ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023

Facial Photo-Guided Head Anatomy Modeling Based on Deep Learning and 2D/3D Shape Prior Model Registration

Meng Wang [1, a], Hongkai Wang [2,3, b]

[1] School of Biomedical Engineering, Dalian University of Technology, Dalian 116024, China;

[2] School of Biomedical Engineering, Dalian University of Technology, Dalian 116024, China;

[3] Liaoning Key Laboratory of Integrated Circuit and Biomedical Electronic System, Dalian University of Technology, Dalian 116024, China.

[a] wmm123456@mail.dlut.edu.cn, [b] wang.hongkai@dlut.edu.cn

Abstract. Three-dimensional (3D) models of individualized head anatomy are often used in the preoperative planning of plastic surgery. Most current methods use Computed Tomography (CT) to acquire the head anatomy, but CT scan exposes the patients to X-ray radiation. Alternatively, artificial intelligence-based estimation of 3D head anatomy from the facial photo is a safer and more efficient choice, especially for scenarios that do not require high modeling accuracy of internal structures (e.g., doctor-patient communication and surgery procedure demonstration). We develop a method for constructing a personalized 3D model of complete head anatomy (including skin and internal structures) from the front and side view photos. We first detect key facial landmarks from the photos with a convolution neural network (CNN), then perform 2D/3D registration to morph a statistical shape model (SSM) of complete head anatomy to match the dual-view photos. Our method uses deep learning to achieve accurate facial landmark detection and employ the anatomy shape prior to yielding reasonable internal structure estimation. We evaluated the method based on 12 subjects (including 7 males and 5 females). The facial surface reconstruction error was assessed using the 3D surface scan of the subjects as the ground truth. Our method's root-mean-square-error (RMSE) was 3.60±0.49 mm, which was 1.06±0.39 mm lower than the state-of-the-art (SOTA) CNN-based face reconstruction method. Our method also predicts all the internal head structures (bones, muscles, vessels, nerve fibers, fat, glands, and brain structures) which are not provided by the SOTA method. To the best of our knowledge, this is the first study modeling complete head anatomy from facial photos toward clinical applications.

Keywords: Facial photos; Head anatomy modeling; 2D/3D registration; Statistical shape model.

1. Introduction

Three-dimensional (3D) face model reconstruction has been extensively used in the game development, film industry, clinical plastic surgery, etc. After decades of development, current deep learning methods can successfully reconstruct personalized 3D face models based on a single front-view photo [11], [14]-[15]. To make up for the lack of depth information of the single view, multi-view photo reconstruction methods are also proposed [16]-[18] with improved reconstruction accuracy in 3D space.

Despite the progress in photo-based face model reconstruction, current methods only achieve visually similar modeling of the facial outer look. The accuracy of face models reconstructed from photos cannot meet the requirements of clinical surgery, and quantitative assessment of face reconstruction accuracy is also lacking in the current studies. Moreover, besides the reconstruction of the face surface, successful surgery planning requires modeling internal anatomical structures including multiple head anatomical structures, including bones, muscles, vessels, nerve fibers, fat, glands, and brain structures.

Tomographic imaging devices are usually used to acquire both external face shape and internal structure anatomy, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). However, tomographic imaging is tedious and expensive, patients may not accept this option at the preoperative communication stage. In contrast, facial photos are much more efficient to acquire. Therefore, the estimation of a 3D personalized head anatomy model based on facial photos or

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023
optical surface scanning can significantly simplify the preoperative image acquisition procedure and can be useful for scenarios that do not require high modeling accuracies, such as surgery procedure demonstration and approximate surgery outcome simulation.

