

Using the Orthographic Projection Model to Approximate the Perspective Projection Model for 3D Facial Reconstruction

Jin-Yi Wu and Jenn-Jier James Lien

Department of Computer Science and Information Engineering,
National Cheng Kung University, Tainan, Taiwan 70101, ROC
{curtis, jjlien}@csie.ncku.edu.tw

Abstract. This study develops a 3D facial reconstruction system, which consists of five modules, using the orthographic projection model to approximate the perspective projection model. The first module identifies a number of feature points on the face and tracks these feature points over a sequence of facial images by the optical flow technique. The second module applies the factorization method to the orthographic model to reconstruct a 3D human face. The facial images are acquired using a pinhole camera, which are based on a perspective projection model. However, the face is reconstructed using an orthographic projection model. To compensate for the difference between these two models, the third module implements a simple and efficient method for approximating the perspective projection model. The fourth module overcomes the missing point problem, commonly arising in 3D reconstruction applications. Finally, the fifth module implements a smoothing process for the 3D surface by interpolating additional vertices.

Keywords: 3D reconstruction, factorization, orthographic projection, and perspective projection.

1 Introduction

The goal of 3D facial reconstruction, which has been studied for decades, is to reconstruct the 3D face model from either a single image or a set of images taken from known or unknown camera viewpoints. Lee et al. [9] developed a technique for constructing 3D facial models by using laser scanners, and the method acquires a very high accuracy reconstruction result. However, their method is time-consuming and the equipments are expensive. Hence its applicability in the public domain is limited. Various researchers present different approaches in the use of single or multiple cameras for 3D reconstruction. We organize those approached into several categories.

One category is to apply the Bundle Adjustment (BA) approach, such as [5], [11], [12], [14], [15], [22], and [27]. They model the 3D reconstruction problem as a minimization problem between the 2D ground-truth feature point locations and a 2D location estimating function which consists of a 3D-2D projection function (the intrinsic and extrinsic parameters of the camera motion) and the 3D shape of the object. By using the Levenberg-Marquardt (LM) algorithm, which takes the advantages of both

the Gauss-Newton algorithm and the steepest descent algorithm, they can solve both the parameters of the 3D camera motion and the shape of the 3D object simultaneously. However, because applying LM algorithm to this case will suffer the problem of calculating the inverse Hessian matrix, whose size is dominated by the estimated parameters, LM will take a long time for the computational converge. In addition, the BA approach itself is also a large sparse geometric parameters estimation problem. There are many sparse optimization methods to accelerate the BA algorithm, but it still needs a long time to solve the parameters, especially for very long image sequences. Therefore, [11], [12], [27] use local optimization to accelerate the speed to solve the problem.

Another category is the shape from shading (SfS) approach [17], [28], which can reconstruct the 3D shape as long as the scene or object satisfy the Lambertian reflection model. However, not all scene or object can satisfy this constrain. Therefore, a number of researches [8], [16], [20] turn to seek some Bidirectional Reflection Distribution Functions (BRDFs), which can more generally model the reflection model to reconstruct the 3D structure.

The other category is the shape from motion (SfM) approach, which can be solved by using the factorization approach [13], [18], [21]. By introducing the rank constraint and factorizing by singular value decomposition (SVD), the factorization approach with the orthographic projection models [18] or the para-perspective projection models [13] can factorize the location matrix of the 2D feature points in the image plane to the 3D rotation matrix of the 2D image frame and the 3D shape matrix of the object. Moreover, the work in [21] generalizes works in [18] and [13] to recover both the 3D rotation information and the 3D shape of the object by applying the perspective projection model.

All above-mentioned methods are used to reconstruct the static or rigid objects. Recently, some researches focus on the 3D reconstructions of non-rigid objects over an image sequence [1], [2], [3], [19], [25], [26]. They model the 3D shapes as a linear weighted combination of a set of shape vectors, so that they can represent different shapes correspond to different images by giving different weights.

