LETTER

3D Face Landmarking Method under Pose and Expression Variations

Yuan HU[†a)], ***Student Member***, **Jingqi YAN**[†], **Wei LI**[††], ***and*** **Pengfei SHI**[†], ***Nonmembers***

SUMMARY A robust method is presented for 3D face landmarking with facial pose and expression variations. This method is based on Multi-level Partition of Unity (MPU) Implicits without relying on texture, pose, orientation and expression information. The MPU Implicits reconstruct 3D face surface in a hierarchical way. From lower to higher reconstruction levels, the local shapes can be reconstructed gradually according to their significance. For 3D faces, three landmarks, nose, left eyehole and right eyehole, can be detected uniquely with the analysis of curvature features at lower levels. Experimental results on GavabDB database show that this method is invariant to pose, holes, noise and expression. The overall performance of 98.59% is achieved under pose and expression variations.

key words: *3d face, landmark, face localization, MPU*

1. Introduction

Automatic face recognition has achieved great progress in the last two decades. Most of the studies have focused on 2D intensity images, but the performance of 2D face recognition [1] still suffers from pose, illumination, occlusion and expression variations. The potential of 3D face recognition [2] to alleviate these limitations is one of the main reasons for the significant attention given to this area. 3D data contains more geometrical information of shape and size and is immune to illumination variations. Furthermore, 3D face data is robust to pose changes.

Face landmarking plays an important part in face recognition. The most frequently used approach in the detection of 3D facial landmarks is based on curvature analysis. Gordon [3] presents an accurate analysis of the face based on curvatures. Moreno [4] segments the faces using surface curvatures. The idea of curvature segmentation is also adopted by Colombo [5]. Although curvature analysis method is robust to translation and rotation, it usually generates too many candidates for one landmark and does not work on 3D data with noise, holes and occlusion. However, 3D data obtained from laser scanners usually contain many holes around eyes, eyebrows, nosewing, hair and so on.

In this paper, we propose a novel robust face landmarking method to identify the location of the nose and eyeholes in 3D facial mesh data. The proposed algorithms are based on Multi-level Partition of Unity (MPU) Implicits [4] that relax the constraints on feature map thresholding and do not need the prior knowledge of orientation and pose of the data. The 3D facial mesh is reconstructed by MPU Implicits in a hierarchical way and the holes are filled simultaneously. The curvature analysis is carried out on different level reconstructed facial surfaces. Candidate regions, including nose and eyeholes, are isolated using mean and Gaussian curvature features maps that highlight the curvature properties. This proposed method is invariant to pose, holes, noise and expression and only one candidate can be detected for one landmark.

This paper is organized as follows. In Sect. 2, we describe the method for 3D face landmarking based on hierarchical MPU Implicits. Experimental results are presented in Sect. 3 and Sect. 4 concludes the paper.

2. 3D Face Landmarking Based on MPU Implicits

2.1 MPU Implicits

The multi-level partition of unity implicits surface (MPU) [6] allows us to construct surface models from very large sets of points. There are three key ingredients to MPU: 1) piecewise quadratic functions that capture the local shape of the surface, 2) weighting functions (the partitions of unity) that blend together these local shape functions, and 3) an octree subdivision method that adapts to variations in the complexity of the local shape.

For a bounded domain Ω in a Euclidean space, given a set of nonnegative compactly supported functions $\{w_i(x)\}$, an approximation of a function $f(x)$ defined on Ω is given by:

$$f(x) = \frac{\sum w_i(x) Q_i(x)}{\sum w_i(x)} \tag{1}$$

where, $Q_i(x)$ is the local approximation set of functions with each subdomain.

For approximation purposes we use the quadratic B-spline $b(t)$ to generate weight functions

$$w_i(x) = b\left(\frac{3|x - c_i|}{2R_i}\right) \tag{2}$$

Given a set of scattered points P equipped with normals N, we approximate the signed distance function $f(x)$ from P.

$$w_i(x) = \left[\frac{(R_i - |x - c_i|)_+}{R_i|x - c_i|}\right]^2 \tag{3}$$

Manuscript received September 21, 2010.
Manuscript revised November 23, 2010.

