3D Clothed Human Parametric Model from a Single Scan with Joint Optimization

Sohyun Park
Department of IT Engineering
Sookmyung Women's University
Seoul, Korea
so209@sookmyung.ac.kr

Jiwoo Kang
Department of IT Engineering
Sookmyung Women's University
Seoul, Korea
jwkang@sookmyung.ac.kr

***Abstract*—We introduce an avatar generation method from a single three-dimensional (3D) scan. Through our method, the scan is parameterized with a parametric 3D body model, e.g., a skinned multi-person linear model (SMPL), without loss of vertex resolution and changes in a mesh topology. The parameterized avatar can be animated freely with motions, given by capturing human poses or manually manipulating them. Due to its ease of use and simplicity, we believe that the proposed method can be used in various metaverse applications.**

***Index Terms*—parametric model, 3D human body, avatar, 3D animation, blend skinning**

I. Introduction

Parametric Three-dimensional (3D) human models have been used widely to synthesize and analyze shapes, poses, and motions of the human body. Once we capture 3D scans, the scans need to be parameterized so that they can be animated based on the poses and motions we want to demonstrate. These procedures are commonly performed by fitting the existing parametric models like a skinned multi-person linear model (SMPL) [1] to the 3D scans. However, the fitting procedure [2], [3] produces a posed human body mesh, as described in Fig. 1b, not a clothed human mesh, as illustrated in Fig. 1a. Therefore, an additional procedure is required to deform the fitted human body mesh to the clothed mesh, called the non-rigid or deformable registration. Although the registration procedure can make the scan parametric, the process is prone to error. Also, the resolution loss is accompanied by this procedure because the number of vertices of the existing parametric models is far less than the number of vertices of scans captured with 3D scanners. Therefore, the degradation of 3D mesh quality is inevitable. Our method efficiently resolves these problems by taking s single raw 3D scan of a clothed human and turning them into an animatable avatar.

II. Method

Our objective is to make it easy to construct a realistic 3D avatar of a clothed human that can be easily reposed and animated as existing parametric models such as a skinned multi-person linear model (SMPL) [1]. Similar to the previous methods [2], [3], the proposed method first fits the SMPL model to a single scan, that is, the posed scan. Especially, we find a minimally clothed shape (MCS) for the given scan [2]. This procedure enables us to obtain pose parameters of the SMPL model [1], the minimally clothed mesh in a canonical pose (T-pose) in Fig. 1c, and geometric transformations of the scan. Then, we approximate the blend weights of the scan by matching the nearest ones of the posed MCS. The scan's pose parameters and blend weights enable the scan to be inversely transformed into the canonical pose, as seen in Fig. 1d. Here, if the approximated pose parameters are accurate, the inversely transformed scan in the canonical pose becomes matched well with MCS in the canonical pose. Using the relationship, we can correct the misalignment fitting error and blend weights of the scan by minimizing the errors between scans and MCSs in canonical and posed spaces concurrently. Thus, using the Levenberg–Marquardt method, we optimize pose parameters while the blend weights of the nearest model vertices determine the ones of the scans. It significantly helps refine the errors in the fitted pose and obtain accurate blend weights in the canonical pose. As a result, the scan mesh in the canonical pose and the accurately refined blend weights enable the scan to be reposed realistically, as described in Fig. 1d.

III. Discussion and Conclusion

Recently, many applications require realistically clothed bodies. In line with this trend, we proposed a 3D mesh parameterization method, which can be applied even for a single scan. The avatar can be animated with given motions. Due to the ease of use and simplicity of the proposed method, someone can effortlessly build avatars, and we hope that our approach will be used in diverse metaverse applications.

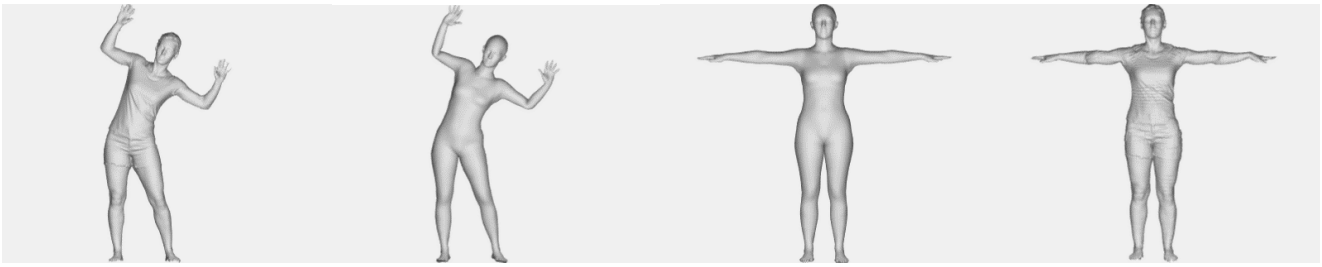
This work was supported by the National Research Foundation of Korea (NRF) funded by the Korea Government (Ministry of Science and ICT, MSIT) under Grant NRF-2022R1F1A1068704.