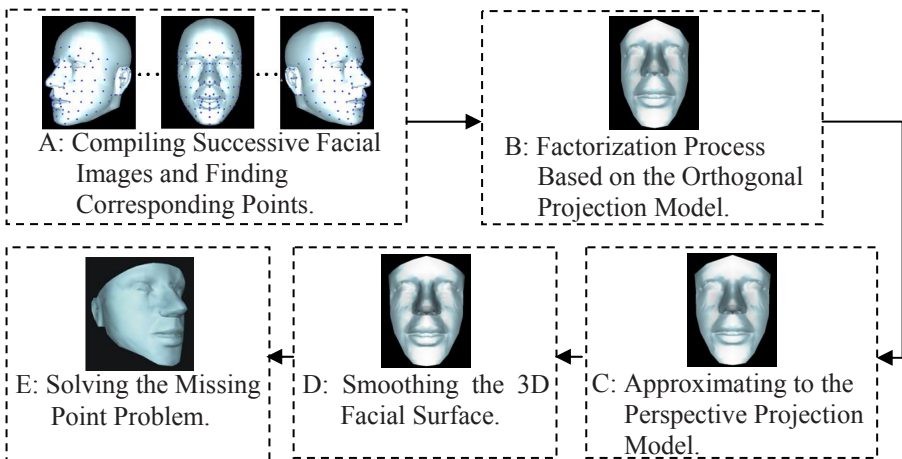


Fig. 1. Workflow of the proposed 3D facial reconstruction system

Regarding above existing works, our goal is to reconstruct a human face with a common PC camera and a regular PC, and it is easy for human to do pan-rotation without any non-rigid motion. Hence, we choose the method easily to reconstruct 3D rigid object.

In all kinds methods mentioned above, the factorization method is a relative simple method and can acquire a good reconstruction result. Thus, current study develops a straightforward and efficient 3D reconstruction approach in which the perspective projection model is approximated by applying the factorization method to the orthographic projection model. A solution is also presented for solving the missing point problem when the face moves through large pan-rotation angles. Finally, a smoothing method is presented to interpolate additional 3D vertices in order to give the surface of the reconstructed 3D face a smoother and more realistic appearance.

2 System Description

Fig. 1 shows the major five modules in the proposed 3D facial reconstruction system. We will discuss each module in the following sections, and the difference between the results of modules C and D will be shown more clearly in Section 3.

2.1 1st Module: Compiling Successive Facial Images and Finding Corresponding Points

A conventional PC camera is fixed at a stationary position in front of the subject and is used to capture a sequence of facial images as the subject gradually turns his or her head from left to right or right to left in the pan-rotation direction. In the first frame of the facial image sequence, N various facial feature and contour points $p_m(u, v)$ are automatically located using the method proposed in [23]. The optical flow technique [10] is then used to track the corresponding points in the remaining image frames. However, a tracking error may occur for some of the feature points located in textureless regions of the face. Therefore, a manual correction procedure is applied to remedy those particular feature points. Subsequently, Delaunay Triangulation [4], [6] is applied to construct the 2D triangles from the features and contour points.

2.2 2nd Module: Factorization Process Based on the Orthographic Projection Model

In the proposed system, the 3D face is reconstructed using the factorization approach based on the orthographic projection model [18], which is a simple model compared with the perspective projection model [13], [21]. Based on the locations of N corresponding 2D feature points over F facial image frames, a $2F \times N$ point matrix W is created. By applying the factorization method [18], the point matrix W becomes:

$$W_{2F \times N} \xrightarrow{\text{Factorization}} R_{2F \times 3} \times S_{3 \times N} \quad (1)$$

where R is the 3D rotation matrix of x and y axes and S is the 3D shape matrix. Then the 2D triangles are used to construct corresponding 3D polygons, where the position of each vertex s_m is defined as (x_m, y_m, z_m) .

2.3 3rd Module: Approximating to the Perspective Projection Model

The reconstruction method used in this study is based on the orthographic projection model. However, the pinhole camera used to acquire the facial images is based on a perspective model. Therefore, if the orthographic model is to provide an accurate approximation of the perspective model, the ratio of the object depth, Δd , to the distance between the camera and the object, d , should be very small, i.e. $\Delta d \ll d$. Since d is determined when the facial images are recorded, the parameter can be adjusted is Δd . In other words, if Δd can be reduced, the ratio $\Delta d/d$ is also reduced, and hence the reconstruction result obtained for the vertex's depth z_m , which corresponds to the origin of the 3D facial coordinate system, will be improved.

