Functional Data Analysis as an Alternative for the Automatic Biometric Image Recognition: Iris Application

Dania P.-Muñoz, Francisco José Silva Mata, Noslen Hernández,
and Isneri Talavera Bustamante

Centro de Aplicaciones de Tecnologías de Avanzada, CENATAV,
Cuba

{dpmunoz, fjsilva, nhernandez, italavera}@cenatav.co.cu

Abstract. Functional data analysis has been a novel option for representing images, since the continuous nature of images is preserved. Image representation using functional data provides significant advantages, being the appreciable reduction of the dimensionality one of the most significant. This paper gives a detailed description of the entire imaging process using the proposed approach. As an example, the representation of iris images through functional data for recognition tasks was used. The paper presents experiments and results of applying this approach to the recognition of iris images, demonstrating its effectiveness.

Keywords. Functional data analysis, Zernike bases, biometric image recognition.

El análisis de datos funcionales como alternativa para el reconocimiento automático de imágenes biométricas: aplicación en el iris

Resumen. El análisis de datos funcionales ha sido una opción novedosa para la representación de imágenes, ya que su naturaleza continua se conserva. La representación de imágenes usando datos funcionales proporciona muchas ventajas, donde una de las más significativas es la reducción apreciable de la dimensión de los datos. En este trabajo se propone una descripción detallada de todo el proceso de representación de imágenes mediante el enfoque propuesto. Además, la representación de las imágenes del iris por medio de los datos funcionales para tareas de reconocimiento, se utilizó como un ejemplo. El artículo presenta algunos experimentos y resultados de la aplicación de este enfoque para el reconocimiento de imágenes del iris, lo que demuestra la eficacia de la misma.

Palabras clave. Análisis de datos funcionales, bases de Zernike, reconocimiento de imágenes biométricas.

1 Introduction

A system for the automatic recognition of biometric images consists of several components, which typically include detection, tracking (if necessary), normalization (in some cases may be a simple alignment), feature extraction; object representation, dimensionality reduction (if possible) and finally a matching process. Some of these steps contain some particularities depending on the used representation. One way of representation can be the Functional data analysis (FDA) approach. The basic unit of information in FDA is the function as a whole, instead of a set of individual features. A representation based on functions allows the analysis of some functional characteristics as smoothness, monotony and the derivatives of various orders [2].

Moreover, the functional representation is a natural way of dimensionality reduction of the data, which is often accomplished through the representation itself.

The origin of this approach is the result of the development of instruments and measurement methods, that in many occasions generate large sets of data, representing time series, continuous spectra or records of events or random variables [1]. Some of the data that comply with the above said are those obtained by electrocardiograms, seismographs, gas chromatographs and mass spectrometers [14].

Although FDA is a young discipline, there have been developed several applications in different fields such as biological sciences, specifically in genetics. In medicine it has also been used in studies related to hepatic cirrhosis, in assessing

antiviral therapies against AIDS. Other applications are appreciated in geology, psychology, economics, etc. Despite the rise shown by the FDA in various fields, this has been recently introduced in the field of image processing [5].

According to [3], an image is a mathematical representation of a physical observation as a function over a spatial domain. Therefore, digital images can be seen as functions of two variables, where for each coordinate (x = row, y = column) of the pixel matrix, a sample of grey tones or intensity has been taken. Finding such function that approximates the observed discrete data is one of the key issues of this approach.

The human iris is the only internal organ truly visible from the outside [6]. It has been demonstrated that the eye texture has practically unrepeatable patterns among human beings, even twins [7], which makes the iris a biometric entity usable for person identification. Unlike other body parts used for biometric purposes, such as the face, the iris patterns are difficult to modify by actions of makeup. These features, together with the emerging development of technologies for capturing images of high resolution, make the iris recognition a promising branch of biometrics.

The automatic processing of the iris image contains the essential steps of image processing. This includes innovative solutions for its detection and segmentation [8, 9, 28], as well as obtaining a description that reduces the high dimensionality of the data, maintaining its discriminative capacity while reducing the computational cost of the process.

This paper proposes a scheme based on FDA approach for representing images to be used in biometric analysis, where the iris recognition is presented as an example.

The obtained solution for the particular case of iris images, demonstrates the validity of the approach and traces a new general methodology for analyzing biometric images. The general diagram presented in Fig. 1 shows the necessary steps to perform an image recognition task based on Functional data analysis.

