3D Shape–based Face Recognition using Automatically Registered Facial Surfaces

M.O. İrfanoğlu, B. Gökberk, L. Akarun
Computer Engineering Dept.
Boğaziçi University
{irfanogl, gokberk, akarun}@boun.edu.tr

Abstract

In this paper, we address the use of three dimensional facial shape information for human face identification. We propose a new method to represent faces as 3D registered point clouds. Fine registration of facial surfaces is done by first automatically finding important facial landmarks and then, establishing a dense correspondence between points on the facial surface with the help of a 3D face template–aided thin plate spline algorithm. After the registration of facial surfaces, similarity between two faces is defined as a discrete approximation of the volume difference between facial surfaces. Experiments done on the 3D RMA dataset show that the proposed algorithm performs as good as the point signature method, and it is statistically superior to the point distribution model–based method and the 2D depth imagery technique. In terms of computational complexity, the proposed algorithm is faster than the point signature method.

1. Introduction

Automatic recognition of human faces from 2D intensity images has been studied extensively in the computer vision community. Facial variations due to illumination conditions, pose changes and different facial expressions make the recognition problem very challenging. Due to recent advances in both 3D acquisition systems and computational power of computers, it has become worthwhile to examine the advantages of 3D facial information in recognition.

Most of the initial attempts in 3D–based face recognition try to solve the identification task by fitting a deformable 3D model to 2D images [1, 2]. For example, in [1], 3D shape and texture of faces are estimated from a single face image by fitting a statistically morphable 3D face model. Model parameters of 3D shape and texture are then used to represent and recognize faces. Similarly, in [2], a 3D head model is generated from a 2D input face image, and once a 3D head is found, 2D face images can be rendered under different pose and lighting effects. Their fractal-based recognition algorithm uses rendered front view face images.

Several studies treat the face recognition problem as a 3D shape recognition problem of free-form curved surfaces. In [3] Extended Gaussian Image (EGI) representation is used to code faces using principal curvatures and their directions. Recognition is performed using Fisher's spherical correlation on EGI's of faces. Their approach does not require facial feature extraction and surface segmentation prior to recognition. In another work, point signatures which were proposed as free form surface representation technique, have been used for 3D face recognition [4].

Facial profile has also been considered as an important 3D feature for identification. In [5], both central profile and lateral profile are used for 3D face authentication. Once the profiles are extracted from faces, curvature values extracted along profile curves are used to compare two faces.

Various studies compare and combine both 2D intensity and 3D shape information. In [6], it has been shown that combining 2D Gabor wavelet based image intensity features with point signature-based 3D shape features has a superior performance than using each modality alone. Combination of facial texture and shape information also used in a stereo-based system [7]. 3D image meshes which contain the position and disparity of facial features with texture information generated by Gabor Kernels are used to represent faces. At the recognition phase, each model mesh is projected onto the test image pair, and the most similar model mesh is selected according to the Gabor kernel outputs.

In this paper, we propose the Point Set Distance (PSD) technique, which uses registered 3D point sets for face representation. We propose a novel technique to automatically register facial surfaces. Similarities between different faces are calculated by Euclidean norm between registered 3D point sets.

In the next section, we give details of the proposed method. In Section 2.2, we will briefly describe our 3D registration procedure which is a crucial preprocessing step for the PSD technique. Experimental results are presented in Section 3 by comparing our results with different state–of–the–art 3D face recognition techniques.

2. Proposed Algorithm

2.1. Dense Correspondence Algorithm

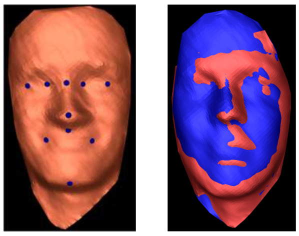
One of the main problems in a typical face data set obtained with stereo vision is the fact that different faces contain different portions of the face. Also considering the variations on pose and expressions, it is not possible to directly establish a point-to-point correspondence of each face. In [8], a similar problem is dealt with by first manual landmarking and then using the TPS warping algorithm. There, the aim is to establish a dense correspondence between each faces' points in order to extract the eigenvectors of Point Distribution Models to learn the shape variations present in the training set. This is done by landmarking few important points in each face in the training set and then obtaining a dense correspondence. The test data is processed by extracting the transformation and shape parameters in an iterative manner based on the Iterative Closest Points (ICP) algorithm [9]. In our case, we wish to automatically locate these landmarks to obtain a dense point-to-point matching with the test data.

The initial step in our algorithm is exactly the same as in [8] but differs in later stages in that we do not need to manually landmark all the faces in our training set. Instead, we only need to have a base mesh that contains the points that are present in all faces. The algorithm to obtain the base mesh can be summarized as follows:

1. Select the face with the least number of points as base.
2. Manually landmark 10 fiducial points (see Figure 1a).
3. Compute mean landmarks with Procrustes [10].
4. Warp landmarked faces onto the mean landmarks using Thin Plate Spline warping [11].
5. Delete a base mesh vertex if the distance to the closest point to any warped face is greater than a threshold.

