3D FACE RECOGNITION USING LOCAL SHAPE MAP

Zhaohui Wu, Yueming Wang and Gang Pan

Department of Computer Science and Engineering
Zhejiang University, Hangzhou, 310027, P.R.China
{ymingwang,gpan}@zju.edu.cn

ABSTRACT

This paper firstly proposes a new scheme for description of 3D local shape, termed *local shape map* (LSM). The LSM for a point on the surface is a 2D histogram constructed from mapping 3D coordinates of surface's points within a sphere centralized at this point into a 2D space. The similarity metric for LSM is given. Secondly, application of LSM to 3D face recognition is presented, which is performed by incorporating LSM into an ad-hoc voting method. It does not require to register two face models. With a dataset of 31 facial range images for six subjects, in presence of significant change in pose and slight variation in expression, the equal error rate of 2.98% is obtained. The experimental results also demonstrate the performance of 3D face recognition using LSM outperforms that using *spin image*.

1. INTRODUCTION

Automatic face recognition has been studied extensively over the past decade. Most efforts have been made for face recognition from 2D images[1], but only a few approaches exploited 3D information [2, 3, 4, 6, 7, 9, 10, 11, 12]. Although the 2D face recognition system has good performance under constrained conditions, since the 2D image essentially is a projection of the 3D human face, it is still challenged by changes in illumination, pose and expression [1, 16]. Utilizing 3D information can improve the system performance[16, 9] due to its explicit representation of facial surface.

The previous work utilizing 3D information for face recognition can be categorized into four groups.

- Curvature analysis-based: most of the early studies concentrate on curvature analysis. In [3], the sensed surface regions are classified as convex, concave and saddle by calculating the minimum and maximum principal curvature, then locations of facial features are determined, which are used for template comparison. Lee's method[2] detects corresponding regions in two range images by graph matching based on Extended Gaussian Image(EGI). Tanaka et al [4] also use EGI. For each face, two EGIs are constructed from maximum principal direction and minimum principal direction. The EGI similarity is measured by Fisher's spherical correlation.
- Spatial matching: recognizing via matching facial surface or profile directly in 3D Euclidean space[10, 11].
- Recover-and-synthesis: for this kind of methods, their probe still is 2D image but not 3D data. From the 2D image the 3D information is recovered, then the facial image in virtual view is synthesized for recognition [12, 7] or the recognition task is accomplished with the recovered parameters[9].
- Shape representation-based method: at first 3D facial shape is converted to another shape representation such as *point signature*[6], *spin image* [14], *shape distribution*[15], thus, the recognition task can be achieved in the representation domain.

This work focuses on face recognition from 3D models, i.e. the data in the database and the probe are both 3D face models. We assume that the face models have been acquired.

This paper, firstly, introduces a new approach of 3D local shape representation, called *local shape map* (LSM), which is motivated by *spin image* [13, 14] and *point signature* [6]. The LSM of a point is a 2D histogram generated from mapping 3D coordinates of surface's points within a sphere centralized at this point into a 2D space. The mapping process employs this point and its normal as reference point and reference direction. Secondly, the application of LSM into 3D face recognition is performed.

This paper is organized as follows: The LSM scheme is described in Sec. 2. Sec. 3 applies LSM to 3D face recognition. The experimental results and conclusion are in Sec. 4 and Sec. 5 respectively.

2. LOCAL SHAPE MAP — A NEW REPRESENTATION

2.1. Definition of Local Shape Map

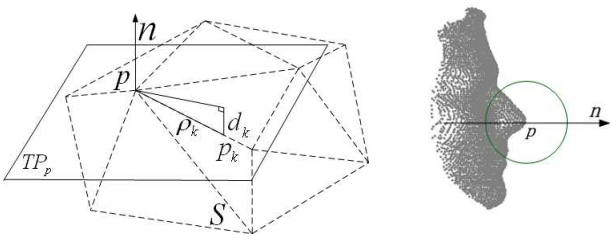


Fig. 1. Illustration of Local Shape Map.

Local Shape Map is a new representation of 3D local shape for a point. Its definition is as follows.

