NEURAL INFORMATION PROCESSING SYSTEMS

3D Pose Transfer with Correspondence Learning and Mesh Refinement

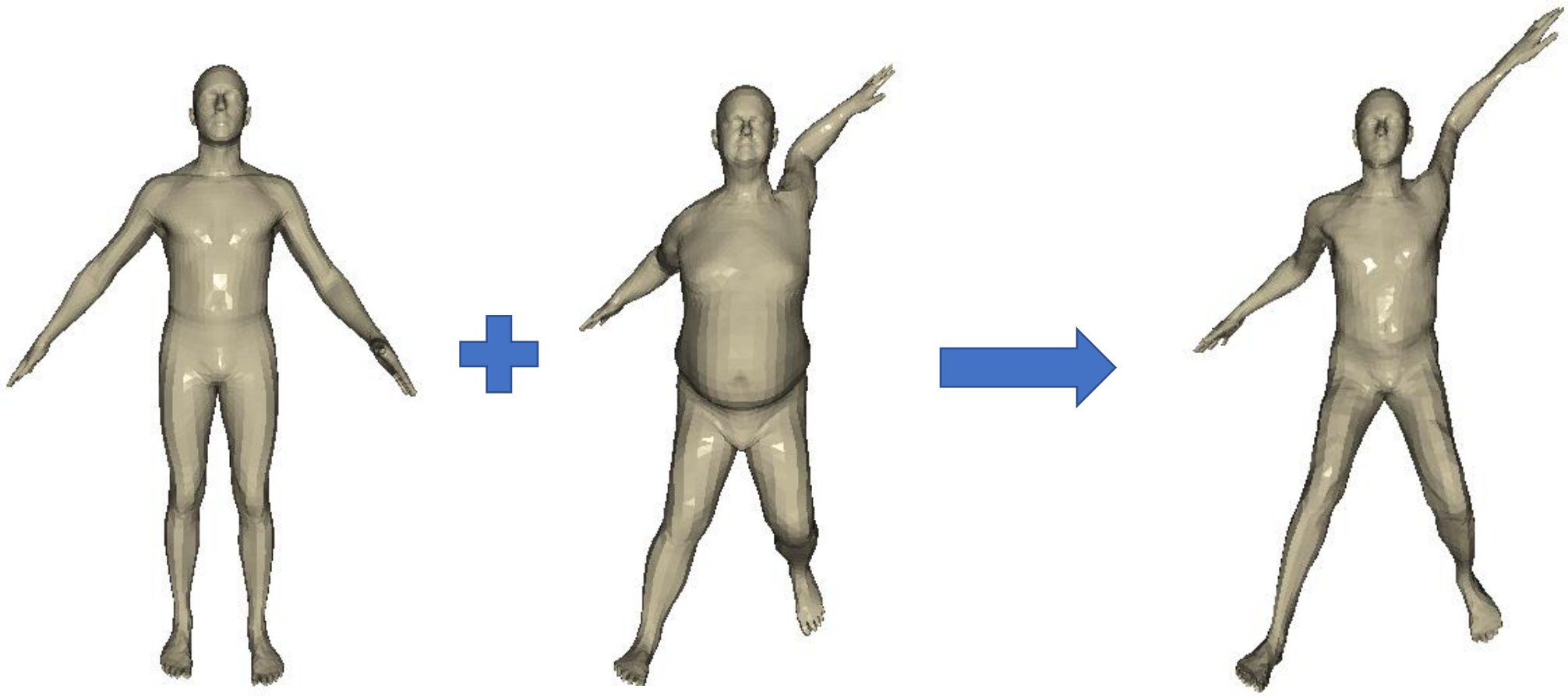
Chaoyue Song[1], Jiacheng Wei[1], Ruibo Li[1], Fayao Liu[2], Guosheng Lin[1]

[1]Nanyang Technological University, [2]A*STAR

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

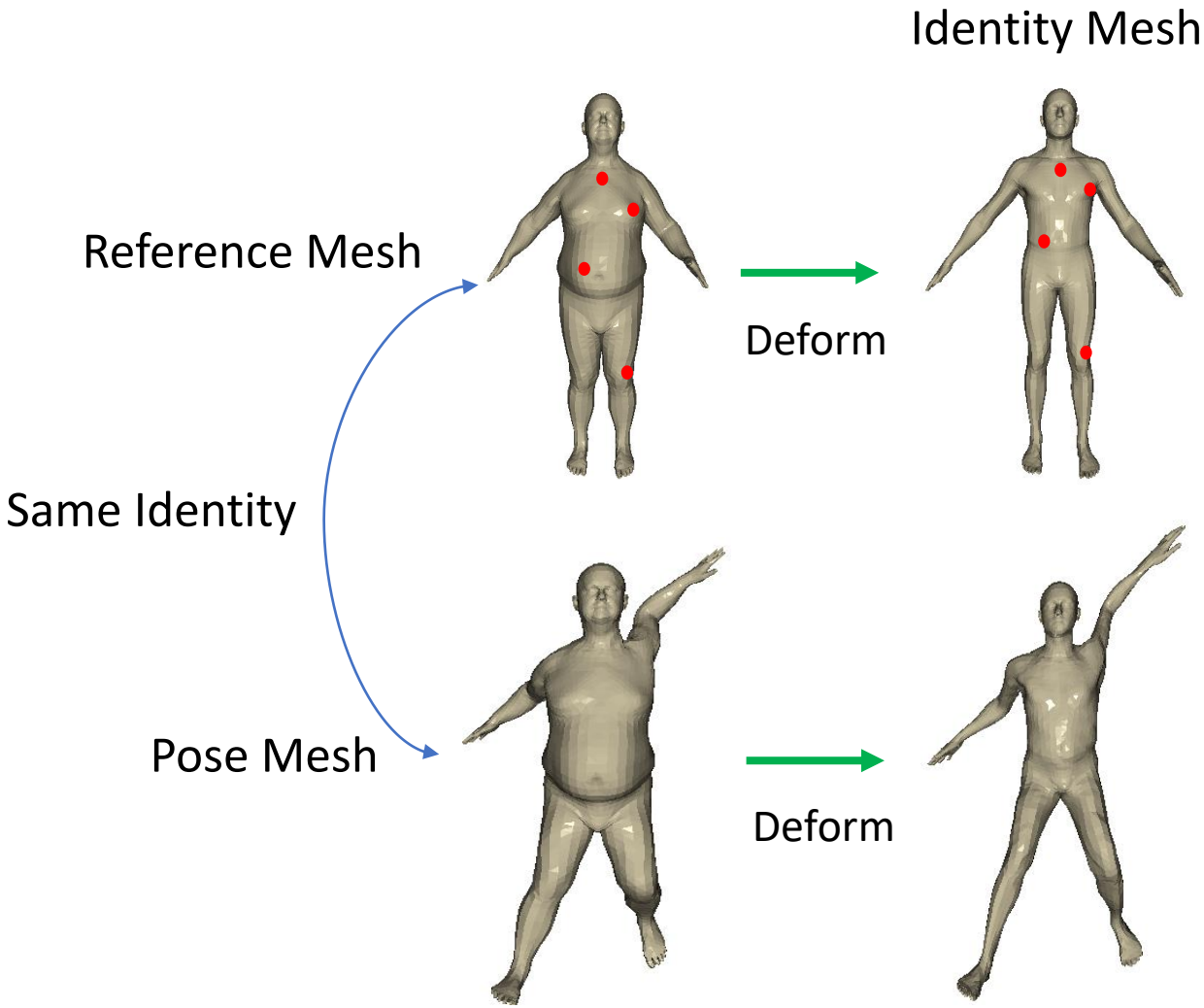
Agency for Science, Technology and Research SINGAPORE

