Geometry Sampling for 3D Face Generation via DCGAN

Guoliang Luo
East China Jiaotong University and
Jiangxi Booway New Technology Co. Ltd
Nanchang, China
luoguoliang@ecjtu.edu.cn

Xin Zhao
East China Jiaotong University
Nanchang, China
2018578081203001@ecjtu.edu.cn

Yang Tong
East China Jiaotong University
Nanchang, China
tongyang@ecjtu.edu.cn

Qiang Chen
East China Jiaotong University
Nanchang, China
qiangchen@ecjtu.edu.cn

Zhiliang Zhu
East China Jiaotong University
Nanchang, China
rj_zzl@ecjtu.edu.cn

Haopeng Lei
Jiangxi Normal University
Nanchang, China
leihaopeng@163.com

Juncong Lin
Xiamen University
Xiamen, China
jclin@xmu.edu.cn

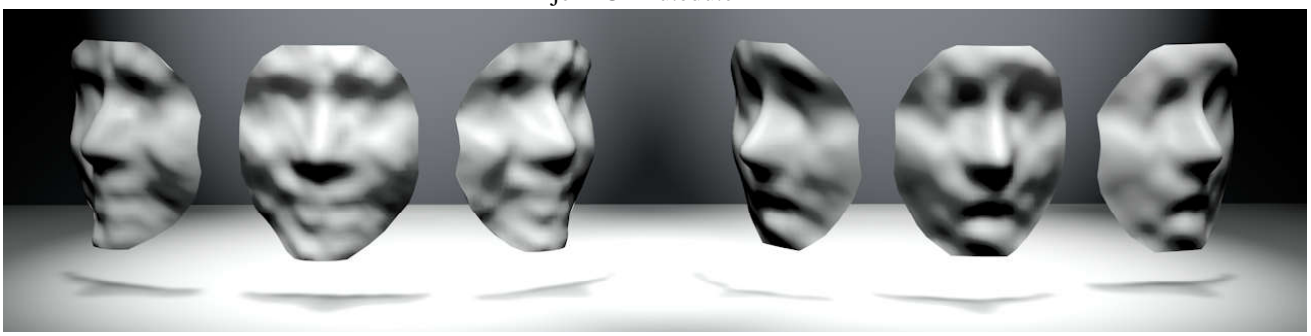

Fig.1. The generated 3D faces with 'Happy' and 'Surprise' expressions by using our approach.

***Abstract*—Despite numerous progresses in the past decades, 3D shape acquisition techniques remain a threshold for various 3D face based applications. Moreover, advanced 2D data generative models based on the deep networks may not be directly applicable for 3D objects. In this work, we propose a geometry sampling approach to bridge the gap between unstructured 3D face models and the powerful deep networks towards an unsupervised 3D face generative model. Specifically, we devise a geometry sampling approach to obtain a structured representation of 3D faces, which enable us to adapt the 3D faces to the Deep Convolution Generative Adversarial Network (DCGAN) for 3D face generation. We have demonstrated the effectiveness of our generative model by producing a large variety of 3D faces with different facial expressions.**

***Index Terms*—Geometry sampling, 3D face generation, DCGAN**

I. Introduction

With the rapid advancements of display equipments and the growing network bandwidths, 3D data is becoming another popular media due to its fullness of realistic information. On the other hand, nowadays 3D shape acquisitions either rely on expensive equipments or require the expertized knowledge and skills. For this reason, 3D data acquisition techniques remain a threshold for the broaden applications of 3D data.

3D shape synthesis techniques provide an alternative means for 3D data generation, which can be useful for various purposes without copyright infringement [1, 2]. Specifically, the 3D synthetic faces can be broadly used in the video games, the beauty applications, virtual reality and so on. However, although various synthetic approaches for 2D images and video have been intensively studied previously, it remains a challenging task for 3D shape synthesis because: 1) A 3D shape is a set of 3D vertices in the space, which is an unstructured data, compared to 2D images that can be stored in a structured 2D matrix. 2) 3D shapes are sensitive to noises, while one or even a few outlier pixels may not be visually noticeable in a 2D image.

