Mesh-based Autoencoders for Localized Deformation Component Analysis

Qingyang Tan[1,2], Lin Gao[1], Yu-Kun Lai[3], Jie Yang[1,2], 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, yangjie01, xsh}@ict.ac.cn, LaiY4@cardiff.ac.uk

1 Abstract

Spatially localized deformation components are very useful for shape analysis and synthesis in 3D geometry processing. Several methods have recently been developed, with an aim to extract intuitive and interpretable deformation components. However, these techniques suffer from fundamental limitations especially for meshes with noise or large-scale deformations, and may not always be able to identify important deformation components. In this paper we propose a novel mesh-based autoencoder architecture that is able to cope with meshes with irregular topology. We introduce sparse regularization in this framework, which along with convolutional operations, helps localize deformations. Our framework is capable of extracting localized deformation components from mesh data sets with large-scale deformations and is robust to noise. It also provides a nonlinear approach to reconstruction of meshes using the extracted basis, which is more effective than the current linear combination approach.

2 Feature Representation

$$\arg\min_{\mathbf{T}_{m,i}} \sum_{j\in N(i)} c_{ij}\|(\mathbf{p}_{m,i}-\mathbf{p}_{m,j})-\mathbf{T}_{m,i}(\mathbf{p}_{1,i}-\mathbf{p}_{1,j})\|_2^2 \longrightarrow \mathbf{T}_{m,i}$$ *deformation gradient*

polar decomposition

$$\mathbf{T}_{m,i}=\mathbf{R}_{m,i}\mathbf{S}_{m,i}$$ *rotation matrix, scale and shear matrix*

rewritten as

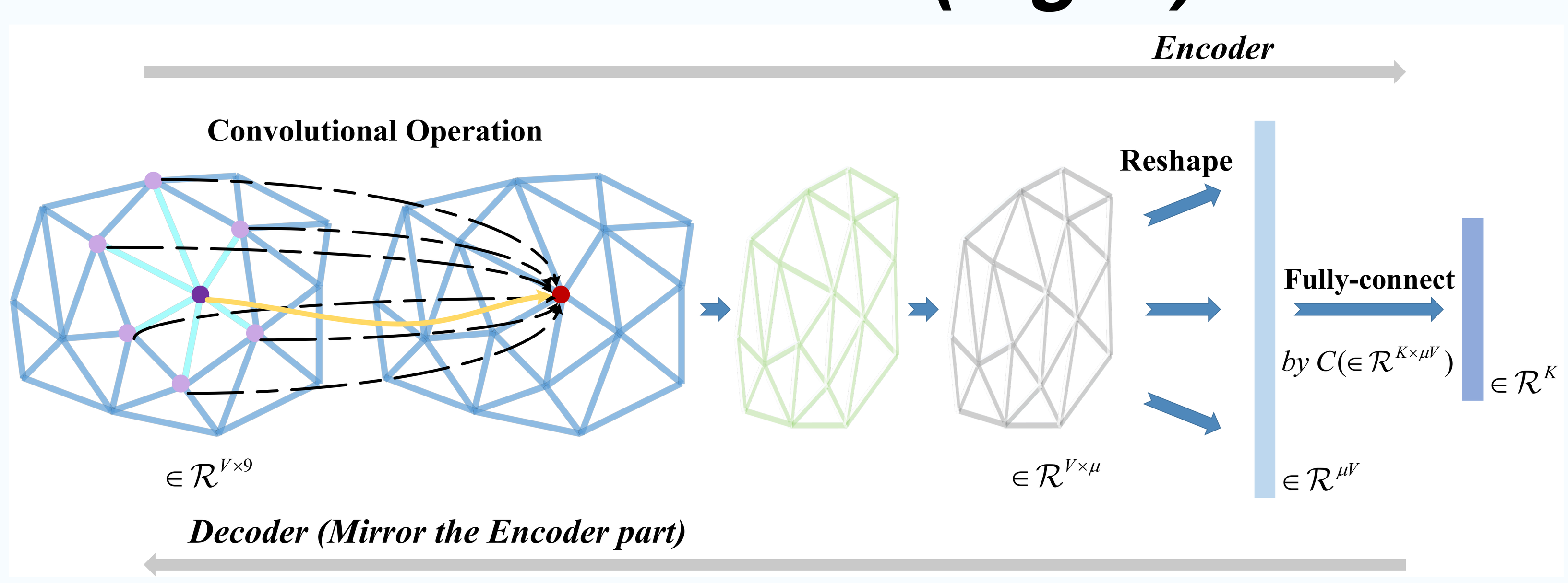
rotating around an axis $\boldsymbol{\omega}_{m,i}$ *by an angle* $\theta_{m,i}$

