An Error Driven 3D Face Modeling Scheme based on Partial Differential Equations

Chuanjun Wang[1], Zhenfeng Shi[2],Xiamu Niu[1,2]
1. Harbin Institute of Technology Shenzhen Graduate School, 518055 Guangdong China
2. School of Computer Science and Technology at Harbin Institute of Technology, 150080 Harbin China
{chuanjun.wang,zhenfeng.shi,xiamu.niu}@ict.hit.edu.cn

***Abstract*—With the rapid development of computer graphics, the Partial Differential Equations (PDE) based geometry reconstruction technique has recently been considered as a powerful tool for 3D face modeling. However, the challenge on the accuracy of the reconstructed 3D face still exists due to the diversity of facial expressions. In this paper, an error driven 3D face modeling scheme based on PDE is proposed. Given a required error threshold, the proposed 3D face modeling scheme will dynamically select the boundary curves and iteratively approach to the required PDE based 3D face model under the control of the error distribution strategy. Experiments have shown the feasibility of the proposed scheme. And the generated 3D face model features the higher flexibility and the descriptive ability.**

3D face modeling; 3D face represntation; Partial differential equations;

I. INTRODUCTION

3D face models can find extensive application in face recognition and authentication [1], facial animation [2], video coding [3], etc. In most application scenarios, surface fitting and reconstruction are the two indispensible techniques in describing the geometry of human face. However, due to the diversity of facial expressions, face surface representation is still fraught with great technical challenges. It is worth noting that different application scenarios usually impose different error tolerance on face surface representation techniques. As a consequence, many kinds of models have to be built for different applications. For example, the first parameterized face model has been proposed in [4] in 1972, in which Parke tried to used a relative low number of triangles to simulate the facial expression animations. On the other hand, researchers still have to build the new models to deal with the expressions for the specific applications of recognition [1]. Is there any model that can provide the flexibility in terms of the different descriptive accuracy? Because the PDEs based method owes the ability to adjust the descriptive accuracy by the variations on the properties of the corresponding boundaries, it has been considered as a powerful tool to describe the face surface [5].

Elyan in [5], Yusheng in [2] have pioneered in the development of parametric 3D face model using the PDEs based method. In their parametric 3D face model, 28 boundary curves are first extracted from the scanned range image of the 3D face, and then nine PDEs are solved by taking these 28 curves as boundary conditions. The resulting parameters of the nine PDEs are the parametric representation of the reconstructed 3D face. This parametric representation scheme of the 3D face model based on PDE is of great efficiency and it provides some kind of flexibility of descriptive accuracy by using the different sampling ratios. However, the 28 boundary curves and the 9 PDEs are fixed, which constraints its flexibility in providing the accuracies for different individuals. In this paper, we start from the acceptable error tolerance of the reconstructed 3D face model and propose an error driven 3D face modeling scheme, in which the output models can provide a wide range of flexibility in terms of descriptive accuracy and efficiency. Given the requirements on error tolerance, the proposed modeling scheme dynamically selects the boundary curves of the PDEs and the reconstructed 3D face iteratively approach to the required descriptive accuracy. By this means, the efficiency of the reconstructed 3D face model depends on the descriptive accuracy, and the flexibility between the efficiency and accuracy is provided in the same 3D face model.

The rest of this paper is organized as following. Details of the proposed scheme are introduced in section II. Section III illustrates and analyzes some of the experiment results. And the paper closes with the conclusion and discussion on the future work in section IV.