The contributions of this work are two folds:

- First, we propose a new geometry sampling approach, which enables us to generate 3D shapes with deep networks. Specifically, with our geometry sampling approach, we can represent the 3D faces into a structured matrix-like structure.

978-1-7281-6026-2/20/$31.00 ©2020 IEEE

- Second, we present a straight-forward unsupervised 3D face generative model, which does not require any pre-processing steps such as the extraction of facial feature points or pre-computing the correspondences.

The remainder of the paper is organized as follows. We first briefly review the state-of-the-art in Section II. Then, we present our geometry 3D face generation model in Section III, followed with the results and the discussions in Section IV. Finally, we conclude the work in Section V.

II. Related Work

Although the Deep Neural Network (DNN) based data-driven synthetic methods have been intensively studied in the computer vision field, it remains a challenging topic for data-driven 3D shape generation. In this section, we first briefly review the 3D shape reconstruction techniques in the computer vision domain. Then, we summarize the recent related works on the 3D shape representations and data-driven 3D shape modeling, respectively.

3D shape reconstruction is important to theuser interaction, auto piloting, etc. Izadi et al. present a GPU-based pipeline to achieve the 3D pose reconstruction in real-time by using the low-cost handheld scanning depth camera, which is demonstrated with the object segmentation and user interaction [3]. In order to facilitate the learning based algorithm for 3D shape reconstruction, Song et al. present a large-scale benchmark with 3D annotations and 3D evaluation metrics of RGB-D images for the learning based 3D scene understanding [4]. Similarly, Handa et al. present a dataset of RGB-D sequences with perfect ground truth pose and the corresponding ground truth surface model that enables of quantitatively evaluating the final map or surface reconstruction accuracy [5]. The work in [6] take one input image as a guide to "mold" a single reference model to reach a reconstruction of the sought 3D shape, based on the assumption of Lambertian reflectance and harmonic representations of lighting. To improve the efficiency of the learning methods, Zhu et al. present an actor-critic model for the fast-converge learning that can be applied to target-driven visual navigation [7]. In [8], the Recurrent Reconstruction Neural network based model learns a mapping from images of objects to their underlying 3D shapes from a large collection of synthetic data. Furthermore, Fan et al. present a learning paradigm with a conditional shape sampler that is capable of predicting multiple plausible 3D point clouds from an input image [9]. Recently, Garrido et al. present a coarse-to-fine scheme for 3D face rigging from a monocular video [10]. They first compute a coarse-scale face reconstruction with a novel variational fitting approach. Then, the fine-scale skin detail, such as wrinkles, are obtained from video via shading-based refinement. By following the similar coarse-to-fine scheme, Jiang et al. achieve the 3D face reconstruction with a single image with the fine details obtained by using the photometric consistency constraints and the shape-from-shading method.

Shape representations are fundamental to 3D models, as the 3D model's primitives are un-structured compared to the pixels of a 2D image in a matrix form. In computer graphics community, researchers have proposed a variety of classical feature descriptors for 3D shapes. Recently Soltanpour et al. summarized various local feature descriptors for 3D face shape recognition [11], including the Gaussian curvatures [12], the radial curves [13], and so on. A number of parameterization methods have been proposed to flatten 3D shapes into 2D, including the Mobiüs method [14], woven mesh fitting method [15], and the geometry image/video [16, 17]. However, most of the existing 3D shape representations cannot be directly applicable to data-driven 3D shape synthesis because they are either un-structured [11], or irreversible of the geometry properties [14, 15], and the geometry video requires extra operations on the eyes/mouth removal [17].