[†]The authors are with Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, 200240 China.
[††]The author is with Shenzhen Institutes of Advanced Technology, China Academy of Sciences, 518055 China.
a) E-mail: huyuan@sjtu.edu.cn

DOI: 10.1587/transinf.E94.D.729

where $(a)_+ = \begin{cases} a & if\ a > 0 \\ 0 & otherwise \end{cases}$, c_i is the center of a cubic cell that was generated during the subdivision process, and R_i is a spherical support of radius.

First, we use an octree-based adaptive space subdivision of Ω to control the error of the approximation while adapting the complexity of the representation to the complexity of the shape.

Second, we use piecewise quadratic functions resulting from Boolean operations for the accurate representation of sharp features. One of three local approximations is used:

(a) A general 3D quadric, which is used to approximate larger parts of the surface. A local shape function is given by:

$$Q(x) = x^T Ax + b^T x + c \tag{4}$$

(b) A bivariate quadratic polynomial in local coordinates, which is used to approximate larger parts of the surface. A local shape function is given by:

$$Q(x) = w - (Au^2 + 2Buv + Cv^2 + Du + Ev + F) \tag{5}$$

(c) A piecewise quadric surface that fits an edge or a corner. For the surface P with an edge, we subdivide P into two clusters P_1 and P_2 according to normals. The quadratic fit procedure is applied separately to P_1 and P_2 and a non-smooth local shape function approximate P is constructed via the max/min Boolean operations. For the surface N with a corner, we subdivide N into three sets, N_1 and N_2 are constructed as above. For the corners of degree four, the third cluster is subdivided into two pieces. If the resulting four clusters of normals correspond to either a convex or concave corner, it is reconstructed via Boolean operations. Otherwise, we go to (a) and a general quadric is fitted to P. More complex types of sharp features are approximated by smooth functions.

2.2 Hierarchical Reconstruction

The MPU Implicits is used for reconstruction of face surface in a hierarchical way. The coarse to fine strategy is adopted to reconstruct face surfaces with increasingly subdivision using MPU Implicits. For the first level L0, a quadratic polynomial is applied to fit the face surface. For the second level L1, L0 is subdivided into 8 parts; a quadratic polynomial is applied to fit each cell. For the third level L2, the 8 parts of L1 is subdivided and fitted continually. Figure 1 shows the hierarchical reconstruction of 3D face surface. L0 reconstruct face surface as one smooth surface with no organs. With the hierarchical process, the contours of organs emerge successively. The nose appears in L1, eyes present in L2, and mouth emerges in L3. L4 is quiet sufficient for reconstruction of features of surface. L6 can describe fine features of face surface.

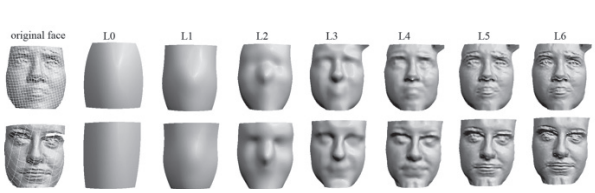


Fig. 1 The hierarchical reconstruction of 3D face surface.

This method also fit the face surface with holes, by reconstructing it, which fills up the holes automatically. The second row of Fig. 1 shows the original sample with holes in the right eyebrow. With MPU reconstruction, it fills up the holes automatically.

2.3 Curvature Analysis

To analyze the curvature of surface of different level, we let P be the surface defined by a twice differentiable real valued function

$f : U \to R$, defined on an open set $U \subseteq R^2$:

$P = \{(x, y, z)|(x, y) \in U; z \in R; f(x, y) = z\}$

For every point $(x, y, z) \in P$ we consider two curvatures, the mean curvature (H) and the Gaussian curvature (K), which can be calculated by [7]:

$$H = \frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{2(1 + f_x^2 + f_y^2)^{3/2}} \tag{6}$$

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2} \tag{7}$$

Where, $f_x, f_v, f_{xv}, f_{xv}, f_{vv}$ are the first and the second derivative of f in (x, y).

