3D Face Recognition using iso-Geodesic Surfaces

Stefano Berretti, Alberto Del Bimbo and Pietro Pala
Dipartimento di Sistemi e Informatica
Università degli Studi di Firenze
via S.Marta 3, 50139 Firenze, Italy
{berretti,delbimbo,pala}@dsi.unifi.it

Abstract

In this paper, we propose an original framework for description and matching of three dimensional faces for recognition. Basic traits of a face are encoded by extracting iso-geodesic stripes from the surface of a face model. A compact representation is constructed through a modeling technique capable to express the basic shape of iso-geodesic stripes and quantitatively measure their spatial relationships in the 3D space. This information is encoded in an attributed relational graph. Experimental results on a 3D face database and baseline comparison show that the proposed solution attains high face recognition accuracy and is reasonably robust to facial expression changes.

1. Introduction

Face recognition has been an active research area in the last years, with a major emphasis targeting detection and recognition of faces in still images and videos. More recently, the increasing availability of three-dimensional (3D) data, has paved the way to the use of 3D face models to improve the effectiveness of face recognition systems (see [1] for a recent survey). In fact, solutions based on 3D face models feature less sensitivity—if not invariance—to lighting conditions and pose. This is particularly relevant in real contexts of use, where face images are usually captured in non-controlled environments, without any particular cooperation by human subjects.

The recognition task requires the comparison between an individual face (probe) and all the model faces stored in a database (gallery). Since this task is usually performed on-line, efficiency issues are extremely relevant for recognition. In general, face recognition poses several issues which are closely related to those typically involved in retrieval applications: effective representation models and similarity measures, and efficient solutions to handle a one-to-many model comparison in a short time.

However, only few works have proposed 3D face recognition by jointly addressing expression variations and matching efficiency. One of the main reasons for this is related to the use, in most of previous works, of the iterative closest point (ICP) algorithm for surface matching [4]. Due to the computational complexity of this algorithm recognition under expression variations cannot be well coped with. In addition, although some preliminary results become available, the study of 3D face recognition under expression variations is just at the beginning and it is difficult to draw any conclusions about the robustness of available algorithms [3, 2].

Based on these considerations, in this paper we propose a 3D face recognition approach capable to account for expression variations and efficiently operate on large face galleries. In the proposed solution, iso-geodesic stripes of the face are identified by measuring distances of surface point to a fiducial point located on the nose tip. Facial information captured by iso-geodesic stripes is then represented in a compact form by extracting the basic 3D shape of each stripe and evaluating the spatial relationships between every pairs of stripes. To this end, we propose a modeling technique capable to quantitatively measure the spatial relationships between three dimensional entities. Then, we show how to extend the model so as to capture relationships between 2D surfaces in a 3D space. Finally, surfaces and their relationships are cast to a graph-like representation, where graph nodes are the representations of the iso-geodesics, and graph edges are their spatial relationships. Defining a distance measure between three-dimensional weighted walkthroughs (3DWW), and a matching algorithm for the extracted face representations allows the effective and efficient comparison of face models.

The rest of the paper is organized in four Sections. In Sect.2, a face model is decomposed into a set of iso-geodesic stripes by considering the distance of every mesh vertex to a fiducial point (the nose tip) automatically detected on the mesh. In Sect.3, the theory of 3DWW is developed, and a similarity measure for their comparison is also provided. Then, in Sect.4, the use of 3DWW to capture the salient shape information of iso-geodesic stripes and their mutual spatial relationships in the 3D space is expounded. This enables the effective representation and matching of a face model through an attributed relational graph accounting for the iso-geodesic stripes and their relationships. Face recognition results under expression variations and a baseline comparison are reported in Sect.5.
2. Iso-Geodesic Surfaces of Face Models

With respect to 2D face recognition, the use of 3D models entails the potential for supporting recognition through structural information about the face. In the proposed approach, structural information of a face model is captured through the 3D shape and relative arrangement of iso-geodesic stripes identified on the model surface. Iso-geodesic stripes are loci of surface points characterized by the same value of a function computed with respect to a fiducial point. The nose tip is identified as the reference point for the computation of the function. The value of the function on a generic point on the model surface is denoted as the normalized geodesic distance of the point to the nose tip. This normalization guarantees invariance to face expressions, this normalization factor doesn’t bias values of the function under expression changes.