choose unique representation from
$$\Omega_{m,i}=\{(\boldsymbol{\omega}_{m,i},\theta_{m,i}+t\cdot 2\pi),(-\boldsymbol{\omega}_{m,i},-\theta_{m,i}+t\cdot 2\pi)\}$$ [1]

extracting non-trivial elements & normalized

$$X_{m,i}=\{\widetilde{r_{m,i}},\widetilde{s_{m,i}}\}$$ *deformation feature*

3 Network Architecture (Fig. 1)



4 Loss

convolutional operation (extended from [2])
$$y_i = W_{point}x_i + W_{neighbour}\frac{\sum_{j=1}^{D_i} x_{n_{ij}}}{D_i} + b$$

sparsity constraints (following [3])
$$\Omega(C)=\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{V}\Lambda_{ik}\|C_k^i\|_2$$

regularization term (avoid trivial solution)
$$\mathcal{V}(Z)=\frac{1}{K}\sum_{j=1}^{K}(\max_m|Z_{jm}|-\theta)$$

Total Loss
$$\mathcal{L}=\frac{1}{N}\sum_{m=1}^{N}\|\widehat{X}_m-X_m\|_2^2+\lambda_1\Omega(C)+\lambda_2\mathcal{V}(Z)$$

autoencoder structure (detailed pipeline in Fig. 1)