Accordingly, in the third module of the proposed reconstruction system, the 3D facial shape S is divided into n groups (layers) g_1, g_2, \dots, g_n according to the z value by K-means classification. In other words, the 3D facial shape S is described by $S = g_1 \cup g_2 \cup \dots \cup g_n$ and $|g_i \cap g_{i+1}| \geq 4$, where $1 \leq i \leq n-1$. Each group g_i has a set of corresponding 2D projection points p_i . Registering two 3D coordinate systems requires the presence of at least 4 common points in these two coordinate systems in order to solve the 12-parameter rotation and translation transformation matrix. Then applying the factorization approach to each set of p_i , we can have $p_i \xrightarrow{\text{factorization}} r_i' g_i'$, for $i=1$ to n . Since each group g_i' has its own 3D coordinate, it is necessary to register all of the groups into the same 3D coordinate system. To achieve this, the coordinate of g_{i+1}' is aligned with that of g_i' . This procedure, which is called the group registration procedure, is employed to register the different coordinates of all the groups g_i' into the same world coordinate system to form a complete reconstructed 3D model. The basic steps of the group registration procedure can be summarized as follows: Using the overlapping regions $o_{i(i+1)}$ in g_i' and $o_{(i+1)i}$ in g_{i+1}' , the 12-parameter transformation matrix τ_{i+1} between $o_{i(i+1)}$ and $o_{(i+1)i}$ can be found by $o_{i(i+1)} = \tau_{i+1} o_{(i+1)i}$, where $\tau = [R_{3 \times 3} T_{3 \times 1}]$ and R, T are the rotation matrix and the translation vector. Based on the transformation matrix τ , the new coordinate of $g_{i+1}' = \tau_{i+1} g_{i+1}'$ has the same coordinate as g_i' . By finding all transformation matrix τ between $n-1$ successive group pairs, we can transform all coordinates of all groups into the same coordinate. Then the registered 3D facial model (or facial shape) S' is described by $S' = g_1' \cup g_2' \cup \dots \cup g_n'$. Therefore, the depth z_m of each vertex in the 3D object in the world coordinate system is more accurate.

After improving the depth value z_m of each vertex in the 3D object, the accuracies of the x and y values of each vertex are also needed to be improved. Analyzing the reconstructed 3D facial model S' , it is found that the x -axis and y -axis components of the 3D coordinate (x'_m, y'_m, z'_m) of every vertex s'_m in S' are virtually the same as the corresponding 2D projection point (u_m, v_m) in the image coordinate system (Fig. 2.a).

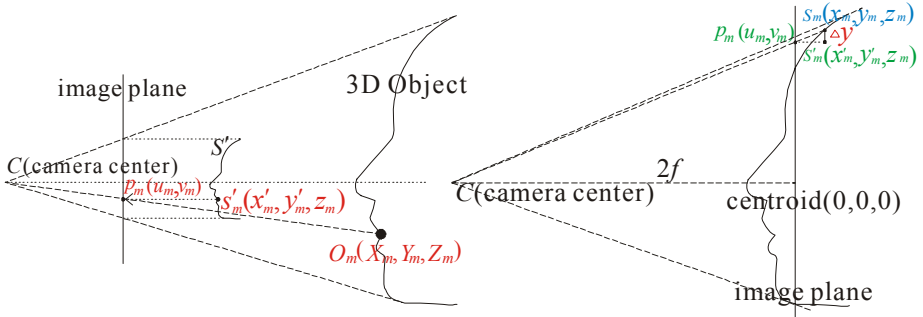


Fig. 2. (a) Point p_m in the image plane represents the 2D projection point of both the 3D reconstructed point s'_m in the orthographic model and the 3D object point O_m in the perspective model. (b) The difference Δy between the reconstruction results in the y-axis direction obtained from the orthographic model and the perspective model, respectively.

This is understandable because the factorization approach uses the orthographic projection model. However, the situation, $x'_m = u_m$, and $y'_m = v_m$, under the perspective model occurs when the 3D object and 2D image plane are both located at a distance of twice the focal length f , which can be known in advance, from the origin of the camera coordinate. Therefore, approximating the perspective projection model using the orthographic projection model requires an assumption that the x-y plane of the 3D object's coordinate system overlaps the 2D u - v image plane at a distance of $2f$ from the origin of the camera coordinate system, as show in Fig. 2.b. And it can be seen that the reconstructed vertex s'_m and the real vertex s_m have an error Δy in y-axis, where:

$$\Delta y = y'_m \frac{z_m}{2f} \tag{2}$$

Therefore, the improved value y_m for vertex s_m is given by:

$$y_m = y'_m + \Delta y = y'_m + \frac{y'_m z_m}{2f} \tag{3}$$

By the same logic, the improved value x_m for vertex s_m is expressed as:

$$x_m = x'_m + \Delta x = x'_m + \frac{x'_m z_m}{2f} \tag{4}$$

Applying (3) and (4), an improved reconstruction result for the x_m and y_m components of the vertex s_m in S can be obtained. Note that hereafter, the result of this “improved reconstruction process” is referred to as “the improved result.”

