mfp0438

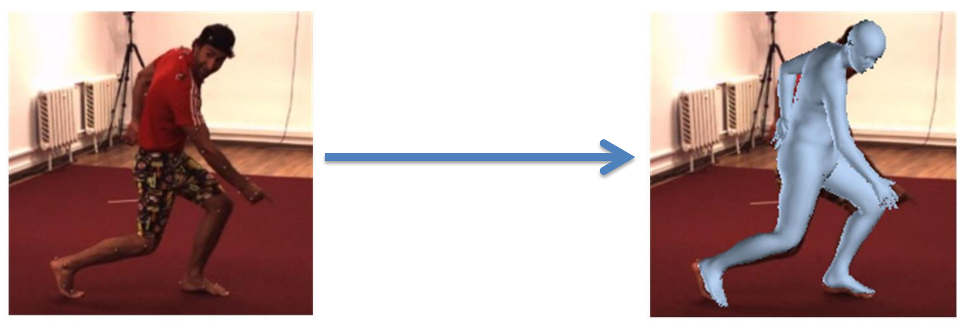
DC-GNet: Deep Mesh Relation Capturing Graph Convolution Network for 3D Human Shape Reconstruction

Shihao Zhou, Mengxi Jiang, Shanshan Cai, Yunqi Lei
Department of Computer Science, Xiamen University

Introduction

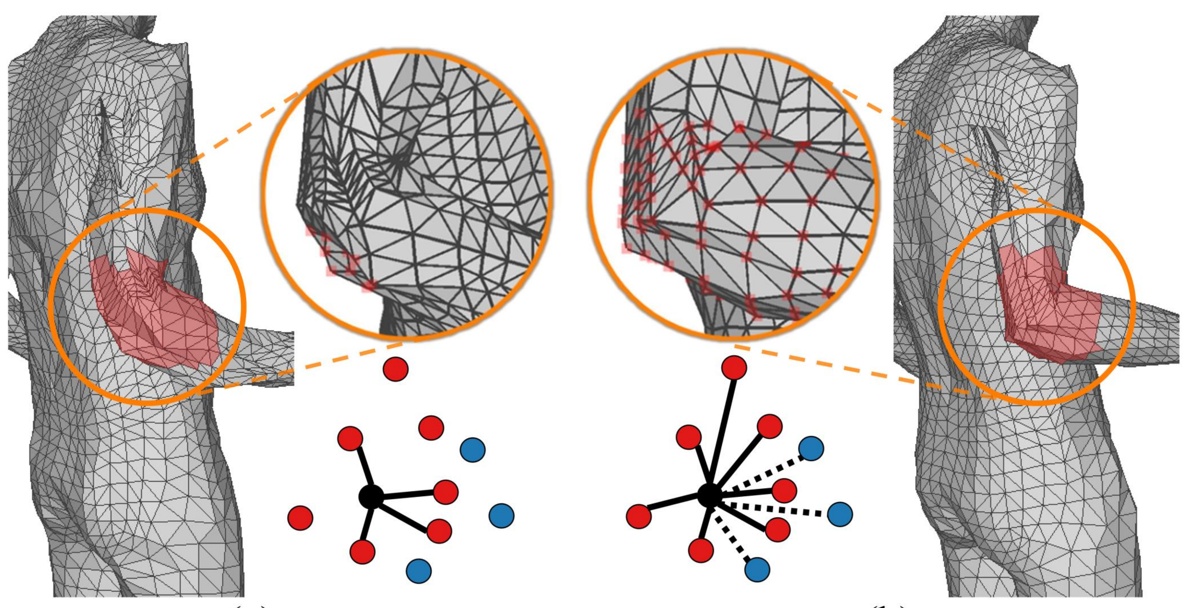
- We propose a Deep Mesh Relation Capturing Graph Convolution Network, DC-GNet, with a shape completion task for 3D human shape reconstruction

