Variational Autoencoders for Deforming 3D Mesh Models

Qingyang Tan[1,2], Lin Gao[1], Yu-Kun Lai[3], Shihong Xia[1]
[1]Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences
[2]School of Computer and Control Engineering, University of Chinese Academy of Sciences
[3]School of Computer Science & Informatics, Cardiff University
tanqingyang14@mails.ucas.ac.cn, {gaolin, xsh}@ict.ac.cn, LaiY4@cardiff.ac.uk

CVPR 2018 SALT LAKE CITY • JUNE 18-22

1 Abstract

3D geometric contents are becoming increasingly popular. In this paper, we study the problem of analyzing deforming 3D meshes using deep neural networks. Deforming 3D meshes are flexible to represent 3D animation sequences as well as collections of objects of the same category, allowing diverse shapes with large-scale non-linear deformations. We propose a novel framework which we call mesh variational autoencoders (mesh VAE), to explore the probabilistic latent space of 3D surfaces. The framework is easy to train, and requires very few training examples. We also propose an extended model which allows flexibly adjusting the significance of different latent variables by altering the prior distribution. Extensive experiments demonstrate that our general framework is able to learn a reasonable representation for a collection of deformable shapes, and produce competitive results for a variety of applications, including shape generation, shape interpolation, shape space embedding and shape exploration, outperforming state-of-the-art methods.

2 Feature Representation

Rotation-invariant mesh difference (RIMD) feature[1]
- Linearly combinable
- Mean reconstruction error extremely small

$$E(\mathbf{T}_i) = \sum_{j\in N_i} c_{ij}\|\mathbf{e}'_{ij} - \mathbf{T}_i\mathbf{e}_{ij}\|^2 \xrightarrow{\text{minimize energy}} \mathbf{T}_i = \mathbf{R}_i\mathbf{S}_i$$

rotation & scaling/shear; deformation gradient

RIMD representation

$$\mathbf{f} = \{\log dR_{ij}; \mathbf{S}_i\} \quad (\forall i, j \in N_i) \xleftarrow[\text{logarithm}]{\text{matrix}} dR_{ij} = \mathbf{R}_i^T\mathbf{R}_j$$

cancel out global rotation

uniform normalization

$$\widetilde{\mathbf{f}}_i^j = 2a \times \frac{\mathbf{f}_i^j - \min_j(\mathbf{f}_i^j)}{\max_j(\mathbf{f}_i^j) - \min_j(\mathbf{f}_i^j)} - a$$

fit activation function tanh → avoid gradient vanishing problem

3 Mesh VAE & Conditional Mesh VAE

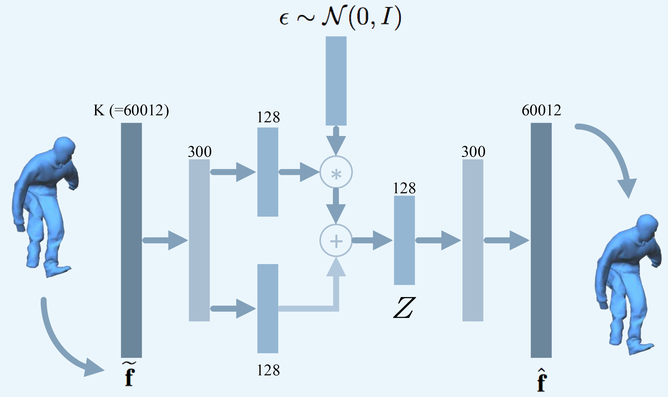
Reconstruction Loss + KL divergence[2]

$$L_{meshVAE} = \alpha\frac{1}{2MK}\sum_{j=1}^{M}\sum_{i=1}^{K}(\hat{\mathbf{f}}_i^j - \widetilde{\mathbf{f}}_i^j)^2 + D_{KL}(q(z|\widetilde{\mathbf{f}})\|p(z))$$