When this base mesh is constructed, we can establish a point-to-point correspondence with our training and test set data. We do this by automatic landmarking of the data that is being processed, with 10 points as in Figure 1a, and then applying step 4 of above and reconstructing the data according to its closest points to base mesh vertices as in [8].

Automatic landmarking of these landmark points is the key point in our analysis. Let L_i define the i^{th} guess for the 10 landmark points. L_i's are computed as follows:



(a) (b)

Figure 1. (a)Salient facial landmark positions (b) TPS warped and registered facial surfaces

1. Compute surface normals, mean and Gaussian curvatures for all points in the test face.
2. Match the test face with the base mesh using ICP.
3. Obtain L_1 by computing the closest points of the base face landmark points to the test faces.
4. Find the closest point of the face onto the projection plane, which is again computed via the ICP transform. This is the tip of the nose.
5. Compute the symmetry plane based on the symmetry plane of base face and the ICP transformation.
6. Fine–tune the location of the tip of the nose using curvature and surface normal information and symmetry plane (L_2).
7. Fine–tune the locations of upper nose, lower nose and chin according to surface normals, symmetry plane and curvature information (L_3).
8. Compute the eye plane and mouth plane as in step 5.
9. Perform fine–tuning operations for eyes and mouth corners using their corresponding planes (L_4).

Proposed automatic landmark finding algorithm produces sufficiently good results for dense landmarking. When automatic landmarking of these sparse points are made, point-to-point dense matching is established with TPS and closest point matching as in [8]. Figure 1b shows the result of the algorithm in the form of two registered facial surfaces.

2.2. Point Set Distance Approach (PSD)

Suppose that we have a base face mesh Φ_{base} of n points. After the dense correspondence is established, when we are given the i^{th} individual's 3D face $\Phi_i = \{p_1^i, p_2^i, ...p_{N_i}^i\}$, where N_i is the number of points in the face and p^is are

3D coordinates, we generate a new 3D point cloud for i^{th} individual that has n points with our automatic landmarking procedure. Let Φ_i^n represent this new 3D face. Since all face images will have the same correspondence with the mean 3D face, we can define the distance between two faces Φ_i and Φ_j as:

$$D(\Phi_i^n, \Phi_j^n) = \sum_{k=1}^{n} ||p_k^i - p_k^j|| \quad (1)$$

where $||.||$ denotes Euclidean norm. The identity of the test image is then found by selecting the training image having the smallest distance.

3. Experimental Results

In our experiments, we used 3D RMA dataset[1] [5]. Specifically, the first session of noise free faces were used in experiments, which consists of 30 people each having 3 shots. The data is obtained with a stereo vision assisted structured light system. On the average, faces contain 4000 3D points, and they cover different portions of the faces and the entire data is subject to expression and rotation changes. For all the features except implicit polynomials, we applied a Delaunay triangulation onto the point cloud followed by a Laplacian 3D smoothing filter in order to obtain a continuous surface. In all of the experiments, 1–nearest neighbor with Euclidean distance is used as the classifier. To be able to statistically compare the algorithms, we designed three experimental sessions. Session ID indicates the shot number of the person which is placed in the test set, and the remaining two images are put to the training set.

We will present the results of the recognition experiments for the PSD technique. Additionally, in order to compare the performance of the proposed approaches, we have implemented point signatures, point distribution models, implicit polynomial–based technique, and 2D depth imagery–based PCA technique.

Recognition results of the PSD technique are presented at the second and third rows in Table 1 for both automatic and manual landmarking. Please note that although we did not have to manually landmark all the faces, we did so to assess the correctness of our automatic landmarking algorithm. Results show that automatic landmarking performs as good as full manual landmarking on this dataset.

Point Distribution Models (PDM) learns the shape variability of the data based on the manually landmarked points using principle component analysis (PCA) [12]. PDM represents shapes with a few parameters that can be used as features in a classification system. After establishing a dense correspondence between faces, 99.8% of shape variability is explained with approximately 40 shape parameters that yields recognition results indicated at the third row of Table 1, when used as features.

Table 1. Comparative analysis of recognition performances on three different experimental sessions. Second column denotes feature dimensionalities

Method	Comp.	S1	S2	S3
PSD (Automatic L.)	N_p	96.66	96.66	96.66
PSD (Manual L.)	N_p	90.00	96.66	96.66
PDM	40	76.66	83.33	86.66
Point Signature	$N_p \times 35$	100.00	96.67	96.67
IP coeffs. (d=6)	84	46.67	46.67	63.33
IP invariants (d=4)	11	13.33	16.66	23.33
PCA	60	76.67	70.00	80.00

Point signature (PS) technique is a commonly used local feature–based surface representation method that has been applied to the 3D face recognition problem [4]. PS encodes a surface point p using the minimum distances of its neighbors in a predefined periphery to a tangential plane, P passing through p. Neighbors of point p are found by intersecting the original surface with a sphere. P is formed by fitting a plane to the neighboring points, and by translating it to the original point p. Signed distances are sampled by $\Delta\theta$ degree intervals, thus forming a 1D parametric curve, $d(\theta)$. In the recognition phase, each point signature extracted from the test image is compared with each model image's point signatures, and a total similarity between the model and scene image is computed according to the sum of individual point signature distances. In our experiments, point signatures were computed for all vertices in the face data. In a 3D point cloud of average image resolution R_a, we have chosen the sphere radius, r as $5 \times R_a$, which is found empirically to be the most suitable local shape descriptor. Fifth row in Table 1 shows the recognition accuracy of the PS technique.