The paper is organized as follows. In Section 2 the steps of the proposed flow diagram are explained in a general way. In this section, the basis selection process and the approach to
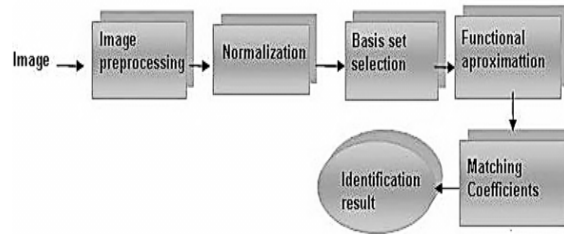

Fig. 1. General diagram of FDA approach for image recognition task

determine the number of coefficients are also described. In Section 3, the summary diagram of the entire process, until iris image identification, is shown. Section 4 is dedicated to experimental setup and results. Conclusions and future research topics are shown in Section 5.

2 Flow Diagram Steps

It is necessary to emphasize that digital images are commonly sampled in a two dimensional space using a sampling grid in Cartesian coordinates $I(x, y)$. As it was mentioned before, digital images can be represented by functions. Thus, we propose and explain in detail a methodology for representing images using FDA and how further analysis can be performed with this approach.

The subsections below expose a detailed description of the steps that explain the complete process. The flowchart represented in Fig. 1 contains the particularities of the overall biometric image process using FDA to perform automatic recognition tasks.

2.1 Image Pre-processing and Geometric Normalization

Some blocks of the diagram, as image pre-processing and normalization are usual in general diagrams of image analysis, presented in specialized handbooks [10]. The result of these processes is an image with the object of interest correctly segmented.

In the case of iris images, raw iris images must be converted to grey scale and then subjected to

a segmentation process, obtaining only the iris region of interest. This region is extracted and then re-scaled to an image with new dimensions, which fulfills the requirement that iris radius must contain at least 70 pixels [11].

Each original image must be subjected to histogram equalization, in order to highlight the patterns of iris texture. The iris and pupil borders are detected through the method proposed by Silva and Alvarez [12].

The goal in iris normalization has been to convert the iris texture region to a rectangular shape by the "rubber sheet" method, proposed by Daugman [7]. Instead, we propose to segment the iris and the pupil image such that its original shape is maintained. Pupils may vary in size and position in the eye (as its normal function in the ocular system), introducing variations that can cause differences in the iris representation. For that reason, the performed normalization in this research is focused on obtaining the same size for all irises and the same size for each pupil with the same image center coordinates.

This guarantees the same domain for all the functional representations. For this normalization, it is proposed to assess 25 reference points: 12 equidistant on the iris circle, 12 equidistant on the pupil circle and 1 corresponding to the pupil center. The normalization consists in a warping transformation (interpolation and sampling) in relation to a unique pattern iris image. The dimensions of the iris and pupil for the image pattern were obtained by the statistical mode of the whole set of iris training samples. The mode was calculated by considering similar values (according to a given threshold) as the same value. In this way, we tried to reduce the possible number of cases that needed normalization. It is only done if the difference between the reference points from the pattern and the samples is greater than a given threshold. Fig. 2 shows an example of an iris image normalization, where the pupil is smaller than in the pattern image. It was not aligned with the iris center.

2.2 Iris Representation by Functional data analysis

Functional data analysis is a relatively new and growing research field that deals with the
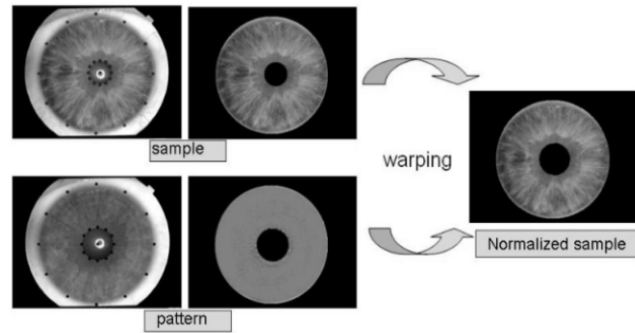

Fig. 2. Normalization process, 25 referent points are marked in the image

statistical description and modeling of such kind of (functional) data [2, 13].

The idea of the FDA methods is to obtain an adequate representation of continuous data that allows exploring the intrinsic properties of the data (smoothness) [14]. During the past years, many methods have been reported for the one-dimensional case, but for the two-dimensional case there are few proposed methods in the literature.

As it was mentioned before, digital images are recorded discretely; however, their nature is continuous. Therefore, instead of representing it by vectors, it might be more appropriated to represent the data by an underlying continuous function, in a way that this representation can be closer to the original nature of the data.

A common strategy for achieving this representation is using functional basis expansion [2]. The basis function selection is an essential step to achieve a good functional representation of the data. A function $I_f(x, y)$ can be represented in a smooth way, expressing it as a linear combination of basis functions as it is shown in Eq. 1.

$$I_f(x, y) = \sum_{k=1}^{K} c_k b_k(x, y), \qquad (1)$$

where $\{b_k(x, y)\}_k$ denotes the set of K basis functions and $\{c_k\}_k$ represents the coefficients of the expansion. Assuming the representation illustrated in Eq.1, the aim is to obtain, given the discrete observations $I(x, y)$ (i.e., the digital

image), its functional representation $\hat{I}_f(x,y)$ through estimates $\{\hat{c}_k\}_k$ of the coefficients $\{c_k\}_k$ [15].