Input **Output**

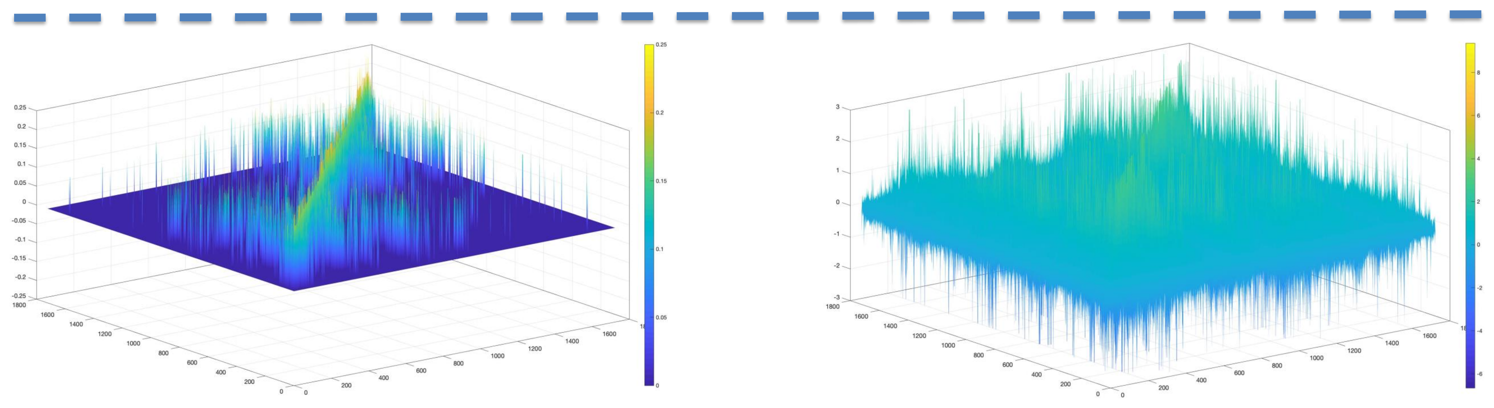


Motivation

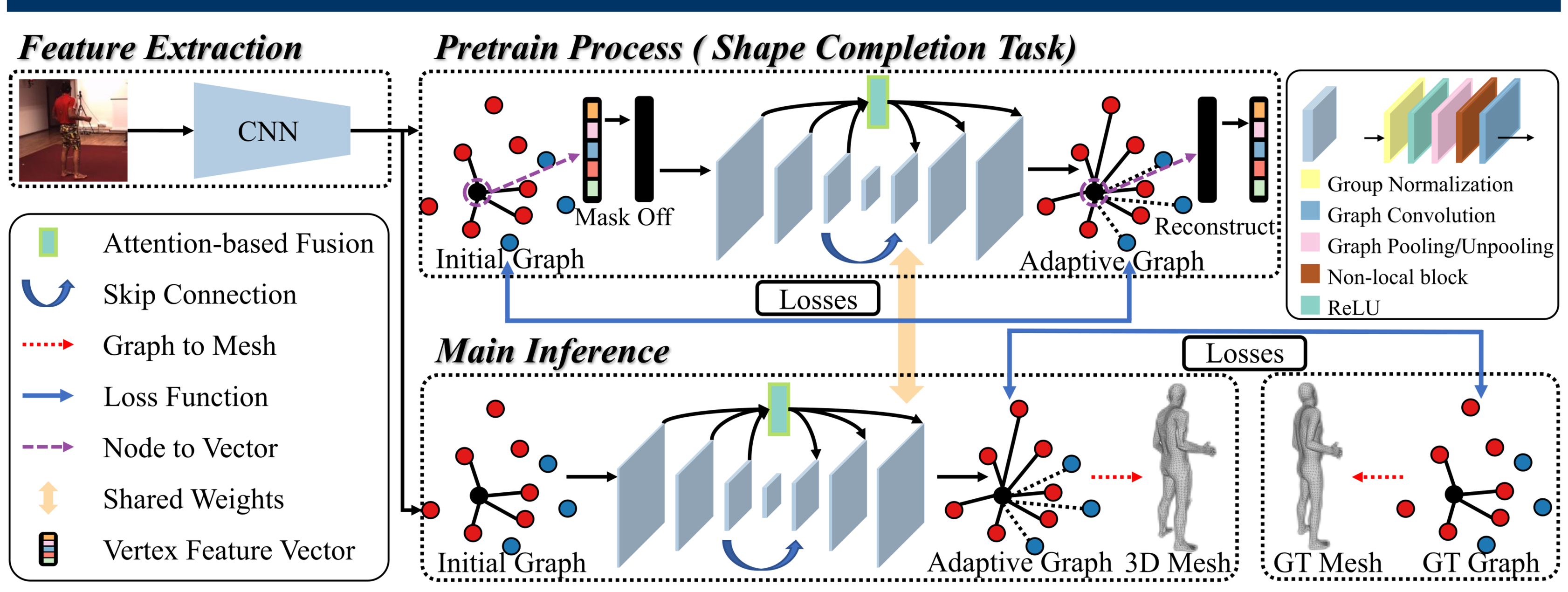
- A good mesh representation for human shape needs to capture or learn:
  - Deep relations within mesh vertices
  - prior about various kinds of occlusion cases
- However:
  - Existing human shape reconstruction methods focus on a fixed adjacency matrix to encode the inherent shape nodes relations.
  - Most approaches trained on limited indoor datasets meet the performance degradation under real-world scenes.



