3D FACE REPRESENTATION AND RECOGNITION BY INTRINSIC SHAPE DESCRIPTION MAPS

Zhe Guo1,2, Yanning Zhang1,*, Yong Xia2,3, Zenggang Lin1, Dagan Feng2,3

1 Shaanxi Provincial Key Lab of Speech& Image Information Processing (SAIIP), School of Computer Science, Northwestern Polytechnical University, Xi’an, China;
2 Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, Australia;
3 Center for Multimedia Signal Processing (CMSP), Department of Electronic & Information Engineering, Hong Kong Polytechnic University, Hong Kong

ABSTRACT

We present a novel method for 3D face recognition, in which the 3D facial surface is first mapped into a 2D domain with specified resolution through a global optimization by constrained conformal geometric maps. The Intrinsic Shape Description Map (ISDM) is then constructed through a modeling technique capable to express geometric and appearance information of the 3D face. Hence the 3D surface matching problem can be simplified to a 2D image matching problem, which greatly reduces the computational complexity. Finally, the Intrinsic Shape Description Feature (ISDF) of ISDM and the discrimination analysis can be calculated. Experimental results implemented on GavabDB demonstrate that our proposed method significantly outperforms the existing methods with respect to pose variation.

Index Terms—Face recognition, Conformal Geometric Maps, Intrinsic Shape Description Map

1. INTRODUCTION

The increasing availability of 3D facial data has paved the way to use 3D models to improve the accuracy of face recognition. However, as a relatively new topic, a number of challenges still exist, which limit the performance and application of current 3D face recognition algorithms[1]. Two biggest challenges are precise alignment before matching stage and the computational efficiency.

Since most existing methods measure the individual similarity under uniform correspondence, model registration is the necessary requirement. Iterative Closest Point (ICP)[2], which is the most popular alignment algorithm, and widely used for 3D models registration. However, the iterative process makes ICP computationally expensive and this registration must be done for each model in database. Moreover, the precision of ICP will be decreased severely by unsuitable initialization and noise.

It has been proved that more information in 3D facial model leads to higher computational cost. Recently, several methods have been proposed based on 3D facial information and produced high identification rate[3,4,5]. Nevertheless, these methods all focus on computation in 3D space, computational complexity and efficiency become the major limitation for the practical application. Furthermore, shape representations using local shape signatures made these methods unstable for noise, occlusion and cluster.

This paper proposes a new 3D face recognition approach to deal with the problems described above, shown as Fig. 1. Our method is based on conformal geometric maps which does not need 3D models registration, and also maps 3D facial shape to a 2D domain which is a diffeomorphism through a global optimization. The 2D maps integrate geometric and appearance information and have the ability to describe the intrinsic shape of the 3D facial model, called Intrinsic Shape Description Maps (ISDMs). The identification experiments have been tested on the multi-pose subset of GavabDB[14]. The main contributions of our work are listed as follows,

1) To our knowledge, there are no try to using conformal geometric maps for 3D face recognition on public 3D face database. Although method in [6] used the similar strategy, the 3D face are self-collected and amount is quite small, only 10 people. 2) Our method does not need the 3D face posture alignment procedure.

2. CONFORMAL GEOMETRIC MAPS

2.1. Constrained discrete conformal maps

In addition to the notable advantages, such as sound mathematical basis, preservation of continuity of the underlying surfaces and angle preserving, which leads to less distortion and robust to noise, discrete conformal maps directly utilize geometric information of 3D surfaces and implement the fixed border surface mapping where surface unfolding with constrained conformal embedding to get the isometric 2D mapping results[7].

It can be proven that there exists a mapping from any surface with a disk topology to a 2D planar domain, which is
one-to-one, onto, and angle preserving[8]. As one of the most efficient methods for mapping disk-like surfaces into the plane, discrete conformal map parameterization was introduced by Eck et al. [9]. It attempts to lower angle deformation by minimizing a discrete version of the Dirichlet energy[10] using the finite element method based on linear elements.

For the fixed boundary mapping, the largest border of the 3D facial shape is first found as the fixed 2D planer domain border and arc-length border parameterization is then used as the border parameterization to define a set of constraints. Fig.2 shows one example of the homeomorphic 2D discrete structure of the 3D face.