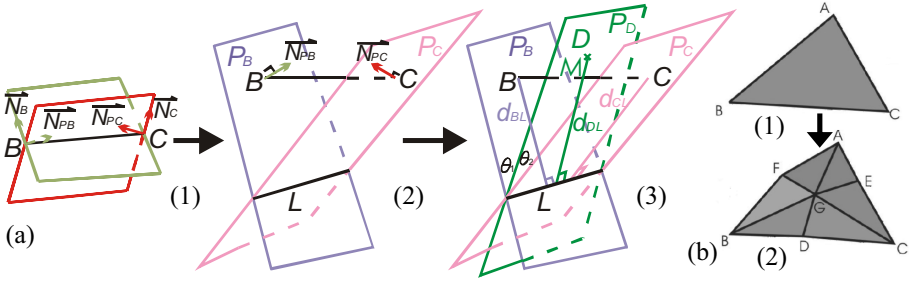


Fig. 3. (a) Application of smoothing method. (b) Smoothing result

2.4 4th Module: Smoothing the 3D Facial Surface

The higher density of the 3D feature points in the facial image, the smoother the appearance of the reconstructed 3D facial shape. However, a high feature point density may cause a tracking problem in that some points may be erroneously matched with one of their neighbors. Therefore, in the proposed 3D facial reconstruction system, a small number of 2D feature points are used initially to create a draft 3D facial model containing only a limited number of polygon vertices. Then, a smoothing method is implemented in which additional vertices (*D*, *E*, *F*, and *G* in Fig. 3.b.2) are interpolated to create a mesh of finer polygons.

For each vertex *v*, the normal direction \overline{N}_v can be found by calculating the mean of the normal directions of those meshes which contain *v* as one vertex. As shown in Fig. 3.b.1 and 3.a.1 on edge \overline{BC} , vertices *B* and *C* have normal directions \overline{N}_B and \overline{N}_C ; and vectors \overline{BC} and \overline{CB} can be calculated. Therefore, a vector \overline{N}_{PB} , which lies on the plane formed by \overline{N}_B and \overline{BC} , can be found by solving the following equations:

$$\begin{cases} \overline{N}_{PB} \cdot \overline{N}_B = 0 \\ \overline{N}_{PB} \cdot (\overline{N}_B \times \overline{BC}) = 0 \end{cases} \tag{5}$$

Similarly, vector \overline{N}_{PC} can be found in the same way. Having found \overline{N}_{PB} and \overline{N}_{PC} , the two corresponding planes, P_B and P_C , can be determined. As shown in Fig 3.a.2, P_B and P_C are planes, which intersect at line *L*, with normal directions \overline{N}_{PB} and \overline{N}_{PC} passing through *B* and *C*, respectively. An arbitrary point *M* on edge \overline{BC} (see Fig. 3.a.3) and line *L* form a plane P_D with a normal direction \overline{N}_{PD} . The vector \overline{N}_M , the normal direction of point *M* lying on P_D , can be found by solving:

$$\begin{cases} \overline{N}_M = \alpha \overline{N}_B + (1 - \alpha) \overline{N}_C \\ \overline{N}_M \cdot \overline{N}_{PD} = 0 \end{cases} \tag{6}$$

where α is a scaling factor used to control the weights of \overline{N}_B and \overline{N}_C . Finally, the interpolated vertex *D* can be determined from:

$$\begin{cases} \overline{MD} = \beta \overline{N_M} \\ d(D, L) = d(B, L) \frac{\theta_1}{\theta_1 + \theta_2} + d(C, L) \frac{\theta_2}{\theta_1 + \theta_2} \end{cases} \quad (7)$$

