ROBUSTNESS AND EXPRESSION INDEPENDENCE IN 3D FACE RECOGNITION

Shun Miao, Hamid Krim

North Carolina State University
ECE Department
Raleigh, NC-27606, USA

ABSTRACT

We describe a robust method for 3D face recognition under variance of facial expressions. The method utilizes the identical areas on two facial images as the measurement of distance. To segment the identical area, partial shape matching is performed using closest point registration and level set method. The segmentation problem is formalized into an Eikonal equation, which can be efficiently solved by fast marching method. The presented 3D face recognition method shows a very promising recognition rate.

Index Terms—Face Recognition, Level Set Method, Fast Maching Algorithm, Iterative Closest Point (ICP)

1. INTRODUCTION

With the rapid progress of 3D scanning techniques, and the steadily decreasing cost of cameras, research interest in 3D face recognition has greatly increased over the last ten years. In contrast to 2D face recognition techniques, 3D approaches enjoy the capacity of reducing the limiting impact of uncontrolled parameters, such as lighting and pose, as it accounts for the face intrinsic geometric information. Facial expressions which may generate a non-rigid deformation on a human face, however, remain a significant limitation for face recognition applications.

As surfaces, 3D faces may be compared using surface matching techniques, which invoke rotation and translation transformations to establish alignment and measure their similarity. The Iterative Closest Point (ICP) [1] algorithm is a very efficient and popular surface matching technique, and also widely used in 3D face recognition. ICP algorithm registers two surfaces by closest point and iteratively optimizes the alignment. It has, however, two major drawbacks: a. the optimization will for instance converge to a local minimum in some cases. While numerous improvements have reduced the likelihood of such an event [2], their successes remain heavily dependent on the initial alignment. The Robust Point Matching (RPM) algorithm [3] has addressed this issue and has improved the dependency on initial alignment by an exhaustive search which is reduced over time with an appropriate annealing schedule. b. When face images are modified by facial expressions, any rigid surface matching/registration method will inevitably lead to a misalignment. The recognition rate of ICP-based face recognition techniques will thus be significantly affected by facial expressions.

This challenge has recently led researchers to study the characterization of face expressions, and to attempt to reduce their impact. Luet. al [4] have improved on previous attempts using a Thin-Plate Spline (TPS) approach for deformation analysis to result in a deformable model able to synthesize different facial expressions from a given neutral face. Bronstein et. al [5] proposed an isometric model of facial expressions to allow for deformation related to facial expressions. In [6], Li et. al proposed to use multiple intrinsic geometric features, including angles, geodesic distance and curvatures for face recognition, and a training procedure to obtain weights for extracted features, according to their sensitivity to facial expressions.

This paper describes a method of partial shape matching which in turn forms the basis for our proposed robust 3D face recognition. Intuitively speaking, facial expressions only affect part of a human face, leaving the other invariant under expressions transformations which can be used for face recognition. To that end, we introduce a 3D surface segmentation approach to adaptively extract identical parts between two facial images which are in turn compared for partial similarity. The latter is small between different subjects, and large between images from the same subject. Based on this argument, the identical facial area may be used as a distance measure between two facial images.

The balance of this paper is organized as follows. In section 2, we discuss in detail the surface registration step, the alignment and the region growing scheme. In section 3, we will show some promising experimental results on FRGC II database [7]. In section 4, we summarize and provide some concluding remarks.

2. METHOD

In this section, we describe a segmentation method to detect partial shape matching and for comparison. Given a reference 3D face and a 3D input face, we grow a facial area on the 3D input face to perform partial shape matching and to delineate the partial area similarity. The 3D input face is converted to a depth image and the delineation (segmentation) is performed on the parameter space. We first extract an initial area followed by a coarse alignment achieved by registering this initial area and the reference face. We subsequently and iteratively grow the matched area and adjust the alignment between the matched area and the reference face until the mean square distance between the two, reaches a predefined threshold. The initialization of a coarse alignment is chosen and the alignment refinement is achieved by the Iterative Closest Point (ICP) algorithm [1], which results in a local minimum of the mean square distance between the matched area and the reference face. To perform the evolution of the delineating curve of a partial match, a level set method is applied on the parameter space of the input face, and is implemented by the fast marching algorithm [8].