Deep learning based 3D shape synthesis is becoming popular in the computer graphic community in recent years [1, 2]. For example, Li et al. proposed a learning based facial expression transfer method to drive an example model with the learned expression [18], and Chen et al. apply the convolutional network for the synthesis of 3D cloth wrinkles [19]. In the computer vision community, there are two widely studied DNN models for 2D image and video synthesis: the Variational Auto-encoder (VAE) [20, 21] and the Generative Adversarial Network (GAN) [22]. Between them, the GAN model contains a Generative model and a Discriminative model. The Generative model keeps updating the generated data until the discriminative model cannot distinguish the difference between the generated data and the original training data. In [23], Radford et al. successfully integrate the GAN model with the convolutional network for a Deep Convolutional GAN (DCGAN) model, which significantly improves the potential of GAN for image synthesis. In [24], Jean et al. incorporated the shape geometry properties to enhance the performance of DCGAN for 2D object generation. While both of VAE and GAN techniques apply the learned features to an existing 3D shape [18, 19], in this work we propose a DCGAN based approach to directly generate a synthetic 3D face.

III. Geometry Sampling and 3D Face Generation

Figure 2 shows the overview of the proposed 3D face generation approach. In this section, we first present a geometry sampling approach for 3D face models, which outputs a structured representation for an input 3D face. See details in Section III-A. Based on the structured representation, we further present the adaption of DCGAN for 3D face generation in Section III-B.

A. Geometry Sampling of 3D Faces

Given a 3D face $\mathscr{F}$, the objective of geometry sampling is to obtain a matrix-like, structured representation for any input 3D face. Our geometry sampling approach is based on two 3D face feature curves, i.e., the iso-geodesic curves and the radial curves [25]. Figure 3 shows the pipeline of our geometry sampling approach, which can be described in details as follows:

1) *Iso-geodesic curves*. Given the detected nose tip **O**, an iso-geodesic curve contains a sequence of the vertices on

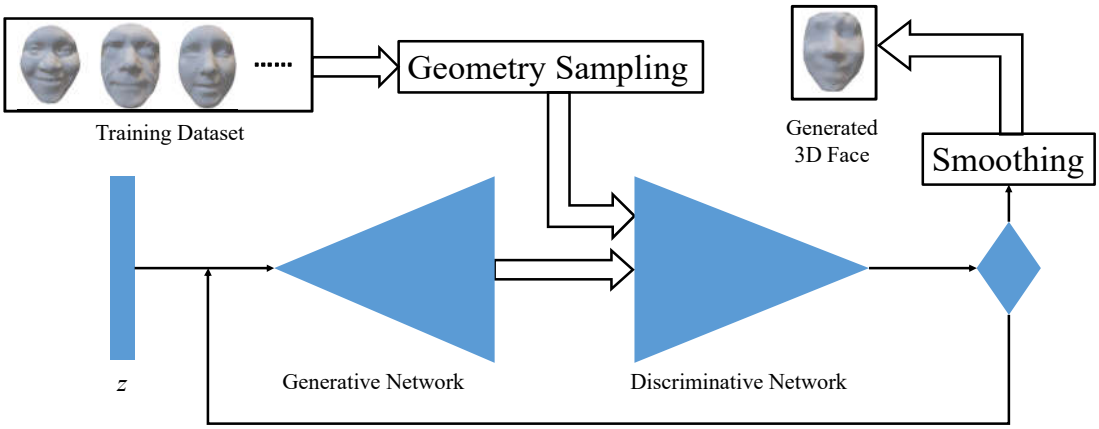

Fig. 2. Overview of the proposed 3D face generation approach. First, we adapt the 3D point cloud faces for the deep neural network by using the proposed geometry sampling approach. With the random input z, the Generative and Discriminative adversarial networks repetitively update the generated 3D face, until it is recognized as 'real'. The generated 3D point cloud model is further smoothed for the final 3D face mesh. Note that both the Generative and Discriminative Networks comprise a fully-connected layer and 3 transpose-convolution layers, but in the reverse order. Detailed specifications are described in Section III-B. Note that the double-headed arrows denote the facial data flow.