II. THE PROPOSED SCHEME

In the 3D face modeling based on PDE, the parametric 3D face is represented as a series of parameters that represent the PDE solutions. Thus, the processes of fitting and reconstructing the 3D face are to find the parametric solution of the corresponding partial differential equations, through a set of boundary conditions extracted from the original scanned 3D face. Figure 1 illustrates the process of the proposed error driven 3D face modeling scheme. The main step of the proposed modeling scheme includes the preprocessing, the available curve extraction, the boundary curve selection and the PDE face generation. Firstly, we normalize and transfer the scanned face into the predefined coordinate for normalized sampling. We then extract all the available curves as the candidates of boundary curves from the normalized sampling data. After that, the boundary

978-0-7695-4222-5/10 $26.00 © 2010 IEEE
DOI 10.1109/IIHMSP.2010.115

curves are dynamically selected from the available curve set under the control of error distribution strategy. These boundary curves are used to determine the parameters of the corresponding PDEs. Simultaneously, the resulting PDE face is generated and iteratively converges to the required error thresholds.

In order to extract the available curves from the 3D face range image, we need to preprocess the scanned face first.

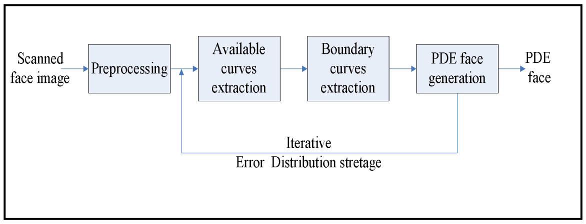

Figure 1. Flow chart of the error driven scheme.

A. *Preprocessing*

The original scanned 3D face is supposed to be in the form of point cloud, which can be represented as a set of points $p_i = \{x_i, y_i, z_i, i = 1, 2 \ldots n\}$, n is the number of the points. Usually, the scanned face image contains some noise data, or holes in terminology, as well as the different pose variation. So, we first have to preprocess the data and the following items describe the details.

- Apply a median filter on the scanned 3D face image with the window size of 20.
- Identify the landmarks of the 3D face data using the method proposed in [6]. The landmarks used include the top line of the eyebrow, the nose tip, the nose peaks, the inner eye corner, etc.
- Transfer the 3D face data into the normalized coordinate by calculating the mid-line plane of the face using the method presented in [7].
- Interpolate the Z axes of 3D face with an acceptable resolution of $M \times N$ in the X axes and Y axes. The reason why we interpolate the data is for normalized sampling. We can also use some other methods to extract the available curves without interpolating the data. The commonly used interpolation methods include the linear interpolation, the cubic interpolation and the nearest neighbor interpolation. In our scheme, we use the triangle based cubic interpolation.

After the step of preprocessing, the facial data is aligned and normalized in the predefined coordinate. And they will be considered as the goal of the reconstruction as well as the input of the proposed 3D face modeling scheme.

B. *Curve extraction*

Because the reconstructed 3D face is represented as the solution of the corresponding PDEs, the properties of the boundary curves play the most important role in the PDE based 3D face reconstruction. The boundary curves can affect the results either by the allocation of the curves or the number of curves used to reconstruct the 3D face. Therefore, we first extract all the available curves and then dynamically select the boundary curves from the available curve set. Many kinds of curve extraction methods can be adopted, such as the geodesic distance based method used in [5]. In our scheme, we directly extract the available curves from the normalized data. In the preprocessing step, the face data has already been represented in the form of $p_i = \{x_i, y_i, z_i, i = 1, 2 \ldots M \times N\}$, M is the number of the rows and N is the number of column in the X-Y plane. So we can extract M curves, each with N points.

Inspired by the MPEG-4 standard [3], where the landmarks that will be affected by the facial action are defined, we divide the 3D face into four semantic regions and exploit the optimization on the allocation of the boundary curves. The four regions are the forehead, the eye area, the middle of the face (including the nose and the face area), and the mouth area. Figure 2(a) is an example of the divided face. Specifically, the adjacent curve between the forehead and the eyes area is chosen as the top line of the eye brow. The adjacent curve between the eye area and the middle of the face is chosen as the curve that has the same distance from the top of the eye brow to the line of the two eye corners. And the adjacent curve between the middle of the face area and mouth area is chosen as the line that is close to the nose peak.

