Supplemental Materials of Beyond 3DMM Space: Towards Fine-grained 3D Face Reconstruction

Xiangyu Zhu[1,2], Fan Yang[3], Di Huang[4], Chang Yu[1,2], Hao Wang[1], Jianzhu Guo[1,2], Zhen Lei[1,2], and Stan Z. Li[4]

[1] CBSR & NLPR, Institute of Automation, Chinese Academy of Sciences
[2] School of Artificial Intelligence, University of Chinese Academy of Sciences
[3] College of Software, Beihang University
[4] Beijing Advanced Innovation Center for BDBC, Beihang University
[5] School of Engineering, Westlake University
{xiangyu.zhu, jianzhu.guo, zlei, szli, chang.yu}@nlpr.ia.ac.cn,
{fanyang, dhuang}@buaa.edu.cn, haowang7308@gmail.com

1 Registration Results of T-ICP

In this section, we demonstrate several registration results by our Texture constrained ICP (T-ICP) in Fig. 1, which utilizes the face texture, from both the RGB-D image and the face model, as a constraint during closest point matching. The T-ICP not only completes the depth image to a full 3D face but also gets the topology-uniformed 3D geometry. The registration results are further used in out-of-plane pose augmentation in the next section.

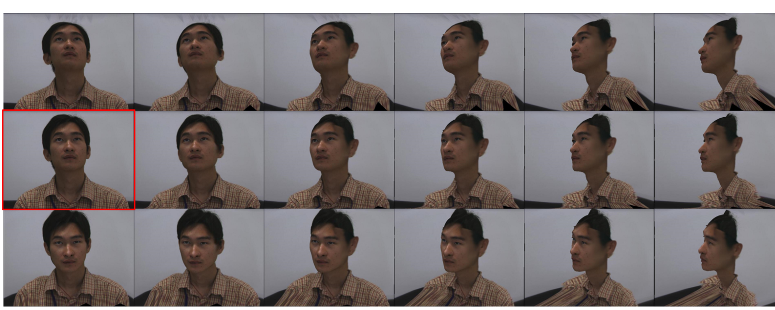
Fig. 1. The registration results of T-ICP from FRGC [4], BP4D [6] and CASIA-3D [3].

2 Results of Out-of-plane Pose Augmentation

In this section, we demonstrate the synthesized samples of our out-of-plane pose augmentation, shown in Fig. 2. Since the augmentation is constrained by real depth information, the synthesized images are close to the ground-truth profile.

(a) Augmentation results of FRGC samples

(b) Augmentation results of BP4D samples



(c) Augmentation results of CASIA-3D samples

Fig. 2. The synthesized samples of our out-of-plane pose augmentation, from (a) FRGC [4], (b) BP4D [6], (c) CASIA-3D [3]. The red contour indicates the base image. We augment each image by enlarging the yaw angle until 90°, and randomly enlarging the pitch angle within 25°.

3 Visualization of the Attention in the Model-view

In this section, we provide visualizations about the attention map from the UV-visibility in Fig. 3, illustrating the learned knowledge from this map. We can see that the weights on the self-occluded regions are close to zero to eliminate the unreliable textures, which also implicitly encode the pose information of the original face image.

Fig. 3. The attention map learned from the UV-visibility.

4 Registering Florence

In this section, we detail how we register the samples of the Florence [2] dataset for evaluation, shown in Fig. 4. In the Florence, we aim to register the "frontal1" 3D mesh of each subject. Firstly, we render the 3D mesh to an image (Fig. 4(a)) and detect 240 landmarks (Fig. 4(b)). The landmarks are manually checked to ensure the accuracy. Secondly, we fit 3DMM to the image with Multi-Features Framework (MFF) [5]. With the dense landmarks as a solid constraint throughout the fitting process, the MFF can always converge to a good initial shape (Fig. 4(c)). Thirdly, we rigid transform the initial shape to the 3D mesh with 10 manually labelled 3D landmarks, providing a close enough initialization for non-rigid registration (Fig. 4(d)). Finally, we perform registration by the Optimal Non-rigid ICP [1] method, where the large missing regions are constrained by the initial shape. The parameters in Optimal Non-rigid ICP including the landmark weight, the stiffness weight and the distance threshold are manually adjusted for each sample to ensure the robustness. Note that, we use ICP rather than our T-ICP during registering, which aims to avoid the over-fitting in topology for fair comparison. Some registration results are shown in Fig. 5.