3D Pose Transfer



Identity mesh Pose mesh Result

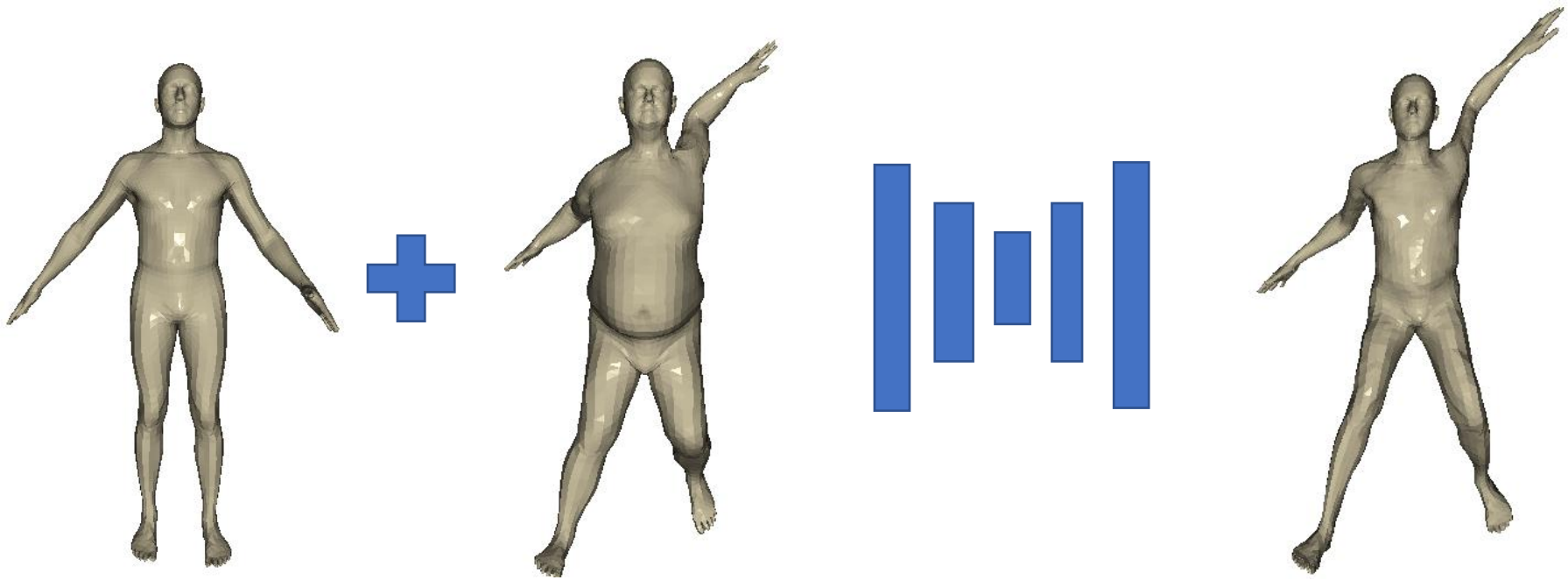
Previous work: Deformation Transfer



- Need user Labeling
- Need reference mesh
- Time-consuming

Deformation Transfer
[Sumner, et al., TOG ’04]

