Eurographics Workshop on 3D Object Retrieval (2008)
I. Pratikakis and T. Theoharis (Editors)

Face Recognition by SVMs Classification and Manifold Learning of 2D and 3D Radial Geodesic Distances

Stefano Berretti[1], Alberto Del Bimbo[1], Pietro Pala[1] and Francisco Josè Silva Mata[†2]

[1]University of Firenze, Firenze, Italy
[2]Center for Advanced Technological Applications, Havana, Cuba

Abstract

An original face recognition approach based on 2D and 3D Radial Geodesic Distances (RGDs), respectively computed on 2D face images and 3D face models, is proposed in this work. In 3D, the RGD of a generic point of a 3D face surface is computed as the length of the particular geodesic that connects the point with a reference point along a radial direction. In 2D, the RGD of a face image pixel with respect to a reference pixel accounts for the difference of gray level intensities of the two pixels and the Euclidean distance between them. Support Vector Machines (SVMs) are used to perform face recognition using 2D- and 3D-RGDs. Due to the high dimensionality of face representations based on RGDs, embedding into lower-dimensional spaces using manifold learning is applied before SVMs classification. Experimental results are reported for 3D-3D and 2D-3D face recognition using the proposed approach.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Curve, surface, solid, and object representations

1. Introduction

Person identification based on facial data has been largely addressed in the last years mainly focussing on the detection and recognition of faces in 2D still images and videos [ZCPR03]. However, the success of solutions based on 2D imaging is jeopardized in real application contexts, where invariance to pose and illumination conditions remains a largely unsolved problem. As a result, the accuracy of these solutions is not satisfactory to be used to support automatic person recognition in real world application scenarios.

Recently, three-dimensional (3D) facial data has been exploited as a means to improve the effectiveness of face recognition systems [BCF06]. Since 3D face models are less sensitive, if not invariant, to lighting conditions and pose variations, recognition based on 3D facial data entails the potential for better accuracy and robustness.

However, a common drawback of solutions that perform recognition by matching 3D facial data is that, despite recent advances in 3D acquisition technologies and devices, acquisition of 3D facial data of a person can be accomplished only in controlled environments and requires the person to stay still in front of a 3D scanning device for a time that ranges from some seconds up to a few minutes. In addition, multiple scans from slightly different acquisition view-points are typically necessary in order to reconstruct parts of the face that can be self-occluded when acquired from a particular view. The 3D face model is then constructed from multiple scans by post-processing steps which include registration and merging of the scans, holes filling, smoothing, regularization, etc. Accordingly, the adoption of *pure* 3D face recognition solutions is cast to a set of specific applications.

A viable solution is to adopt hybrid 2D-3D matching schemes in which 3D facial data is compared against 2D face images. In this case, the operational cycle of the recognition system includes two distinct steps: acquisition and recognition. Acquisition of 3D facial data of persons to be recognized is performed only once. Instead, recognition takes place by comparing acquired 3D facial data to images (video frames) taken on the fly as people transit through surveilled areas, monitored by video-camera systems.

† This work was partially done when the author was at the University of Firenze, Italy.

Some of these 2D-3D matching solutions operate by transforming (projecting) the 3D geometry to 2D images so as to exploit well established representation techniques developed in 2D in order to perform recognition. As an example, in [PCJ05], face recognition in videos is made invariant to pose and lighting by using 3D face models. 3D database models are used to capture a set of projection images taken from different points of view. Similarity between a target image and 3D models is computed by matching the query with the projection images of the models. In other solutions, 3D models are mostly used to obtain 2D images representing particular views of the model to be used in the 2D-2D match with face images [BV03].

Approaches based on conformal transformations exploit the property which allows any surface homeomorphic to a disc to be mapped to a 2D planar domain [HAT*00]. The conformal mapping is one-to-one, onto, and angle preserving thus simplifying the 3D surface-matching to a 2D image-matching problem. In [WCT05], this solution has been applied with an experimentation limited to a few face and brain models. In [WWJ*06], least squares conformal geometric maps are applied to 3D faces, and results are provided for a relatively small database comprising 100 scans.

Use of the eigenface approach based on Principal Component Analysis (PCA)–first introduced for 2D faces in [TP91]–has been reported in [HPA04], [PHWW05]. In this latter approach, first a region of interest is defined as the intersection between a sphere centered on the nose tip with the face surface. Then, the region is parameterized into an isomorphic 2D planar circle trying to preserve the intrinsic geometric properties of the surface, and also mapping its relative depth values. Eigenface analysis is finally performed on the mapped relative depth image and used to compare faces.