Given a surface $\boldsymbol{S}$ and a point $\boldsymbol{p}$ on the surface, suppose that $\boldsymbol{n}$ is the unit normal at point $\boldsymbol{p}$ and $\boldsymbol{TP_p}$ is the tangent plane of point $\boldsymbol{p}$, shown in Fig. 1(a). Thus, each point $\boldsymbol{p_k}$ on the surface $\boldsymbol{S}$ can be associated with two variables:

0-7803-8554-3/04/$20.00 ©2004 IEEE.

1) the distance between $\boldsymbol{p}$ and $\boldsymbol{p_k}$:

$$\rho(k) = \|p - p_k\| \tag{1}$$

2) the displacement from p_k to the tangent plane $\boldsymbol{TP_p}$:

$$d(k) = n \cdot (p_k - p) \tag{2}$$

where $d(k)$ can be both positive and negative while $\rho(k)$ always is non-negative. The ρ_k and d_k is shown in Fig. 1(a).

With the two variables ρ and d, the distribution of all points of surface $\boldsymbol{S}$ on the plane ρ-d can form a 2D image or a 2D histogram, called ρ-d histogram. To achieve a local shape descriptor for one point on the surface, We place a sphere with radius r centrally at $\boldsymbol{p}$, and only count those points within the sphere statistically for the ρ-d histogram. The constrained 2D ρ-d histogram is termed *local shape map* (LSM) for point $\boldsymbol{p}$. It can be formalized as:

$$\boldsymbol{LSM}\text{: } S^r(\boldsymbol{p}) = (S^r_{ij}(\boldsymbol{p}))_{m\times n, 1\le i\le m, 1\le j\le n} \tag{3}$$

$$S^r_{ij}(\boldsymbol{p}) = \sum_k s^k_{ij} \tag{4}$$

where

$$s^k_{ij} = \begin{cases} 1, \text{ if } \boldsymbol{p_k} \in S \text{ and } \|\boldsymbol{p_k} - \boldsymbol{p}\| \le r \\ \quad \text{and } i * \Delta b_1 \le \rho_k < (i+1) * \Delta b_1 \\ \quad \text{and } j * \Delta b_2 \le d_k + r < (j+1) * \Delta b_2 \\ 0, \text{ otherwise} \end{cases} \tag{5}$$

where Δb_1 and Δb_2 are the size of vertical bin and horizonal bin for LSM, and m and n are the total number of bin in vertical and horizonal axis respectively. Obviously, LSM is insensitive to rotation and translation. Some LSMs at different locations on facial surface are shown in Fig. 2, notice the difference in the LSMs.

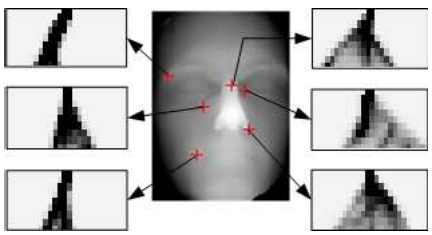


Fig. 2. Examples of LSMs at different points on facical surface.

The LSM representation scheme is motivated by *spin-image* [13] and *point signature* [5]. However, there are several differences in them.

1. The LSM of a point $\boldsymbol{p}$ uses points within the sphere centrally at $\boldsymbol{p}$ to construct 2D histogram so that it describes local shape at that point. The spin image of a point on the surface uses support distance and support angle to control the amount of global information covered. Both methods can reduce the chances of clutter and self occlusion corrupting the images(LSM and spin image).
2. LSM employs the distance from $\boldsymbol{p}$ to neighboring point as a dimension, rather than the distance from neighboring point to the normal of point $\boldsymbol{p}$ as a dimension in *spin image*. The point-to-point distance is consistent with the use of sphere to define point's neighborhood.
3. In *point signature*, the reference direction $\mathbf{n}_2$ is defined as a start edge of angle θ, whereas determination of $\mathbf{n}_2$ is ambiguous and not robust. LSM uses the normal which could be determined more reliably.
4. LSM encodes the whole 3D shape within the sphere while *point signature* only exploits a 3D curve, the intersection of the surface and the sphere.

To deal with the different resolution of original 3D data, before building LSM, we may resample all the models with same method to construct triangular meshes. In addition, we simply let the 2D histogram divided by the total number of those points in the sphere, to achieve normalization. In our implementation, the bin size are predetermined to be two to three times the depth image resolution, and several values for radius r are tested. As for LSM size, m and n could be specified by Δb_1 , Δb_2 and r.