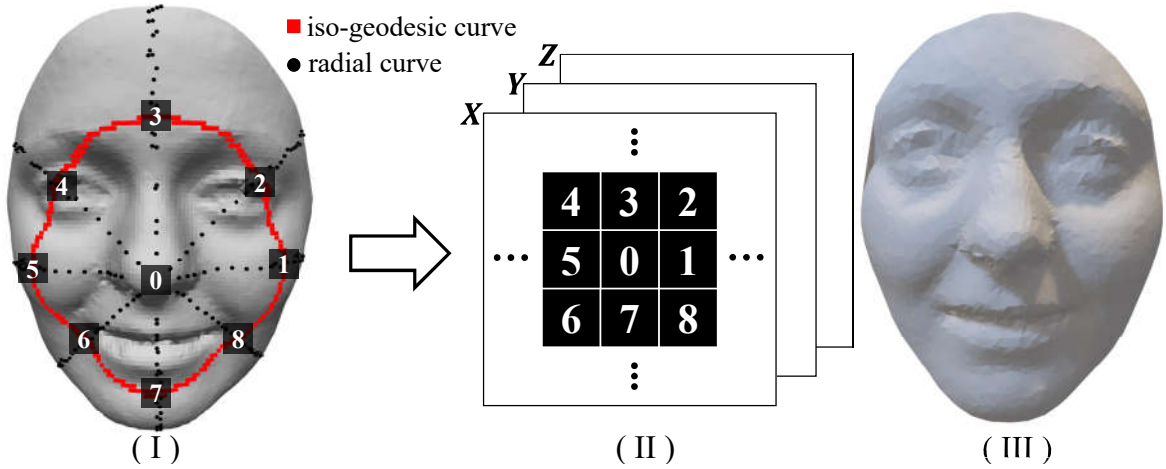

Fig. 3. The pipeline of our geometry sampling approach. (I) The original 3D face, with the detected nose-tip, the iso-geodesic curve, the radial curves, and the sampled vertices. (II) The sampled vertices are stored in a structured matrix. (III) The sampled face.

the face surface that have the same geodesic distance to the nose-tip. We denote an iso-geodesic curve as follows:

$$\mathbf{G}_i(d) = (v_1^d, v_2^d, \ldots, v_{n_d}^d), d \in [0, D], \quad (1)$$

where n_d denotes the total number of the vertices on the iso-geodesic curve $\mathbf{G}(d)$, and D denotes the maximal geodesic distance from the nose-tip. In our experiments, we set the same D that is large enough to cover the chin and the eyebrows for all the faces. See an example of the iso-geodesic curve in Figure 3(I).

2) *Radial curves*. We first align the 3D face model to the **XOY** plane. Then, given the nose tip **O**, a radial curve contains a sequence of the vertices whose projections on the **XOY** plane have the same angle to the **X** axis, i.e., $\angle \mathbf{XOv_i} = \theta$. We denote a radial curve as follows:

$$\mathbf{R}_j(\theta) = (v_1^\theta, v_2^\theta, \ldots, v_{n_\theta}^\theta), \theta \in [0, 360), \quad (2)$$

where n_θ denotes the total number of the vertices on the radial curve $\mathbf{R}(\theta)$. See the examples of the radial curves in Figure 3(I).

3) *Geometry sampling based on the intersections*. Our objective of the geometry sampling step is to sample the vertices on a 3D face and save them into a squared matrix of the size $2R+1$, where R denotes the total number of the iso-geodesic curves. In order to obtain the averaged sampling, we compute the iso-geodesic curves $\mathbf{G}(d)$ with the linearly increased d, i.e.,

$$d = k \cdot (D/K), k = 1, \ldots, K, \quad (3)$$

where K denotes the total number of the iso-geodesic curves. Figure 3 depicts the geometry sampling based on the intersections between the iso-geodesic curves and the radial curves, which can be described as follows:

a) First, starting from the detected nose tip $\mathbf{O}$, we assign it to the center of the sampling matrix, i.e., $\mathbf{M}(K+1, K+1) = \mathbf{O}$.
b) Then, we compute the r-th iso-geodesic curve, and $8K$ radial curves $\mathbf{R}(\theta), \theta = t \cdot \frac{360}{8K}, t = 1, \ldots, 8K$.
c) After that, by computing the intersections between the k-th iso-geodesic curve and the newly computed $8K$ radial curves, we obtain $8k$ intersected vertices in order, which can be stored into the k-th ring within the sampling matrix $\mathbf{M}$.
d) By repeating the steps (b)-(c) until $k = K$, we can obtain the full sampling matrix M.