Implicit polynomials (IPs) are powerful shape descriptors suitable for use in 2D and 3D computer vision and image analysis [13]. Their power lies in their ability to express every possible shape, their invariants and their interpolation abilities. However, their use in classification problems is mainly dependent on the robustness and stability of the polynomial fitting method applied and the correctness of the invariants extracted. In our 3D face recognition analysis, IPs are fitted to each face's point cloud data with the 3L method [14]. Based on the coefficients extracted, invariants of the polynomial are computed using symbolic computation. We used both the coefficients and a number of invariants, specifically, 6^{th} degree IP coefficients and 11 in-

[1]We would like to thank Charles Beumier from Ecole Royale Militaire, Belgium, for allowing us to use their face dataset

dependent invariants of a 4^{th} degree IP are used as features. Recognition accuracies of IP coefficients and invariants are shown in the sixth and seventh row of Table 1, respectively. In our experiments, we see that low degree IPs (d=2,3) are not powerful enough to capture shape variations, and high degree IPs ($d \geq 8$) learn the noise present in the data.

PCA of range images is widely used in 3D object recognition systems. As our last comparative algorithm, we have extracted 2D depth images of 3D faces using the z coordinates as brightness values. Since original faces are 3D point clouds, pixel values are computed using uniform sampling of the facial surface. Recognition performance of the PCA technique is shown at the last row of Table 1.

We made a t–test analysis with 95% confidence level in order to compare the algorithms. The results can be summarized as follows: Best recognition performances are obtained by PSD and PS methods, which are statistically indifferent. Among all the algorithms, implicit polynomials perform worst. Computational complexity of algorithms depends mainly on the dimensionality of the features extracted. Second column in Table 1 depicts feature dimensionality of each method where N_p denotes the number of 3D points in face data, which is 2300 in our data.

4. Conclusion

In this work, we have proposed a novel 3D face recognition system based only on the shape information of human faces. Proposed algorithm (PSD) first establishes a dense correspondence between faces. Correspondence finding algorithm consists of two steps: in the first step, a number of salient facial points are found automatically, and in the second step, these points are used to find dense correspondences of points on the facial surface using TPS warping algorithm and a surface reconstruction with a base mesh. Recognition is carried out by computing a discrete approximation of a volume between registered facial surfaces.

In order to compare the recognition performance of our algorithm, we have implemented point signature–based representation, point distribution model parameters, implicit polynomial–based technique, and range image based 2D PCA algorithm. Statistical test results on three experimental sessions reveals that both the proposed PSD technique and point signatures perform best, whereas the computational complexity of the proposed PSD technique is significantly less than that of point signatures.

References

[1] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9):1063–1073, 2003.

[2] A. Z. Kouzani, F. He, and K. Sammut. Towards invariant face recognition. *Information Sciences*, 123(1-2):75–101, 2000.

[3] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation - principal directions for curved object recognition. In *Proceedings of Int. Conf. on Automatic Face and Gesture Recognition*, pages 372–377, 1998.

[4] C-S. Chua, F. Han, and Y-K. Ho. 3d human face recognition using point signature. In *Proceedings of Int. Conf. on Automatic Face and Gesture Recognition*, pages 233–237, 2000.

[5] C. Beumier and M. Acheroy. Automatic 3d face authentication. *Image and Vision Computing*, 18(4):315–321, 2000.

[6] Y. Wang, C-S. Chua, and Y-K. Ho. Facial feature detection and face recognition from 2d and 3d images. *Pattern Recognition Letters*, 23(10):1191–1202, 2002.

[7] A. Eriksson and D. Weber. Towards 3-dimensional face recognition. In *Proceedings of the 5th IEEE AFRICON*, pages 401–406, 1999.

[8] T. J. Hutton, B. F. Buxton, and P. Hammond. Automated registration of 3d faces using dense surface models. In *Proceedings of the British Machine Vision Conference*, pages 439–448, 2003.

[9] P. Besl and N. McKay. A method for registration of 3d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–256, 1992.

[10] J. C. Gower. Generalized procrustes analysis. *Psychometrika*, 40:33–51, 1975.

[11] F. L. Bookstein. Shape and the information in medical images: A decade of the morphometric synthesis. *Computer Vision and Image Understanding*, 77(2):97–118, 1997.

[12] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision, Graphics and Image Processing*, 61(1):38–59, 1995.

[13] J. Subrahmonia, D. B. Cooper, and D. Keren. Practical reliable bayesian recognition of 2d and 3d objects using implicit polynomials and algebraic invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):505–519.

[14] Z. Lei, M. M. Blane, and D. B. Cooper. 3l fitting of higher degree implicit polynomials. In *Proceedings of Third IEEE Workshop on Applications of Computer Vision*, 1996.