Although these solutions show that 3D information can boost the recognition rates, none of them use the actual 3D geometry either as direct input or in the match. Rather, the 3D model is an intermediate source to render 2D views of a 3D model from different viewpoints and under different illumination conditions so as to best match 2D facial data represented in 2D images. In contrast, direct comparison of 2D facial data to 3D geometric information would enable more reliable matching as 3D geometric information is by its nature invariant to lighting and pose.

A possible way to extract 3D geometric information of a face model is to measure distances among 3D points of the model surface. The use of distances to capture 3D facial information is directly motivated by the relevance that face metrology has in studies conducted in medical disciplines. In particular, the form and values of these measurements are defined in *face anthropometry*, the biological science dedicated to the measurement of the human face. This field has been largely influenced by the seminal work of Farkas [Far94]. In particular, Farkas proposed a total of 47 landmark points to describe the face, with a total of 132 measurements on the face and head. In these measurements, geodesic, Euclidean, and angular distances between facial landmarks are used. Until recently, the measurement process could only be carried out by experienced anthropometrists by hand, but recent works have investigated 3D scanners as an alternative to manual measurements.

In 2D, shading plays an important role in the human perception of surface shape. Artists have long used lighting and shading to convey vivid illusions of depth in paintings. Researchers in human vision have attempted to understand and simulate the mechanisms by which our eyes and brain actually use the shading information to recover the 3D shapes. In computer vision, the idea of using the gradual variation of shading in an image to recover 3D shape dates back to the first studies on shape-from-shading [Hor77]. A vast literature exists on this subject, and interesting results have been obtained [ZTCS99]. However, in these solutions the final objective is the reconstruction of the 3D shape of the entire object. More related to our work are researches on computing geodesic distances in 2D images. For example, in [LJ05] geodesic sampling is used treating a 2D image as a surface embedded in a 3D space. In this framework, image intensity is weighted relative to the distance in the x-y plane of the image. It is shown as this weight increases, geodesic distances on the embedded surface are less affected by image deformations so that, in the limit, distances are deformation invariant. Geodesic distance measures have also been used in object recognition. For example, in [EK03] they are used to build bending invariant signatures for real surfaces.

Grounding on previous considerations, in this paper we propose an original framework to represent 2D and 3D facial data using Radial Geodesic Distances (RGDs) computed with respect to a reference point of the face (i.e., the nose tip). The objective is to define a face representation that can be extracted from 2D face images as well as from 3D face models and used to directly compare them in order to perform recognition. In 3D, the RGD of a point on the face surface is computed as the length of the particular geodesic that connects the point to the nose tip along a radial direction. In 2D, the RGD from a pixel to the fiducial point is computed based on the differences of the image gray level intensities along a radial path on the image. Matching between 2D- and 3D-RGDs result into feature vectors which are classified by a set of Support Vector Machines (SVMs). Since the feature vectors lay in a high-dimensional space, dimensionality reduction methods are applied before SVMs classification. Results on 3D-3D face recognition using 3D-RGDs, and preliminary results on 2D-3D face recognition, show the viability of the approach.

The paper is organized as follows. In Sect.2, 3D-RGDs are defined and used to capture geometric characteristics of a face. In Sect.3, a 2D face representation based on the computation of 2D-RGDs in the intensity domain of an image is presented. Issues related to the matching of 2D- and 3D-

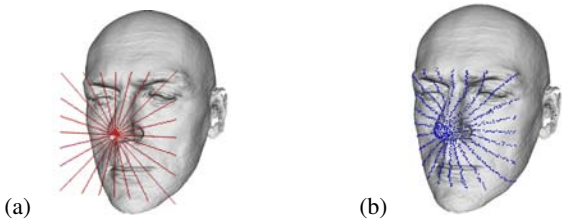
RGDs are also addressed in this Section. In Sect.4, the face recognition process is encompassed into a framework which includes dimensionality reduction and SVMs classification of the matching between 2D- and 3D-RGDs. Based on this framework, in Sect.5 3D-3D face recognition results using 3D-RGDs are presented, and preliminary results for 2D-3D face recognition based on SVMs classification of RGDs are reported. Conclusions and future research directions are outlined in Sect.6.

2. 3D radial geodesic distances

The *Morse*'s theory [Mil63], first proposed the idea of defining smooth real valued function on the surface of a 3D model in order to capture its characteristics. In this theory, differential properties of the function are used to make explicit the topological properties of a surface, and different characteristics of the surface can be evidenced depending on the choice of the function.

Following this idea, in our case the function in a generic point of a 3D model surface is defined as the 3D *radial geodesic distance* (3D-RGD) between the point and the *nose tip*. A *radial geodesic* is defined as the particular geodesic that connects one point of the model to the nose tip along the radial direction connecting the two surface points (see Fig.1(b)). According to this, a 3D face representation is constructed by considering K radial geodesics taken at fixed angular steps in the $[0,360]$ interval, with N sampling points per radial geodesic. In this way, a 3D face model is represented by a feature vector of size $K \cdot N$, whose elements are the 3D-RGDs computed for the sampling points.