The iris has almost a circular shape, so it is intuitive to work with polar coordinates [16]. The process of converting an image $I(x; y)$ to polar coordinates $I(\phi; \rho)$ is a sampling and interpolation process [11]. To achieve compliance with the Nyquist-Shannon sampling theorem and to avoid the overlapping phenomenon [17] (known as aliasing) during this conversion, it must be ensured a total of $\sqrt{2}r$ (where r is the radius in pixels) samples in the radial direction and $2\sqrt{2}\pi r$ samples in the angular direction [18]. With images in polar coordinates, the functional approximation of Eq.1 can be written as $I_f(\rho,\theta) = \sum_{k=1}^{K} c_k b_k(\rho,\theta)$; where $b_k(\rho,\theta)_k$ is a suitable basis. For obtaining the functional representation $\hat{I}_f(\rho,\theta)$ of an image $I(\rho,\theta)$, there are some hyper-parameters that need to be selected with some care. Examples of these parameters are the basis function to be used and the number of coefficients in the expansion.

2.2.1 Basis Selection and Determination of the Number of Coefficients

The number of basis functions K, determines the degree of accuracy of the representation. As a general rule, an accurate representation can be obtained, since the number of basis functions is equal to the number of sampling points. Without going deeper into the mathematical aspects of the problem, it is clear that the choice of the basis functions is crucial to achieve a good representation of the original image information. The selection of the set of basis functions is of critical importance for the good performance of the process, especially with respect to its computational efficiency. Functional analysis methods provide the tools needed to determine which set of basis functions is the most convenient and other important issues such as the number of terms needed. Basis selection is conditioned by multiple criteria like the geometry of the domain (which is a determinant reason according to a number of authors), the computational complexity, the differentiability, the periodicity of the event to model, the ratio of speed of convergence, the completeness

understood as the ability to represent any function with high precision and with enough term coefficients and even to the behavior of the classifiers or regression function depending on the case. In classification problems, bases are often selected such that the relative entropy time-frequency energy between classes is maximized; in regression problems bases are selected to minimize the regression error [19].

Heuristics based on the results obtained by the evaluation of the mentioned criteria help to make the decision of which basis to select from the candidates. Iris images have a circular geometric appearance that in conjunction with its apparent rotational symmetry suggest a functional basis expansion in a circular domain. Although there are multiple sets of basis functions for functional approximation on a circular domain or disc [20], only a few have been evaluated for this research. The most widely used in computational vision and ophthalmologic research are the Zernike polynomials, the polynomials of Logan-Shepp, the Fourier Chebyshev basis, the Fourier-Bessel basis, and the radial basis functions [20]. We mentioned some criteria for the selection of the best basis, nevertheless, the actual basis selection process not only includes the evaluation of these aspects but also a detailed analysis of the particular problem we want to model and a correct balance of the advantages and disadvantages of each basis. It is impossible to mention here all the details we considered to choose a basis, however, we will describe briefly why we chose the Zernike polynomials as the basis functions to work with.

The Zernike representation needs half of the functions $O(N^2/2)$ which, for example, are used in the Chebyshev-Fourier base $O(N^2)$ for representing the data with the same accuracy. Furthermore, the number of basis functions needed for exactly representing Cartesian polynomials of degree N using Zernike is lower than if the Chebyshev-Fourier basis is used. The Zernike basis also outperforms the Chebyshev-Fourier basis with respect to the rate of convergence because it converges more quickly with the same number of coefficients. Regarding the computational orders, Chebyshev-Fourier leads to $O(N^2 \log_2 N)$ operations and $O(N^2)$ storage in memory, while Zernike leads to $O(N^3)$

operations and $O(N^3)$ memory usage [20]. In spite of the fact that the use of the Fast Fourier Transform accelerates the calculations for the Chebyshev-Fourier, this is not enough to decide in favor of this basis and it was decided to use Zernike. In this case, more importance was given to the number of terms and the remaining elements mentioned before. The paper of Boyd and Yu (2011)[20] offers an extensive explanation of this topic.