2.1. Closest Point Registration

For most 3D surface matching, comparison and segmentation techniques, surface registration is a necessary and important step. In this paper, the proposed surface growing technique is highly dependent on local similarity measures, which require registration between two 3D surfaces. We adopt a simple and very efficient registration technique- Closest Point Registration, which establishes correspondence between vertices in one 3D mesh to the closest (minimum Euclidean distance) point in the other 3D mesh. The computational complexity of Closest Point Registration is N2.

2.2. Front Propagation

We initialize a closed curve in the parameter space $x-y$ plane of the reference image, and evolve the curve to detect the identical partial shape. The constraint of the curve evolution is that the interior of the closed curve be matched to a subset of the input face image. To perform this curve evolution, an initial curve needs to be defined. As shown by Chang, K.I. et al. [9], the nose region is the most stable part of a human face, and remains invariant under extreme facial expressions. In other words, given two 3D facial images A and B of the same subject, a small area around the nose tip in A can be perfectly matched to B. Therefore, the initial curve is chosen to be a small rectangular near the nose tip in A.

As noted earlier, to perform the front propagation, we use a level set method. A level set curve defined by a level set function $\phi(x,y,t)$ is a path $(x(t),y(t))$ on which the level set value must be zero. The evolution equation for this level set curve is [8]

$$\phi_t + F|\nabla\phi| = 0 \quad (1)$$

Given an interface identified as a level set curve of the level set function ϕ, Equation 1 describes the time evolution of the interface. Consider a special and simple case where the level set function $\phi(x,y)$ is defined as the time at which the interface crosses the point (x,y), denoted as $T(x,y)$. In this special case, the derivative of the level set function with respect to time t is always equal to 1, thus Equation 1 becomes a partial differential equation,

$$|\nabla T|F = 1 \quad (2)$$

This is a form of the well-known Eikonal equation, which can generally be solved by a fast marching method [8]. The boundary condition in this case is $T(x,y) = 0$ on the initialization curve.

A discrete form of Equation 2 is introduced by Rouy and Tourin [10] as

$$\frac{1}{F^2} = max\ (max(D_i^{-x}j,0)^2\ , max(D_i^{+x}j,0)^2 \\ + max\ (max\left(D_i^{-y}\ j,0\right)^2, max\left(D_i^{+y}j,0\right)^2 \quad (3)$$

To solve the Eikonal equation, a boundary condition is needed. The boundary condition of T is determined by the initialization of the curve, on which $T(x,y)$ is set to be zero. Based on the above argument, we know that if the ICP algorithm is performed on the nose region of the same subject, the mean square distance between them should be almost zero. A zero mean square distance between two 3D surfaces clearly indicates that these two surfaces are identical. Denote S_{in} as the region inside level set curve. The basic idea is to grow the region S_{in} with the constraint that the mean square distance between S_{in} and B is close to zero, to indicate the identity property. Intuitively, we would like the front of the level set curve to propagate faster in those regions which are invariant to the presented facial expression and slower in those regions which are affected by a non-rigid deformation. Thus, the speed vector is chosen to be $(x,y) = 1/D(x,y)^{\lambda}$, where $D(x,y)$ is the distance to surface B based on closest point registration. This is a monotonically decreasing function of $D(x,y)$, and it satisfies the notion that a higher speed should be assigned to the regions which are more similar, and lower speed to the regions which are more dissimilar.

2.3. Algorithm

fast marching algorithm was proposed by Sethian [8] as a efficient numerical solution of the Eikonal equation. For this discussion, we limit ourselves to a positive speed function $F(x,y)$, which makes the level set curve monotonically propagate off the initial line.

The key of the fast marching algorithm is that the upwind different structure means the information propagates "one way", from smaller values of $T(x,y)$ to larger ones, which clearly satisfies the assumption of a positive speed function. Since the idea of fast marching method is to iteratively update the value $T(x, y)$ around the existing front, the selection of which grid to update around the front is the key of the whole algorithm. Because of the assumption that the front propagates “one way”, the smallest value around the front is always correct; other points around the front with larger $T(x, y)$ cannot affect it. The technique is algorithmically explained as follows,

1. Definition

   a) *Alive* points: points whose value of $T(x, y)$ is known for sure. Letting (i, j) denote the index of the grid, the set of alive points is represented by $\Omega = \{(i, j)\}$.

   b) *Narrow band* points: points which are not alive points but adjacent to at least one alive point. (For the grid (i, j), the adjacent grids are $(i-1, j)$, $(i+1, j)$, $(i, j-1)$, $(i, j+1)$. The values of T at the *Narrow Band* points are estimated by Equation 3.

   c) *Far away* points: points which are neither *alive* points nor *narrow band* points. The value of T at the *Far Away* points are set to be infinity.