Computation of the geodesic distance on the piecewise planar mesh is accomplished through the Dijkstra’s algorithm and approximates the actual geodesic distance with the length of the shortest piecewise linear path on mesh vertices. Once values of the function are computed for every surface points, iso-geodesic stripes can be identified. For this purpose, the range of Morse function values is quantized into $n$ intervals $c_1, \ldots, c_n$. Accordingly, $n$ level set stripes are identified on the model surface, the $i$-th stripe corresponding to the set of surface points on which the value of the Morse function falls within the limits of interval $c_i$.

![Figure 1. Iso-geodesic stripes for two face models with neutral facial expression.](image)

As an example, Fig.1 shows levels set stripes identified on the face models of two different individuals with neutral facial expression. In this figure, different colors are used to represent different iso-geodesic stripes (i.e., the red stripe corresponds to the face region nearest to the nose tip; the other stripes represent intervals of geodesic distance progressively farther from the nose tip).

3. Modeling Iso-geodesic Surfaces

In a three dimensional Cartesian reference system, with coordinate axes $X, Y, Z$, projections of two points, $a = (x_a, y_a, z_a)$ and $b = (x_b, y_b, z_b)$ on each axis, can take three different orders: before, coincident, or after (e.g., projection $x_a$ of point $a$ on the $x$-axis can precede, be coincident or follow—according to the positive versus of the $x$-axis—the projection $x_b$ of point $b$ on the same axis). The combination of the three projections results in 27 different three-dimensional displacements (primitive directions), which can be encoded by a triple of indexes $(i, j, k)$:

$$i = \begin{cases} -1 & x_b < x_a \\ 0 & x_b = x_a \\ +1 & x_b > x_a \\ \\ -1 & y_b < y_a \\ 0 & y_b = y_a \\ +1 & y_b > y_a \\ \\ -1 & z_b < z_a \\ 0 & z_b = z_a \\ +1 & z_b > z_a \end{cases}$$

In general, given two 3D extended sets of points $A$ and $B$, the pairs of points $(a, b)$ with $a \in A$ and $b \in B$ can be connected by different primitive directions. According to this, the triple $(i, j, k)$, is a walkthrough from $A$ to $B$ if it encodes the displacement between at least one pair of points $(a, b)$ with $a \in A$ and $b \in B$. In order to account for its perceptual relevance, each walkthrough $(i, j, k)$ is associated with a weight $w_{ijk}(A, B)$ measuring the number of pairs of points belonging to $A$ and $B$, whose displacement is captured by the direction $(i, j, k)$.

The weight is evaluated as an integral measure over the six-dimensional set of point pairs in $A$ and $B$ (see Fig.2):

$$w_{ijk}(A, B) = \frac{1}{K_{ijk}} \int_A \int_B f(x, y, z) \, dx_a \, dx_b \, dy_a \, dy_b \, dz_a \, dz_b$$

where $f(x, y, z) = C_i(x_b - x_a)C_j(y_b - y_a)C_k(z_b - z_a)$

where $K_{ijk}(A, B)$ acts as dimensional normalization factor, and $C_{\pm 1}()$ are the characteristic functions of the positive and negative real semi-axis $(0, +\infty)$ and $(-\infty, 0)$, respectively. In particular, $C_0() = \delta()$ denotes the Dirac’s function, and acts as a characteristic function of the singleton set $\{0\}$.

The 27 weights between $A$ and $B$ are organized in a $3 \times 3 \times 3$ matrix ($w(A, B)$), of indexes $i, j, k$ (see Fig.2), that we call the the 3D weighted walkthroughs (3DWW) between $A$ and $B$. Note that, as particular case, Eq.(1) also holds if $A$ and $B$ are coincident (i.e., $A \equiv B$).