In order to make the 3D face representations comparable among them, and with face representations extracted from 2D images, the sampling points along radial geodesics are selected based on a $2N \cdot 2N$ grid of points (see Fig.1(a)). By projecting points of the grid onto the 3D model surface a set of sampling points in 3D is identified. Values of the 3D-RGD are computed for each sampling point in 3D.

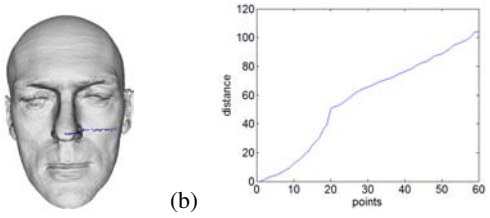


(a) (b)

Figure 1: *(a) The $2N \cdot 2N$ grid laying on the XY plane; (b) radial geodesics corresponding to the grid. Along each radial geodesic, the 3D-RGD is computed for the vertices of the model which are the nearest to the projections of the sampling points of the grid.*

In practice, the surface S of a face model is approximated through a discrete mesh M with n vertices $v_1,\ldots,v_n$, with the fiducial vertex v_f located at the nose tip. As a consequence, 3D-RGDs are computed for the vertices of the mesh that are the nearest to the projections of the grid points on the model. Computation of the 3D-RGD for a sampled vertex v_i along a radial geodesic R^k is obtained as the length of the shortest piecewise linear path on mesh vertices connecting the vertex v_i with the nose tip vertex v_f along R^k: $\mu_3^k(v_i,v_f) = L(P^k(v_i,v_f))$ (we will refer to this as $\mu_3^k(i)$). In this expression, $P^k(v_i,v_f)$ is the sequence of vertices along R^k from v_i to v_f, defined as an ordered sequence of adjacent vertices, and $L(P^k(v_i,v_f))$ is the length of the path measured as the sum of the Euclidean distances between adjacent vertex pairs. Furthermore, all the vertices of the path are constrained by the additional condition: $P^k(v_i,v_f) = P^k(v_{i-1},v_f) \bigcup v_i$ for $i = 2,\cdots,N$. This ensures that the set of vertices of a radial geodesic is repeatedly extended by adding the new vertex v_i to the current set.

The 3D-RGDs values allow to capture the differences occurring on the model for points along the radial geodesics. Fig.2(a), shows a 3D face model where the radial geodesic originated from the nose tip and oriented along the direction at 0 degrees is evidenced. In Fig.2(b), the 3D-RGD computed along this direction is shown for 60 sampling points of the grid. It can be observed that the 3D-RGD values capture information on the profile and the extent of the self-occlusion occurring at the base of the nose.



(a) (b)

Figure 2: *(a) A 3D face model with the radial geodesic at 0 degrees; (b) The 3D-RGD computed for the radial geodesic shown in (a).*

The final objective of this representation is to prove the 3D-RGDs capture salient face information, and can be used to perform face recognition through the comparison with 2D-RGDs computed for 2D images. As we will show in Sect.5, the 3D face representation based on 3D-RGDs can be also used to directly perform 3D face recognition.

3. 2D radial geodesic distances

2D face images capture intensity variations of the light reflected by the face surface. Therefore, pixel values are related to the reflectance properties and to the 3D geometry of the face. According to this, we aim to define a 2D face representation based on the adjacency and intensity variations

of image pixels that can be directly compared against the 3D face representation based on 3D-RGDs.

To this end, a face representation is constructed in 2D which has the same basic structure of the 3D representation. Similarly to 3D, the nose tip of the face image is used as fiducial point and radial paths on the image are considered. These originate from the nose tip and are extended along radial directions up to a fixed number of points in the image plane (see Fig.3). *2D radial geodesic distances* (2D-RGDs) are computed according to the intensity variations and proximity of image pixels. Considering a particular radial path R^k, the following equation is used to compute the 2D-RGD:

$$\mu_2^k(j) = \sum_{i=1}^{j} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + |\Delta I|^{w^k(\Delta I)}} \quad (1)$$

where: j is the index of the pixel along the radial path; x_i and y_i are the coordinates of the pixels along the radial path, and $\Delta I = |I(x_i, y_i) - I(x_{i-1}, y_{i-1})|$ is the difference of intensity between adjacent image pixels. The graph in the middle of Fig.3, provides a visual representation of Eq.(1).