To address the above concerns, this paper explores the solution of complete head anatomy modeling based on non-tomographic imaging. We select facial photos as the first choice since optical surface scanning is not suitable for remote doctor-patient communication. We detect the facial landmarks with a deep-learning method and then perform 2D/3D registration to match a deformable statistical shape model (SSM) of full head anatomy to the front and side view photos. Since the SSM was previously constructed based on segmented volumetric CT images, it contains the shape prior to full head anatomy and provides a reasonable estimation of the internal structures based on the external face shape. Our method not only reconstructs a visually similar surface model with the color texture from the personal photo; but also provides 3D models of the internal structures to assist preoperative planning of plastic surgery.

2. Related Work

With the rapid development of deep learning technology, many monocular face reconstruction methods have been proposed. Various network models were designed to regress the parameters of 3D face geometry models, and supervise them according to synthetic data [1]-[3], pre-calculated 3D Morphable Model (3DMM) fitting [4], RGB images [5]-[9] and feature tags [10]. To deal with complex facial geometry more flexibly, the method [12]-[13] regresses geometric residuals to restore fine-scale details. So far, most deep learning methods have used CNN to reconstruct the shape details of the face from 2D photos. Guo et al. [14] proposed a coarse-to-fine framework to reconstruct fine 3D faces from monocular videos and single images in real-time.

Chen et al. [15] proposed a conditional generation countermeasure network (CGAN) based deep detail network (DFDN) which uses a combination of geometry and appearance loss functions. However, this method cannot process low-resolution images, which limits its application.

A series method based on multi-view reconstruction was recently proposed to take advantage of multi-view photos for more accurate face reconstruction [16]-[17]. For example, Wu et al. [16] proposed a multi-view image 3D modeling method based on end-to-end trainable CNN, but their method only focuses on shape modeling without color texture details. Later, Bai et al. [18] introduced traditional multi-view geometry into CNN-based face reconstruction. By imposing multi-view geometric constraints to solve the 3D reconstruction problem, they effectively capture shape details and improve the generalization ability of unknown data. Photometric stereo [22] is also used in [19]-[21] for facial reconstruction, but this method can only have a good reconstruction effect for images with a face deflection angle within ± 30 degrees away from the frontal view. Therefore, this method is not suitable for side profile photos.

For the modeling of internal head anatomy based on external surface shape, Xiao et al. [23] studied the estimation of bone shape based on facial reconstruction. They proposed a facial bone shape estimation framework through 3D facial reconstruction and deformable shape model, using pre-trauma conventional portraits and post-trauma computed tomography (CT) of the head. However, in his research, only facial bones were estimated, and other important anatomical structures, such as nerves and blood vessels, were missing.

At present, there are several common problems in the current research. First, the 90 degrees profile photo is rarely used, leading to the inaccurate reconstruction of the side profile. Second, most methods only focus on the external surface or bony structure modeling without complete internal anatomy, and they reconstruct the 3D shape without a color texture map, limiting their value in preoperative doctor-patient communications. Last, the current face modeling accuracy cannot meet the precision requirements of plastic surgery, barely any study has evaluated such accuracy quantitatively.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023

Considering the above problems, this paper proposes a method to register a head anatomy SSM to both frontal and lateral view photos. The registration result provides a personalized 3D model of complete head anatomy with face texture mapped from the color photos, and the face modeling accuracy is also evaluated for the reference of surgery planning.

3. Method

Fig. 1 shows the workflow of the proposed algorithm. Our method composes of three steps. The first step is to input dual-view photos to the CNN to detect 2D facial landmarks. The second step is to initialize the SSM of head anatomy according to the facial landmarks and maps the color texture from the photos to the SSM surface. The third step conducts refined matching of the SSM to the dual-view photos via 2D/3D registration, in which the detected landmarks guide accurate matching of key facial features, and the intensity-based registration of the SSM projection images leads to the accurate matching of the texture details.