2.2. Comparing LSMs

Correlation is frequently used to measure the similarity of images, it is a good choice to use correlation to match LSMs. The correlation coefficient between two LSMs S^r and S'^r are:

$$Corr(S^r, S'^r) = \frac{Cov(S^r, S'^r)}{\sqrt{D(S^r) * D(S'^r)}} \tag{6}$$

where

$$Cov(S^r, S'^r) = \frac{1}{N}\sum_{i,j} S^r_{ij} S'^r_{ij} - \frac{1}{N^2}\sum_{i,j} S^r_{ij} \sum_{i,j} S'^r_{ij} \tag{7}$$

and

$$D(S_r) = \frac{1}{N}\sum_{i,j} (S^r_{ij})^2 - \left(\frac{1}{N}\sum_{i,j} S^r_{ij}\right)^2 \tag{8}$$

Actually, the correlation is the LSMs' covariance normalized by the standard variance of two LSMs.

From Fig. 2, it can be seen that the majority of bins in the LSM are zeros (the white means value zeros) because surface's points within the related sphere only cover little part of the LSM. Zeros in corresponding bins of LSMs are meaningless in LSMs comparison and make LSMs more correlative and less comparable. To reduce this effect, are there two ways:

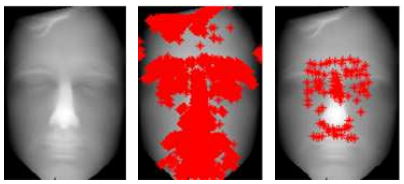
AND-rule: for two LSMs, those corresponding bins with both value zero are both eliminated before correlation calculation;

OR-rule: for two LSMs, the corresponding bins with either value 0 are both removed before correlation calculation;

3. FACE RECOGNITION USING LSM

In this work, face recognition is implemented by a LSM-comparison-based voting method after constructing the model database.

The first step is to build up the model database. For each model, in order to reduce the time cost, we only select a few *candidate points* for future online LSM comparison and compute their LSMs in advance. For the selection, we simply threshold the minimum principal curvature of surface points, $k_{min} \le T_s$, yield regions of extreme curvature. The points in these regions are chosen as the candidate points, shown in Fig. 3(b) where the candidate points are in red color.



(a) (b) (c)

Fig. 3. (a) The original range image. (b) The selected candidate points, marked in red color. (c) 100 voting points, marked in red.

During comparing two 3D face model, this paper employs a ad-hoc voting method. Suppose that a model in database:

$$S_m = \{(p_i, \Psi_i), i = 1, ..., N_m\} \tag{9}$$

And the probe model (for the probe, the preprocessing of selection of candidate points is also performed):

$$S_p = \{(q_j, \Phi_j), j = 1, ..., N_p\} \tag{10}$$

Where p_i's LSM is Ψ_i and q_j's LSM is Φ_j.

For each point q_j in S_p, compute the correlation coefficients of its LSM Φ_j with all LSMs in S_m. If one of these coefficients is larger than the threshold T_f, then S_m receives a voting made by Φ_j or point q_j. This voting is repeated through all points in S_p. Assuming the model S_m receives n votings by S_p, a voting rate of the model S_m is n/N_p. The voting procedure is similar to the reference [6]. The voting rate represents the likelihood of each model in database being correctly matched with the probe. Our voting method does not require to register two face models.

The number of voting points in a probe can be reduced further from the candidate points, because, for many pairs of neighboring candidate points in the probe model, their LSMs are quite similar to each other. In our experiment, to "vote" for likely model faces in database, only 100 points are automatically selected from the candidate points of the probe such that the resulting points' LSMs differ widely from each other. It means that N_p is 100. Fig. 3(c) shows the automatically selected 100 voting points (in red color) from the candidate points in Fig. 3(b).

4. EXPERIMENTAL RESULTS

The experimental data are downloaded from SAMPL at Ohio State University [17]. Our experiments involve with 31 facial range images of six human subjects and two human subjects have several images with significantly different pose and slightly different expression. All the range images were captured by Minolta 700 range scanner, and are with resolution 200 $\times$ 200. Fig. 4 shows some samples of the range images.