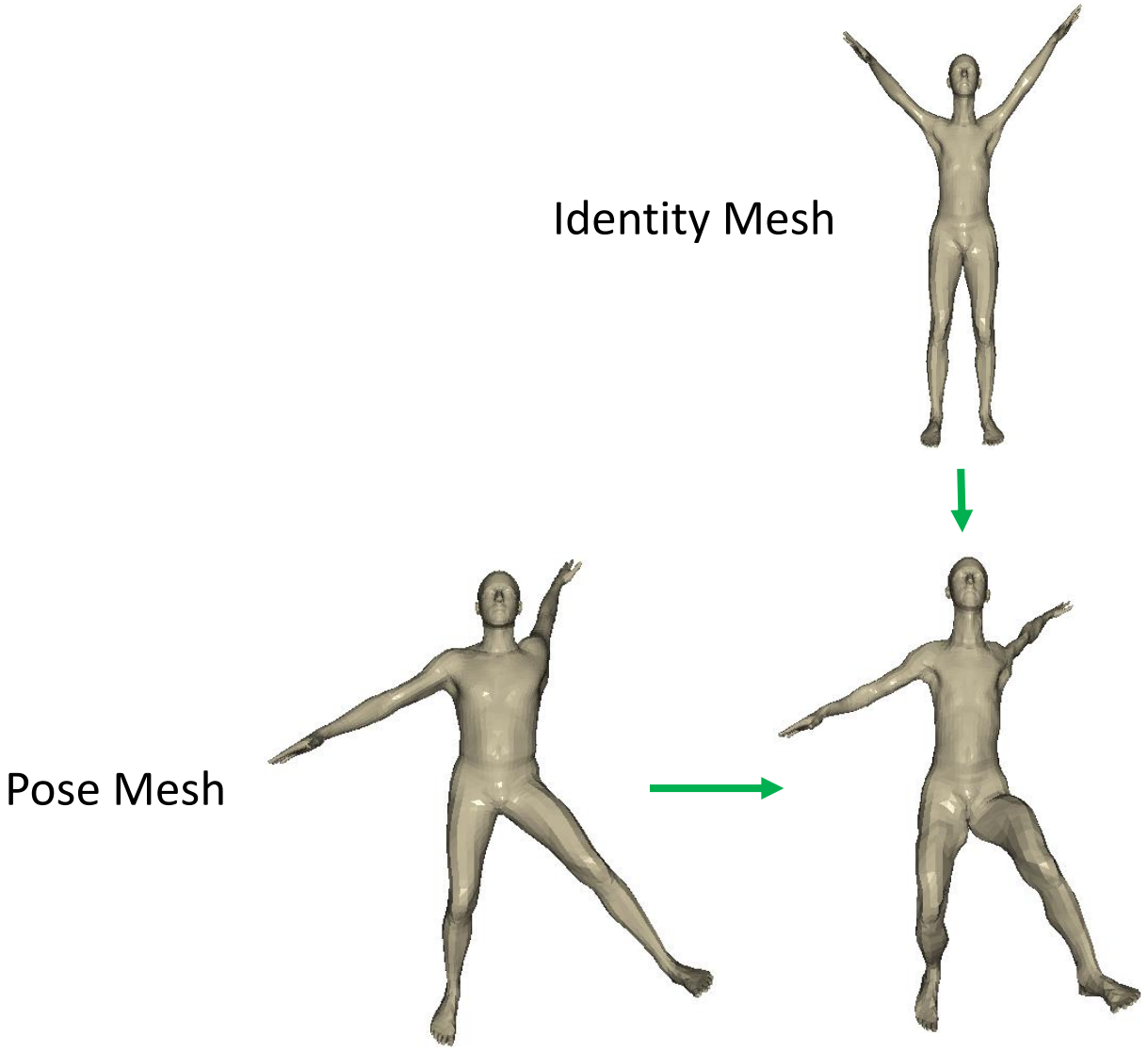
Previous works: Deep Learning-based



Identity I Pose P Neural Network f Ground truth G

Training: $\text{argmin loss}(f(I, P), G)$