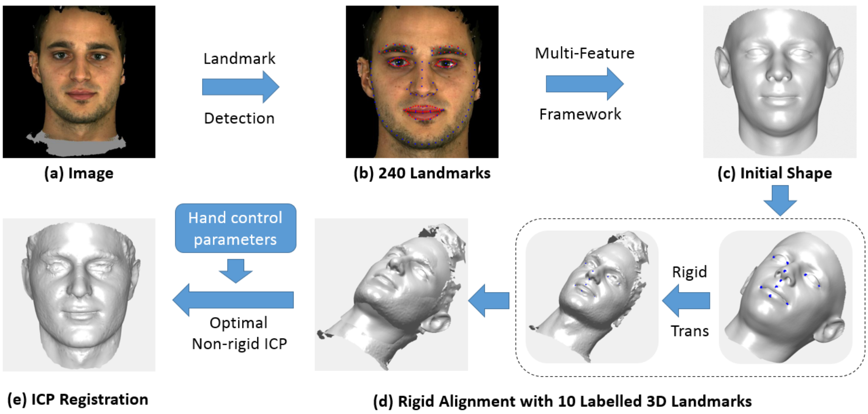


Fig. 4. The process to register Florence [2] samples.

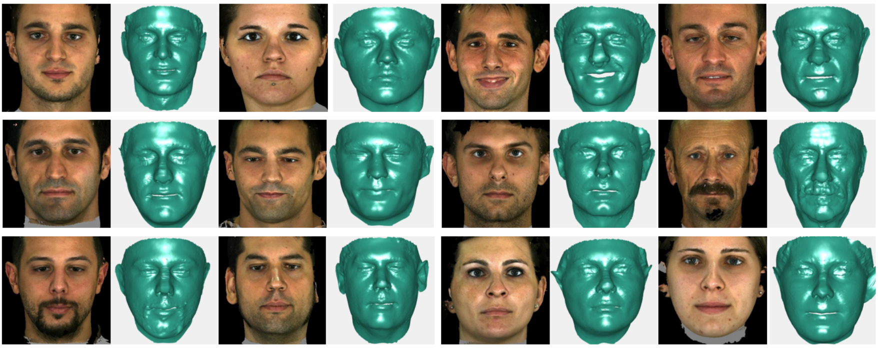


Fig. 5. The registration results of Florence [2].

References

1. Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid icp algorithms for surface registration. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8. IEEE (2007)
2. Bagdanov, A.D., Del Bimbo, A., Masi, I.: The florence 2d/3d hybrid face dataset. In: Proceedings of the 2011 joint ACM workshop on Human gesture and behavior understanding. pp. 79–80. ACM (2011)
3. Li, S.: Casia 3d face database - center for biometrics and security research (2004)
4. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). vol. 1, pp. 947–954. IEEE (2005)
5. Romdhani, S., Vetter, T.: Estimating 3D shape and texture using pixel intensity, edges, specular highlights, texture constraints and a prior. In: Computer Vision and Pattern Recognition (CVPR), 2005 IEEE Conference on. vol. 2, pp. 986–993. IEEE (2005)
6. Zhang, X., Yin, L., Cohn, J.F., Canavan, S., Reale, M., Horowitz, A., Liu, P., Girard, J.M.: Bp4d-spontaneous: a high-resolution spontaneous 3d dynamic facial expression database. Image and Vision Computing **32**(10), 692–706 (2014)