

Mesh Variational Autoencoders with Edge Contraction Pooling (Supplementary Material)

Yu-Jie Yuan^{a,b}, Yu-Kun Lai^c, Jie Yang^{a,b}, Qi Duan^d, Hongbo Fu^e and Lin Gao^{a,b}

^a*Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, CAS*

^b*University of Chinese Academy of Sciences*

^c*School of Computer Science and Informatics, Cardiff University, UK*

^d*SenseTime Research*

^e*City University of Hong Kong*

1. Overview

This supplementary document provides further results that accompany the paper. It contains three parts, including two qualitative generalization comparison and more interpolation results.

- Section 2 provides more qualitative comparison with CoMA [3].
- Section 3 provides some qualitative comparison with mesh sampling proposed in [3].
- Section 4 provides more interpolation comparison with other methods.

2. More Comparison with CoMA [3]

In Fig. 1, we show more qualitative comparison results with [3] on the Swing dataset [5]. It can be seen that our method has better reconstruction results on unseen data.

3. Comparison with Mesh Sampling

We replace our pooling operation with sampling operation [3] to compare reconstruction errors in the submitted paper. Here, we present several qualitative comparisons with sampling operation on the SCAPE dataset [1] in Fig. 2.

4. More Interpolation Results

In Fig. 3, we show a comparison with the method of [2], another VAE-based method, which leads to artifacts especially in the synthesized human hands. We also compare our method on the SCAPE dataset [1] with MeshVAE as shown in Fig. 4. We can see that Mesh VAE [4] produces interpolation results with obvious artifacts.

References

- [1] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: shape completion and animation of people. *ACM transactions on graphics*, 24(3):408–416, 2005. 1
- [2] Or Litany, Alex Bronstein, Michael Bronstein, and Ameesh Makadia. Deformable shape completion with graph convolutional autoencoders. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 1, 3
- [3] Anurag Ranjan, Timo Bolkart, Soubhik Sanyal, and Michael J. Black. Generating 3D faces using convolutional mesh autoencoders. In *European Conference on Computer Vision (ECCV)*, pages 725–741. Springer International Publishing, 2018. 1, 2, 3
- [4] Qingyang Tan, Lin Gao, Yu-Kun Lai, and Shihong Xia. Variational autoencoders for deforming 3d mesh models. In *CVPR*, June 2018. 1, 4
- [5] Daniel Vlasic, Ilya Baran, Wojciech Matusik, and Jovan Popović. Articulated mesh animation from multi-view silhouettes. *ACM Transactions on Graphics*, 27(3):97, 2008. 1