where β is a scaling factor and $d(X, L)$ is the distance from X to L , where X is the vertex B , C , or D , respectively. Similarly, the interpolated vertices E and F can be obtained from (5) to (7) for edge \overline{CA} and \overline{AB} , respectively, as shown in Fig. 3.b.2. In this work, the point M in Fig. 3.a.3 represents the mid-point for the interpolated vertices D , E , and F . Similarly, to locate vertex G in Fig. 3.b.2, lines AD , BE , and CF are employed to find the interpolated vertices G_{AD} , G_{BE} , and G_{CF} , respectively, from (5) to (7). Because G should be located at the centroid of the ABC triangular polygon, the M position for G_{AD} , G_{BE} , and G_{CF} should be as $d(X, M) : d(M, Y) = 2:1$, where (X, Y) is (A, D) , (B, E) , or (C, F) . Vertex G is then given by $G = (G_{AD} + G_{BE} + G_{CF}) / 3$.

After the smoothing process, the original polygon m_{ABC} is transformed into six individual sub-polygons in the 3D space. Hence the 3D surface has a smoother appearance. As a result, when all the polygons in the original 3D model S have been processed using the smoothing method, the reconstructed facial surface is far smoother and more lifelike than the original reconstructed model.

Table 1. Reconstruction Error (%) of Each Axis Using The Factorization Method With the Orthogonal Model And the Improved Method (sepetated by “/”)

Angles	-1° to +1°	-5° to +5°	-10° to +10°	-15° to +15°
X-axis	2.60% / 0.74%	2.57% / 0.70%	2.54% / 0.63%	2.47% / 0.57%
Y-axis	3.49% / 1.15%	3.49% / 1.13%	3.47% / 0.93%	3.46% / 0.92%
Z-axis	7.49% / 6.19%	7.47% / 6.17%	7.40% / 6.14%	7.29% / 6.05%

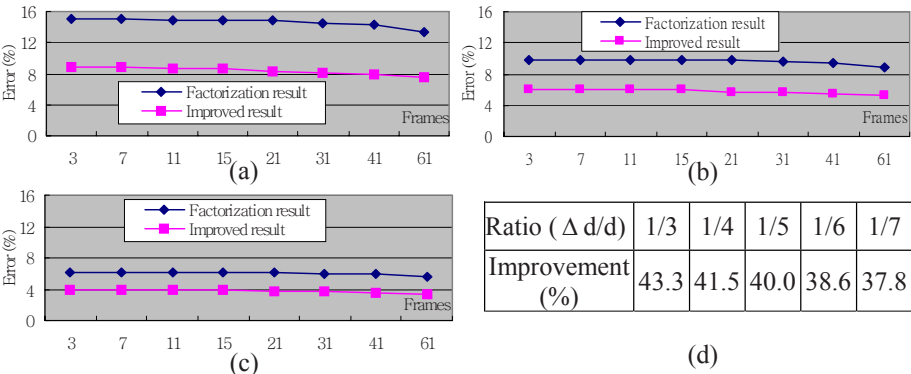


Fig. 4. Average errors for different $\Delta d/d$ ratios: (a) 1/3, (b) 1/5 and (c) 1/7. (d) is the average improvement for different $\Delta d/d$ ratios.

2.5 5th Module: Solving the Missing Point Problem

When the pan-rotation angle of the face is large, some facial feature points will be occluded, i.e. some of the feature points will disappear. This presents problems when applying the factorization method because some of the elements in 2D position matrix W will be missing. From observation, it is found that when the face turns to the left or right through an angle of approximately around 10° to 15° , the feature points adjacent to the nose are occluded by the nose. It is also observed that the ears of the face become clearly visible when the head is rotated to the left or right by more than 30° . Therefore, the captured video sequence is segmented into three parts, namely the left side-view V_1 (from -45° to -30°), the frontal-view V_2 (from -10° to 10°), and the right side-view V_3 (from 30° to 45°), according to [7] and [24]. The improved reconstruction process is then used to reconstruct these three parts, V_1 , V_2 , and V_3 , individually. The side-views, V_1 and V_3 , are then registered to the frontal-view using the same registration procedure as that described in the previous section.