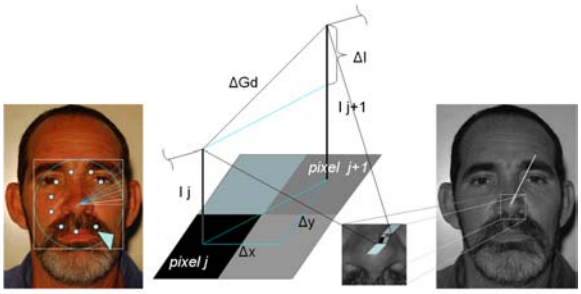


Figure 3: *The 2D-RGD computed according to the intensity values of the image. On the right, a particular radial path is shown. The central graphic in the figure illustrates the terms appearing in Eq.(1). On the left, multiple radial paths are shown.*

The exponent $w^k(\Delta I)$ of Eq.(1) is varied with the value of the intensity differences in order to establish the best correspondence between 2D and 3D measures. In particular, if the difference of intensity between two image pixels is zero (i.e., $\Delta I = 0$), $\mu_2^k(j)$ is reduced to the Euclidean distance between the pixels.

In order to determine the function $w^k(\Delta I)$ of Eq.(1), which maps 2D- to 3D-RGDs, we rely on a reference image. To this end, during the enrollment of a new subject into the gallery, a frontal face image is also acquired (3D scanners usually take this image) and used as reference. The mapping determined for the reference image is then applied to generic images of any subject that must be compared with the 3D model. To this end, for any radial geodesic R^k, the function $w^k(\Delta I)$ is determined which best maps the N-dimensional vector μ_2^k of the 2D-RGDs computed for the reference image, into the N-dimensional vector μ_3^k of the 3D-RGDs. Function $w^k(\Delta I)$ is found as solution of the following minimization problem:

$$w^k(\Delta I) = \min_{\Delta I} \sum_{j=1}^{N} |\mu_3^k(j) - \mu_2^k(j, \Delta I)|^2 \quad (2)$$

being j the index of corresponding points along the 3D radial geodesic and the 2D radial path.

3.1. Normalization of face images

In order to compare 2D- and 3D-RGDs, geometric normalization of 2D face images with respect to 3D face models is necessary. Normalization requires that at least one correspondence between a pair of image pixels and a pair of 3D points is identified, and that the distance computed between the pair of 3D points and the distance computed between the pair of 2D points are equal. If this is not the case, the image must be re-scaled accordingly. We used the Euclidean distance between the two *endocanthions* (i.e., the points at the inner commissure of the left and right eye fissure) computed in 3D to provide information of the real dimension of the face. These points have been verified to be easily detectable using curvature information, and stable with respect to face variations [BBK06], [CBF06]. We assumed this measure as an intrinsic characteristic of a 3D face model, and computed it during the enrollment of 3D face models into the *gallery* of known subjects. The algorithm in [CCS06] has been used to detect endocanthions and the nose tip in 3D. These points have been also used to align the models with respect to a global 3D reference system.

In 2D, image processing techniques have been used for the automatic detection of the face and for the identification of the two endocanthions and the nose tip. The face and eyes regions are first detected using a *Haar-cascade* detector [VJ04]. Endocanthions are automatically identified by processing the eyes region of the face in order to extract corner points (a *Harris* corner detector has been used). Many corner points are usually identified on the border of the eyes, the irises and the pupils. The corners corresponding to the endocanthions are selected using heuristics on their reciprocal positions, and on their positions with respect to the eyes region. To validate the position of the two endocanthions, an iris detector based on the *Hough* transform is used. This estimates the circles that best fit the irises and uses their position to validate the endocanthions extracted by the corner detector. As an example, Fig.4(a)-(b) show the eyes region, and the detected corners for the face images of three different subjects. The corners identified at the two endocanthions are highlighted in black in the figure.

When a face image must be compared against a 3D face model, the distance between the two endocanthions of the image is computed and the image is re-scaled according to

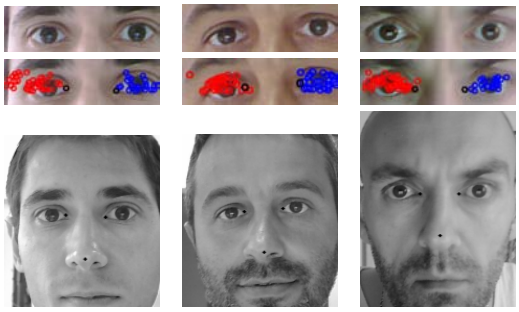


Figure 4: *Detection of the fiducial points of 2D face images of three sample subjects: (a) eyes regions identified using a Haar cascade detector; (b) corners detected in the eyes region. The corners identified as the two endocanthions are highlighted in black; (c) normalized images with the nose tip and the two endocanthions highlighted in black.*