Figure 2 shows the mean and the Gaussian curvature maps of L0 to L4.

Based on the signs of the mean and the Gaussian curvatures, an HK classification [4], [5], [8] is adopted to obtain the description of local surface shape, as show in Table 1.

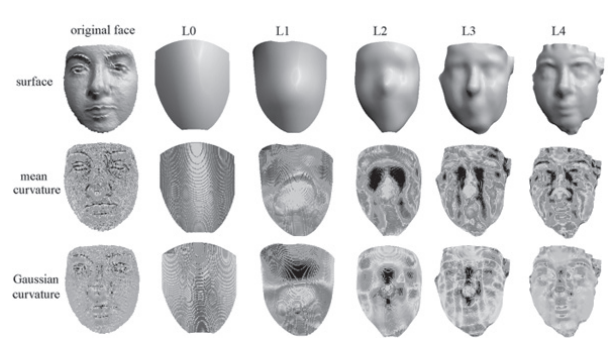
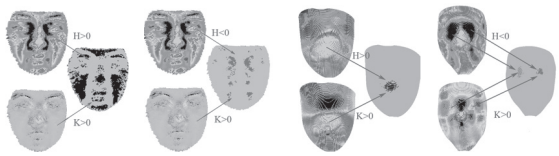


Fig. 2 Curvature analysis for different level faces reconstructed by MPU Implicits. The first row is the result of reconstruction from L0 to L4, the second row is the mean curvature map and the third row is the Gaussian curvature map.

Table 1 Surface shape classification.

Shape	Hyperbolic	Convex cylindrical	Convex elliptical	Planar	Concave cylindrical	Concave elliptical
H	/	>0	>0	0	<0	<0
K	<0	0	>0	/	>0	>0



(a) (b) (c) (d)

Fig. 3 (a) The convex elliptical regions extracted from an original face after smoothed. (b) The concave elliptical region extracted from an original face after smoothed. (c) The convex elliptical region extracted from L1 reconstructed face by MPU implicits. (d) The concave elliptical regions extracted from L2 reconstructed face by MPU implicits.

2.4 Landmarks Detection

On the face, the nose has the most distinct geometry property and eyeholes are typical elliptical concave. The nose and eyeholes are well-isolated and immune to expression and hair. As other regions, like the mouth, or forehead, or cheeks, do not present particular or simple curvature characteristics and easily influenced by expression and beard.

Applying HK-classification on original face surface, a lot of candidates of nose tip and eyeholes can be obtained. The convex elliptical regions, as candidates of nose regions, are extracted by HK-classification on the original face after smoothed, as shown in Fig. 3 (a). The concave elliptical regions, as candidates of eyehole regions, are extracted by HK-classification on the original face after smoothed. It is noted that many candidates will be detected on the original face surface. Some methods [4], [5] based on curvature analysis have to use prior knowledge or classifier to detect landmarks. However, these regions are exclusive on the lower level surfaces reconstructed by MPU implitics. The nose region can be found in L1, which is elliptical convex with high value of the mean and Gaussian curvature, as shown in Fig. 3 (c). The eyehole regions are found from the high negative mean curvature and high Gaussian curvature in L2, as shown in Fig. 3 (d). The number of the candidates will enhance with increasing level. In our experiment, only one candidate nose can be detected on L1 as well as two candidate eyeholes on L2.

3. Experimental Results

The experiment has been performed using models from the GavabDB database [9]. This includes 3D facial surfaces of 61 individuals (45 males and 16 females). The total of the individuals are Caucasian and their age is between 18 and 40 years old. Each image is given by a mesh of connected 3D points of the facial surface in the VRML format without texture. For each person, 7 different models are taken, differing in terms of viewpoint, resulting in 427 facial models. In particular, for each subject there are 1 looking up, 1 looking down, 2 frontal, 1 random gesture, 1 laugh and 1smile models. Due to errors in the acquisition steps, some of the faces contain significant amount of noise or holes on the face. Some artifacts such as hair, beard and moustache are also presented.