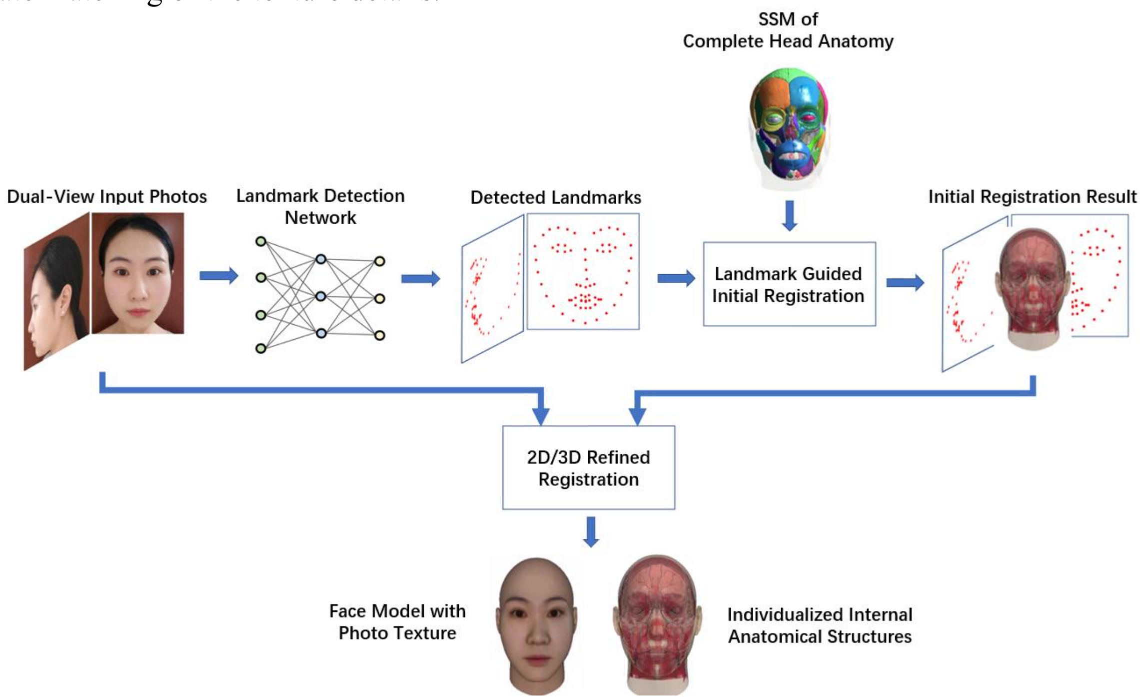
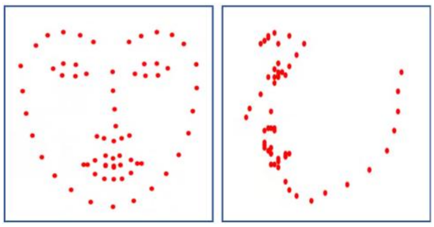


Fig. 1 Workflow of the proposed method for head anatomical structure modeling.

3.1 Landmark Detection Network

Our method detects 68 landmarks from the frontal view photo and 68 landmarks from the side view photo, using a network structure similar to [18] which solved the problem of 3D face reconstruction from multi-view images with different expressions. We use the 2D landmarks to determine a bounding box for face region clipping. Fig. 2 shows the definition of front and side face landmarks.



(a) (b)

Fig. 2 The front and side face landmarks. (a) The front face landmarks. (b) The side face landmarks.

Of the 68 landmarks obtained, eleven face landmarks can be customized, including not only the main feature information parts such as facial pupil and nose wing; but also, the face contours

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023
information parts such as the zygomatic arch and mandibular angle, which can effectively extract the face feature information. By using these face landmarks, we can carry out the next step of model initialization.

In addition, the neural network also estimates the model projection angle which can be used in subsequent 2D/3D registration.

3.2 Section Headings

The SSM method [24] was used to construct the atlas based on 65 computer tomography (CT) images (including 46 male and 19 female images). The deformable head atlases (DHAs) include 218 sub-structures of facial muscles, skull bones, vessels, fat, nerves, glands, and brain structures.