the endocanthions distance associated to the model. This assumes that, given a 3D model, images of the same subject of the model are re-scaled accordingly and correspondence between 3D and 2D distances can be rightly established. In the same way, images of subjects different from the model are re-scaled with the endocanthions distance associated to the model, but this likely determines wrong correspondence between 3D and 2D measures. Automatic detection of the nose tip in 2D is also necessary to extract the 2D-RGDs. Differently from the endocanthions, accurate detection of the nose tip is difficult to be performed in the image. We solved this problem by using information on the position of the nose tip in the 3D model. After normalization, the position of the nose tip in the image is estimated by using the Euclidean distances between the two endocanthions and the nose tip measured on the 3D model. As an example, Fig.4(c) shows the face regions of three persons with evidenced (in black) the nose tip and the two endocanthions that have been automatically detected.

4. Manifold embedding of RGDs

In matching 2D- and 3D-RGDs the difference $e^k(j) = \mu_2^k(j) - \mu_3^k(j)$, is computed for every radial geodesic R^k. According to this, an *error vector* of size $K \cdot N$ is constructed and used to characterize the correspondence between a face image and a 3D face model.

In the experimented solution, $K = 72$ radial geodesics at intervals of 5 degrees are used, each with 50 points, thus resulting in an error vector of size 3600. Since 2D-3D face recognition is based on SVMs classification of the error vectors, operating in this high dimensional space can not be effective due to the *curse of dimensionality*. To avoid this difficulty, error vectors undergo to a dimensionality reduction before to be used for face recognition.

The process of transforming data residing in a high dimensional space to a low dimensional subspace is based on the assumption that the data actually lies, at least approximately, on a *manifold* of smaller dimension than the data space. The goal is to find a representation of that manifold that allows the projection of the data vectors on it and obtains a low-dimensional, compact representation of the data. Ideally, the reduced representation should have a dimensionality that corresponds to the intrinsic dimensionality of the data. Based on the type of the transformation function that performs the mapping between the high and the low dimensional space, *linear* and *non-linear* techniques can be distinguished. Linear techniques assume that the data lies on or near a linear subspace of the high-dimensional space. Non-linear techniques do not rely on the linearity assumption as a result of which more complex embedding of the data in the high-dimensional space can be identified. Among linear techniques, we experimented the *Principal Component Analysis* (PCA), while we considered *Multidimensional Scaling* (MDS), *Isomap*, *Locally Linear Embedding* (LLE) and *Laplacian Eigenmaps* (LE) as non-linear methods [VPV07].

4.1. SVMs classification

Once face representations are embedded into a low-dimensional space, face recognition/authentication is managed as a classification problem using SVMs with a *radial basis function* kernel [Vap98] (the *libsvm* package [CL01] through the *Weka* environment has been used: `http://www.cs.waikato.ac.nz/~ml/weka/`).

SVMs belongs to the class of maximum margin classifiers [Vap98]. In a *binary* classification problem, they find a decision surface that has maximum distance to the closest points in the training set (called *support vectors*). Given a set of points $\mathbf{x}_i \in \Re^n$, $i = 1,\ldots,L$, let us suppose each point $\mathbf{x}_i$ belongs to one of two classes identified by the label $y_i \in \{-1,1\}$. Assuming for simplicity that data is linearly separable, the goal of maximum margin classification is to separate the two classes by a hyperplane such that the distance to the support vectors is maximized. This *optimal separating hyperplane* has the form:

$$f(\mathbf{x}) = \sum_{i=0}^{l} \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x} + b \tag{3}$$

where α_i and b are the solutions of a quadratic programming problem [Vap98]. Classification of a new data point $\mathbf{x}$ is performed by computing the sign of the right side of Eq.(3).

The construction can be extended to the case of nonlinear separating surfaces. Each point in the input space is mapped to a point $\mathbf{z} = \Phi(\mathbf{x})$ of a higher dimensional space. In this *feature space*, the data is separated by a hyperplane. The main property of this construction is that the mapping $\Phi(.)$ is subject to the condition that the dot product of two points in the

feature space $\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})$ can be rewritten as a *kernel function* $K(\mathbf{x},\mathbf{y})$. The decision surface has the equation:

$$f(\mathbf{x}) = \sum_{i=0}^{l} \alpha_i y_i K(\mathbf{x},\mathbf{y}) + b \qquad (4)$$

where, similarly to Eq.(3), α_i and b are the solutions of a quadratic programming problem. In both Eq.(3) and Eq.(4), it is relevant to note that $f(\mathbf{x})$ does not depend on the dimensionality of the feature space.