3 Experimental Results

To evaluate the performance of the improved reconstruction method (excluding the smoothing process), 10 different pre-known 3D head model was used to evaluate the reconstruction results obtained for different values of the $\Delta d/d$ ratios. Fig. 4 shows the estimation errors. Figs. 4.a, 4.b, and 4.c present the errors obtained for different numbers of frames in the reconstruction process for $\Delta d/d$ ratios of $1/3$, $1/5$, and $1/7$. It can be seen that as the $\Delta d/d$ ratio decreases, i.e. the perspective projection model more closely approximates the orthographic projection model, the estimation errors under both methods reduce. Furthermore, the improved method provides a better reconstruction performance than the original factorization method with an orthographic model. Fig. 4.d shows that the improvement in the reconstruction results obtained by the improved method reduces as the ratio decreases. This result is to be expected since the orthographic model used to create the reconstruction results more closely approximates the perspective projection model at lower values of the $\Delta d/d$ ratio.

Table 1 shows the reconstruction errors using a $\Delta d/d$ ratio of $1/5$. Comparing the errors of factorization method with orthographic model and of the improved method, the reconstruction results of the latter are 78%, 70%, and 17% better than those of the factorization method with the orthographic model in the x-axis, y-axis, and z-axis directions. Therefore, the improved method yields an effective improvement in the reconstruction results in the x-axis and y-axis directions. However, it fails to provide an obvious improvement in the z-axis direction. The reason for this is that the ratio of $\Delta d/d$ used in the layered reconstruction approach is insufficiently small. Nonetheless, if an adequate number of feature points are assigned such that the ratio of $\Delta d/d$ can be reduced, it is reasonable to expect that the improved reconstruction method will provide a more accurate reconstruction result.

The robustness of the proposed reconstruction system to tracking errors was evaluated by adding white Gaussian noises with various standard deviations to the tracking results, the 2D feature point positions, to simulate the tracking errors. The reconstruction results obtained using the improved method are shown in Figs. 5.b, 5.c, and 5.d

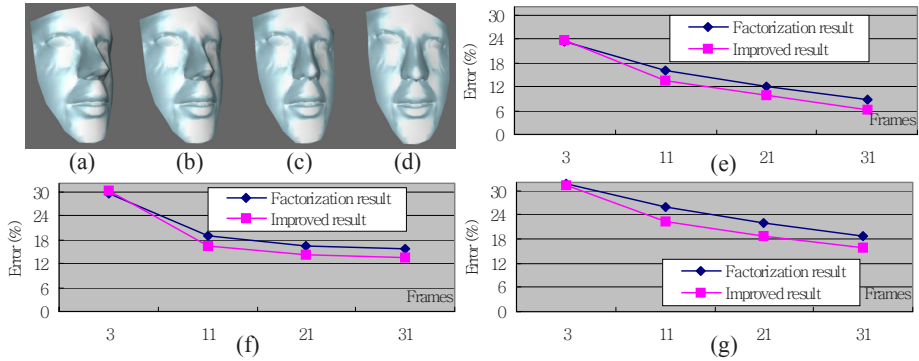


Fig. 5. (a) Reconstruction result obtained using the improved method in the absence of noise. Note that the $\Delta d/d$ ratio is $1/5$. (b), (c), and (d) show the reconstruction results obtained when white Gaussian noise is added with a standard deviation of 1, 3, and 5, respectively. (e), (f), and (g) show the reconstruction errors of the 3D facial model after applying the factorization method with the orthogonal model and the improved method, when white Gaussian noise is added to the tracking results with a standard deviation of 1, 3, and 5, respectively.

for white Gaussian noises with standard deviations of 1, 3, and 5, respectively. The estimation errors obtained from the factorization method with the orthogonal model and the improved method are shown in Figs. 5.e, 5.f, and 5.g for Gaussian noises with standard deviations of 1, 3 and 5, respectively. Figs. 5.e, 5.f, and 5.g show that the reconstruction performance of both methods are degraded as the standard deviation of the Gaussian noise adding to the tracking results increases. In addition, Figs. 5.b, 5.c, and 5.d can be seen that the reconstruction errors in the z-axis direction (i.e. the depth direction) increases more significantly than the errors in the x- and y-axis directions as the standard deviation of the Gaussian noise adding to the tracking results is increased. In other words, the reconstructed 3D face becomes flatter as the tracking errors increase.