(a) (b)



Methods



- ***A: Feature Extraction***
  - *An image-based convolutional network is applied as a feature extractor and outputs a 2048-D feature vector for every single vertex in the human mesh representation.*
- ***B: Deep Relations Capture***
  - *Instead of using a pre-defined adjacency matrix, an adaptive adjacency matrix is introduced to learn subtle relationships between nonadjacent nodes.*

$$X_{out} = \sigma(AX_{in}W)$$

A : *a pre-defined adjacency matrix of the human mesh*

$$X_{out} = \sigma(\hat{A}X_{in}W)$$

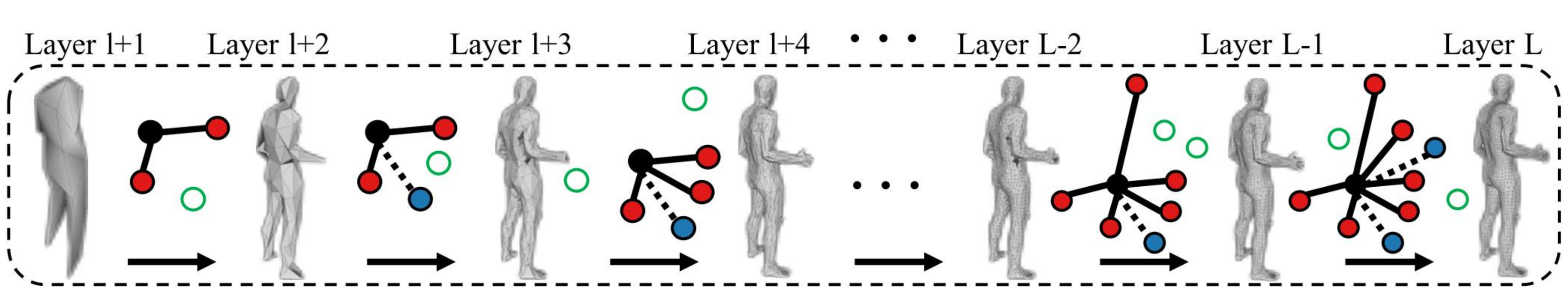
$\hat{A}$: *a learnable adjacency matrix*

  - *Encoder: the high redundancy and memory requirements are both reduced.*

$$Y_l = f(Y_{l-1})$$

  - *Decoder: the body shape is gradually refined when more features are fused.*

$$Y_{l+1} = f([f(Y_l); f(m(Y_1, \dots, Y_l)); Y_{L-l}])$$



Methods

- ***C: Shape Completion Task***
  - *In order to enable the network to learn a generic adjacency matrix for various occlusion cases, a shape completion task in which a mask off and a reconstruction part are included, is proposed.*

C.1 Mask Off Part

  - *Simulating the occlusion cases by fabricating artificial holes on the surface of the initial human mesh.*