References

[1] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: A skinned multi-person linear model," *ACM Transactions on Graphics*, vol. 34, no. 6, pp. 1–16, 2015.

[2] C. Zhang, S. Pujades, M. J. Black, and G. Pons-Moll, "Detailed, accurate, human shape estimation from clothed 3D scan sequences," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4191–4200.

[3] G. Pons-Moll, S. Pujades, S. Hu, and M. J. Black, "ClothCap: Seamless 4D clothing capture and retargeting," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–15, 2017.



(a) (b) (c) (d)

Fig. 1. To find an accurate T-posed mesh from a scan mesh, the proposed method jointly minimizes the distance between (a) the raw scan and (b) the posed SMPL mesh. At the same time, the distance between (c) the SMPL mesh in T-pose and (d) the T-posed scan using the inverse kinematics is jointly minimized.

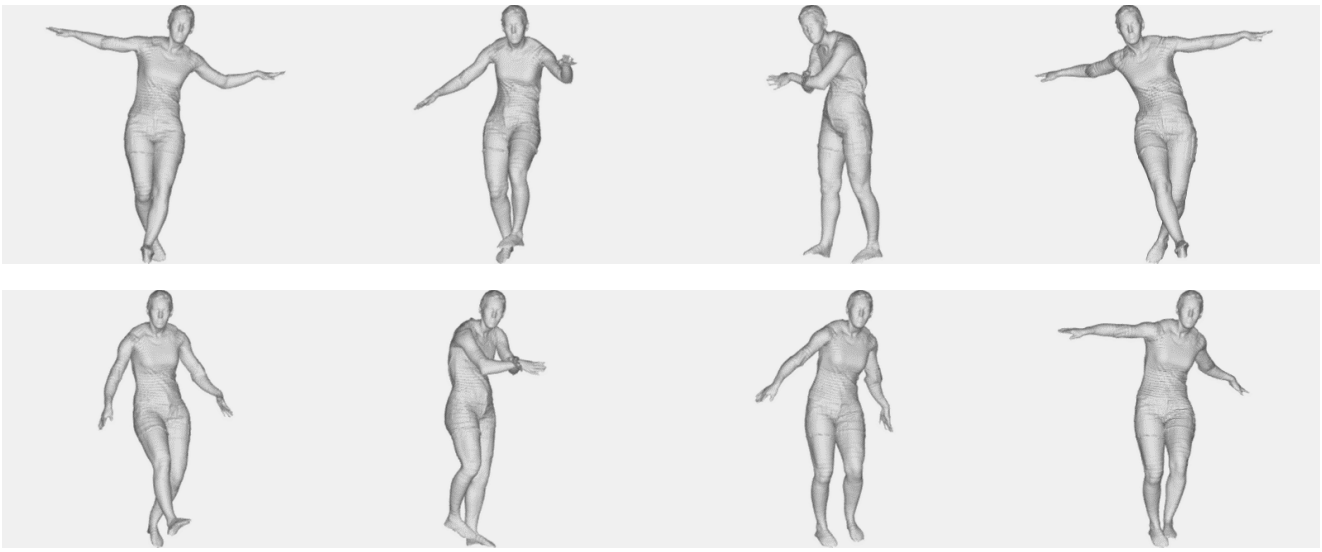


Fig. 2. Animated scan by dance motions of the AIST++ Dataset [4] with a human scan in Cape Dataset [5].

[4] R. Li, S. Yang, D. A. Ross, and A. Kanazawa, "AI choreographer: Music conditioned 3D dance generation with AIST++," in *IEEE International Conference on Computer Vision*, 2021, pp. 13 401–13 412.

[5] Q. Ma, J. Yang, A. Ranjan, S. Pujades, G. Pons-Moll, S. Tang, and M. J. Black, "Learning to dress 3D people in generative clothing," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2020, pp. 6469–6478.