2. Fast Marching Algorithm

   a) Detect a rectangular region near nose tip, $S_{in}(0)$, as the boundary condition of the Eikonal equation. According to the fast marching algorithm, S_{in} is the *Alive* set. Based on this statement, initialize the *Narrow Band* and the *Far Away* sets.

   b) Beginning of the loop: Estimate registration vector $\vec{p} = [\overrightarrow{p_R}, \overrightarrow{p_T}]^T$ between S_{in} and B by the ICP algorithm. Perform the rotation $R_{\overrightarrow{p_R}}$ and the translation $\overrightarrow{p_T}$ on the surface A.

   c) Compute the distance from each node $x, y, z = f(x, y)$ in A to B, denoted as $D(x, y)$.

   d) Set $F(x, y)$ to be $F(x, y) = \frac{1}{D(x,y)^{\lambda}}$, and $f(x, y) = D(x, y)^{\lambda}$

   e) Let (i_{min}, j_{min}) be the point in *Narrow Band* with the smallest value for T. Add the point (i_{min}, j_{min}) to the S_{in}, and remove it from Narrow Band.

   f) Tag as neighbors any points $(i_{min} - 1, j_{min})$, $(i_{min} + 1, j_{min})$, $(i_{min}, j_{min} - 1)$, $(i_{min}, j_{min} + 1)$ that are either in *Narrow Band* or *Far Away*. All these neighbors are then moved to the *Narrow Band* set.

   g) Re-compute the values of T at all neighbors according to the largest possible solution to the quadratic equation (Equation 3).

   h) Return to the top of the loop

3. EXPERIMENT AND RESULT

In this section, we present a partial surface matching-based face recognition technique and test it on FRGC Ver2.0 database [7]. FRGC Ver2.0 is a large public database of frontal facial range images. It contains 4007 face range images from 465 subjects, captured by a Minolta Vivid 900/910 series scanner with the resolution of 640x480. In this database, every subject provides several facial images, including both neutral expression and some non-neutral expressions. Our experiment is performed on a subset of FRGC Ver2.0 database, with 200 images of 50 subjects, and ach subject provides 4 different expressions (neutral, smile, surprise, inflated). In our experiment, every face is downsampled to 320x240, and then cropped by a sphere located on the nose tip with radius of 90mm.

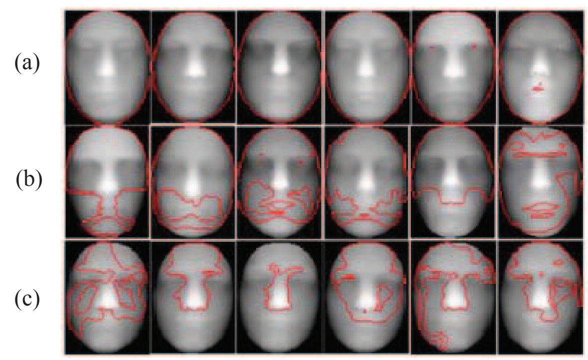

Fig. 1. Partial surface matching results. (a) Matching between the same subject with the same facial expression. (b) Matching between the same subject with different facial expressions. (c) matching between different subjects

3.1. Partial Surface Matching

The proposed partial surface matching method is validated on different subjects and facial expressions. As shown in Fig. 1, for the same subject, the entire facial area can be matched with same facial expression, while expression-insensitive part can be matched across different facial expressions. For different subject, a small facial part is match randomly, and the size of the match area is much smaller.

3.2. Face Recognition

A face verification is a hypothesis test taking two facial images as input, and make decision between two hypothesis: H1 as “given two facial images are from the same subject” and H0 as “given facial images are from different subjects”. In the presented paper, we first perform partial surface matching between two input faces, and then use the size of matched area as the similarity measure. All inputs with matched larger than a pre-defined threshold are classified into H1, and all others go to H0.