For each region, we also extract all the available curves and then dynamically select the boundary curves for each of face regions. This optimization of face division provides the merits to treat the different area with different error distribution strategy, which will be very important in different applications. For example, in the application of facial expression analysis, the eye area is much more important than the forehead area, so we can focus on the accuracy of the eye area, while paying more attention on the efficiency of the forehead area. Figure 2(b) gives an example of the face in which the accuracy basis is applied in the eye area, while the efficiency is paramount in the forehead area. In the following, we will treat each face regions as the unit of reconstruction and the reconstructed PDE face will be represented as the combination of the four PDE face regions.

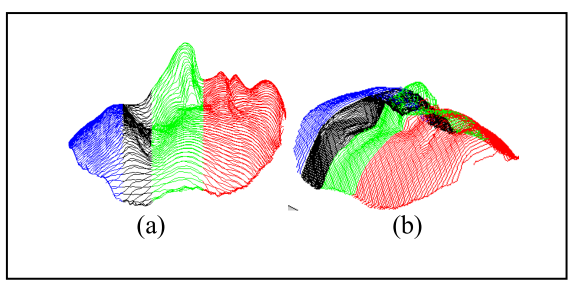

Figure 2. Examples of the face division and curve extraction.

C. *PDE face generation*

The PDEs are solved through a set of boundary curves selected from the available curves. Once the boundary curves of the semantic regions are extracted, we can use these curves as the boundary conditions to solve the fourth order PDEs in (1).

$$\left(\frac{\partial^2}{\partial u^2}+a^2\frac{\partial^2}{\partial v^2}\right)^2 \underline{X}(u,v)=0 \quad (1)$$

The analytic solution of (1) can be represented as (2).

$$X(u,v)=A_0(u)+\sum_{n=1}^{\infty}[A_n(u)\cos(nv)+B_n(u)\sin(nv)] \quad (2)$$

where,

$$\begin{aligned} A_0(u)&=\alpha_{00}+\alpha_{01}u+\alpha_{02}u^2+\alpha_{03}u^3 \\ A_n(u)&=\alpha_{n1}e^{anu}+\alpha_{n2}ue^{anu}+\alpha_{n3}e^{-aun}+\alpha_{n4}ue^{-anu} \quad (3) \\ B_n(u)&=\beta_{n1}e^{anu}+\beta_{n2}ue^{anu}+\beta_{n3}e^{-aun}+\beta_{n4}ue^{-anu} \end{aligned}$$

The solution of equation (1) is represented in the form of Fourier series. And we usually use the first 3~6 modes to represent the solution, which means that the solution of equation (1) is usually chosen as (4).

$$X(u,v)=A_0(u)+\sum_{n=1}^{K}[A_n(u)\cos(nv)+B_n(u)\sin(nv)] \quad (4)$$

where K is always set to be 3~6, and the remainder term is defined as (5).

$$r(u,v)=r_1(u)e^{wu}+r_2(u)e^{-wu}+r_3(u)ue^{wu}+r_4(u)ue^{-wu} \quad (5)$$

where w is chosen as $w=a(K+1)$, and r_1, r_2, r_3, r_4, are the functions denoting the difference between the boundary curves and the reconstructed boundary curves.

Figure 3. Pseudo code of the face generation.

```
Divide the 3D face into four semantic regions
For each region
{   Set the acceptable error threshold
    For min(PDEs):max(PDEs)
    {Dynamically select the boundary curves
     Reconstruct the 3D face region
     Compare the errors distance
     If (the error distance is below the threshold)
        {Output the results
          break
        }
      Else
        {Increase the number of PDE used
         Dynamically select the new boundary curves
        Continue
        }
    }
```

Figure 3 describes the pseudo code that is used to generate the parametric 3D face. Specifically, we first set the error thresholds for each region of the face, and start with the minimum number of PDEs that we can afford to reconstruct the parametric 3D face. Denote the number of the available curves as M and the number of PDEs as S. Then, M is firstly cut into $3S+1$ regions, and the average available curves in each regions are chosen as the $3S+1$ boundary curves used to resolve the S PDEs. The resulting S PDEs describe the reconstructed face surfaces, so we can compare the error distance between the reconstructed 3D face and the normalized 3D face. If the error distance is below the threshold, we use the resulting parameters of the corresponding PDEs as the representation of the 3D face, or if the error distance is above the threshold, we continue to increase the number of PDEs until the maximum number of the PDEs is reached.