![Fig. 1. The framework of our method.](image)

### 2.2. Image regularization

In this section, we present a polygon grids discretization method in the 2D planar domain to get the regular image description of the mapping results (MR). This is a kind of resampling regularization procedure, and can be implemented by interpolating pixel attributes at the mesh vertices [11].

The procedure of the image regularization method can be described as follows,

1) As the range of the $u, v$ value of the MR is $[0, 1]$, the bin size $s$ can be computed as

$$s = \frac{1}{m} \times \frac{1}{n} = s_u \times s_v \quad (1)$$

where $m, n$ is the height and width of the regular 2D image, respectively, which results in $m \times n$ bins.

2) For each bin $(i, j)$, the equations relating the coordinates of point $(\alpha, \beta)$ in 2D discrete image are

$$i = \left\lfloor \frac{1 - \beta}{s_u} \right\rfloor \quad j = \left\lfloor \frac{\alpha}{s_v} \right\rfloor \quad (2)$$

where $\lfloor f \rfloor$ is floor operator which rounds $f$ down to the nearest integer.

3) Given the bin size, the bilinear weights used to increment the bins in the MR can be calculated.

$$a = \alpha - is_u \
\quad b = \beta - js_u \quad (3)$$

4) Suppose the attribute value of the point $(\alpha, \beta)$ is $S(x)$, the contribution of the point is bilinearly interpolated to the four surrounding bins in the 2-D array.

$$SI(i, j) = SI(i, j) + (1 - a) \times (1 - b) \times S(x) \; ;$$

$$SI(i + 1, j) = SI(i + 1, j) + a \times (1 - b) \times S(x) \; ;$$

$$SI(i, j + 1) = SI(i, j + 1) + (1 - a) \times b \times S(x) \; ;$$

$$SI(i + 1, j + 1) = SI(i + 1, j + 1) + a \times b \times S(x) \; . \quad (4)$$

For each point, the bilinear weights are determined and new values are interpolated to the four surrounding bins.

### 3. INTRINSIC SHAPE DESCRIPTION MAPS

Based on the Conformal Geometric Maps, we investigate the 3D facial shape description and convert the 3D face recognition to a 2D image matching problem. In this section, we design a compact and intrinsic shape description method, called Intrinsic Shape Description Map (ISDM). The ISDM aims to depict the 3D facial intrinsic shape using a compact and efficient way.

Once we obtain mapping between 2D planes and surfaces in 3D space, various attributes of the surface like

![Fig. 2. Constrained conformal mapping result (a) original 3D model (b) the mapping result of (a).](image)
depth, normal, curvature or texture can be discretized in the image plane. Here we choose the discrete mean curvature which has the good ability to describe shape variation of 3D model as the attribute, i.e., $S(x)$ in Equation (4), therefore generate the corresponding attributed image.

3.1. Correspondence determination

Before obtaining ISDM, the correspondence between two attributed images should be first determination for the purpose of image analysis and comparison. Our correspondence determination method based on mutual information consists of the following four steps,

1) For each query and reference attributed image, noted as $I_q$ and $I_r$, the mutual information (MI) based method is applied to determine the corresponding points between two images.

The Parzen window is employed for the joint probability computation and the Marquardt-Levenberg optimization method [12] is used to maximize the MI. To speed up the computation, we use spline pyramids [13]. This method works with the entire image data and directly with image intensities.

2) The global mapping model uses bivariate polynomials is applied to estimate the transform model includes rotation and translation parameters.

3) The mapping functions are then used to transform the query image $I_q$ and the result is noted as $I_r$. The bicubic interpolation takes place in the query image on the regular grid.

4) The compact description ISDM is extracted for efficient computation by implying primary facial region from $I_q$.

Since the ISDM encodes the intrinsic shape of 3D face surface, we call it the Intrinsic Shape Description Map and use it as a critical presentation for face identification. Some examples of ISDMs are shown in Fig.3.

![Fig.3. Samples of the ISDMs.](image)

4. EXPERIMENTS

4.1. GavabDB database

GavabDB [14] contains 549 3D scans of facial surfaces corresponding to 61 different individuals (45 male and 16 female). There are 9 different scans for each person, two neutral frontal, four neutral with pose (looking left, right, down and up), and three frontal scans in which the subject presents different and accentuated facial expressions.