Two main strategies can be used to solve *multi-class* problems with SVMs [HHP01] (in the following, q classes are considered, each corresponding to a 3D model in the gallery). In the *one*-vs-*all* approach, q SVMs are trained. Each SVM separates a single class from all the remaining classes. In the *pairwise* approach, $q(q-1)/2$ machines are trained, each separating a pair of classes. The pairwise classifiers are arranged in a tree, where each tree node represents a SVM. Regarding the training effort, the one-vs-all approach is preferable since only q SVMs have to be trained, compared to $q(q-1)/2$ SVMs in the pairwise approach. The run time complexity of the two strategies is similar. For recognition, the one-vs-all approach requires the evaluation of q SVMs, while the evaluation of $q-1$ SVMs is required for the pairwise approach.

5. Experimental results

In the following, we report on experiments of 3D-3D face recognition, and 2D-3D face authentication using the *Radial Geodesic Distance* (RGD) approach.

5.1. 3D-3D recognition experiments

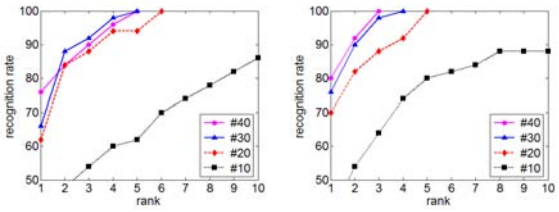
3D-3D face recognition shows the capability of 3D-RGDs to effectively discriminate between 3D face models and provides an indication of the significativeness of the representation. In this work, 3D-RGDs are not directly proposed as 3D face recognition approach. State of the art solutions for 3D face matching [KPT*07], [MBO07], can be addressed for this purpose. In these experiments, we directly used the 3D-RGDs in the space with $K \cdot N$ dimensions, without any dimensionality reduction. This is motivated by the objective to test the intrinsic information that is captured by the 3D-RGD representation.

According to this, to compare 3D-RGDs of two face models A and B, the *Euclidean* distance between 3D-RGDs in the space of size $K \cdot N$ has been evaluated:

$$D(\mu_3(A),\mu_3(B)) = \left(\sum_{j=1}^{K}\sum_{i=1}^{N}(\mu_3^j(i) - \mu_3^j(i))^2\right)^{1/2} \qquad (5)$$

The Gavab face database (publicly available at `http://gavab.escet.urjc.es/`) has been used in these experiments. It includes 3D face models of 61 individuals (45 males and 16 females). The whole set of subjects are Caucasian and most of them are aged between 18 and 40. For each person, 7 different models are taken, differing in terms of acquisition pose or facial expression, resulting in 427 facial models. In particular, for each subject there are 2 neutral frontal and 2 neutral rotated models, and 3 frontal models in which the person laughs, smiles or exhibits a random expression. Models are coded in VRML with resolution of approximately 10000 vertices.

For each individual, one of the two scans with frontal view and neutral expression is used as reference model and included in the gallery. All the other scans of a subject are used as probes. According to this, we conducted a set of recognition experiments using 366 probes (with neutral and non-neutral facial expression) on a gallery of 61 models. Each probe is compared against all the gallery models producing a result list of gallery models ranked in increasing order of scored distance from the probe. The effectiveness of recognition has been measured according to the rank-k recognition rate, and presented with *Cumulative Matching Characteristics* (CMC) curves. In particular, a rank-k recognition experiment is successful if the gallery face representing the same individual of the current probe is ranked within the first k positions of the ranked list. CMC curves measure, for each k value, the corresponding percentage of successful rank-k experiments.



(a) (b)

Figure 5: *CMC curves for 3D face recognition based on 3D-RGDs: (a) $N = 18$ radial geodesics displaced by 20 degrees; (b) $N = 36$ radial geodesics displaced by 10 degrees. In both the cases, curves for different number of points are reported.*

In order to tune the parameters of the 3D-RGD approach, the Gavab database has been first divided into a train set and a test set. In particular, we used the models of 11 subjects as train set (these subjects have been randomly selected in the database), while the models of the remaining 50 subjects have been used as test set. Then, operating on the test set, we performed a preliminary set of tests to investigate the importance of the number N of radial geodesics and of the number of points K along each radial geodesic. Results of these tests are reported in Fig.5(a)-(b) for models with neutral expressions. In Fig.5(a) the CMC curves are reported for radial geodesics taken at intervals of 20 degrees in the

range $[0,360]$, with 10, 20, 30 and 40 points along the radial geodesics, respectively. These points were taken with uniform displacement between each other along the radial geodesics. In Fig.5(b), experiments have been performed using the same number of points, but with radial geodesics displaced by 10 degrees. In general, it can be observed that increasing the number of radial geodesics, the rank-1 recognition rate improves independently from the number of points used along the radial geodesics (compare plots in (a) against plot in (b)). Similarly, it emerges that increasing the number of points along radial geodesics also improves the recognition rates (compare the plots for different number of points). However, though there is a remarkable increase of the performance passing from 10 to 20 points, the relative increase from 20 to 30 points, and from 30 to 40 points are not so relevant. In particular, we can observe a sort of saturation that does not further improve the recognition rates.