Due to limitation in test data, we only experiment with face verification. The recognition performance is evaluated by equal error rate (EER), that means false acceptance rate(FAR) and false reject rate (FRR) are the same, where FAR is the percentage of imposters wrongly matched and FRR is the percentage of valid users wrongly rejected.

Our first experiment attempts to find best combination of radius *r*, voting threshold T_f and comparison rule. We test *r* with 20, 30, 40 and 50 mm in computing LSMs. For each *r*, five T_fs, together with AND-rule or OR-rule, are used to find the best setting. The results are shown in Table 1. For r=20,30, AND-rule is outperformed by OR-rule in the most case. For r=40,50, their performance is close, but AND-rule seems to be better. The EER of 2.98% is reached at r=40, T_f=0.97 and AND-rule.

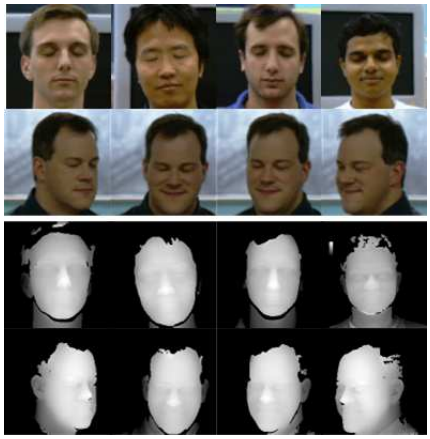


Fig. 4. Sample images. The first two rows are color images which is not used in the experiment, and the last two rows are their corresponding depth images.

Table 1. EERs with different experimental setting

r = 20					
T_f	0.97	0.98	0.985	0.987	0.99
AND-rule	15.96%	12.17%	8.25%	8.78%	9.34%
OR-rule	15.14%	11.08%	7.87%	8.81%	9.17%
r = 30					
T_f	0.965	0.970	0.975	0.980	0.985
AND-rule	7.61%	5.41%	4.78%	4.37%	5.63%
OR-rule	6.25%	5.01%	4.12%	4.60%	6.10%
r = 40					
T_f	0.955	0.960	0.965	0.970	0.975
AND-rule	6.02%	4.02%	3.47%	2.98%	4.32%
OR-rule	5.23%	4.19%	3.66%	4.07%	4.10%
r = 50					
T_f	0.955	0.960	0.965	0.970	0.975
AND-rule	5.34%	4.41%	3.99%	3.14%	4.62%
OR-rule	5.51%	4.39%	3.88%	4.20%	5.49%

We also carry out the comparison experiment with *spin image*. The voting rate matrix by our LSM, also as similarity matrix, with r=40, AND-rule and T_f=0.970 is shown in Fig. 5(a). And the similarity matrix by *spin image* is shown in Fig. 5(b). Their ROC curves is in Fig. 6. It demonstrates that our method LSM significantly outperforms *spin image* for face recognition experiment.

5. CONCLUSION

Local shape map presented in this paper can describe the 3D local shape. Its successful application to 3D face recognition demonstrates LSM's capability in effectively capturing the discriminative information for classification.

The pose and illumination variation are two challenges in 2D image-based face recognition[16]. Recognizing faces across pose

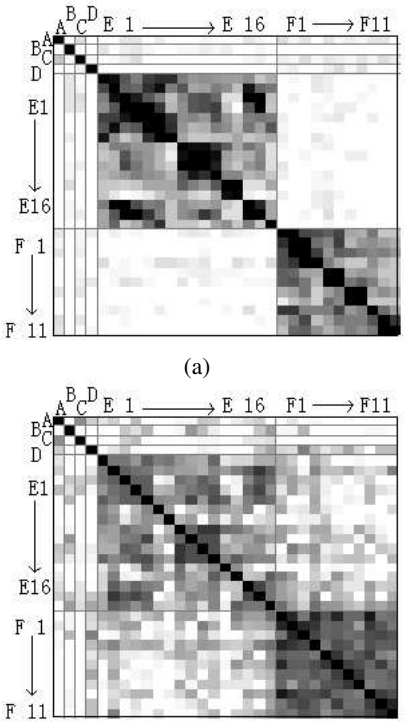


Fig. 5. (a) Similarity matrix by LSM. (b) Similarity matrix by *spin image*. The black indicates high similarity. E1-E16 are the same subject, and F1-F11 are the same individual.

and illumination is still a big problem. 3D face recognition is likely to solve these two problems. Our face recognition method is insensitive to pose variation, and does not require to register two face models.