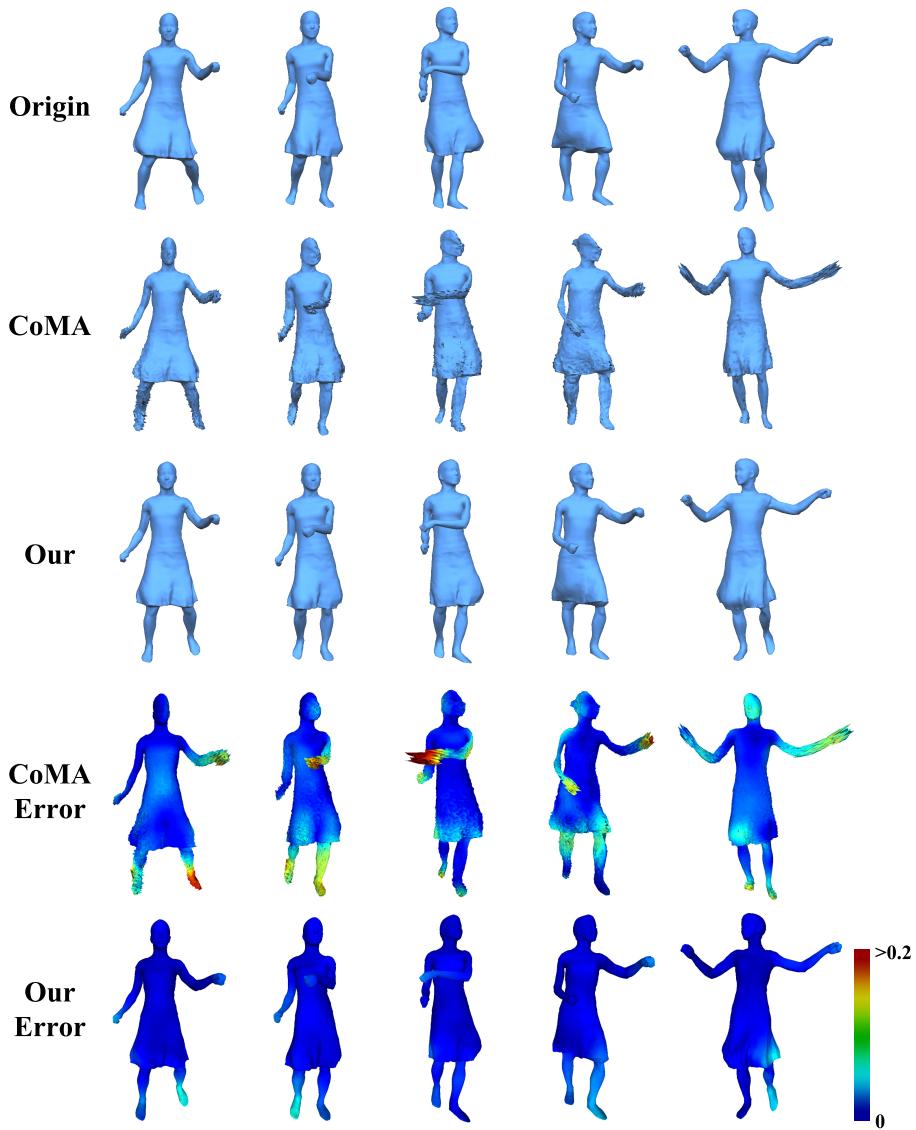


Figure 1. More qualitative comparison with [3]. It can be seen that our method leads to more accurate reconstructions.