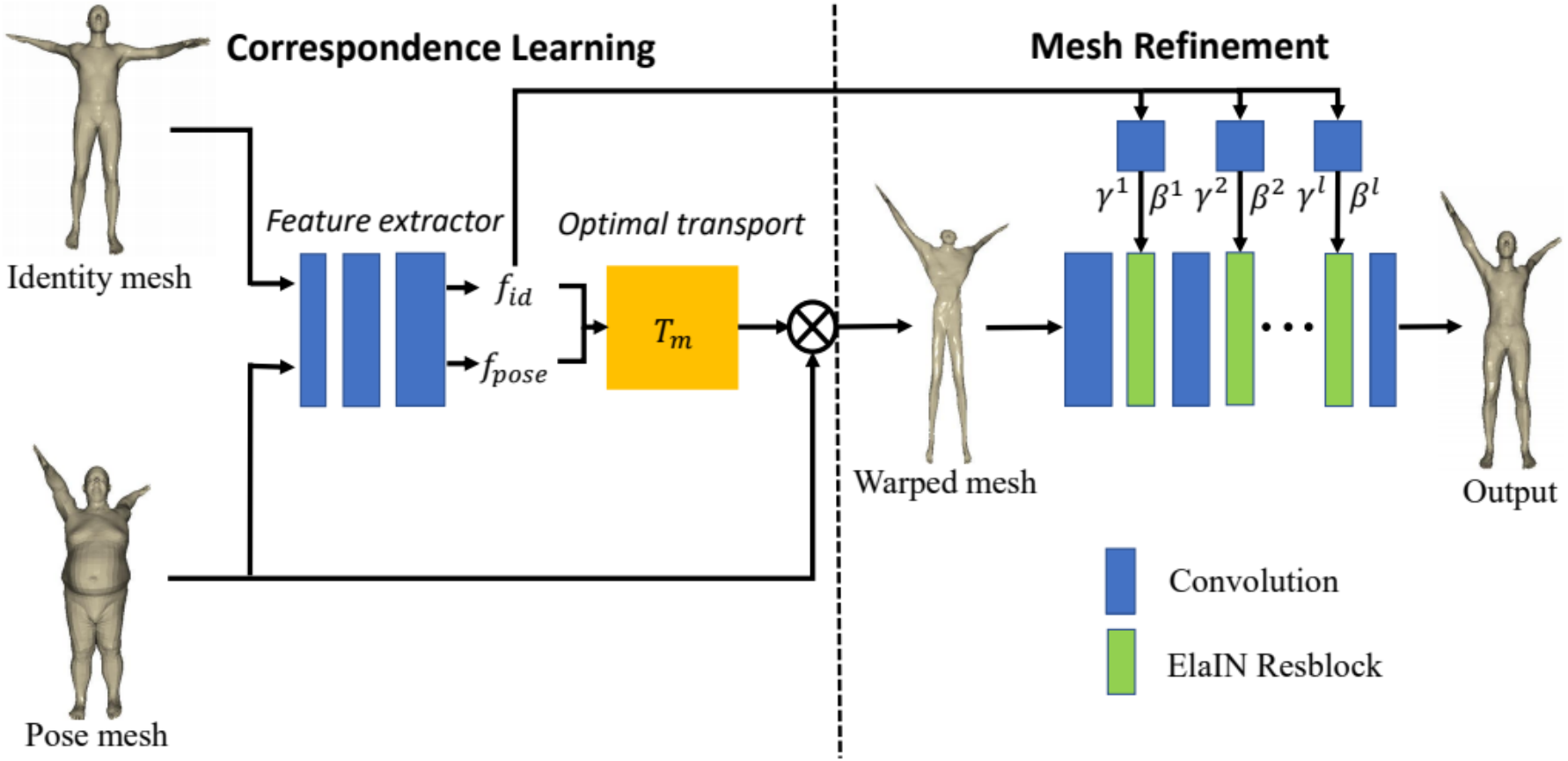
Previous works: Neural Pose Transfer



Without correspondence
→ poor generated results

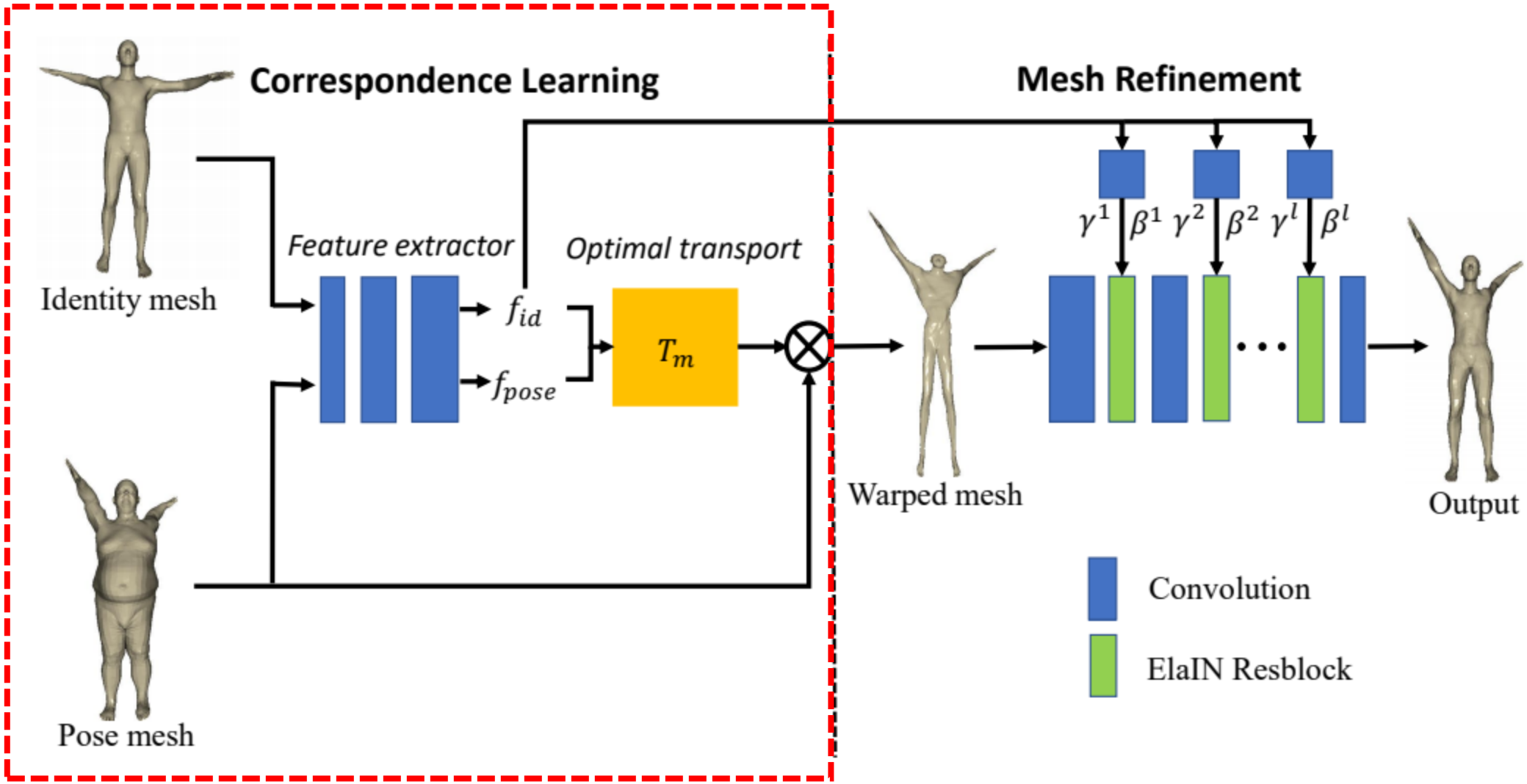
NPT [Wang, et al., CVPR ’20]