The SSM is expressed as a linear combination of average shape plus different shape variation patterns:

$$X = \overline{X} + \sum_{i=1}^{n} a_i \Phi_i \tag{1}$$

where X is an instance of the shape generated by the model. X represents the shape vector $(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_k, y_k, z_k)^T$ containing the 3D coordinates of k mesh vertices. $\overline{X}$ is the average shape vector of all training objects. $\{a_i\}, i = 1, \ldots, n$ are shape parameters that are used as weights for various modes. Different values of $\{a_i\}$ generate different instances of the model. The continuous change of $\{a_i\}$ value causes the continuous deformation of the head anatomical structure to simulate the anatomical changes between subjects.

To use the projection angle obtained in the first step, we need to match the SSM to the coordinate system where the neural network reconstructs the model. The method adopted is to mark 57 3D face landmarks on the corresponding positions of the neural network average model and the average SSM, and apply them to the SSM by calculating the similarity transformation matrix from the landmark set of SSM to the landmark set of the neural network model, to match the SSM to the target coordinate system.

The complete set of head anatomical structures in the SSM [24] is obtained through soft tissue mapping in the standard digital human template. The mapping process uses nonlinear thin plate spline (TPS) transformation, and the control points are the surface vertices on the skin and bones.

The deformation field of the external surface of the head is interpolated to the internal anatomical structure using 3D space interpolation to realize the personalization of the internal anatomical structure and achieve the accurate coincidence between the internal anatomical structure and the external surface model.

3.3 2D/3D Registration

The 2D/3D registration algorithm consists of two steps. The first step is to project the SSM mean shape into both views and then perform 2D registration between the projected model image and the facial photo, obtaining the deformation vector field in the 2D projection space. The second step is to back-project the 2D deformation vector to 3D space and to obtain the deformation of the 3D SSM. Fig. 3 shows the principle of the back-projection:

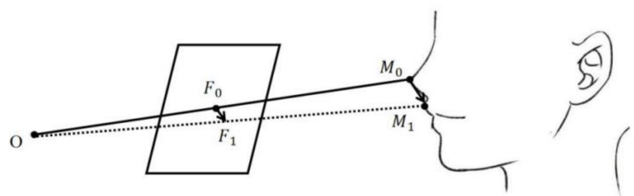


Fig. 3 Schematic diagram of the similar triangle.

where, O is the position of the viewpoint, F_0 is the position of the 2D point before deformation in the world coordinates system on the model photo, F_1 is the position of the 2D point after deformation in the world coordinates system on the model photo, M_0 is the 3D point

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023
corresponding to the 2D pixel point of the model photo before deformation, and M_1 is the 3D point corresponding to the two-dimensional pixel point of the model photo after deformation.

Let the distance between O and F_0 be m, the distance between O and M_0 be n, and the distance between F_0 and F_1 be p, then the distance between M_0 and M_1 can be calculated according to the side length relationship of similar triangles, as shown in (2).

$$|M_0M_1| = \frac{n\times p}{m} \tag{2}$$

Let the coordinates of F_0 be (X_0, Y_0, Z_0), the coordinates of F_1 are (X_1, Y_1, Z_1), and the unit direction vector of $\overrightarrow{F_0F_1}$ is shown in (3).

$$\vec{p} = \frac{\overrightarrow{F_0F_1}}{|F_0F_1|} = \frac{(X_1-X_0, Y_1-Y_0, Z_1-Z_0)}{\sqrt{(X_1-X_0)^2+(Y_1-Y_0)^2+(Z_1-Z_0)^2}} \tag{3}$$

According to the distance between M_0 and M_1 in (2) and the unit direction vector of $\overrightarrow{F_0F_1}$ in (3), the coordinates of M_1 can be obtained. Because $\overrightarrow{M_0M_1}$ and $\overrightarrow{F_0F_1}$ have the same unit direction vector, M_1 coordinates are shown in (4).

$$M_1 = M_0 + \vec{p} \times |M_0M_1| \tag{4}$$

Using the 2D/3D registration algorithm, the front photo is registered to the initial model obtained in the second step to obtain model F, and then the lateral photo is registered to model F to obtain model L. The final reconstructed model is obtained by smoothly deforming model L onto the initial model.