D. Discussion on the descriptive threshold

Given an arbitrary error threshold T, can the proposed 3D face modeling provide the reconstructed 3D face with a error distance less than T? The answer depends on the specific applications. So we want to know what will affect the accuracy of the reconstructed 3D face and how best the proposed 3D face modeling scheme can provide.

Based on the analysis on the proposed 3D face modeling scheme, we know that the following three factors will affect the fidelity of the reconstructed 3D face.

- The number and allocation of the boundary curves. They play the most important roles in the PDE based 3D face modeling, and they are eventually the most important factors that will affect the reconstructed accuracy. However, in the best situation, the error distance caused by the number and the allocation of the boundary curves can be very small, and even be zero in the ideal case.
- The number of points used in each boundary curve, or the sample ratio referred in [5]. In the process of surface reconstruction, the number of the sample points used in each boundary curve will affect the results of the reconstructed surface. Elyan in [5] has demonstrated that when the sample ratio varies from 5% to 60%, the error distance varies from $1.251*10^{-3}$ to $7.408*10^{-3}$.
- The number of the Fourier modes. According to the properties of the Fourier modes, the error distance caused by the number of the modes is the remainder term of the Fourier modes, which has been illustrated in (5).

All the above three factors will affect the fidelity of the reconstruct 3D face. And eventually, the number and allocation of the boundary curves play the most important role, that is why we divide the face into different regions and apply different error distribution strategies on the different regions of the face. As a conclusion, the flexibility of accuracy and efficiency basis of the error driven modeling can be achieved through the different strategies applied on the above factors.

III. EXPERIMENT RESULTS AND ANALYSIS

In this section, we evaluate the descriptive accuracy and efficiency of the 3D face models generated by the proposed 3D face modeling scheme. We use the Bosphorus 3D face database [8] as the scanned 3D face image in our experiments. The expression variations in this database not only contain the six universal emotional expressions (happiness, surprise, fear, sadness, anger and disgust) but also face action units of the facial action coding system (FACS). Based on the properties of the Bosphorus 3D face database, we use 297 faces with neutral pose, 1535 faces with lower facial action units, 428 faces with upper facial action units, 166 faces with action unit combinations and 445 faces with emotional expressions in our experiments. Firstly, we test the descriptive accuracy and then the efficiency of the 3D face models generated by the proposed error driven modeling scheme. Figure 4 and 5 are illustrated as an example of the results.

In both figure 4 and 5, the number of the sample points P is set to 24, the number of the available curves M is 500, and the Fourier mode is chosen as 3. Empirically, we set the threshold for each of the four face regions as 3:3:3:4 (forehead: eye area: nose area: mouth area). In this case, all the faces that accept the error distance of 0.6mm can be represented by the generated parametric 3D face model. And if the acceptable error threshold is above 0.8mm, the proposed 3D face modeling scheme can automatically generate the corresponding parametric 3D face for all the five type of different facial pose. Noting that the available error threshold is affected by the properties of the boundaries, the number of the sample points and the Fourier modes, we can adjust these factors and get the different descriptive abilities of the generated model. Figure 5 shows the corresponding number of PDEs used to describe the generated parametric 3D face.