Our approach: 3D-CoreNet



√ Without labeling: learn correspondence
Fast inference
High-quality results

3D-CoreNet: Correspondence Learning



Optimal matching matrix

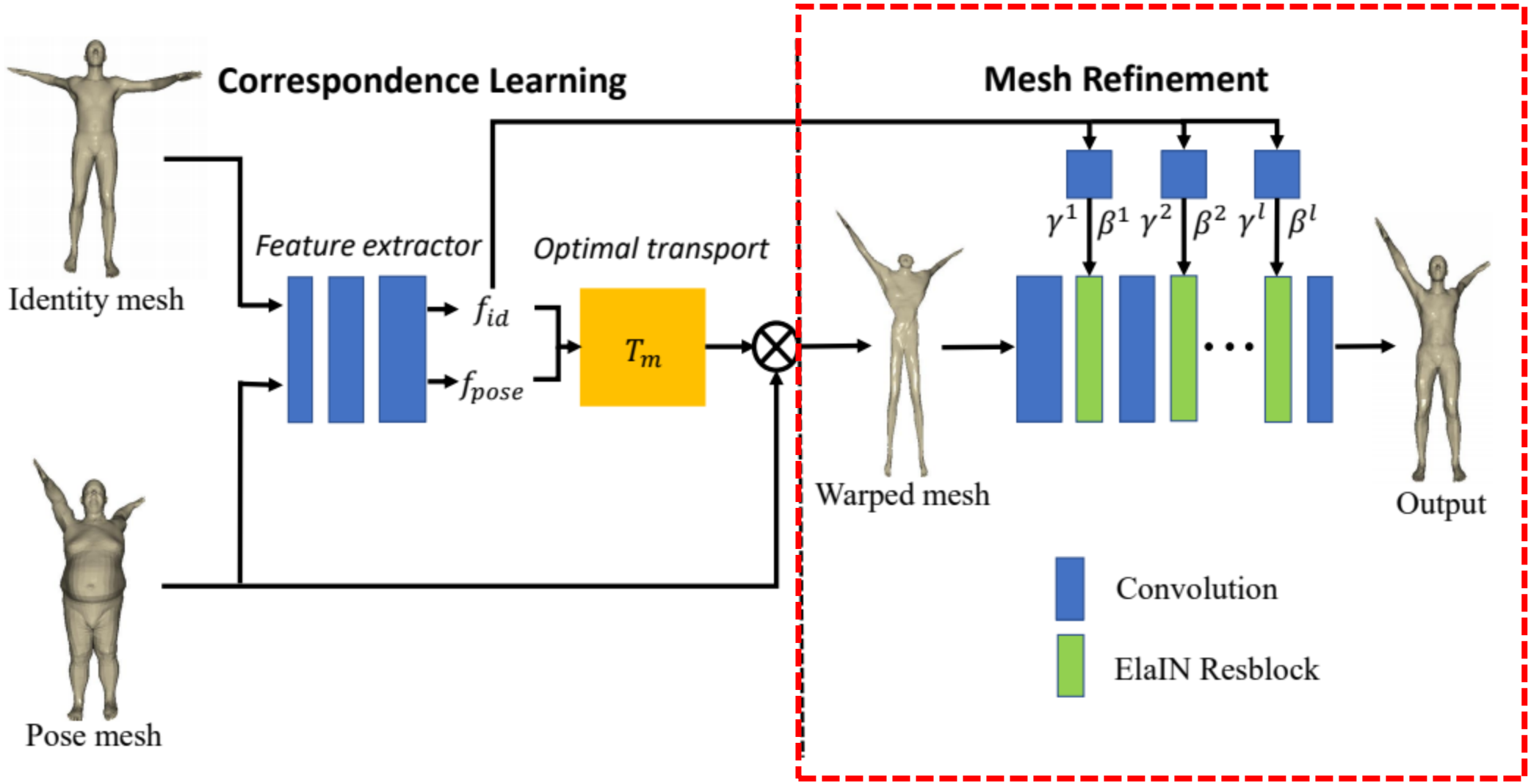
Correlation matrix $$\mathbf{C}(i,j) = \frac{f_{id}(i)^\top f_{pose}(j)}{\|f_{id}(i)\| \, \|f_{pose}(j)\|}$$

Optimal matching matrix $$\mathbf{T}_m = \underset{\mathbf{T} \in \mathbb{R}_+^{N_{id} \times N_{pose}}}{\arg\min} \sum_{ij} \mathbf{Z}(i,j)\mathbf{T}(i,j)$$

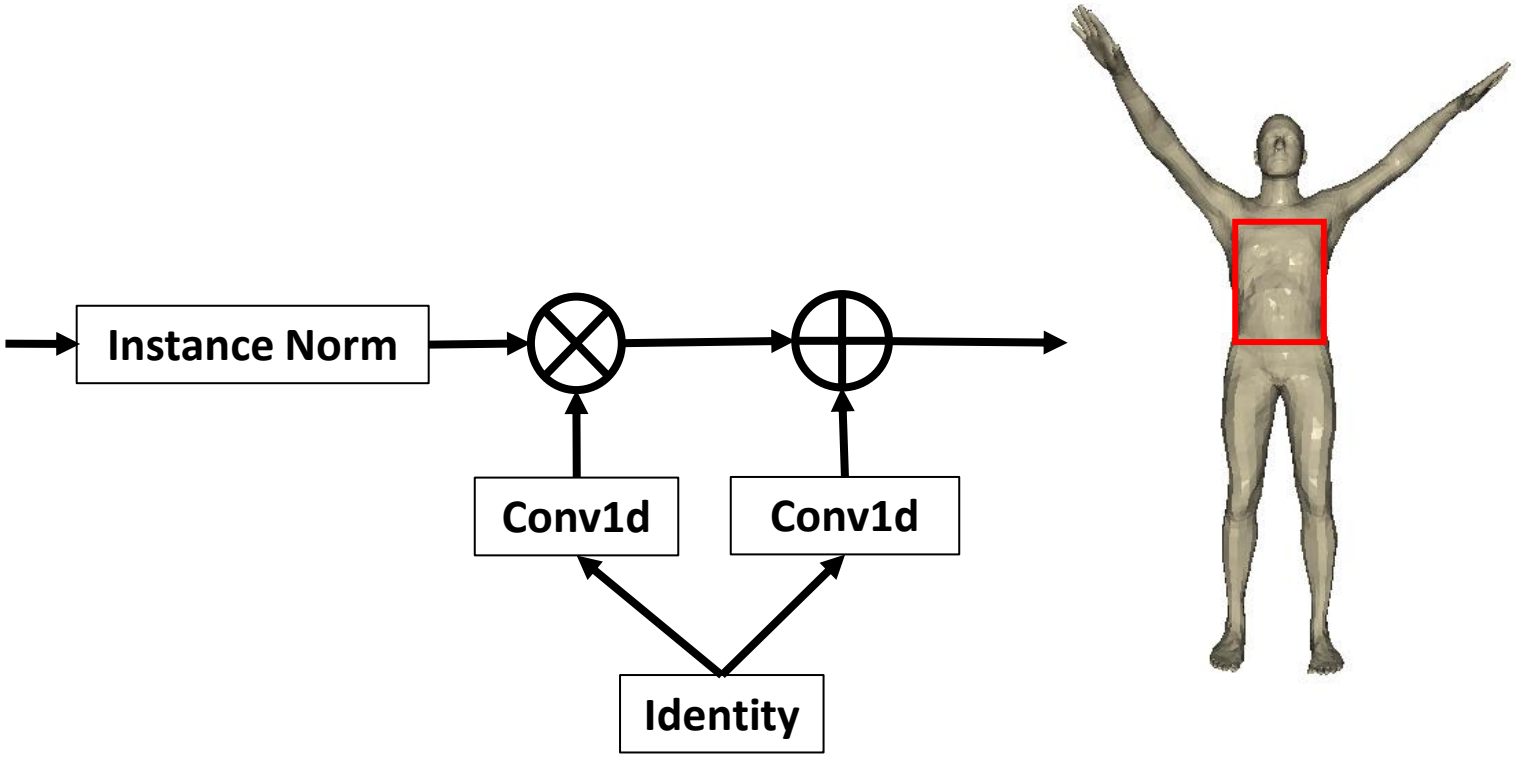
$$s.t. \quad \mathbf{T}\mathbf{1}_{N_{pose}} = \mathbf{1}_{N_{id}} N_{id}^{-1}, \quad \mathbf{T}^\top \mathbf{1}_{N_{id}} = \mathbf{1}_{N_{pose}} N_{pose}^{-1}.$$