The Zernike polynomials basis is a tensor product of the Fourier basis in the angular direction and a special Jacobi polynomial in the radial direction, defined as in [21]:

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

$$(2)$$

where ρ represents the radial coordinate and 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 $-n; -n+2; -n+4, \ldots, n$ [21]. The normalization factor is given by N_n^m:

$$N_n^m \sqrt{\frac{2(n+1)}{1+\delta_{m_0}}} \quad with \quad \delta_{m_0} = \begin{cases} 1, & m=0 \\ 0, & m \neq 0 \end{cases} \quad (3)$$

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

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

$$(4)$$

For the specific case of the iris, the value ρ is comprised between the pupillary boundary and the limbus boundary. Then, any function $I_f(\rho,\theta)$ defined over a two dimensional space can be approximated using the Zernike polynomials basis in the following way [22]:

$$I_f(\rho,\theta) = \sum_n \sum_m C_n^m Z_n^m(\rho,\theta) \quad (5)$$

where C_n^m represents the vector of coefficients and Z_n^m represents the Zernike basis functions. In order to obtain an estimate $\hat{I}_f(\rho,\theta)$ of $I_f(\rho,\theta)$ for each image $I_f(\rho,\theta)$, we need to estimate the coefficients in the expansion. This will be done by least squares fitting. Due to the fact that Zernike polynomials are orthonormal in a unit circle,

$$\iint Z_n^m Z_j^i \, d\rho d\theta = 0 \text{ if } i \neq m \text{ or } j \neq n, \quad (6)$$

the operations (e.g., inner products and norms) between functions expressed on this basis are reduced to operations between their corresponding coefficients, which makes the computations easier.

There is still one hyper-parameter that needs to be established before performing the least squares estimation of the coefficients, and it is the number of coefficients (or number of basis functions) in the expansion. This parameter has the role of a smoothing parameter. Statistically, keeping a few coefficients in the expansion is equivalent to conducting a heavy amount of smoothing for the original data. It also determines the dimensionality reduction achieved with this representation. One way of determining this parameter is through a bootstrapping strategy [23]. Bootstrap is a statistical technique for assessing the accuracy of a parameter estimator in situations where conventional techniques are not valid. This method randomly reassigns the observations and re-computes the estimates. The main benefit in using bootstrap is that knowledge of the distribution of the measurement and modeling error is not necessary.

The pseudo-code below seeks to explain it briefly. For notational convenience let us assume that in Eq. 5, a given combination of n and m determines a basis Z_k with the corresponding coefficient C_k. In this way, Eq. 5 can be expressed as

$$I_f(\rho, \theta) = \int_{k=1}^{p} C_k Z_k(\rho, \theta), \qquad (7)$$

where p is the number of coefficients to be tuned.

1. Select a maximum model order P_{max}. Then estimate the coefficients $\hat{C}_{Pmax}$ of C_{Pmax} using the method of least squares.
2. Compute the estimated surface, $\hat{I}_f(\rho_d, \theta_d) = \sum_{k=1}^{Pmax} \hat{C}_k Z_k(\rho, \theta)$, where k is the true but unknown model order, and $d = 1, \ldots, D$ represents the number of evaluation samples.
3. Calculate the residuals from the original image and the estimated image:
   $\hat{r} = I(\rho_d, \theta_d)$, where $d = 1, \ldots, D$.

4. Rescale the empirical residuals:

$$\tilde{r}_d = \sqrt{\frac{D}{L_D}} \left(\hat{r}_d - \frac{1}{D} \sum_{d=1}^{D} \hat{r}_d \right).$$

5. For all model orders $1 \leq p \leq Pmax$:

   a) Calculate $\hat{C}_p$ and $\hat{I}_f(\rho_d, \theta_d)$ as in step 1.
   b) Draw random values $\tilde{r}_d^*$ with replacement from the original $\hat{r}$.
   c) Using those random values, define the bootstrap surface:

   $I_f^*(\rho_d, \theta_d) = \hat{I}_f(\rho_d, \theta_d) + \tilde{r}^*.$

   d) Using the surface calculated in c), estimate $\tilde{C}_p^*$.
   Using least squares method, calculate

   $$I_f^*(\rho_d, \theta_d) = \sum_{k=1}^{P} \hat{C}_k^* Z_k(\rho, \theta) \text{ and}$$

   $$SSE_{D,L_D}^*(p) = \frac{1}{D} \sum_{d=1}^{D} (I(\rho_d, \theta_d) - \hat{I}_f^*(\rho_d, \theta_d))^2.$$

   e) Repeat steps (b)–(d) a big number of times, (in our case we chose $B = 100$) to obtain a B bootstrap statistics
   $SSE_{D,L_D}^*(p)_1, \ldots, SSE_{D,L_D}^*(p)_B$, and estimate the bootstrap mean-square error

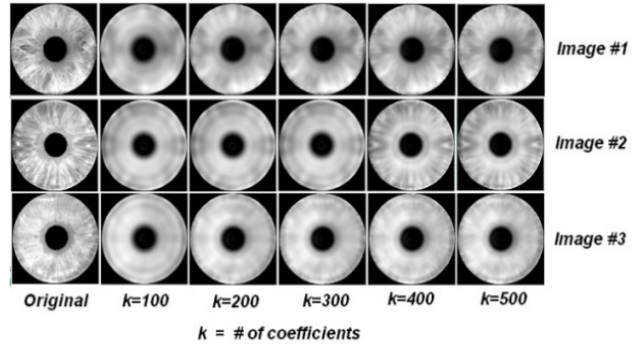 $$\hat{I}_{D,L_D}(p) = \frac{1}{B} \sum_{b=1}^{B} SSE_{D,L_D}^*(p)_b.$$

Fig. 3. Original images and the correspondent recovered images by Zernike basis approximations with *K* coefficients from UPOL database

f) Choose p for which $\hat{I}_{D,L_D}(p)$ is a minimum.