$X\in\mathcal{R}^{V\times 9}$

$z=Cf$

deformation components $C\in\mathcal{R}^{K\times\mu V}$

$\hat{f}=C^T z$

$\widehat{X}\in\mathcal{R}^{V\times 9}$

stacked to combinational weights

Fig. 2 Errors of applying our model to generate unseen data

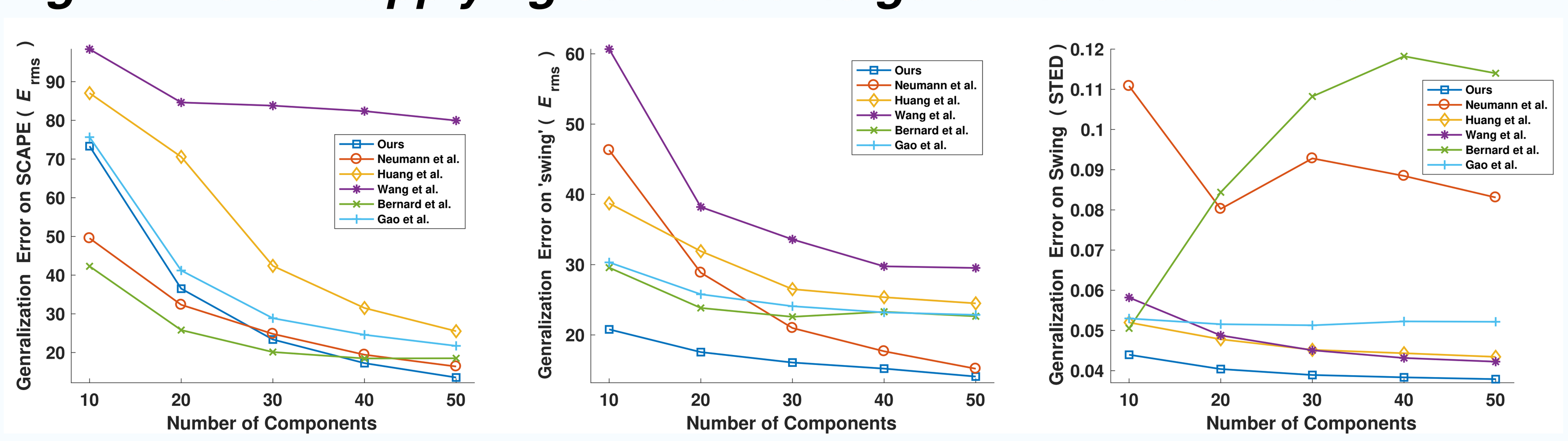


Fig. 3 Deformation components extracted from SCAPE

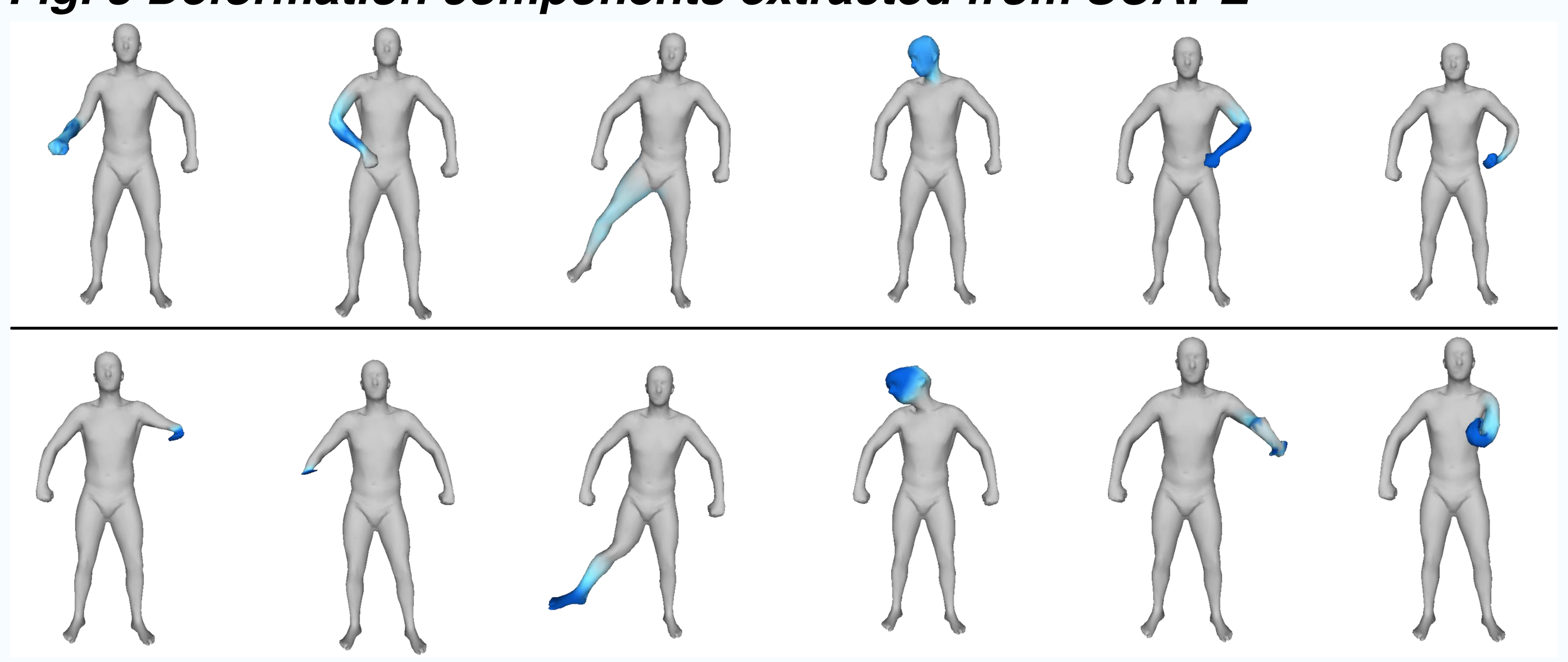
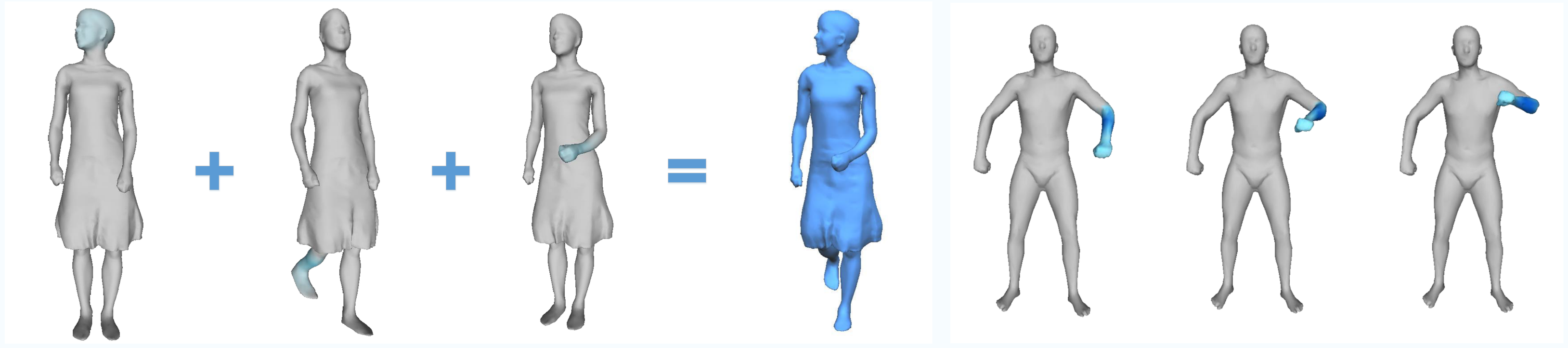