Due to limitation in getting 3D face data, the experiment only were carried out on a publicly available small data set. The tests on a larger data set are our future efforts. To cope with the variation in data resolution and data acquisition devices is still in our future work.

6. ACKNOWLEDGEMENTS

This work is in part supported by NSF of China (No. 60273059), Zhejiang NSF for Outstanding Young Scientist (No.RC01058) and Doctoral Fund (No.20020335025).

7. REFERENCES

[1] W.Zhao, R.Chellappa, P.J.Phillips, A.Rosenfeld. Face recognition: a literature survey. ACM Computing Surveys, 35(4):399-458, 2003

[2] J.C.Lee, E.Milios. Matching range images of human faces. *Proc. IEEE ICCV*, p.722-726, 1990.

[3] G.G.Gordon. Face recognition from depth maps and surface curvature . *SPIE Conf. on Geometric Methods in Computer Vision*, 1570:234-247, 1991.

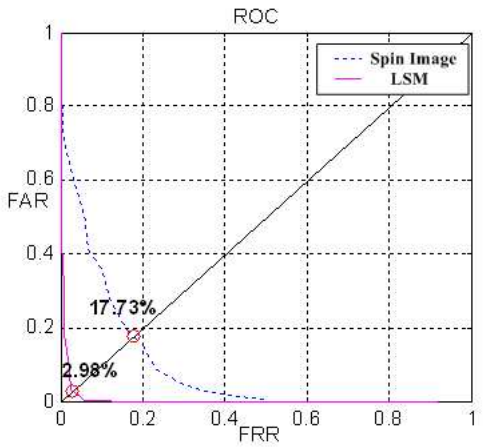


Fig. 6. ROC curves for comparison with Spin Image.

[4] H.T.Tanaka, M.Ikeda, H.Chiaki. Curvature-based Face Surface Recognition Using Spherical Correlation-Principal Directions for Curved Object Recognition-. *Int'l Conf. on Automatic Face and Gesture Recognition*, p.372-377, 1998.

[5] C.S.Chua, R.Jarvis. Point signatures: A new representation for 3D object recognition. *IJCV*, 25(1):63-85, 1997.

[6] C.S.Chua, F.Han, Y.K.Ho. 3D Human Face Recognition Using Point Signature. *Int'l Conf. on Automatic Face and Gesture Recognition*, p.233-238, 2000.

[7] M.W.Lee, S.Ranganath. Pose-invariant face recognition using a 3D deformable model. *Pattern Recognition*, 36(8):1835-1846, 2003.

[8] R.J.Campbell, P.J.Flynn. A Survey Of Free-Form Object Representation and Recognition Techniques. *CVIU*, 81(2):166-210, 2001.

[9] V.Blanz, S.Romdhani, T.Vetter. Face identification across different poses and illumination with a 3D morphable model. *Int'l Conf. on Automatic Face and Gesture Recognition*, p.202-207, 2002.

[10] C.Beumier, M.Acheroy. Automatic 3D face authentication. *Image Vision Computing*, 18(4):315-321, 2000

[11] G.Pan, Y.J.Wu, Z.H.Wu, W.Y.Liu. 3D Face Recognition by Profile and Surface Matching. *Proc. IEEE IJCNN'03*, p.2168-2174, 2003.

[12] W. Zhao, R. Chellappa. Illumination-insensitive face recognition using symmetric shape-form-shading. *Proc. IEEE ICCV*, 1:286-293, 2000.

[13] A. E. Johnson, M.Hebert. Surface matching for object recognition in complex three-dimensional scenes. *Image Vision Computing*, 16:635-651, 1998.

[14] A.E.Johnson, M.Hebert. Using spin images for efficient object recognition in cluttered 3D scenes.*IEEE Trans. PAMI*, 21(5):433-449, 1999.

[15] R.Osada, T.Funkhouser, B.Chazelle, D.Dobkin. Shape distributions. *ACM Trans. Graphics*, 21(4):807-832, 2002.

[16] Face Recognition Vendor Test 2002, http://www.frvt.org/.

[17] SAMPL Range Imagery, http://sample.eng.ohio-state.edu