2.2.2 Final Representation and Functional Principal Components Analysis

Once the number of coefficients has been determined, the next step is to obtain the coefficients vector of the expansion by fitting data using the least squares method.

Fig. 3 shows some examples of original iris images and their functional approximations using K basis functions.

The smoothing effect of the representation can be appreciated in the recovered images. However, the differences between images coming from different subjects can also be noticed. This is a fact that intuits their identifying capabilities.

Although the dimension of the new representation (K coefficients) is commonly smaller than the dimension of the original data, some methods for reducing dimensionality can be used in the functional space [14]. One of the techniques widely used to reduce dimensionality extracting relevant information from data is Principal Component Analysis (PCA). For the applications in the functional data context, PCA have been extended to functional spaces with the formulation of Functional principal component analysis (FPCA) [24]. The FPCA finds eigenfunctions instead of eigenvectors. The basic idea of FPCA is the same as in PCA, but as we are working now in functional spaces, FPCA replaces vectors by functions, matrices by linear operators, covariance matrices by covariance
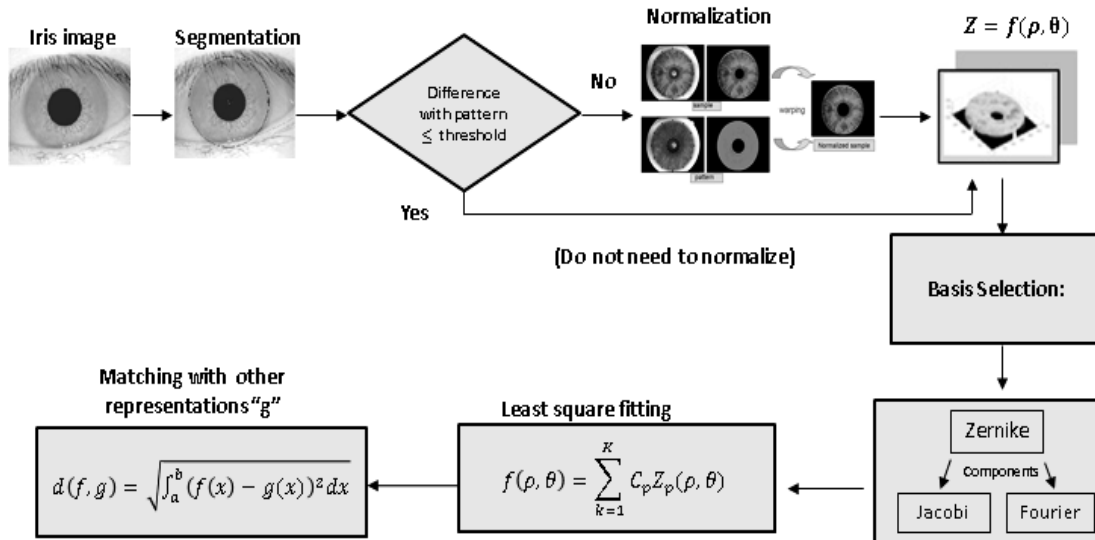
Fig. 4. Summary diagram of iris recognition process by functional data analysis approach

operators and scalar products in vector space by inner products in functional space (usually the space of square-integrable functions) [2]. Then, FPCA finds the set of orthogonal principal component functions that maximizes the variance along each component.

3 Summary Diagram of the Whole Process

The diagram in Fig. 4 shows the whole process needed to identify an iris image based on functional data analysis. The basis selection is conducted only once for a database. The last block, devoted to the coefficients matching, must determine the similarity (or distance) between the objects being represented and the representations previously calculated. The calculation of distances between the functions is a direct way of computing the similarity between two functions, although it is not the only one. In fact, there are different ways of defining a distance between two functions, for example $d(f; g) = max |f(x) - g(x)|$. If f and g are continuous functions on a closed and bounded interval $[a; b]$ and there is a maximum, the

distance induced by the inner product of functions can be $(f, g) = \sqrt[n]{\int_a^b (f(x) - g(x))^n dx}$, which corresponds to the L_n norm. In our case we have used the above distance with $n = 2$ (L_2 norm).