Table 1: *Rank-1 recognition rates for probes with neutral and non-neutral expression.*

	rank-1 recognition rate	
	neutral frontal	non-neutral expressions
3D-RGD	86.9%	75.4%
ICP	70.5%	62.3%

Based on the results of the previous tests, matching of 3D-RGDs, has been performed using 72 radial geodesics displaced by 5 degrees, with 50 points each (3600 points total). Using this setting, rank-1 recognition rates are reported in Tab.1 for the 3D-RGD approach, and for the 3D face matching solution that uses the Iterative Closest Point (ICP) registration [BM92]. Results evidence that RGDs are able to improve results of the ICP, for 3D face models with both neutral and non-neutral facial expressions.

5.2. 2D-3D authentication experiments

Preliminary authentication experiments have been performed to prove the viability of SVMs classification of the error vectors between 2D- and 3D-RGDs projected into an embedding subspace. Separate SVMs have been trained on five different projection subspaces, obtained using PCA, MDS, Isomap, LLE and LE, respectively (see Sect.4). Pairwise and one-vs-all classifiers have also been tested using different dimensions of the projection subspace. Percentage of correctly authenticated persons are reported in Tab.2 for three experiments. Each experiment has been conducted using 10-fold cross validation of frontal face images acquired for each subject under controlled illumination conditions (10 images per subject).

In the first experiment (Exp.1), SVMs *binary* classifiers are trained in a three dimensional projection subspace, one classifier for each ordered pair of gallery models. According to this, person authentication is obtained by a cascade of binary classifiers that use the same dimensionality reduction approach.

Table 2: *For each experiment and projection method, the percentage of correctly classified persons using SVMs is reported.*

	PCA	MDS	Isomap	LLE	LE
Exp.1	94.8	94.9	93.5	**97.4**	94.6
Exp.2	91.9	96.2	95	95.6	**97.5**
Exp.3	**99.4**	96.8	84.3	96.2	93.1

In the second experiment (Exp.2), SVMs *one*-vs-*all* classifiers are trained in a three-dimensional projection subspace, one classifier for gallery model. In this case, person authentication is obtained as response of one classifier. Following the same approach, in the third experiment (Exp.3), a SVMs *one*-vs-*all* classifier is trained in a projection subspace with fifteen dimensions, one classifier for gallery model. Also in this case, person authentication is obtained by the response of one classifier.

In general, it can be observed that the percentage of correctly classified persons is quite high. In particular, using *binary* classifiers (Exp.1), the dimensionality reduction methods have similar performance (maximum difference equal to 3.9%), with LLE scoring the best result (bold entries in the table evidence the approach that scores the highest percentage of correct classification in each experiment). Binary and one-vs-all classifiers have similar results (compare Exp.1 and Exp.2). The effect of increasing dimensionality of the embedding subspace does not emerge clearly from Exp.2 and Exp.3. In fact, the absolute maximum is obtained by the PCA approach in Exp.3, but ISOMAP and LE decrease their performance passing from Exp.2 to Exp.3. As final observation, though PCA performs reasonably well in the three experiments, non-linear methods like MDS and LLE seem able to provide better combination with the SVMs classification in several different conditions.

6. Conclusions and future work

In this paper, an original approach has been proposed for representing 2D face images and 3D face models, and to compare them for recognition and authentication purposes. 2D and 3D face representations are based on geodesic distances computed along radial directions originated from the nose tip. Experiments are reported for 3D-3D face recognition, together with preliminary experiments on 2D-3D face authentication using SVMs classification of the RGDs. Dimensionality reduction is applied to the RGDs before to perform classification using SVMs.

Future work will address different *feature selection* methods to perform dimensionality reduction for classification,

and a larger experimentation for 2D-3D face authentication in a real application context. The 3D-3D face recognition approach will be evaluated on larger benchmark data set, like the Face Recognition Grand Challenge database.

References

[BBK06] BRONSTEIN A., BRONSTEIN M., KIMMEL R.: Robust expression-invariant face recognition from partially missing data. In *Proc. European Conference on Computer Vision* (Gratz, Austria, May 2006), pp. 396–408.

[BCF06] BOWYER K., CHANG K., FLYNN P.: A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition. *CVIU 101*, 1 (Jan. 2006), 1–15.

[BM92] BESL P., MC KAY N.: A method for registration of 3-d shapes. *IEEE Trans. on PAMI 14*, 2 (February 1992), 239–256.