Figure 3(III) shows the sampled face from the original face shown in Figure 3(I). Note that we extract more samples from the larger iso-geodesic curves, which is important to keep the visual facial features in the regions far away from the nose tip. Additionally, the sampled vertices can be represented by a structured matrix.

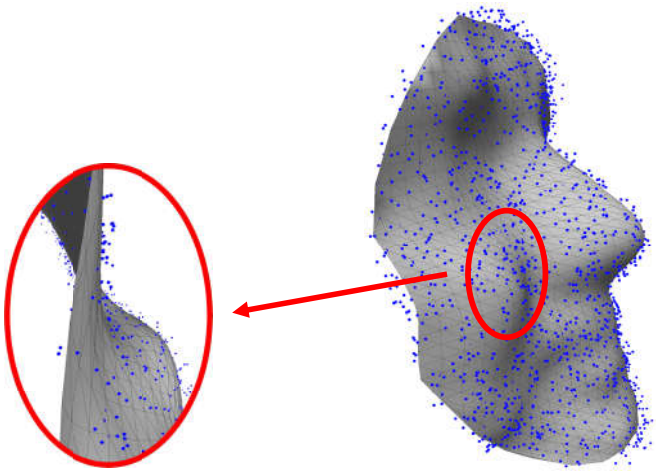

Fig. 4. Surface fitting of the generated 3D face as the point cloud.

B. 3D Face Generation via DCGAN

Now that we have obtained the geometry sampling for all the 3D faces in the training set, we proceed to train a 3D face generative model using the deep networks. In [22], Geoodfellow et al. proposed a GAN model, which contains two deep networks, i.e., a Generator (G) and a Discriminator (D). The GAN model works in the way that the G model keeps updating the output until the D model cannot distinguish the generated output from the training data. Recently in [23], Radford et al. proposed a DCGAN model to improve the performance of the GAN model with the following network settings:

- Apply the transposed convolutions for G and the stride convolutions instead of the pooling layers.
- Apply the fully Convolutional Networks instead of the fully connected hidden layers.
- Apply the ReLU activation [26] for all the convolution layers and the tanh activation for the output layer in G, and apply the LeakyReLU activation [27, 28] in D.
- Apply the batch normalization [29] in both G and D.

Figure 1 shows the generated 'Happy' and 'Surprise' faces using the proposed 3D face generative model. Our experiments on the training and the generated results are presented in details in Section IV.

Unlike 2D images, the smoothness of a 3D shape surface can be easily tainted by noise while often a noisy pixel in a 2D image is hardly noticeable. For this reason, we can easily foresee that the generated 3D face of point clouds require a post-processing step (e.g., smoothing). A large number of previous efforts have been focused on the 3D shape reconstruction from dense point clouds [30, 31, 32]. Unlike them, our generated models are sparse that contain $(2K+1)^2$ points. In our implementation, we apply the linear interpolation method to fit a smooth surface to the obtained 3D points. An example of the generated 3D face of point clouds and its smoothed surface is shown in Figure 4.

IV. Results and Discussion

In our experiments, the geometry sampling approach and the DCGAN model were implemented with Matlab and Python, respectively. All the experiments were conducted on an off-the-shelf desktop with an Intel Core with 3.4GHz CPU and 16GB memory.

In order to evaluate the proposed 3D face generative model, we have experimented with the 'BU-3DFE' dataset from the Binghamton University [33]. This face dataset contains 100 subjects (56 females and 44 males with various ethnicities), each of which performs 7 different expressions, shown in Table I.