Table 2 The error number and correctly localized rates for different scans.

	Total number	Error number	Correctly localized rate
Looking down scans	61	1	98.36%
Looking up scans	61	0	100%
Frontal scans	122	1	99.18%
Random gesture scans	61	2	96.72%
Laugh scans	61	2	96.72%
Smile scans	61	0	100%
Total scans	427	6	98.59%

Fig. 4 Three sample subjects showing 3D face registration of seven scans (looking down, looking up, frontal 1, frontal 2, random gesture, laugh and smile). The first, third and fifth row are the original scans, the second, forth and sixth row are the registered faces with detected feature points. The black point is the detected nose tip and the red points are the detected eyeholes.

Based on MPU Implicits, each faces are reconstructed into the second level (L1) and the third level (L2). Surface curvature is exploited respectively to L1 and L2. Depending on an HK classification, the nose region can be found in L1 and the eyehole regions are found in L2 uniquely. After applying the landmarking method introduced in Sect. 2, the results show a global 98.59% success rate over the 427 scans. Table 2 shows the error number and correctly localized rates for different scans.

Figure 4 shows the registration results of three subjects, each subject contains seven models (looking up, looking down, 2 frontal, random gesture, laugh and smile models). The first, third and fifth row are the original scans, the second, forth and sixth row are the registered faces with detected feature points. The black point is the detected nose tip and the red points are the detected eyeholes. Figure 5 shows some challenging examples which can be correctly localized by our method, such as pose variations, holes, noise, hair occlusion, beard, expression. Our method can overcome these challenges effectively. The six failures are caused by null information in nose region, hair occlusion seriously and exaggerated expression, as shown in Fig. 6.

Table 3 Comparing our work with other's work.

	Database	Method	Correctly localized rates(%) on Neutral Frontal Faces			Correctly localized rates(%) on faces with pose and expression		
			nose tip	eyehole	total	nose tip	eyehole	total
Our method	GavabDB(61 individual)	MPU+ curvature	99.18	100	99.18	98.83	99.53	98.59
	FRGC 1.0		100	100	100	/	/	/
A. Moreno [4]	GavabDB(60 individual)	Curvature+ prior knowledge	100	95.15	/	99.8	94.09	/
AA. Salah [11]	FRGC 1.0	3D	96.7	98.4	/	/	/	/
		3D+GOLLUM	98.2	99.3	/	/	/	/
		3D+BILBO	96.9	98.2	/	/	/	/
	FRGC 2.0	3D	/	/	/	96.7	97.2	/
		3D+GOLLUM	/	/	/	98	97.1	/
		3D+BILBO	/	/	/	96.8	96.3	/
C. Conde [12]	FRAV	Spin image	/	/	98.65	/	/	/
X. Dong [13]	BJUT_3D	3DRLS+profile	99.2	97(2D)	/	/	/	/

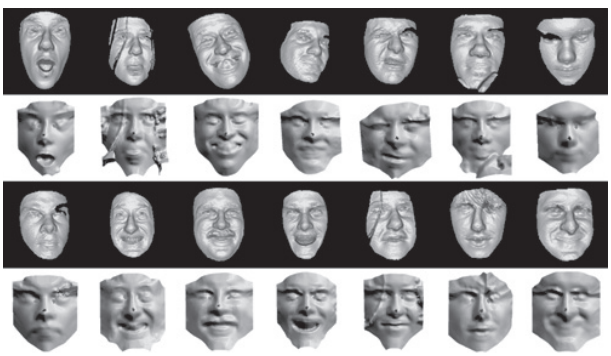


Fig. 5 Some challenging examples.

Fig. 6 Six incorrect localized scans.