Comparing the errors of the 3D reconstruction results shown in Figs. 5.e, 5.f, and 5.g with those shown in Fig. 4.b for the same $\Delta d/d$ ratio of $1/5$, it is observed that the noise factor increases the reconstruction error considerably. In other words, as the noise increases, the tracking accuracy of the feature points decreases, and the 3D reconstruction performance deteriorates accordingly. However, unlike the estimation errors shown in Fig. 5.b, when noise factors are taken into consideration, the error reduces markedly as the number of frames considered in the reconstruction process increases. This is because more feature point information is included in the point matrix W , and hence the overdetermined optimization process of the factorization approach has an improved ability to reduce the effects of noise interference. Nonetheless, the reconstruction error is still greater than that obtained when the tracking results are free of noise.

Fig. 6.a is the comparison of the improved results and the smoothed results. It is clearly seen that the boundaries of the facial features such as the forehead, nose, mouth, and philtrum after the smooth process are much smoother than before ones. Hence, the effectiveness of the proposed smoothing method is confirmed. However,

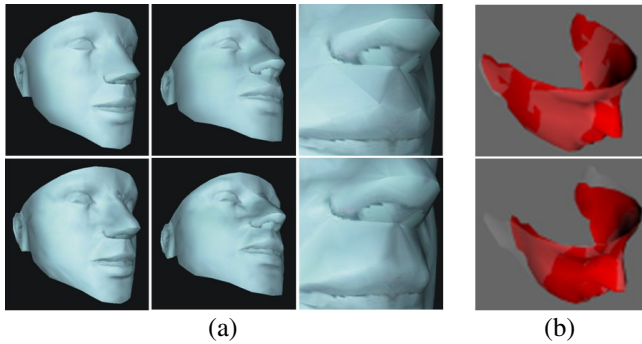


Fig. 6. (a) Upper row is the improved results and the lower row is the smoothed results. (b) Upper one is the result with proper register part (place with small depth change) while lower one is with bad register part (place with larger depth change, such as cheek region).



Fig. 7. 3D facial reconstruction results obtained from three image sequences by using our improved method and smoothing process.

in creating the complete model, an appropriate choice of the overlapping regions is essential. Specifically, the chosen regions should have a small depth Δd , such as the face near the nose, to ensure that the group registration procedure provides the optimum reconstruction results, as shown in Fig. 6.b. And Fig. 7 gives some reconstruction results of real data.

3 Conclusions

This study has developed a straightforward and efficient system for reconstructing 3D faces by using the factorization method and the orthographic projection model to provide a simple approximation to the perspective projection model. The experimental results have shown that the proposed method provides a promising technique for

reconstructing 3D faces. However, some manual refinements are required to compensate for the tracking errors. The missing point problem occurs frequently in 3D reconstruction applications. Accordingly, this study has proposed a solution in which the facial image is divided into three discrete parts in accordance with the rotation angle of the head. This study has also developed a smoothing method based on the linear interpolation of additional 3D vertices to improve the smoothness of the reconstructed facial surface. However, linear interpolation fails to provide an optimum result for regions of the face with large curvatures, e.g. the surface of the nose. In the future, the authors tend to explore the feasibility of using various non-linear curvature functions, e.g. nature spline or B-spline, to improve the smoothing result in the future study.