TABLE I
THE TIMING STATISTICS (IN SECONDS) OF THE GEOMETRY SAMPLING AND TRAINING FOR EACH SET OF THE FACIAL EXPRESSIONS FROM THE 'BU-3DFE' DATASET.

Facial Expressions	Geometry Sampling	1 epoch	5 epochs	10 epochs
Angry	73.2	35.4	164.2	348.2
Disgust	69.9	35.2	163.0	345.8
Fear	70.5	39.3	163.5	356.5
Happy	74.7	40.4	163.5	355.0
Neutral	77.0	38.1	164.0	353.1
Sad	73.8	35.3	166.6	333.8
Surprise	79.9	35.7	167.6	323.8

Given the computational cost can increase exponentially with the number of the iso-geodesic curves K, we set $K = 29$ by balancing between the quality of the 3D face model and the computational costs. This results in the dimension of M as 59. That is, each sampled face contains $59^2 = 3481$ vertices. Table I shows the averaged per-mesh timings of the geometry sampling of the training faces with different expressions. On average, it took about 74.2 seconds to sample a face with around $8,000$ vertices. Furthermore, several sampled faces are

shown in the top row of Figure 6. As can be seen, although the smoothness of the sampled faces have been broken, they are sufficient to observe visual facial features. More importantly, the sampled faces can be represented with a structured matrix.

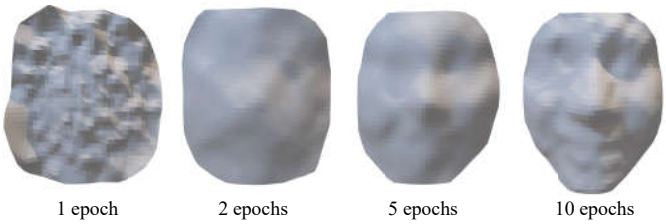

Fig. 5. The generated 3D faces with different training epochs.

Table I shows the training timings of the DCGAN model with different numbers of training epochs for different expressions, which is increased linearly and the timing of each epoch ranges in [35.2 40.4]. As an example for the 'Happy' face, Figure 5 shows the intermediate progress of the proposed generative model. Starting with a matrix of random noises, the generative model gradually improves the quality of the output, until we obtain a 3D face reasonably close to the faces in the training dataset. As can be seen in Figure 5, we can obtain an easily recognizable 'Happy' face using the generative model based on 10 epochs of training. More results are also shown in Figure 1.

The bottom row of Figure 6 shows the generated 3D faces using our approach. From this figure, we can easily observe the facial components including the nose, mouth, cheek, etc. In some of the generated faces, although the details of the eyes may be not clear, we still can easily recognize the facial expressions. Additionally, it is interesting to mention that different subjects may show different facial movements for the same expression, due to the difference of culture, race, or character. Using our approach, we can generate faces with the same expression, but having local surface variances (refer to the mouth and cheek regions of the generated 'Fear' faces).

V. Conclusion

We have presented an unsupervised data-driven model for the generation of 3D faces. Specifically, we first propose a geometry sampling approach to adapt an un-structured 3D models for the classical DCGAN model, which is a competitive data generation model. Our method requires neither explicit face feature extraction nor pre-computed face alignments. Our current method is effective for 3D faces, because the geometry affinities of 3D faces are high, especially the geodesic distance. However, our method can be easily extended for more complex shapes using reversible parameterization. As the future works, we will further investigate the potentials of our new 3D face sampling approach for generation by adapting to the recent advanced deep network techniques. Moreover, we are also interested in exploring the direction of automated data-driven generation of 3D faces with textures.

Acknowledgment

This work has been jointly supported by the National Natural Science Foundation of China under Grant 61962021, 51978271, 61763011 and 61762050, the Key Research and Development Program of Jiangxi Province under Grant 20192BBE50079 and 20181BBE50024, the Natural Science Foundation of Guangdong Province under Grant 2019A1515011793 and 2017A030313347 and the China Postdoctoral Science Foundation under Grant 2019M662261.