Cost matrix $$\mathbf{Z} = 1 - \mathbf{C}$$

3D-CoreNet: Mesh Refinement



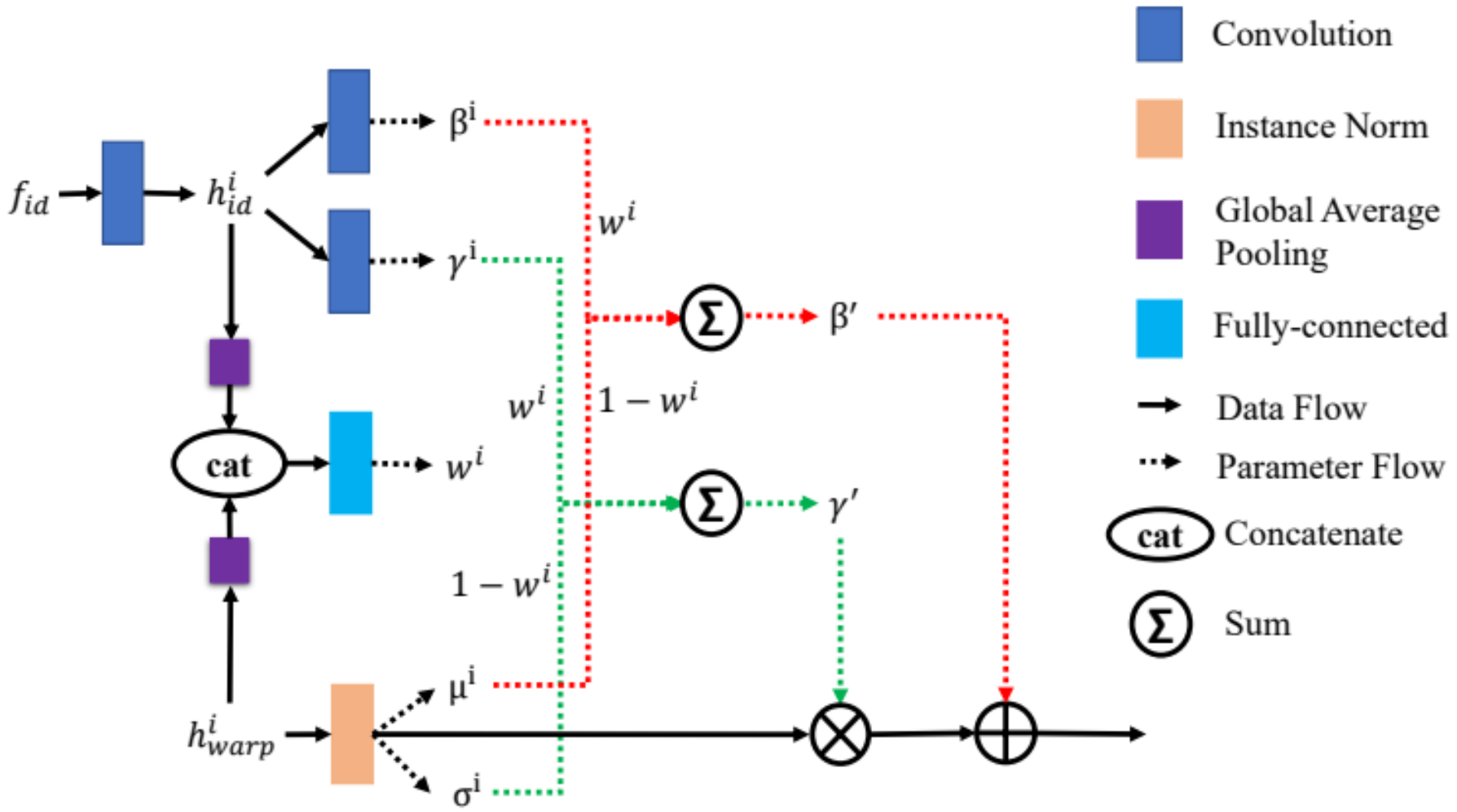
Existing conditional normalization



SPAdaIN in NPT
[Wang, et al., CVPR ’20]

AdaIN [Huang, et al., ICCV ’17]
SPADE [Park, et al., CVPR ’19]

Elastic Instance Normalization (ElaIN)