Because the neural network accurately predicts the landmark recognition in the front photo, the thin-plate splines (TPS) transform is adopted for 2D/2D matching of front photos. The TPS transform formula is defined as follows,

$$T_\mu(x) = x + Ax + t + \sum_{x_k^{fix}} c_k G(x - x_k^{fix}) \tag{5}$$

where $G(r)$ is the basic function and c_k are the coefficients corresponding to each landmark. Calculate the coefficients c_k and the elements of A and t according to the landmark displacements $d_k = x_k^{mov} - x_k^{fix}$.

Based on the front reconstruction results, lateral photo registration is carried out. Because the neural network lacks side view training images and does not perform robust enough for recognizing the side view landmarks, the lateral photo 2D/2D registration uses B-splines transform. B-splines registration is a nonlinear deformation registration based on pixel intensity, which serves as a good alternative when registration based on landmarks is not feasible.

4. Result and Aanlysis

If you follow the “checklist” your paper will conform to the requirements of the publisher and facilitate a problem-free publication process.

Our test set includes 12 young subjects (7 males and 5 females) with an average age of 25 years. The average modeling time of our method is 51.56 s per subject. To evaluate the external face shape reconstruction error, we use a 3D optical scanner to acquire the ground truth surface point cloud of each subject’s face. We use average surface distance root mean square error (RMSE) as an error indicator. The calculation formula of RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(Y_i - f(x_i))^2} \tag{6}$$

where Y_i is the accurate value and $f(x_i)$ is the estimated value.

We compare the facial reconstruction error of our method with one of the state-of-the- (SOTA) methods also art uses CNN to reconstruct facial models from multi-view photos but lacks the refined 2D/3D registration step [18] combined the traditional optimization based on multi-view geometry with modern neural network to design a new non-rigid multi-view stereo (NRMVS) optimization framework, which solved the 3D face reconstruction problem of multi-view images with different expressions. Input the test photos to [18] and our method respectively to calculate the