[BV03] BLANZ V., VETTER T.: Face recognition based on fitting a 3d morphable model. *IEEE Trans. on PAMI 25*, 9 (September 2003), 1063–1074.

[CBF06] CHANG K. I., BOWYER K. W., FLYNN P. J.: Multiple nose region matching for 3d face recognition under varying facial expression. *IEEE Trans. on PAMI 28*, 6 (October 2006), 1695–1700.

[CCS06] COLOMBO A., CUSANO C., SCHETTINI R.: 3d face detection using curvature analysis. *Pattern recognition 39*, 3 (March 2006), 444–455.

[CL01] CHANG C.-C., LIN C.-J.: *LIBSVM: a library for support vector machines*, 2001. Software available at: http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[EK03] ELAD (ELBAZ) A., KIMMEL R.: On bending invariant signatures for surfaces. *IEEE Trans. on PAMI 25*, 10 (October 2003), 1285–1295.

[Far94] FARKAS L.: *Anthropometry of the Head and Face*. Raven Press, New York, NY, 1994.

[HAT*00] HAKER S., ANGENENT S., TANNENBAUM A., KIKINIS R., SAPIRO G., HALLE M.: Conformal surface parameterization for texture mapping. *IEEE Trans. on Visualization and Computer Graphics 6*, 2 (April-June 2000), 181–189.

[HHP01] HEISELE B., HO P., POGGIO T.: Face recognition with support vector machines: Global versus component-based approach. In *Proc. ICCV* (July, Vancovuer, Canada 2001), pp. 688–694.

[Hor77] HORN B.: Understanding image intensities. *Artificial Intelligence 8*, 2 (April 1977), 1250–1267.

[HPA04] HESELTINE T., PEARS N., AUSTIN J.: Three-dimensional face recognition: an eigensurface approach. In *Proc. International Conference on Image Processing* (Singapore, October 2004), vol. 2, pp. 1421–1424.

[KPT*07] KAKADIARIS I. A., PASSALIS G., TODERICI G., MURTUZA N., LU Y., KARAMPATZIAKIS N., THEOHARIS T.: Three-dimensional face recognition in the presence of facial expressions: An annotated deformable approach. *IEEE Trans. on PAMI 29*, 4 (April 2007), 640–649.

[LJ05] LING H., JACOBS D.: Deformation invariant image matching. In *Proc. International Conference on Computer Vision* (Beijing, China, October 2005), vol. II, pp. 1466–1473.

[MBO07] MIAN A. S., BENNAMOUN M., OWENS R.: An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *IEEE Trans. on PAMI 29*, 11 (November 2007), 1927–1943.

[Mil63] MILNOR J.: *Morse Theory*. Princeton University Press, Princeton, NY, 1963.

[PCJ05] PARK U., CHEN H., JAIN A. K.: 3d model-assisted face recognition in video. In *Proc. Canadian Conference on Computer and Robot Vision* (Victoria, Canada, May 2005), pp. 322–329.

[PHWW05] PAN G., HAN S., WU Z., WANG Y.: 3d face recognition using mapped depth images. In *Proc. Conference on Computer Vision and Pattern Recognition* (San Diego, CA, June 2005), vol. 3, pp. 175–181.

[TP91] TURK M., PENTLAND A.: Eigenfaces for recognition. *Journal of Cognitive Neuroscience 3*, 1 (March 1991), 71–86.

[Vap98] VAPNIK V.: *Statistical Learning Theory*. John Wiley and Sons, New York, NY, 1998.

[VJ04] VIOLA P., JONES M. J.: Robust real-time face detection. *International Journal of Computer Vision 57*, 2 (May 2004), 137–154.

[VPV07] VAN DER MAATEN L., POSTMA E., VAN DEN HERIK H.: Dimensionality reduction: A comparative review. *Technical Report* (2007).

[WCT05] WANG Y., CHIANG M.-C., THOMPSON P. M.: Mutual information-based 3d surface matching with applications to face recognition and brain mapping. In *Proc. International Conference on Computer Vision* (Beijing, China, October 2005), pp. 527–534.

[WWJ*06] WANG S., WANG Y., JIN M., GU X., SAMARAS D.: 3d surface matching and recognition using conformal geometry. In *Proc. Conference on Computer Vision and Pattern Recognition* (New York, NY, June 2006), vol. 2, pp. 2453–2460.

[ZCPR03] ZHAO W., CHELLAPPA R., PHILLIPS P. J., ROSENFELD A.: Face recognition: A literature survey. *ACM Computing Survey 35*, 4 (Dec. 2003), 399–458.

[ZTCS99] ZHANG R., TSAI P.-S., CRYER J., SHAH M.: Shape from shading: A survey. *IEEE Trans. on PAMI 21*, 8 (Aug. 1999), 690–706.