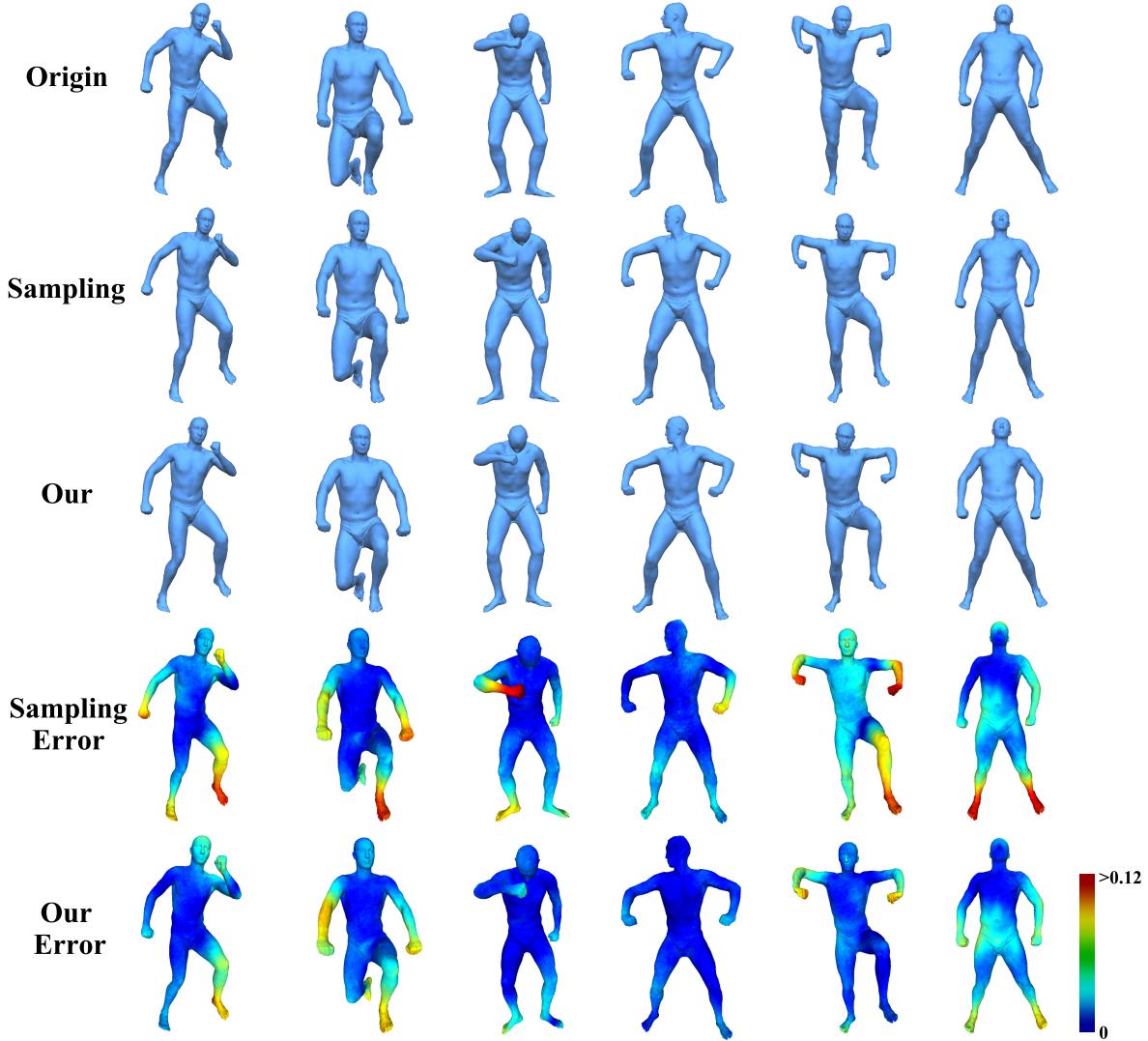


Figure 2. Qualitative comparison with mesh sampling in [3]. Reconstruction errors are color-coded. It can be seen that our method leads to more accurate reconstructions.

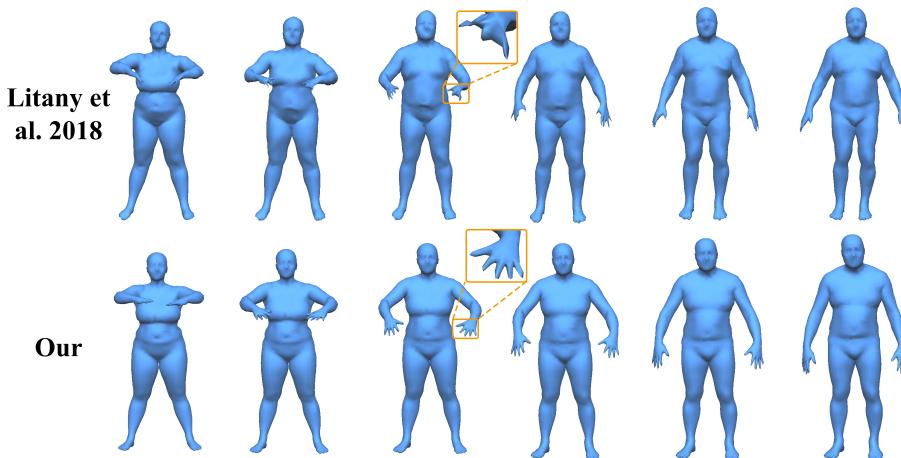


Figure 3. Comparison of mesh interpolation results with [2]. First row is the result of [2], and second row is our result.

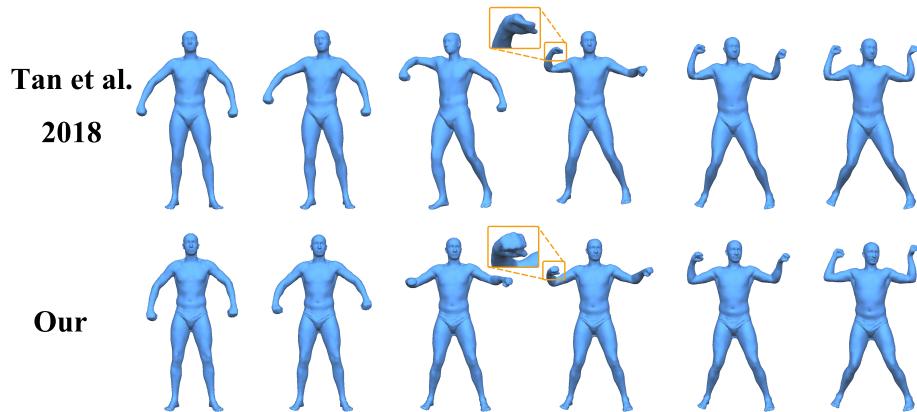


Figure 4. Comparison of mesh interpolation results with [4]. The models in the leftmost and rightmost columns are the input models to be interpolated. 1st row: the results of [4], and 2nd row: our results.