References

[1] E. Richardson, M. Sela, and R. Kimmel, "3d face reconstruction by learning from synthetic data," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 460–469.

[2] S. Z. Gilani and A. Mian, "Learning from millions of 3d scans for large-scale 3d face recognition," *Computer Vision and Pattern Recognition*, pp. 1896–1905, 2018.

[3] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 559–568. [Online]. Available: https://academic.microsoft.com/paper/2099940712

[4] S. Song, S. P. Lichtenberg, and J. Xiao, "Sun rgb-d: A rgb-d scene understanding benchmark suite," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 567–576. [Online]. Available: https://academic.microsoft.com/paper/1923184257

[5] A. Handa, T. Whelan, J. McDonald, and A. J. Davison, "A benchmark for rgb-d visual odometry, 3d reconstruction and slam," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 2014, pp. 1524–1531. [Online]. Available: https://academic.microsoft.com/paper/2058535340

[6] I. Kemelmacher-Shlizerman and R. Basri, "3d face reconstruction from a single image using a single reference face shape," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 394–405, 2011. [Online]. Available: https://academic.microsoft.com/paper/2146566773

[7] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 3357–3364. [Online]. Available: https://academic.microsoft.com/paper/2962887844

[8] C. B. Choy, D. Xu, J. Gwak, K. Chen, and S. Savarese, "3d-r2n2: A unified approach for single and multi-view 3d object reconstruction," in *European Conference on Computer Vision*, 2016, pp. 628–644. [Online]. Available: https://academic.microsoft.com/paper/2342277278

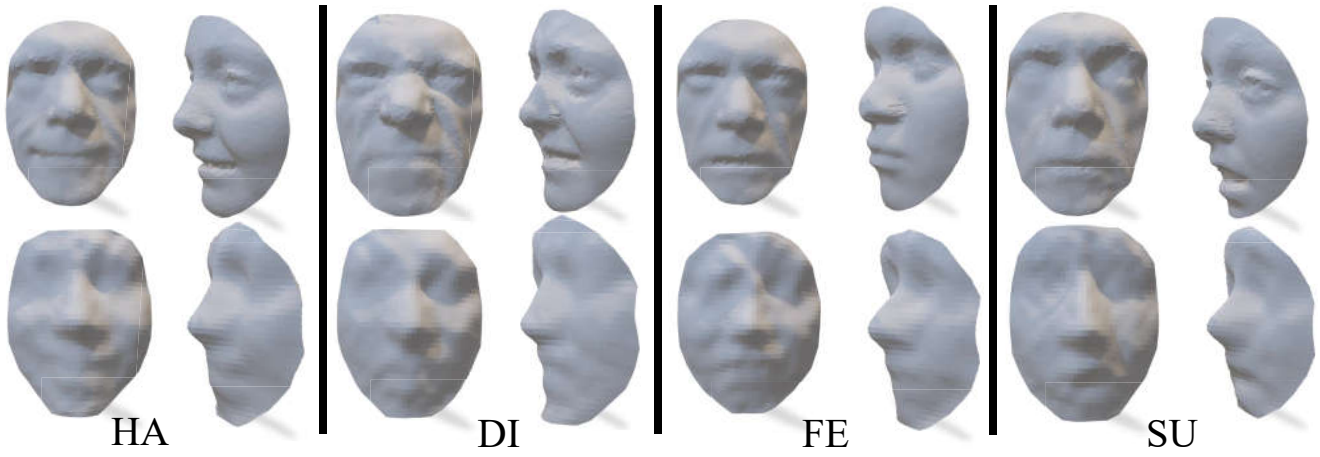

Fig. 6. The sampled faces (top row) and the generated faces by using our approach (bottom row), for the 'Happy'(HA), 'Disgust'(DI), 'Fear'(FE) and 'Surprise'(SU) expressions.