The combination of 200 faces of 50 subjects can provide 300 testing pairs in H1 and 9800 testing pairs in H0. The ROC for the proposed technique and the baseline algorithm (ICP algorithm) generated on these testing data is shown in Fig. 2. From the ROC curve, the improvement of recognition rate at different level of false alarm is obvious and significant.

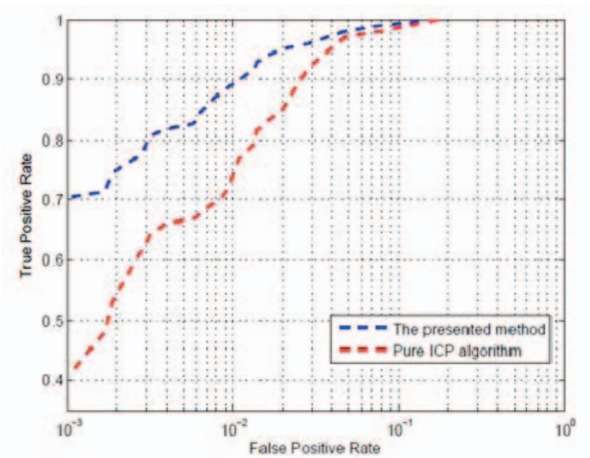

Fig. 2. Receiver Operation Characteristic (ROC) curves for the subset of FRGC Ver.2.0 database

4. CONCLUSIONS AND FUTURE WORKS

We have proposed a framework of robust 3D face recognition based on partial surface matching. In the first step, the ICP algorithm is used to register the nose regions of facial images. In the second step, we apply a level set method to grow the matched area. The level set method is implemented by the fast marching algorithm. Based on the assumption that there is a certain facial area which remains unchanged under facial expressions, the size of matched area is used as the distance measurement between facial images. Experimental results demonstrate the capabilities of the proposed method to partially register and match facial images. As expected, we successfully segmented the identical facial areas between the same subject's two different facial expressions. The size of matched area is shown to be a very effective distance measurement of 3D faces.

We are currently investigating the following issues: (i) exploring other forms of speed function $F(x, y)$ to improve the accuracy of surface matching; (ii) testing the presented framework on other surface registration methods; (iii) reducing the computational cost of matching; (iv) evaluating the performance on a more extensive database.

5. REFERENCES

[1] Zhengyou Zhang, "Iterative point matching for registration of free-form curves and surfaces," International Journal of Computer Vision, vol. 13, no. 2, pp. 119–52, 1994.

[2] Chen Yang and Gard Medioni, "Object modeling by registration of multiple range images," Image and Vision Computing, vol. 10, no. 3, pp. 145 – 155, 1992, Range Image Understanding.

[3] Haili Chui, Anand Rangarajan, Jie Zhang, and Christiana Morison Leonard, "Unsupervised learning of an atlas from unlabeled point-sets," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp. 160–172, 2004.

[4] Xiaoguang Lu and A.K. Jain, "Deformation modeling for robust 3d face matching," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 30, no. 8, pp. 1346–1357, Aug. 2008.

[5] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Expression-invariant representations of faces," Image Processing, IEEE Transactions on, vol. 16, no. 1, pp. 188–197, Jan. 2007.

[6] Xiaoxing Li and Hao Zhang, "Adapting geometric attributes for expression-invariant 3d face recognition," in Shape Modeling and Applications, IEEE International Conference on, June 2007, pp. 21–32.

[7] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min, and W. Worek, "Overview of the face recognition grand challenge," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, June 2005, vol. 1, pp. 947–954 vol. 1.

[8] J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," Proceedings of the National Academy of Sciences of the United States of America, vol. 93, no. 4, pp. 1591–1595, 1996.

[9] K.I. Chang, W. Bowyer, and P.J. Flynn, "Multiple nose region matching for 3d face recognition under varying facial expression," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 10, pp. 1695-1700, oct. 2006.

[10] Elisabeth Rouy and Agns Tourin, "A viscosity solutions approach to shape-from-shading," SIAM Journal on Numerical Analysis, vol. 29, no. 3, pp. 867–884, 1992.