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023

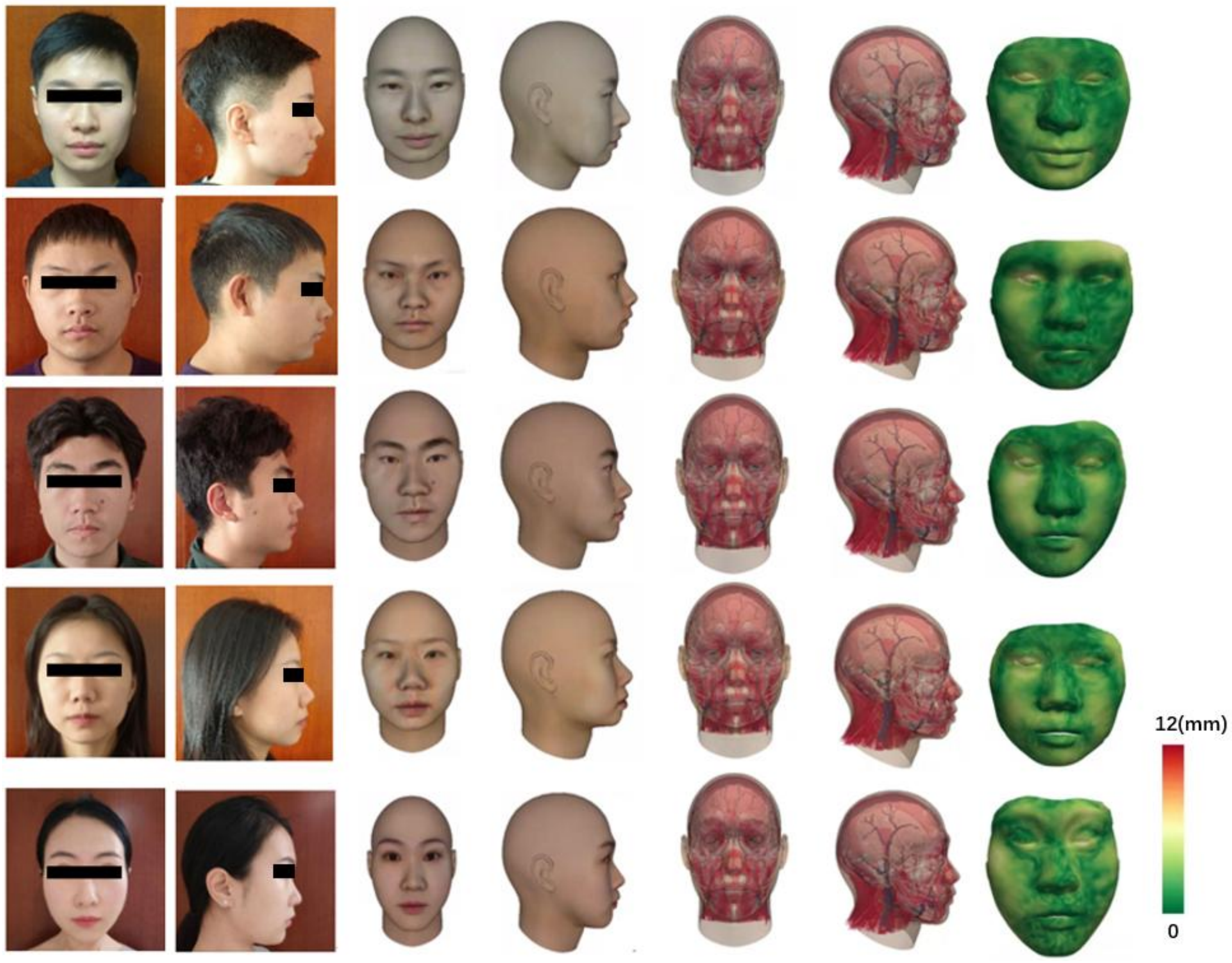
RMSE between the reconstructed model and the ground truth. We compared the mean, standard deviation (STD), maximum and minimum Table 1 shows the comparison results.

Table 1. RMSE comparison (unit: mm)

Method	mean	Std.	max	min
Ours	3.60	0.48	4.51	2.80
Bai et al. [18]	4.66	0.78	6.42	3.24

It can be seen from Table 1 that the model surface error reconstructed by our method is 3.60 ± 0.49mm, and the model surface error reconstructed by [18] is 4.66 ± 0.78mm. The facial reconstruction error of our method is lower than that of [18], which proves that the method combining the 2D/3D registration method and neural network is effective and stable. In addition, it should be noted that possible facial expressions difference between the photo and the surface point cloud scan also affects the error calculation. Therefore, we may expect when lower reconstruction error if there is no expression difference in ideal cases.

Fig. 4 shows partial reconstruction results and error heat maps. From the reconstruction results, we can see that our method can well restore the real 3D head model from the photos. It can be seen from the heat maps that our model has high accuracy in facial reconstruction.



(a) (b) (c) (d) (e) (f) (g)

Fig. 4 Reconstruction results and error thermodynamic diagram. (a) Facial front photos. (b) Facial side photos. (c) Front display of skin model. (d) Side display of skin model. (e) Front display with head anatomical structure model. (f) Side display with head anatomical structure model. (g) Error heat map.

From the semi-transparent rendering of the complete head anatomy, one can see that the personalized modeling of all the anatomical structures can be achieved with our method.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023

We also conducted a qualitative comparison between our method and several popular SOTA deep learning methods for photo-based face modeling from the aspects of whether the reconstructed model includes complete head anatomy and whether it has colored face texture. Table 2 shows the comparison of qualitative between our method and other methods. [16] uses the novel self-supervised view alignment loss to establish a dense correspondence between different views, and merge the multi-view geometric constraints into the network. [17] uses self-supervised training based on a multi-frame video. [18] optimizes 3D face shape by explicitly enhancing the consistency of multi-view appearance It is obvious that our model is the only one with both complete anatomy modeling and face texture mapping.