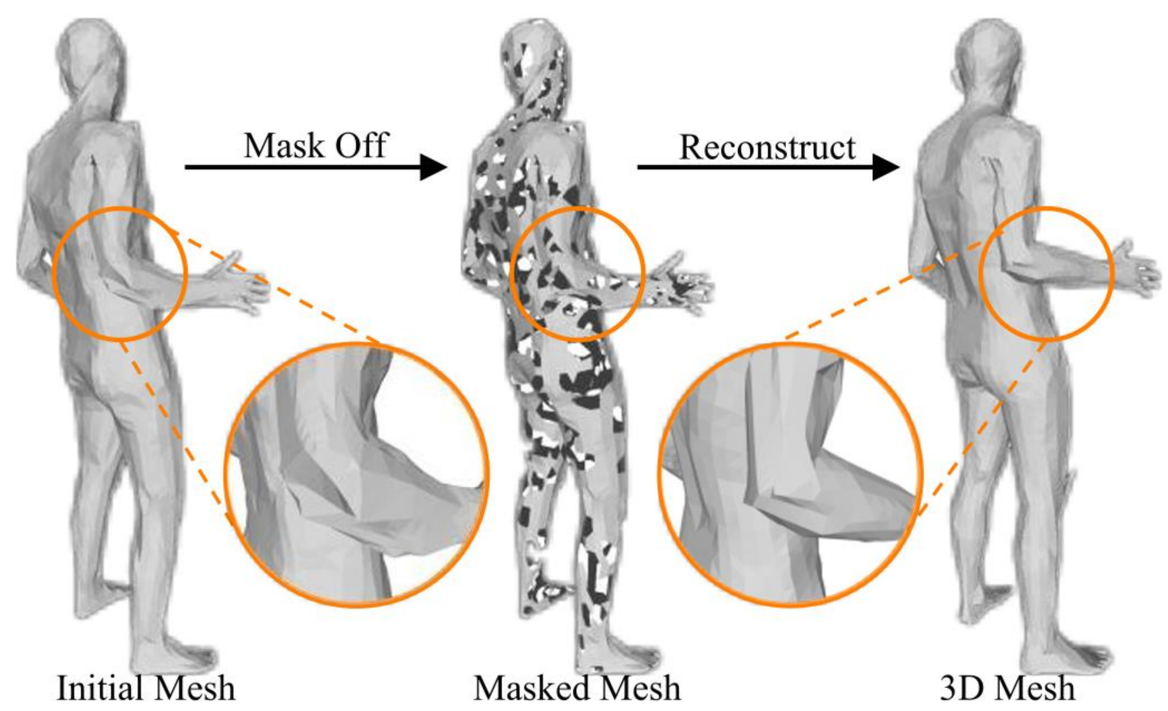
$$\hat{X}_{in} = X_{in} \cdot M$$

M : *a matrix of ones except c row set zeros, and randomly shuffled before dot product.*

C.2 Reconstruction Part

  - *force the network to recover the missing information for the masked human mesh*