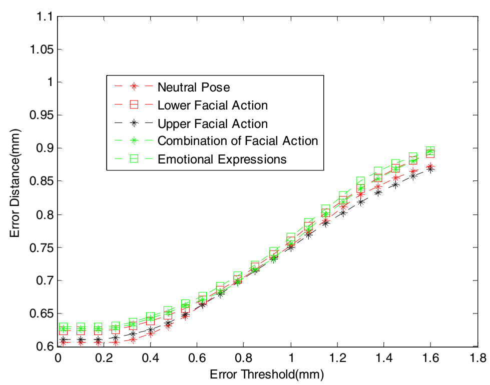

Figure 4. Illustration of the error threshold and the real error distance. P= 24, M = 500,K= 3

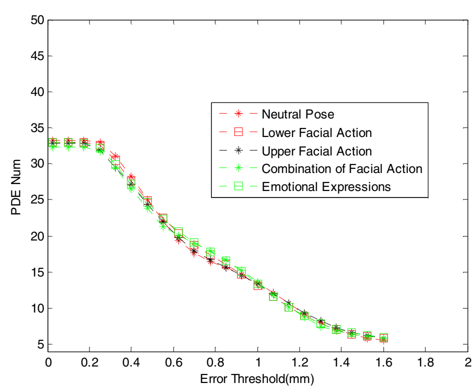

Figure 5. Illustration of the error threshold and the number of PDEs. P= 24, M = 500,K= 3

IV. Conclusion and Futurework

In this paper, an error driven 3D face modeling scheme based on PDEs has been proposed. Starting from the acceptable error threshold, the PDE based 3D face modeling scheme dynamically select the number and the allocation of the boundary curves, and the resulting parametric 3D face model iteratively meet to the requirement. By this means, the proposed 3D face model can describe the geometry of the 3D face with high accuracy or with high efficiency in the same model. In addition to this, the optimization on the allocation of the boundary curves used for 3D face modeling is also exploited. Inspired by the definition of face in MPEG-4, the face is first divided into four regions and then the error driven modeling scheme is applied on each of the regions. Experiment results show the feasibility of the proposed 3D face modeling scheme, and the generated 3D face model features the flexibility in term of accuracy and efficiency.

Based on the proposed error driven 3D face modeling scheme, our future work will focus on the best PDE results that satisfy the requirement of the error thresholds from all the aforementioned three aspects in the purpose of human facial expression analysis [9].

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Project Number: 60703011, 60832010, 60671064), the Chinese national 863 Program (Project Number: 2007AA01Z458), and The Research Fund for the Doctoral Program of Higher Education (RFDP: 20070213047).

References

[1] X. Lu, A. Jain, “Deformation modeling for robust 3D race matching,” IEEE Trans. on Pattern Analysis and Machine Intelligence. vol. 30, NO.8, pp. 1346–1356, August 2008.

[2] S. Yu, P. Willis, G. Castro, and H. Ugail, “PDE-Based Facial Animation:Making the Complex Simple,” Proc. International Symposium on Visual Computing(ISVC2008). Dec. 2008, pp. 723-732.

[3] I. Pandzic, R. Forchheimer, MPEG-4 Facial Animation: the standard, implementation and application, John Wiley, 2002.

[4] F.Parke, “A parametric model for human faces,” Tech. Report UTEC-CSC-75-047, University of Utah, 1974.

[5] E. Elyan, H. Ugail, “Reconstruction of 3D human facail images using partial differential equations,” Journal of computers, vol. 2, No..8, pp. 1–8, October 2007.

[6] H. Dibeklioglu,A. Salah, and L. Akarun, “3D Facial Landmarking under Expression, Pose, Variations and Occlusion,” Proc. IEEE on Biometric: Theory, Application and System (BTAS 2008), Washingto, DC, August. 2008.

[7] D. Colbry, G. Stockman, “Real time identification using a canonical face depth map,” IET Comppput. Vis., vol. 3, Iss.2, pp. 74–92, April 2009.

[8] A. Savran, N. Alyuz, H. Dibeklioglu, et al, “Bosphorus database for 3D face analysis” The first cost 2101 workshop on biometrics and identity management(BIOID 2--8), Roskilde University, Denmark, May. 2008.

[9] S. Krinidis, I. Pitas, “Statistical analysis of human facial expressions,” Journal of Information Hiding and Multimedia Signal Processing.vol. 1, NO. 3, pp. 241-240, July 2010.