Table 2. Comparison of qualitative

Method	complete head anatomy	colored face texture
MVF-Net [16]	×	×
FML [17]	×	√
NRMVS [18]	×	×
Ours	√	√

5. Conclusion

We combine deep learning-based facial landmark detection and 2D/3D SSM registration to reconstruct an individualized 3D model of complete head anatomy from a front photo and a side photo. The experimental results show that our method can improve the accuracy of model reconstruction compared with the SOTA deep learning methods. To the best of our knowledge, this is the first study providing both a textured facial modal and a complete internal anatomical model based on facial photos. Our future research will continue to assess the accuracy of internal structure modeling using head CT images as the ground truth.

Acknowledgment

This work was supported in part by the National Key Research and Development Program No. 2020YFB1711500, 2020YFB1711501, and 2020YFB1711503, the general program of the National Natural Science Fund of China (No. 81971693, 61971445), the funding of Dalian Engineering Research Center for Artificial Intelligence in Medical Imaging, Hainan Province Key Research and Development Plan ZDYF2021SHFZ244, the Fundamental Research Funds for the Central Universities (No. DUT22YG229), the funding of Liaoning Key Lab of IC & BME System and Dalian Engineering Research Center for Artificial Intelligence in Medical Imaging.

References

[1] K. Genova, F. Cole, A. Maschinot, A. Sarna, D. Vlasic, and W. T. Freeman. Unsupervised training for 3D morphable model regression. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol. 1,2, 2018, pp. 8377–8386, doi:10.1109/CVPR.2018.00874.

[2] E. Richardson, M. Sela, and R. Kimmel, “3D face reconstruction by learning from synthetic data,” 2016 Fourth International Conference on 3D Vision (3DV), 2016, pp. 460-469, doi: 10.1109/3DV.2016.56.

[3] M. Sela, E. Richardson, and R. Kimmel, “Unrestricted facial geometry reconstruction using image-to-image translation,” 2017 IEEE International Conference on Computer Vision (ICCV), vol. 2, 2017, pp. 1585-1594, doi: 10.1109/ICCV.2017.175.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023
[4] A. T. Tran, T. Hassner, I. Masi and G. Medioni, “Regressing robust and discriminative 3D morphable models with a very deep neural network,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1,2, 2017, pp. 1493-1502, doi: 10.1109/CVPR.2017.163.

[5] Y. Deng, J. Yang, S. Xu, D. Chen, Y. Jia, and X. Tong, “Accurate 3D face reconstruction with weakly-supervised learning: from single image to image set,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 285-295, doi: 10.1109/CVPRW.2019.00038.

[6] A. Tewari, M. Zollhfer, P. Garrido, F. Bernard, H. Kim, and C. Theobalt, “Self-supervised multi-level face model learning for monocular reconstruction at over 250 Hz,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol. 2,4,6,7, 2018, pp. 2549–2559, doi: 10.1109/CVPR.2018.00270.

[7] A. Tewari, M. Zollhfer, H. Kim, P. Garrido, F. Bernard, and C. Theobalt, “MoFA: Model-based deep convolutional face autoencoder for unsupervised monocular reconstruction,” 2017 IEEE International Conference on Computer Vision (ICCV), vol. 2, 2017, pp. 3735-3744, doi: 10.1109/ICCV.2017.401.

[8] L. Tran and X. Liu, “Nonlinear 3D face morphable model,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7346-7355, doi: 10.1109/CVPR.2018.00767.

[9] L. Tran and X. Liu, “On learning 3D face morphable model from in-the-wild images,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, pp. 157-171, 1 Jan. 2021, doi: 10.1109/TPAMI.2019.2927975.

