, A., Lai, C., Pekalska, E., Duin, R.: Pairwise feature evaluation for constructing reduced representations. *Pattern Analysis and Applications* 10, 58 (2007)
6. Iskander, D.R., Collins, M.J., Davis, B.: Optimal modeling of corneal surfaces with zernike polynomials. *IEEE Trans. Biomed. Eng.* 48, 87–95 (2001)

7. Kolmogorov, A.N., Fomin, S.V.: *Nonparametric Functional Data Analysis Theory and Practice*. Dover Publications (1999)
8. Li, X., Jia, T., Zhang, H.: Expression-insensitive 3D face recognition using sparse representation. In: *Proc. IEEE CVPR*, pp. 2575–2582 (June 2009)
9. Lyche, T., Mørken, K.: *Spline methods*. Tech. rep., Department of Informatics, Centre of Mathematics for Applications, University of Oslo (2008)
10. Mahajan, V.N., Aftab, M.: Systematic comparison of the use of annular and Zernike circle polynomials for annular wavefronts. *Appl. Opt.* 49(33), 6489–6501 (2010)
11. Mahoor, M.H., Abdel-Mottaleb, M.: Face recognition based on 3D ridge images obtained from range data. *Pattern Recognition* 42(3), 445–451 (2009)
12. McAlinden, C., McCartney, M., Moore, J.: Mathematics of zernike polynomials: a review. *J. Clin. Exp. Ophthalmol.* 39, 820–827 (2011)
13. Mian, A.S., Bennamoun, M., Owens, R.: An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 29(11), 1927–1943 (2007)
14. Moreno, A.B., Sánchez, A.: Gavabdb: A 3D face database. In: *Proc. Workshop on Biometrics on the Internet*. pp. 75–80. Vigo, Spain (March 2004)
15. Schumaker, L.L.: *Spline Functions: Basic Theory*, 3rd edn. Cambridge University Press, New York (2007)