![Figure 2. Walkthrough connecting a point in $A$ with a point in $B$. The $3 \times 3 \times 3$ relationship matrix between $A$ and $B$ is represented by three $3 \times 3$ matrices for $k = 1, 0, -1$.](image)
3.1. Distance Measure

Three directional weights, taking values within 0 and 1, can be computed on the eight corner weights of the 3DWW matrix (all terms are intended to be computed between two 3D sets of points A and B, i.e., \( w_{ijk} = w_{ijk}(A, B) \)):

\[
\begin{align*}
    w_H &= w_{1,1,1} + w_{1,-1,1} + w_{1,1,-1} + w_{-1,-1,1} \\
    w_V &= w_{-1,1,1} + w_{1,1,-1} + w_{1,-1,-1} + w_{-1,1,1} \\
    w_D &= w_{1,1,1} + w_{1,-1,1} + w_{-1,-1,1} + w_{-1,1,1}
\end{align*}
\]

These account for the degree by which B is on the right, up and in front of A, respectively. Similarly, seven weights account for the alignment along the three reference directions of the space: \( w_{H_0}, w_{V_0}, w_{D_0}, w_{HV_0}, w_{VD_0} \).

Based on these weights, similarity in the arrangement of pairs of 3D entities (A, B) and \((A', B')\) is evaluated by a distance \( D(w, w') \) which combines the differences between homologous weights in the 3DWW \( w(A, B) \) and \( w(A', B') \). This distance is expressed as:

\[
D(w, w') = \lambda_H |w_H - w'_H| + \lambda_V |w_V - w'_V| + \lambda_D |w_D - w'_D| + H_0 |w_{H_0} - w'_{H_0}| + V_0 |w_{V_0} - w'_{V_0}| + D_0 |w_{D_0} - w'_{D_0}| + H_D |w_{HD_0} - w'_{HD_0}| + V_D |w_{VD_0} - w'_{VD_0}|
\]

where \( \lambda_H, \lambda_V, \lambda_D, H_0, V_0, D_0, H_D, V_D \), and \( \lambda_D D_0 \), are non-negative numbers with sum equal to 1.

Distance \( D \) can be proven to be a metric. In addition, due to the integral nature of weights \( w_{ijk} \), \( D \) satisfies a property of \textit{continuity} which ensures that slight changes in the mutual positioning or in the distribution of points in two sets A and B result in slight changes in their 3DWW.

4. Face Representation

The theory of 3DWW has been developed by focusing on the general case of 3D entities with volumetric extension. However, in that Eq.(1) accounts for the contribution of individual pairs of 3D points, computation of spatial relationships between surfaces in 3D directly descends from the general case. For 3D surfaces, Eq.(1) can be written by replacing volumetric integrals with surface integrals extended to the area of two surfaces in the 3D space.

In practice, the complexity in computing Eq.(1) is managed by reducing the integral to a double summation over a discrete domain obtained by uniformly partitioning the 3D space. In this way, volumetric-pixels \( v_{xyz} \) (voxels) of uniform size are used to approximate entities (i.e., \( A = \bigcup_n A_n \), where \( A_n \) are voxels with a non-null intersection with the entity: \( v_{xyz} \in \{A_n\} \) iff \( v_{xyz} \cap A \neq \emptyset \)). According to this, 3DWW between \( A = \bigcup_n A_n \) and \( B = \bigcup_m B_m \) can be derived as linear combination of the 3DWW between individual voxel pairs \( \langle A_n, B_m \rangle \):

\[
w_{ijk}(\bigcup_n A_n, \bigcup_m B_m) = \frac{1}{K_{ijk}(A, B)} \sum_n \sum_m K_{ijk}(A_n, B_m) w(A_n, B_m)
\]

as can be easily proven by the properties of integrals. Terms \( w(A_n, B_m) \), indicating 3DWW between individual voxel pairs, are computed in closed form in that they represent the relationships occurring among elementary cubes (voxels) and only twenty-seven basic mutual-positions are possible between voxels in 3D.