[10] S. Sanyal, T. Bolkart, H. Feng and M. J. Black, “Learning to regress 3D face shape and expression from an image without 3D supervision,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, 2019, pp. 7755-7764, doi: 10.1109/CVPR.2019.00795.

[11] A. S. Jackson, A. Bulat, V. Argyriou and G. Tzimiropoulos, “Large pose 3D face reconstruction from a single image via direct volumetric CNN regression,” 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 1031-1039, doi: 10.1109/ICCV.2017.117.

[12] L. Tran, F. Liu, and X. Liu, “Towards high-fidelity nonlinear 3D face morphable model,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, 2019, pp. 1126-1135, doi: 10.1109/CVPR.2019.00122.

[13] A. T. Tran, T. Hassner, I. Masi, E. Paz, Y. Nirkin, and G. Medioni, “Extreme 3D face reconstruction: seeing through occlusions,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, vol. 2, 2018, pp. 3935-3944, doi: 10.1109/CVPR.2018.00414.

[14] Y. Guo, j. Zhang, J. Cai, B. Jiang, and J. Zheng, “CNN-based real-time dense face reconstruction with inverse-rendered photo-realistic face images,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 41, no. 6, pp. 1294-1307, 1 June 2019, doi: 10.1109/TPAMI.2018.2837742.

[15] A. Chen, Z. Chen, G. Zhang, K. Mitchell, and J. Yu, “Photo-realistic facial details synthesis from single image,” 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 9428-9438, doi: 10.1109/ICCV.2019.00952.

[16] F. Wu, L. Bao, Y. Chen, Y. Ling, Y. Song, S. Li et al, “MVF-Net: Multi-view 3D face morphable model regression,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 959-968, doi: 10.1109/CVPR.2019.00105.

[17] A. Tewari, F. Bernard, P. Garrido, G. Bharaj and C. Theobalt, “FML: Face model learning from videos,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 10804-10814, doi: 10.1109/CVPR.2019.01107.

[18] Z. Bai, Z. Cui, J. A. Rahim, X. Liu, and P. Tan, “Deep facial non-rigid multi-view stereo,” 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5849-5859, doi: 10.1109/CVPR42600.2020.00589.

[19] A. S. Georghiades, “Recovering 3-D shape and reflectance from a small number of photographs,” Proceedings of the 14th Eurographics Workshop on Rendering Techniques, Leuven, Belgium, June 25-27, 2003, pp. 230-240, doi:10.2312/EGWR/EGWR03/230-240.

[20] M. Klaudiny and A. Hilton, “High-detail 3D capture and non-sequential alignment of facial performance,” 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, 2012, pp. 17-24, doi: 10.1109/3DIMPVT.2012.67.

ISSN:2790-1688 DOI: 10.56028/aetr.4.1.260.2023
[21] J. Roth, Y. Tong, and X. Liu, "Adaptive 3D face reconstruction from unconstrained photo collections," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 11, pp. 2127-2141, 1 Nov. 2017, doi: 10.1109/TPAMI.2016.2636829.

[22] R. J. Woodham, "Photometric Stereo: A reflectance map technique for determining surface orientation from image intensity," Proc of Spie, vol. 155, 1979, pp. 136–143, doi:10.1117/12.956740.

[23] D. Xiao, C. Lian, L. Wang, H. Deng, H. Lin and K. Thung et al., "Estimating reference shape model for personalized surgical reconstruction of craniomaxillofacial defects," in IEEE Transactions on Biomedical Engineering, vol. 68, no. 2, pp. 362-373, Feb. 2021, doi: 10.1109/TBME.2020.2990586.

[24] Z. Chen, T. Qiu, L. Huo, L. Yu, H. Shi and Y. Zhang et al., "Deformable head atlas of Chinese adults incorporating inter-subject anatomical variations," in IEEE Access, vol. 6, pp. 51392-51400, 2018, doi: 10.1109/ACCESS.2018.2869331.