[9] H. Fan, H. Su, and L. Guibas, "A point set generation network for 3d object reconstruction from a single image," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2463–2471. [Online]. Available: https://academic.microsoft.com/paper/2560722161

[10] P. Garrido, M. Zollhfer, D. Casas, L. Valgaerts, K. Varanasi, P. Prez, and C. Theobalt, "Reconstruction of personalized 3d face rigs from monocular video," *ACM Transactions on Graphics*, vol. 35, no. 3, p. 28, 2016. [Online]. Available: https://academic.microsoft.com/paper/2398381847

[11] S. Soltanpour, B. Boufama, and Q. J. Wu, "A survey of local feature methods for 3d face recognition," *Pattern Recognition*, vol. 72, pp. 391–406, 2017.

[12] H. Yamauchi, S. Gumhold, R. Zayer, and H.-P. Seidel, "Mesh segmentation driven by gaussian curvature," *The Visual Computer*, vol. 21, no. 8-10, pp. 659–668, 2005.

[13] H. Drira, B. B. Amor, M. Daoudi, and A. Srivastava, "Pose and expression-invariant 3d face recognition using elastic radial curves," in *British machine vision conference*, 2010, pp. 1–11.

[14] Y. Lipman and T. Funkhouser, "Möbius voting for surface correspondence," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3, p. 72, 2009.

[15] C. C. Wang, K. Tang, and B. M. Yeung, "Freeform surface flattening based on fitting a woven mesh model," *Computer-Aided Design*, vol. 37, no. 8, pp. 799–814, 2005.

[16] H. M. Briceño, P. V. Sander, L. McMillan, S. Gortler, and H. Hoppe, "Geometry videos: a new representation for 3d animations," in *Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer animation*. Eurographics Association, 2003, pp. 136–146.

[17] J. Xia, Y. He, D. Quynh, X. Chen, and S. C. Hoi, "Modeling 3d facial expressions using geometry videos," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 591–600.

[18] T. Li, T. Bolkart, M. J. Black, H. Li, and J. Romero, "Learning a model of facial shape and expression from 4d scans," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, p. 194, 2017.

[19] L. Chen, J. Ye, L. Jiang, C. Ma, Z. Cheng, and X. Zhang, "Synthesizing cloth wrinkles by cnn-based geometry image superresolution," *Computer Animation and Virtual Worlds*, vol. 29, no. 3-4, p. e1810, 2018.

[20] O. Fabius and J. R. van Amersfoort, "Variational recurrent auto-encoders," *arXiv preprint arXiv:1412.6581*, 2014.

[21] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[23] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *CoRR*, vol. abs/1511.06434, 2015.

[24] J. Kossaifi, L. Tran, Y. Panagakis, and M. Pantic, "GAGAN: geometry-aware generative adversarial networks," *CoRR*, vol. abs/1712.00684, 2017. [Online]. Available: http://arxiv.org/abs/1712.00684

[25] M. D. Samad and K. M. Iftekharuddin, "Frenet frame-based generalized space curve representation for pose-invariant classification and recognition of 3-d face," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 4, pp. 522–533, 2016.

[26] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th ICML*, 2010, pp. 807–814.

[27] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models,"

in *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, vol. 1, 2013, pp. 1–6.

[28] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *CoRR*, vol. abs/1505.00853, 2015.

[29] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[30] H. Pfister, M. Zwicker, J. Van Baar, and M. Gross, "Surfels: Surface elements as rendering primitives," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 335–342.

[31] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg *et al.*, "The digital michelangelo project: 3d scanning of large statues," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 131–144.

[32] S. Rusinkiewicz and M. Levoy, "Qsplat: A multiresolution point rendering system for large meshes," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 343–352.

[33] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3d facial expression database for facial behavior research," in *Automatic face and gesture recognition, 2006. FGR 2006. 7th international conference on*. IEEE, 2006, pp. 211–216.