References

1. Blanz, V., Vetter, T.: A Morphable Model for the Synthesis of 3D Faces, pp. 187–194. ACM Press, New York (1999)
2. Brand, M.: A Direct Method for 3D Factorization of Nonrigid Motion Observed in 2D. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 122–128 (2005)
3. Bregler, C., Hertzmann, A., Biermann, H.: Recovering Non-Rigid 3D Shape from Image Streams. In: Proc. IEEE Conf. on CVPR, pp. 690–696. IEEE Computer Society Press, Los Alamitos (2000)
4. Dwyer, R.A.: A Faster Divide-and-Conquer Algorithm for Constructing Delaunay Triangulation. *Algorithmica* 2, 137–151 (1987)
5. Fua, P.: Regularized Bundle-Adjustment to Model Heads from Image Sequences without Calibration Data. *International Journal of Computer Vision (IJCV)*, 153–171 (2000)
6. Hassanpour, R., Atalay, V.: Delaunay Triangulation Based 3D Human Face Modeling from Uncalibrated Images. In: CVPR Workshop, p. 75 (2004)
7. Horprasert, T., Yacoob, Y., Davis, L.S.: Computing 3-D Head Orientation from a Monocular Image Sequence. In: Proc. IEEE International Conference on Automatic Face and Gesture Recognition, pp. 242–247. IEEE Computer Society Press, Los Alamitos (1996)
8. Jin, H., Soatto, S., Yezzi, A.: Multi-view Stereo Beyond Lambert. *IEEE Conf. CVPR 1*, 171–178 (2003)
9. Lee, Y.C., Terzopoulos, D., Waters, K.: Constructing Physics-Based Facial Models of Individuals. In: Proc. of Graphics Interface, pp. 1–8 (1993)
10. Lucas, B.D., Kanade, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. In: Proc. of DARPA Image Understanding, pp. 121–130 (1981)
11. Mouragnon, E., Dekeyser, F., Sayd, P., Lhuillier, M., Dhome, M.: Real Time Localization and 3D Reconstruction. *IEEE Conf. CVPR 1*, 363–370 (2006)
12. Mouragnon, E., Lhuillier, M., Dhome, M., Dekeyser, F., Sayd, P.: 3D Reconstruction of Complex Structures with Bundle Adjustment: an Incremental Approach. In: Proc. IEEE International Conference on Robotics and Automation, pp. 3055–3061. IEEE Computer Society Press, Los Alamitos (2006)
13. Poelman, C., Kanade, T.: A Paraperspective Factorization Method for Shape and Motion Recovery. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)* 19(3), 206–218 (1997)
14. Shan, Y., Liu, Z., Zhang, Z.: Model-Based Bundle Adjustment with Application to Face Modeling. In: Proc. IEEE International Conference on Computer Vision (ICCV), pp. 644–751. IEEE Computer Society Press, Los Alamitos (2001)

15. Shum, H., Ke, Q., Zhang, Z.: Efficient Bundle Adjustment with Virtual Key Frames: A Hierarchical Approach to Multi-Frame Structure from Motion. *IEEE Conf. CVPR 2*, 538–543 (1999)
16. Soatto, S., Yezzi, A., Jin, H.: Tales of Shape and Radiance in Multiview Stereo. In: *IEEE International Conf. on Computer Vision (ICCV)*, pp. 974–981. IEEE Computer Society Press, Los Alamitos (2003)
17. Tankus, A., Sochen, N., Yeshurun, Y.: A New Perspective [on] Shape-from-Shading. *IEEE ICCV 2*, 862–869 (2003)
18. Tomasi, C., Kanade, T.: Shape and Motion from Image Streams under Orthography: a Factorization Method. *IJCV 9*(2), 137–154 (1992)
19. Torresani, L., Yang, D., Alexander, J., Bregler, C.: Tracking and Modeling Non-Rigid Objects with Rank Constraints. In: *IEEE Conf. CVPR*, pp. 493–500. IEEE Computer Society Press, Los Alamitos (2001)
20. Treuille, A., Hertzmann, A., Seitz, S.: Example-Based Stereo with General BRDFs. In: *European Conference on Computer Vision (ECCV)*, vol. 2, pp. 457–469 (2004)
21. Triggs, B.: Factorization Methods for Projective Structure and Motion. In: *IEEE Conf. CVPR*, pp. 845–851. IEEE Computer Society Press, Los Alamitos (1996)
22. Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzibbon, A.W.: Bundle Adjustment - A Modern Synthesis. In: *Proc. of the International Workshop on Vision Algorithms: Theory and Practice*, pp. 298–372 (1999)
23. Twu, J.T., Lien, J.J.: Estimation of Facial Control-Point Locations. In: *Proc. in the Computer Vision, Graphics, and Image Processing (CVGIP)*, pp.E1–3 (2004)
24. Wang, T.H., Lien, J.J.: Rigid and Non-Rigid Motion Separation Using 3D Model. In: *Proc. in the CVGIP*, pp. A2–5 (2004)
25. Xiao, J., Chai, J.X., Kanade, T.: A Closed-Form Solution to Non-Rigid Shape and Motion Recovery. *ECCV*, pp. 573–587 (2004)
26. Xiao, J., Kanade, T.: Uncalibrated Perspective Reconstruction of Deformable Structures. *ICCV 2*, 1075–1082 (2005)
27. Zhang, Z., Shan, Y.: Incremental Motion Estimation through Modified Bundle Adjustment. *International Conf. on Image Processing 3*, 343–346 (2003)
28. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from Shading: A Survey. *IEEE Trans. on PAMI 21*(8), 690–705 (1999)