4.2. Preprocessing

This subsection briefly discusses the processing steps including face denoising, holes filling and facial region extraction as 3D face scans obtained by laser scanner usually contain spikes, holes and other facial accessories. All these operations exclude facial region extraction could be completed off-line using some existing efficient methods. A close face scan cropping scheme [15] was adopted to extract the compact facial region implying primary facial surface information so that the ISDMs can focus on the most significant part of the face. This method is simple and fast, but works very well on GavabDB.

4.3. Feature and classifier

For the ISDMs, the frequency form of the gradient distribution can depict shape variation and is robust to noise as its statistical character. Hence, we extract this kind of feature for identification, called Intrinsic Shape Description Feature (ISDF). The ISDF is composed of certain number of bins and the corresponding cumulative gradient value. By this means, for each ISDM, the ISDF is extracted as the feature description, and included in gallery for identification.

Suppose the number of bins we used is $N$, the $j^{th}$ bin is noted as $b_j$, the corresponding value of $b_j$ is $V_b$, so the ISDM can be parameterized as $N$ dimensional vector $V$, where $V = [V_1, V_2, ..., V_N]$. Consequently, each ISDM can be finally described as a $N$ dimensional feature vector $V$. We then utilize SVM to complete the recognition.

4.4. Experimental results and analysis

4.4.1. Discussion of the frequency form bins

This subsection briefly discusses the selection of the bins amount. The bins amount is the key point in 3D face recognition as it determines the dimension of the feature vector. However, the selection is hard to predict as less dimensions will weaken the discrimination between different individuals whereas more dimensions will increase the computational time. Therefore, the trade-off should be considered according to the experiments.

Here, we use a subset of GavabDB for the experiments, i.e., 61 different individuals, for each person, only two neutral frontal and two neutral with pose down and up are used, as left and right scans have major missing data. One of the neutral frontal scans is regarded as reference model and included in the gallery, and the remaining is used as probes. We change the bins amount from 50 to 400 with the interval of 10. The rank-one recognition rates are shown in Fig.4.

From Fig.4, it can be seen that for bins amount range [50,150], it’s inadequate to reflect the discrimination between individuals. The larger the amount is, the higher discriminative it will be, so the recognition rate increases gradually after 150, just as our analysis before. However, after the number 250, the recognition rate tends to be stable, and the computation time will increase due to the feature dimension. As a result, we select 250 as the optimal bins amount, and use the corresponding ISDFs for the following
identification experiments.

Table. 1 Identification rates of different methods on GavabDB(%).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral frontal</td>
<td>93.5</td>
<td>93</td>
<td>91.1</td>
<td>83.7</td>
<td>N/A</td>
<td>88.1</td>
</tr>
<tr>
<td>Multi-pose</td>
<td>N/A</td>
<td>83</td>
<td>NA</td>
<td>72.6</td>
<td></td>
<td>88.1</td>
</tr>
</tbody>
</table>

Fig.4. Rank-1 recognition rates with different bins amount.

4.4.2. Identification rate

In order to better assess our algorithm and compare with other solutions, the Leave-one-out cross validation strategy is applied to complete the identification on the experimental dataset, which includes 61 different individuals, each person 4 facial scans (2 frontal and 2 multi-pose), i.e., 244 scans, 183 in gallery, 61 as probes. It should be noticed that the multi-pose scans are included in our experimental dataset, and we use the Leave-one-out method, thus our experiments are the multi-pose face identification.

The results of our method and other existing methods for GavabDB are shown in Table1. The rank-one rate of our method for multi-pose face scans is much higher than the existing methods. Furthermore, our method does not need the 3D face posture alignment procedure and has the ability to implement identification totally automatically. In summary, our method has the significant advantage on multi-pose face identification.

5. CONCLUSION

An automatic 3D face recognition system using Intrinsic Shape Description Maps was proposed in this paper. This new shape representation simplified the 3D surface matching problem to a 2D image matching problem. Furthermore, conformal geometric maps do not need the 3D face posture alignment procedure. As the experimental results show, our method achieved the rank-one recognition rate 88.1% on multi-pose dataset, which has the obvious advantages and outperforms the existing methods.

6. ACKNOWLEDGEMENT

The authors would like to acknowledge the use of the GavabDB face database in this paper due to Moreno and Sanchez [14]. This work was partly supported by the “863” Program of China (No.2009AA01Z215), the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (No. 708085), and National Natural Science Foundation of China (No. 60872145).

7. REFERENCES