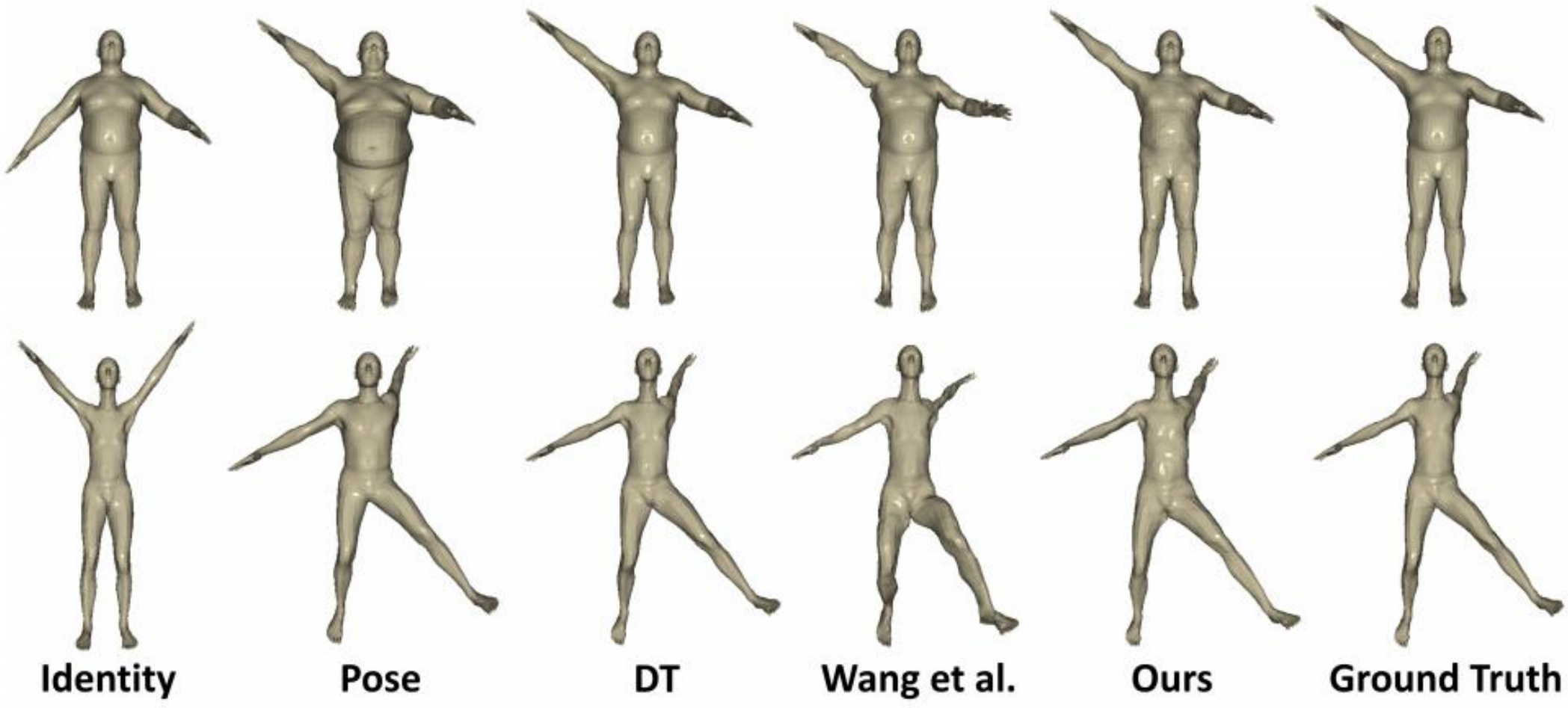
Quantitative comparison with other methods

	Annotation	Dataset	PMD	CD	EMD
DT [34]	Key points		0.15	0.35	2.21
Wang et al. [37]	-	SMPL [23]	0.66	1.42	4.22
Ours	-		**0.08**	**0.22**	**1.89**
DT [34]	Key points		13.37	35.77	15.90
Wang et al. [37]	-	SMAL [44]	6.75	14.52	11.65
Ours	-		**2.26**	**4.05**	**7.28**

Average inference time

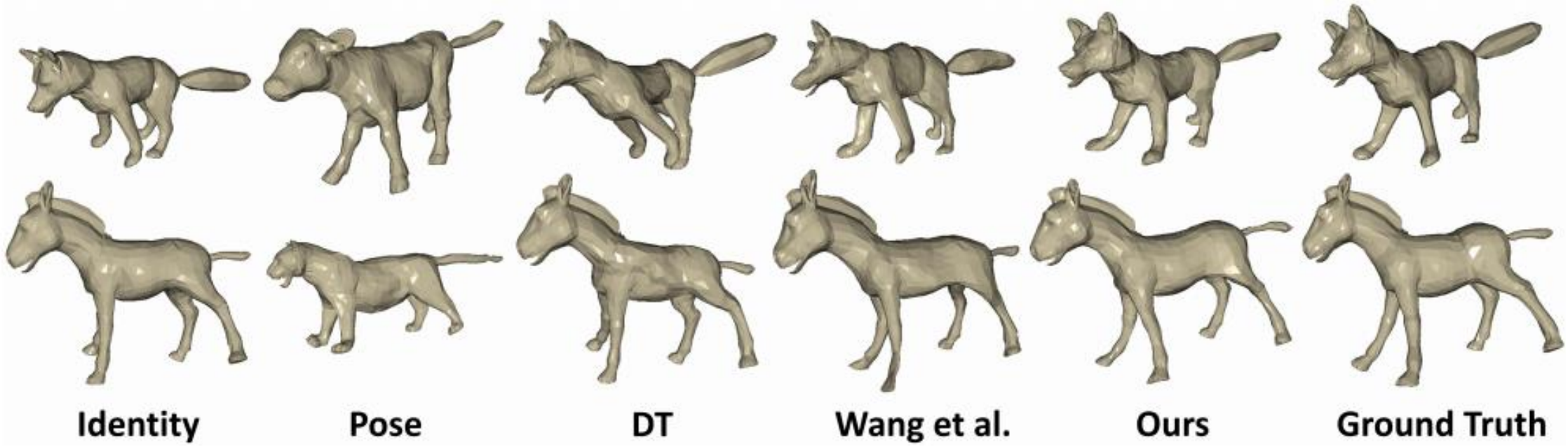
Method	DT	Wang et al.	3D-CoreNet (C)	3D-CoreNet ($\mathbf{T}_m$)
Time	3.3352s	0.0068s	0.0124s	0.0131s