Condition on labels[3] → Control the types of shapes to be generated

$$L_{Conditional\ meshVAE} = \alpha\frac{1}{2MK}\sum_{j=1}^{M}\sum_{i=1}^{K}(\hat{\mathbf{f}}_{ci}^j - \widetilde{\mathbf{f}}_i^j)^2 + D_{KL}(q(z|\widetilde{\mathbf{f}}, c)\|p(z|c))$$

4 Network Architecture



5 Extended model with Improved Low-Dimensional Embedding

Modify the prior distribution of the latent variables

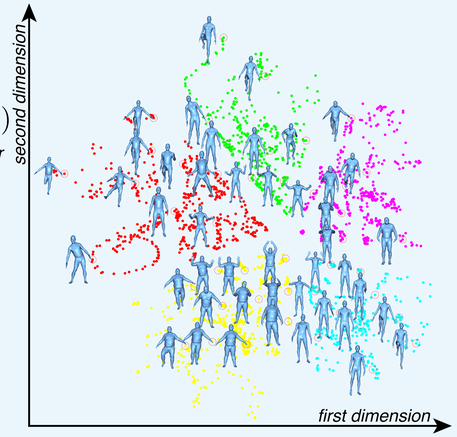
$$p(z) = \mathcal{N}(z; 0, diag(\underbrace{\sigma^2_{object}}_{\text{tunable vector}}))$$

Dimension with small σ_{object_i}

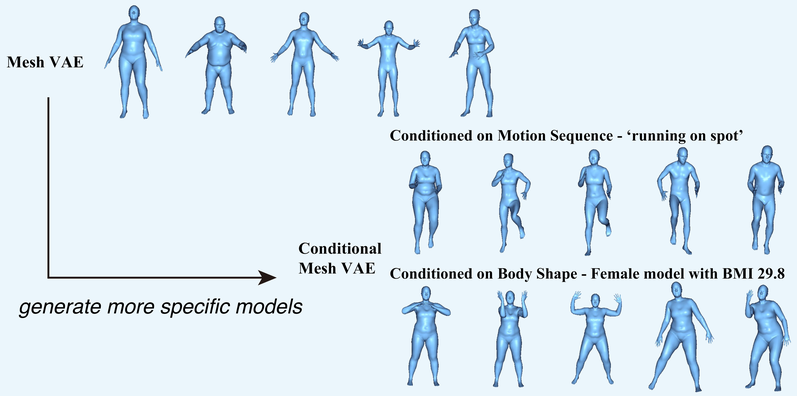
→ Capture dominant changes

Quantitative analyzation

Retrieval task on Dyna dataset
Using styles and poses as labels
AUC:
Our method (0.5168) > t-SNE (0.4961) > PCA (0.2272) > NPE (0.1391)



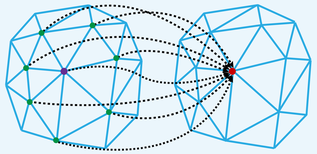
6 Random Generation



7 Graph Convolution

Combine convolutional operation on irregular graph[4]
→ Lower reconstruction errors

Dataset	Ours	With Conv.
Camel	0.0241	0.0167
Horse	0.0603	0.0117
SCAPE	0.1104	0.0486



References
[1] L. Gao, Y.-K. Lai, D. Liang, S.-Y. Chen, and S. Xia. Efficient and flexible deformation representation for data-driven surface modeling. *ACM Trans. Graph.*, 35(5), 2016.
[2] D. P. Kingma and M. Welling. Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114, 2013
[3] K. Sohn, H. Lee, and X. Yan. Learning structured output representation using deep conditional generative models. In *NIPS*, 3483–3491, 2015.
[4] Q. Tan, L. Gao, Y. Lai, J. Yang, and S. Xia. Mesh-based autoencoders for localized deformation component analysis. In *AAAI*, 2018.

Acknowledgments
This work was supported by grants from the National Natural Science Foundation of China (No. 61502453, No. 61772499 and No.61611130215), Royal Society-Newton Mobility Grant (No. IE150731), the Science and Technology Service Network Initiative of Chinese Academy of Sciences (No. KFJSTS-ZDTP-017).

Mesh VAE

convMesh