$$X_{in} = \sigma(\hat{A}\hat{X}_{in}W)$$



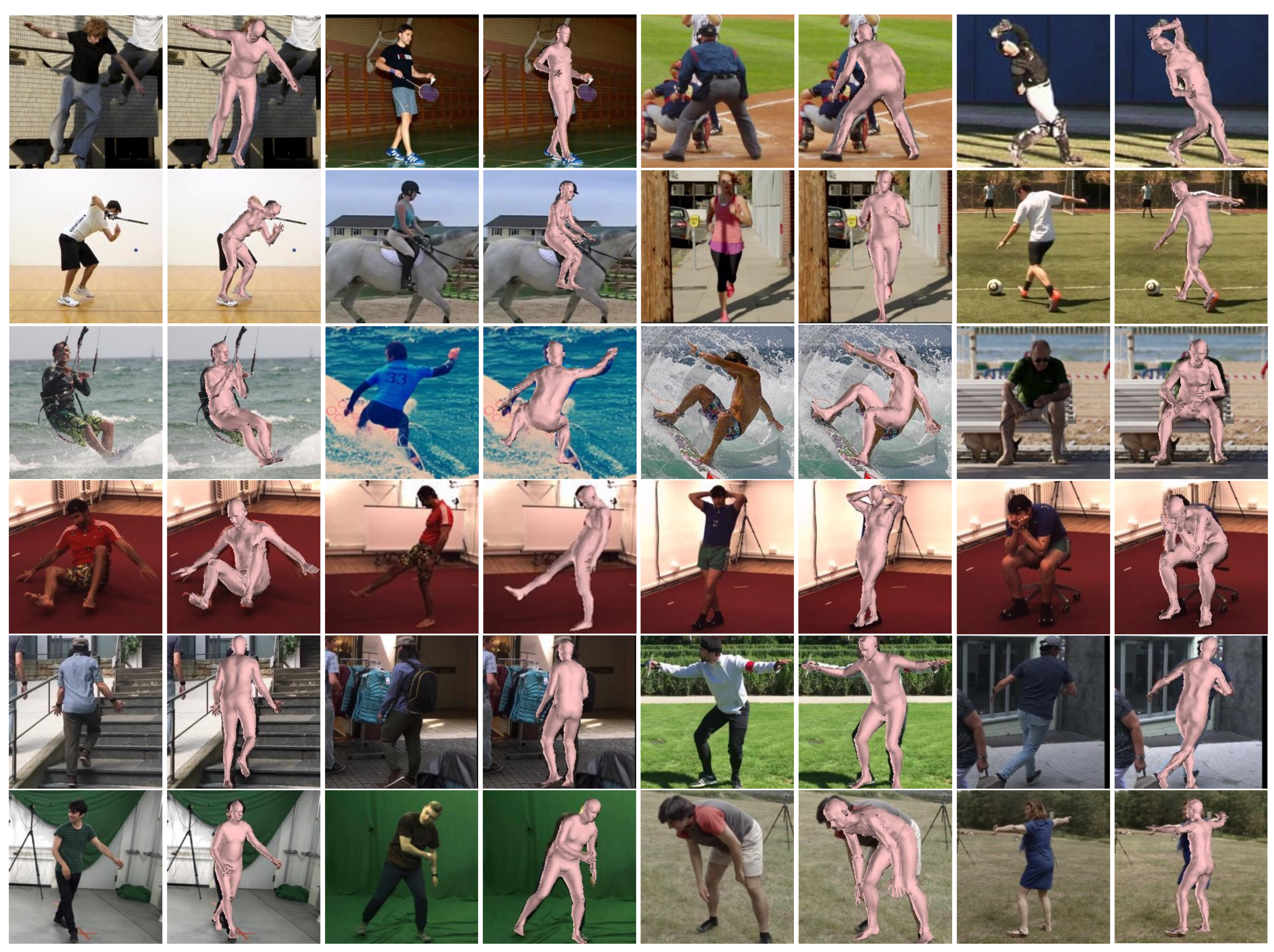
Results

- MPJPE and mean reconstruct error of DC-GNet on Human3.6M, and additionally AUC over a range of 3D-PCK thresholds(150mm) on MPI-INF-3DHP, and mean reconstruct on 3DPW.

Method	MPI-INF-3DHP			Human3.6M	
	AUC ↑	MPJPE ↓	Reconst.Error ↓	MPJPE ↓	Reconst.Error ↓
HMR [17] (CVPR'18)	36.5	124.2	89.8	-	56.8
†HMMR [18] (CVPR'19)	-	-	-	-	56.9
†Arnab *et al.* [3] (CVPR'19)	-	-	-	77.8	54.3
CMR [22] (CVPR'19)	24.3	152.0	83.8	71.9	50.1
†TexturePose [36] (ICCV'19)	-	-	-	-	49.7
SPIN [21] (ICCV'19)	37.1	105.2	67.5	-	41.1
DaNet [56] (ACM MM'19)	-	-	-	61.5	48.6
Jiang *et al.* [15] (CVPR'20)	-	-	-	-	52.7
Kundu *et al.* [24] (ECCV'20)	-	-	-	-	48.1
Pose2Mesh [7] (ECCV'20)	-	-	-	64.9	47.0
†VIBE [20] (CVPR'20)	-	97.7	63.4	65.9	41.5
DecoMR [55] (CVPR'20)	-	102.0	65.9	**60.6**	39.3
DC-GNet	**40.7**	**97.2**	**62.5**	63.9	42.4

Method	Reconst.Error
HMR [13] (CVPR'18)*	81.3
CMR [18] (CVPR'19)*	70.2
Arnab *et al.* [4] (CVPR'19)	72.2
SPIN [17] (CVPR'19)	59.2
Sun *et al.* [25] (ICCV'19)	69.5
DecoMR [28] (CVPR'20	61.7
DC-GNet	**59.1**

- The qualitative analysis of DC-GNet for different datasets, including LSP, MPII, COCO, Human3.6M, 3DPW and MPI-INF-3DHP, are shown.



Conclusions

- A Deep Mesh Relation Capturing Graph Convolution Network, namely DC-GNet, is proposed to reconstruct 3D human shape from a single RGB image. It is the first attempt to learn deep relations among human mesh vertices and consider reasoning from more than partial structure, which boosts the model capturing complex local deformation.
- A shape completion module is proposed as an auxiliary task to alleviate the appearance domain gap issue between indoor and outdoor scenes.
- Extensive experimental results across several benchmarks demonstrate the effectiveness of exploring deep relations among mesh .