Comparison with other methods



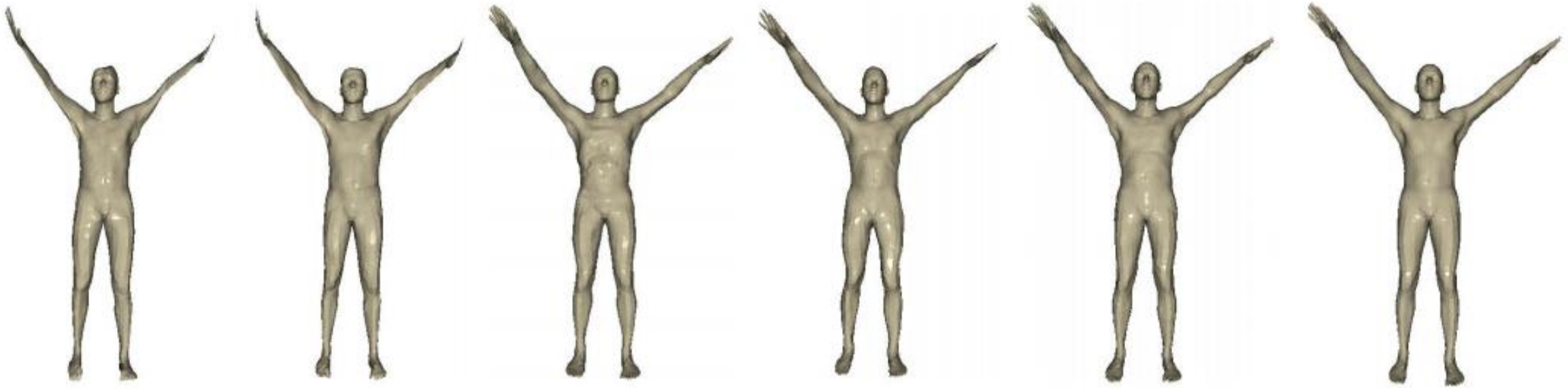
Dataset: SMPL [Loper, et al., TOG ’15]

Comparison with other methods



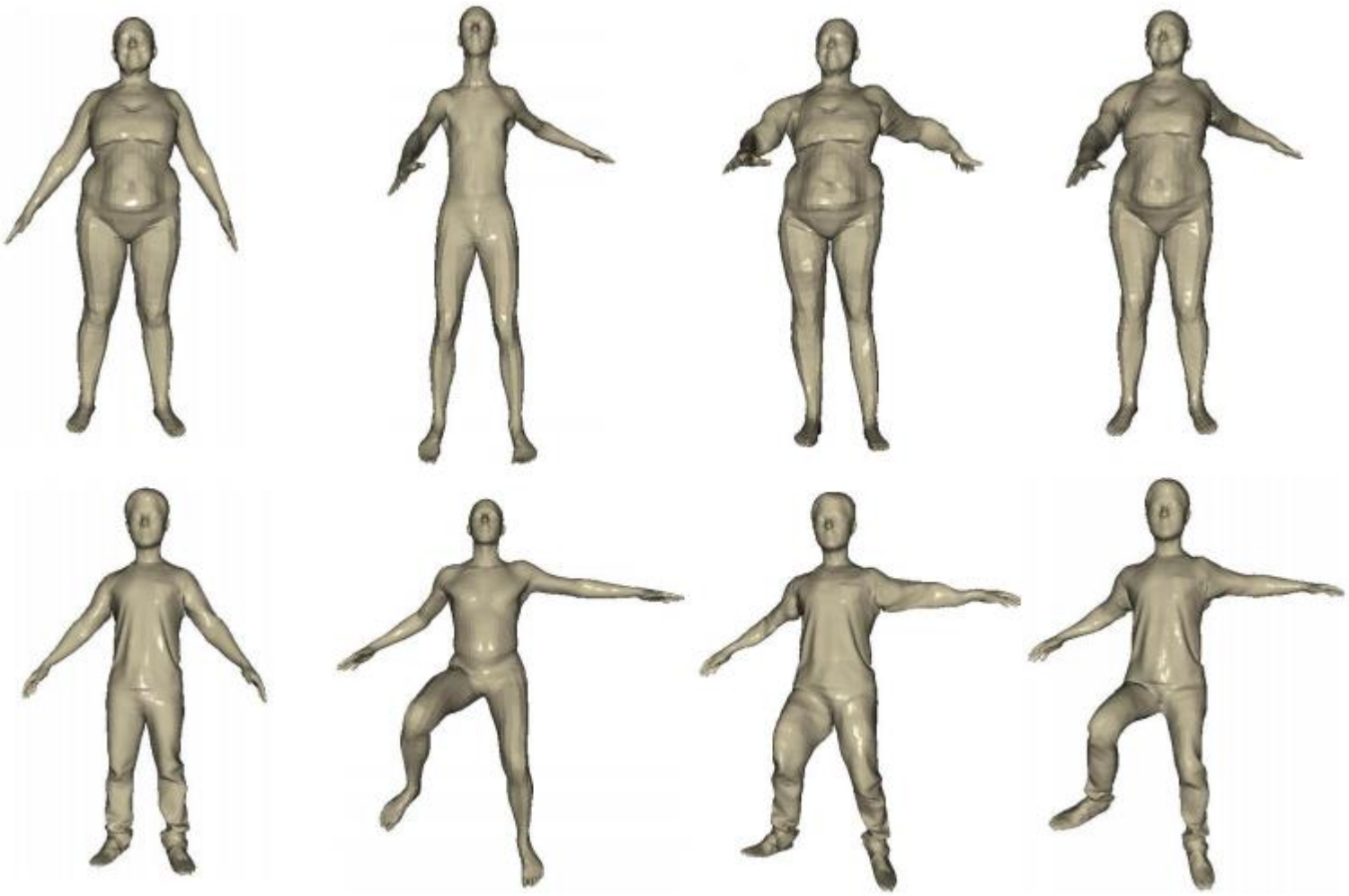
Dataset: SMAL [Zuffi, et al., CVPR ’17]

Ablation study



Dataset		Corr (a)	Corr (b)	w/o ElaIN	w/o $\mathcal{L}_{edg}$	Full model
SMPL [23]	PMD	0.46	0.44	0.15	0.14	**0.08**
	CD	1.39	1.28	0.37	0.34	**0.22**
	EMD	3.49	3.42	2.57	2.28	**1.89**

Generalization capability



Identity Pose Wang et al. Ours

Dataset: FAUST: 6890 vertices
[Bogo, et al., CVPR ’14]
MG-Dataset: 27554 vertices
[Bhatnagar, et al., ICCV ’19]

THANK YOU!

Website: https://chaoyuesong.github.io/3d-corenet

Code will be available at: https://github.com/ChaoyueSong/3d-corenet