Furthermore, we compare our method with four different approaches for 3D face landmarking. To prove the advance of our method, experiments also have been performed on FRGC 1.0 database [10]. A. Moreno [4] segmented the range images into isolated subregions using the traditional curvature analysis and prior knowledge based upon the GavabDB database. They achieved 99.8% correctly localized rates of nose tip for 3D faces with pose and expression, which is better than the result (98.83%) of our method. This is due to: 1) they distinguished the nose tip as the highest number of convex elliptic nodes and we do not use any prior knowledge, our method is more robust to any poses; 2) they deleted one individual of the GavabDB database and we use the whole database; 3) the only one incorrectly landmark of nose tip on neutral frontal face based on our method is due to the missing data of raw data (as the second column of Fig. 6 shown). For eyehole landmarks, our method achieve localized rate of 99.53%, which is a better performance than [4].

For FRGC 1.0 database, we obtain localized rate of 100% which is better than AA. Salah's results [11]. FRGC 2.0 database contains two facial expressions: neutral and smiling. GavabDB database is more complicated than FRGC 2.0 on pose and expression. Our experimental results on GavabDB are better than AA. Salah' work on FRGC 2.0.

For neutral frontal face, we obtain localized rate of 99.18% on GavabDB database and 100% on FRGC 1.0 databse. However, C. Conde [12] achieved localized rate of 98.65% on FRAV database and X. Dong [13] obtained the nose tip localized rate of 99.2% using 3DRLS and profile analysis on BJUT_3D database.

Table 3 summarizes their results as well as ours. As the results shown, our method based on MPU implicits and curvature analysis for landmarking has a better performance for 3D face with pose and expression.

4. Conclusion

In this paper, we have proposed a novel 3D facial landmarking method based on MPU Implicits and curvature analysis. The 3D facial surface is reconstructed by MPU Implicits in a hierarchical way and the holes are filled simultaneously. By analyzing the mean and Gaussian curvature in the lower levels, the nose region can be found in the second level (L1) and the eyehole regions are found in the third level (L2) accurately and uniquely. This proposed method is invariant to pose, holes, noise and expression. Finally, the experimental results with GavabDB database and FRGC 1.0 database have shown its excellent performance.

Acknowledgment

The work is supported by NSF of China (No.60873137) and The National Basic Research Program (973 Program) (No.2010CB723506).

References

[1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," ACM Comput. Surv., vol.35, no.4, pp.399–458, 2003.

[2] K. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition," Comput. Vis. Image Understand., vol.101, no.1, pp.1–15, 2006.

[3] G. Gordon, "Face recognition based on depth and curvature features," CVPR, pp.108–110, 1992.
[4] A. Moreno, A. Sanchez, J. Velez, and F. Diaz, "Face recognition using 3D surface-extracted descriptors," Proc. Irish Machine vision and Image Processing, 2003.
[5] A. Colombo, C. Cusano, and R. Schettini, "3D face detection using curvature analysis," Pattern Recognit., vol.39, pp.444–455, 2006.
[6] Y. Ohtake, A. Belyaev, M. Alexa, G. Turk, and H. Seidel, "Multi-level partition of unity implicits," Trans. Graphics (TOG), vol.22, no.3, pp.463–470, 2003.
[7] M. Do Carmo, Differential Geometry of Curves and Surfaces, Prentice-Hall, Englewood Cliffs, NJ, 1976.
[8] P. Besl and R. Jain, "Invariant surface characteristics for 3D object recognition in range images," Comput. Vis., Graph. Image Process., vol.33, pp.33–80, 1986.
[9] A. Moreno and A. Sanchez, "GavabDB: A 3d face database," Proc. 2nd COST275 Workshop on Biometrics on the Internet, pp.75–80, 2004.
[10] P.J. Phillips, W.T. Flynn, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W.J. Worek, "Overview of the face recognition grand challenge," CVPR, pp.947–954, 2005.
[11] AA. Salah and L. Akarun, "3D facial feature localization for registration," Lect. Notes Comput. Sci., vol.4105, pp.338–345, 2006.
[12] C. Conde and R. Cipolla, "3D facial feature location with spin image," Proc. MVA2005 IAPR Conference on Machine VIsion Applications, 2005.
[13] X. Dong and G. Wang, "Automatic 3D face segmentation based on facial feature extraction," Proc. ICIT, pp.1154–1159, 2006.