4.1. Matching Face Representations

According to the proposed modeling technique, a generic face model \( F \), is represented through a set of \( N_F \) iso-geodesic stripes. In that 3DWW are computed for every pairs of iso-geodesic stripes (including the pair composed by a stripe and itself), a face is represented by a set of \( N_F \cdot (N_F + 1)/2 \) relationship matrices. This model is cast to a graph representation by regarding iso-geodesic stripes as graph nodes and their mutual spatial relationships as graph edges.

As an example, Fig.3 shows the graph derived from the iso-geodesic stripes of a face model. On the top of the figure, the iso-geodesic stripes of a sample face are shown as they appear from a frontal and lateral view. On the bottom, the complete graph constructed from the first six iso-geodesic stripes of the face is shown. To highlight the correspondence between graph nodes and iso-geodesic stripes, we used the same color to fill in homologous stripes and nodes. According to the proposed representation, all nodes and edges are labeled with a 3DWW matrix: in the figure, to preserve the readability, only some of the mutual weights are reported.

![Figure 3. The complete graph constructed on the first six iso-geodesic stripes of the face model. Only some of the labels associated to nodes and edges are shown.](Image 358x337 to 418x425)
an injective function which associates nodes \( p_k \) in the template graph with a subset of the nodes \( \Gamma(p_k) \) in the reference graph, this is expressed as follows:

\[
\mu^\Gamma(P, R) = \frac{\lambda}{N_P} \sum_{k=1}^{N_P} D(w(p_k, p_k), w(r_k, r_k)) + \\
\frac{2(1-\lambda)}{N_P(N_P-1)} \sum_{k=1}^{N_P} \sum_{h=1}^{k-1} D(w(p_k, p_h), w(r_k, r_h))
\]

where the first summation accounts for the average distance scored by matching nodes of the two graphs, and the second double summation evaluates the mean distance in the arrangements of pairs of nodes in the two graphs. In this equation, \( N_P \) is the number of nodes in the probe graph \( P \), and \( \lambda \in [0,1] \) balances the mutual relevance of edge and node distance.

5. Experimental Results

The proposed approach for 3D face recognition has been experimented using models from the GavabDB\(^1\). This includes three-dimensional facial surfaces 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 viewpoint or facial expression—resulting in 427 facial models. In particular, for each subject there are 2 frontal and 2 rotated models with neutral facial expression, and 3 frontal models in which the person laughs, smiles or exhibits a random gesture. Models are coded in the VRML format at one-to-one scale.

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 other scans of a subject are used as probes. According to this models subdivision, 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 the recognition is measured through the cumulative matching characteristic (CMC) curves.

In a preliminary set of tests we experimentally determined the optimal resolution at which iso-geodesic stripes are generated. Specifically, we found a good compromise between effectiveness of recognition and efficiency of comparison by using an extent of 0.08 for the intervals of geodesic distance which determine the iso-geodesic stripes. As discussed in Sect.2, geodesic distances computed on model surfaces are normalized, so that homologous stripes in different models represent equivalent levels of distance from the nose tip. This permits to use the same interval extent in the processing of every face models. The value of 0.08 is used to obtain the experimental results provided in the following. A second parameter which is relevant in comparing the representations of two face models is the number of iso-geodesic stripes used in the match. Of course, this number is strictly related to the resolution used for the extent of the stripes. In our setting, we found that a match comprising the first 8 stripes provides the best results.

![Figure 4. CMC curves for the ISG and the ICP approaches. Curves are separately given for face models with neutral expression and for models with expression changes.](image)

Results are reported in Fig.4 for the proposed approach (Iso-Geodesic Stripes, \( IGS \) for short), and for the ICP solution. Results evidence that the proposed approach is able to improve results of the ICP, especially in the case of expression variations.

6. Acknowledgment

This work is partially supported by the Information Society Technologies (IST) Program of the European Commission as part of the DELOS Network of Excellence on Digital Libraries (Contract G038-507618).

References


---

\(^1\)Publicly available at http://gavab.escet.urjc.es/