4 Experiments and Results

To evaluate the performance of the proposed method, we used the CASIA, UPOL, UBIRIS and MMU databases. The CASIA and MMU databases have been acquired with NIR illumination source, while UPOL and UBIRIS were acquired with visible light. CASIA database version 1.0 consists of 756 iris images from 108 different irises (subjects) with 7 samples for each. All images are stored with a resolution of 320 x 280 pixels. The UPOL database contains 384 images of 576 x 768, extracted from both eyes of 64 subjects (three images per eye). UBIRIS database is comprised of 1877 images collected from 241 subjects, with 6 in two distinct sessions and constituted, at its release date, the world's largest public and free iris database for biometric purposes. Its main characteristic is that, in contrast to the existing public and free databases

(CASIA and UPOL), it incorporates images with several noise factors, allowing the evaluation of robustness in iris recognition methods. Images in this database are of 400 x 300 pixels. The MMU v1.0 database contains 450 iris images, which were collected from 45 subjects. There are 10 images from each subject. These are images of 320 x 240 pixels. This database has images with problems like specular reflections, off-axis and off-angle, blur, focus, no uniform illumination, occlusions such as eyelids, eyelashes, glasses, contact lens, and hair. In our experiments we used Libor Masek´s method [25] for localizing iris and pupils in images. As the result of the localization, the main points were obtained from each region of interest (iris and pupil). With those points, images were normalized as explained in Section 2. After performing those two steps we resized images using a bicubic interpolation, obtaining 221 x 221 images for all databases.

Once we have images normalized we represented them using FDA, as explained below.

a) In order to obtain the optimal number of coefficients for the functional representation, we split the data sets in training (80% of the data) and test (20%) sets. On each training set, a bootstrap method was then applied (explained in Section 2) to find this optimal number. Results are then given by evaluating the test sets with the selected number of coefficients.

b) With the obtained number of coefficients, the basis matrix was calculated. Each column j of this matrix corresponds to the values of the Zernike polynomial with order j.

c) Estimate the coefficients by the least squares method using the above basis matrix and the original image.

d) Apply the steps above for every image in the databases, and perform iris recognition by using KNN classifier.

4.1 Experiment 1

This experiment was performed with the complete UPOL database, using 83 coefficients obtained with the bootstrap method. With this number of coefficients we obtained a 96.83% of accuracy. Afterwards, FPCA was applied to reduce the dimensionality of this dataset. The number of eigenfunctions retained, captured a 95% of the data variance. New experiments were performed using the same images as for the previous experiments. This process generated 17 functional principal components (fpc). Classifiers were built on the resulting fpc representation and results are shown in Table 1.

Table 1. Recognition results on UPOL database with 16 coefficients

Data Base	Dimension	% Classification
UPOL	17	99.94%

4.2. Experiment 2

This experiment consists in evaluating our method on CASIA, UBIRIS and MMU databases, where the main goal is to study the influence of occlusion problems in our approach. The number of coefficients obtained with the bootstrap method for these databases was 320, 450 and 145, respectively. After a feature reduction by using Fisher discriminant analysis [26], the coefficients were reduced to 107, 92 and 89, respectively. Results are reported with Fisher discriminant since they were better than the ones obtained using FPCA for these datasets. The classification results are presented in Table 2.

Table 2. Recognition results on CASIA, UBIRIS and MMU databases

Data Base	Dimension	% Classification
CASIA	107	80.1%
UBIRIS	92	89.97%
MMU	89	83.4%

The difference between the results in both tables might be due to the occlusion level in images. It should be noted that for the CASIA, UBIRIS and MMU databases, the number of coefficients is similar, while a notable difference can be seen with UPOL database. This might be because in the first three databases there are many occlusion problems, while in the UPOL there are very few of them. That is why it is notable how these images can be represented with only 17 coefficients.

The results reported in Table 2 are not competitive with the ones reported in previous works [27]. This can be attributed to the fact that those approaches incorporate a proper preprocessing step to overcome the occlusion problems. The low accuracy of the functional approach is a consequence of the severe occlusion problems in those databases.

In order to improve the results of our research, we are already working with a new basis function. We adopted some new methods for normalization, following Daugman's idea but keeping the functional representation, with the aim of avoiding the occlusion problems.

Future work will be devoted to a study of pre-processing methods to continue tackling the occlusion problems. These methods must be adequate for the functional approach proposed in this paper.

5 Conclusions

This paper introduces a new methodology to cope with recognition tasks on biometric images using functional data analysis. Moreover, a new description of iris texture region by means of functional data representation is proposed, using circular Zernike polynomials as basis functions. The descriptive power of Zernike polynomials as a functional basis for iris representation was shown. This description by functional data provides a significant reduction in the dimensionality of the data. This representation shows good results for images with slight occlusion problems.

According to the experimental results it can be concluded that the proposed representation seems to be a promising approach for iris recognition in images with no occlusion problems. However, results showed that this representation with Zernike polynomials is not robust to deal with occlusion problems without applying suitable pre-processing methods; further studies will be conducted to overcome these problems. Although our paper is mainly focused on iris images, the proposed methodology could be applied to solve recognition problems in other domains.