Fig. 4 Synthesized models



(a) combining deformation components extracted by our method from the Swing dataset with equal weights

(b) components about lifting the left arm extracted by our method with weights 0.3, 0.6 and 0.9

5 Applications

dimensionality reduction, reconstruction, component analysis,

latent vector Z_r | *reference model* | *mesh feature* X_r

replaced with max/min value of the i^{th} dimension of the embedding

$\widehat{Z}_{i_{min}} \longrightarrow \widehat{X_{min}} \longrightarrow \|\widehat{X_{min}}-X_r\|$

$\widehat{Z}_{i_{max}} \longrightarrow \widehat{X_{min}} \longrightarrow \|\widehat{X_{max}}-X_r\|$

applied to the decoder — *worked out differences* — *picked out larger one*

latent vector Z_{i_h} | *i^{th} deformation component* | *mesh feature* X_{i_h}

shape synthesis.

synthesis weight w_s → *i^{th} component* $w_{s_i} \longrightarrow Z_{ri}+(Z_{i_h i}-Z_{ri})\times w_{s_i} \rightarrow Z_{si}$ → Z_s → *applied to the decoder* → $\widehat{X_r}$ → *reconstructed to* → *synthesized shape*

6 Results

Quantitative Evaluation

We compare the generalization ability of our method with several state-of-art methods, and the results are shown in Fig. 2.

Qualitative Evaluation

We demonstrate the components extracted from SCAPE by different methods in Fig. 3, and the synthesized model in Fig. 4.

References

[1] Gao, L.; Lai, Y.-K.; Yang, J.; Zhang, L.-X.; Kobbelt, L.; and Xia, S. 2017. Sparse Data Driven Mesh Deformation. arXiv:1709.01250.
[2] Duvenaud, D. K.; Maclaurin, D.; Iparraguirre, J.; Bombarell, R.; Hirzel, T.; Aspuru-Guzik, A.; and Adams, R. P. 2015. Convolutional networks on graphs for learning molecular fingerprints. In NIPS, 2224–2232.
[3] Neumann, T.; Varanasi, K.; Wenger, S.; Wacker, M.; Magnor, M.; and Theobalt, C. 2013. Sparse localized deformation components. ACM Trans. Graph. 32(6):179.

Metric:
E_{rms} from Kavan, L.; Sloan, P.-P.; and O'Sullivan, C. 2010. Fast and efficient skinning of animated meshes. Comp. Graph. Forum 29(2):327336.
STED from Vasa, L., and Skala, V. 2011. A perception correlated comparison method for dynamic meshes. IEEE Trans. Vis. Comp. Graph. 17(2):220–230.

Compared with:
Neumann, T.; Varanasi, K.; Wenger, S.; Wacker, M.; Magnor, M.; and Theobalt, C. 2013. Sparse localized deformation components. ACM Trans. Graph. 32(6):179.
Huang, Z.; Yao, J.; Zhong, Z.; Liu, Y.; and Guo, X. 2014. Sparse localized decomposition of deformation gradients. Comp. Graph. Forum 33(7):239–248.
Wang, Y.; Li, G.; Zeng, Z.; and He, H. 2016. Articulated-motionaware sparse localized decomposition. Comp. Graph. Forum.
Bernard, F.; Gemmar, P.; Hertel, F.; Goncalves, J.; and Thunberg, J. 2016. Linear shape deformation models with local support using graph-based structured matrix factorisation. In CVPR, 5629–5638.
Gao, L.; Lai, Y.-K.; Yang, J.; Zhang, L.-X.; Kobbelt, L.; and Xia, S. 2017. Sparse Data Driven Mesh Deformation. arXiv:1709.01250.

Acknowledgments

This work was supported by 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)and the NVIDIA hardware donation.