References

1. **Wasserman, L. (2004).** *All of Statistics . A concise course in statistical inference.* New York: Springer.

2. **Ramsay, J.O. & Silverman, B.W. (2005).** *Functional Data Analysis* (2nd ed.). New York: Springer.

3. **Boomgaard, R.V.D. & Dorst, L. (2005).** *Machine vision. An introduction for computer scientists.* University of Amsterdam, Computer Science Department.

4. **Ramsay, J.O. & Dalzell, C.J. (1991).** Some tools for functional data analysis. J*ournal of the Royal Statistical Society*, 53(3), 539–572.

5. **Locantore, N., Marron, J.S., Simpson, D.G., Tripolo, N., Zhang, J.T., & Cohen, K.L. (1999).** Robust principal component analysis for functional data. *Sociedad de Estadística e investigación Operativa*, 8(1), 1–73.

6. **Ross, A. (2010).** Iris Recognition: The Path Forward. *Computer*, 43(2), 30–35.

7. **Daugman, J. & Downing, C. (2001).** Epigenetic Randomness, Complexity,and Singularity of Human Iris Patterns. *Proceedings of The Royal Society B: Biological Sciences*, 268(1477), 1737–1740.

8. **Jan, F., Usman, I., & Agha, S. (2012).** Iris localization in frontal eye images for less constrained iris recognition systems. *Digital Signal Processing,* 22(6), 971–986.

9. **Tuama, A.S. (2012).** Iris image segmentation and recognition. *International Journal of Computer Science & Emerging Technologies*, 3(2), 60–65.

10. **Jain, A.K., Flynn, P., & Ross, A.A. (2008).** *Handbook of Biometrics.* New York: Springer.

11. **Daugman, J. (2004).** How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology,* 14(1), 21–30.

12. **Silva, F., Garea, E., Álvarez, E.M., & Gil, J.L. (2006).** A fast adaboosting based method for iris and pupil contour detection. *11th Iberoamerican conference on Progress in Pattern Recognition, Image Analysis and Applications (CIARP'06),* Cancun, Mexico, 127–136.

13. **Ferraty, F. & Vieu, P. (2006).** *Nonparametric Functional Data Analysis: Theory and Practice.* New York: Springer.

14. **Ramsay, J.O. & Silverman, B.W. (2002).** *Applied Functional Data Analysis.* New York: Springer.

15. **Ramsay, J.O. & Silverman, B.W. (1997).** *Functional Data Analysis.* New York: Springer.

16. **Berggren, L. (1985).** Iridology: A critical review. *Acta Ophthalmologica*, 63(1), 1–8.

17. **Yang, J.W., Woo, D.H., Kim, D.H., & Yi, J. (2008).** The Improvement of the Data Overlapping Phenomenon with Memory Accessing Mode. *Journal of information Display*, 9(1), 6–13.

18. **Wang, Q., Ronneberger, O., & Burkhardt, H. (2008).** *Fourier Analysis in Polar and Spherical Coordinates* (Internal Repor 1/08), Austria: *Albert Ludwigs University Freiburg.*

19. **Boyd, J.P. (2001).** *Chebyshev and Fourier Spectral Methods* (2nd ed., rev.). Mineola, N.Y.: Dover Publications.

20. **Boyd, J.P. & Yu, F. (2011).** Comparing seven spectral methods for interpolation and for solving the poisson equation in a disk: Zernike polynomials, logan-shepp ridge polynomials, chebyshev-fourier series, cylindrical robert functions, bessel-fourier expansions, square-to-disk conformal mapping and radial basis functions. *Journal of Computational Physics*, 230(4), 1408–1438.

21. **McAlinden, C., McCartney, M., & Moore, J. (2011).** Mathematics of zernike polynomials: a review. *Clinical & Experimental Ophthalmology,* 39(8), 820–827.

22. **Goodwin, E.P. & Wyant, J.C. (2006).** *Field Guide to Interferometric Optical Testing.* Bellingham, Wash: *SPIE Press.*

23. **Iskander, D.R., Collins, M.J., & Davis, B. (2001).** Optimal Modeling of Corneal Surfaces with Zernike Polynomials. *IEEE Transactions on biomedical engineering*, 48(1), 87–95.

24. **Benko, M. (2004).** *Functional Principal Components Analysis, Implementation and Applications.* M.Sc. Thesis, Humboldt University, Berlin, Germany.

25. **Masek, L. & Kovesi, P. (2003).** *MATLAB Source Code for a Biometric Identification System Based on Iris Patterns.* The School of Computer Science and Software Engineering, The University of Western Australia.

26. **Martínez, A.M. & Kak, A.C. (2001).** PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 23(2), 228–233.

27. **Birgale, L. & Kokare, M. (2012).** Iris Recognition Using Ridgelets. *Journal of Information Processing Systems*, 8(3), 445–458.

28. **García-Vázquez, M.S. & Ramírez-Acosta, A.A. (2012).** Avances en el Reconocimiento del Iris: Perspectivas y Oportunidades en la Investigación de Algoritmos Biométricos. *Computación y sistemas,* 16(3), 267–276.

Dania P.-Muñoz received her diploma of Software Engineer from the Technical University "José A. Echeverría" (CUJAE), Havana, Cuba, in 2010. She currently works as a researcher and software developer in the Pattern Recognition Department of the Advanced Technologies Application Center (CENATAV), Cuba, pursuing her Ph.D. degree in Functional Data Analysis applied to images. Her main research interests encompass pattern recognition, functional representation of images and its application in iris recognition and face recognition. Dania is a member of the Cuban Society of Mathematics and Computation and the Cuban Association of Pattern Recognition.

Francisco José Silva Mata received his degree in Electronic Engineering at the Polytechnic Superior Institute "Jose A. Echeverría" (ISPJAE), Habana, Cuba, in 1982, and the certificate of Senior Researcher in 2002. He received the title of Master of Science in December 2005. He is currently an expert on Image Processing Department of Pattern Recognition CENATAV. He focuses his main research on the representation and analysis of shapes within biometric images and chemical imaging. He has taught various upper level courses in project management, switching power supplies, electronic measurement, digital image processing, pattern recognition and video processing. He participated in various international conferences: CIARP 2005 Mexico, 2006 ISOEN Spain, ICIAP 2007 Italy, VMDL 2007 Italy. He presented papers as author or co-author at such international events as CIARP 2006 Mexico, ICME'2007 Beijing China, MCAM'2007 China, Tecnolaser 2009 Cuba, IRCDL 2008 Italy, ICME 2008 Germany, Greece Eurographics 2008. He has given lectures and seminars at the University of Havana, the Havana ISPJAE, Burgos University Spain, and the University of Florence, Italy. He published articles in IEEE, LNCS and Analytical Chemistry Acta. He iscoauthor of the following books: *Machine Learning Techniques for Adaptive Multimedia*

Retrieval, published in October 2010, IGI Global, New York USA, and *Computational Intelligence in Digital Forensic*, of the series Studies in Computational Intelligence, Springer, forthcoming. He has been a member of the technical committees of international conferences CIARP2005, ACM Multimedia Information Retrieval (MIR) 2007, MMM2008, 2009, MMM2010, and 2011, IADIS 2009, 2010, 2011, 2012, and 2013 and CIDF2013.

Noslen Hernández graduated in Informatics Engineering (with Honors) in 2005 from the Polytechnic University of Havana (CUJAE), Cuba. In 2011 she received the Ph.D. degree in Mathematics from the Havana University, Cuba. Since 2005, she has been at the Advanced Technologies Application Center (CENATAV), in Havana, Cuba. Her research interests include chemometrics, calibration and regression methods, functional data analysis and digital image and signal processing. She published more than 30 articles on new calibration methods and image processing applications, of which about 10 are published in peer-reviewed journals. She is a reviewer of prestigious journals. She was awarded in 2011 with the Annual Mention to the Best Ph.D. thesis on Natural Science, given by the National Commission for Scientific Degrees in Cuba. She obtained one of the annual prizes given by the Cuban Academy of Sciences in 2011. She received the 2012 D. L Massart Award in Chemometrics, given by the Belgian Chemometric Society to the best Ph.D. thesis defended in the two years preceding the award ceremony. Recently, she won the TWAS-ROLAC Prize for Young Scientist 2012 in the field of Engineering Sciences.

Isneri Talavera Bustamante earned her diploma in chemistry from the University of Havana in 1980. She received the Ph.D. degree in Technical Sciences at Technical University "José A. Echeverría" (ISPJAE) in 1986, Havana, Cuba. She is currently the Vice Director of the Advanced Technologies Application Center, Cuba. Dr. Isneri is also a Senior Researcher and Assistant Professor at the University of Havana. She is the vice-president of the Academy of Sciences of Cuba. Her research has been focused on two main aspects: the study of paper chemistry and degradation processes of these materials as supporting documentation, and from 2003 she has been working on Chemometrics. In this field, her work has been directed towards the processing and multivariate classification of chemical images, incorporating the support vector machine (SVM). At present, she works on obtaining new calibration and multivariate prediction models with SVM and FDA and the classification tools for multiway data. At present, she is responsible for this research line at her institute and its diversification and employment in the country. Dr. Isneri received the Annual Award Academy of Sciences of Cuba in 2000, among others. She is a member of the Cuban Chemical Society and the Cuban Association of Pattern Recognition.