ECCV'20
ONLINE
23-28 AUGUST 2020

16TH EUROPEAN CONFERENCE ON

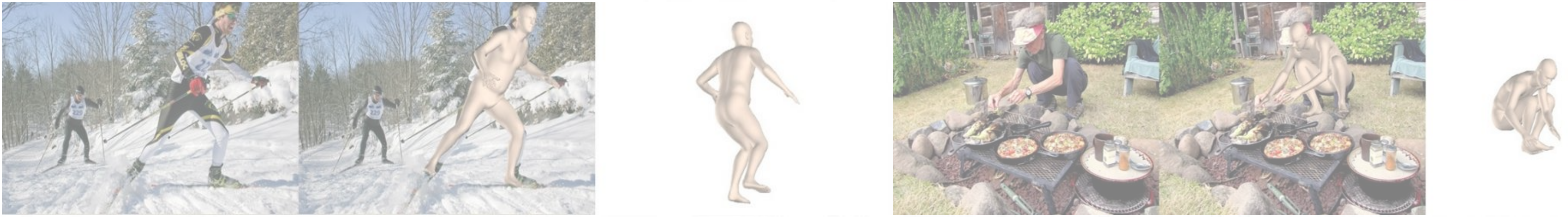
COMPUTER VISION

WWW.ECCV2020.EU

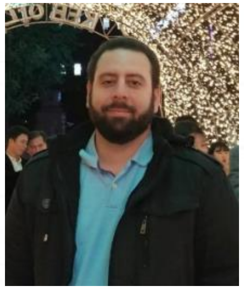
ECCV'20 ONLINE
23-28 AUGUST 2020

UNITED IMAGING | United Imaging Intelligence

GEORGE MASON UNIVERSITY



Hierarchical Kinematic Human Mesh Recovery

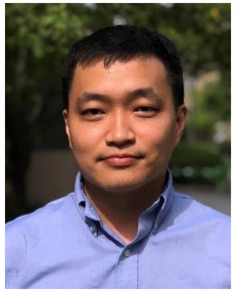

Georgios Georgakis*


Ren Li*


Srikrishna Karanam


Terrence Chen

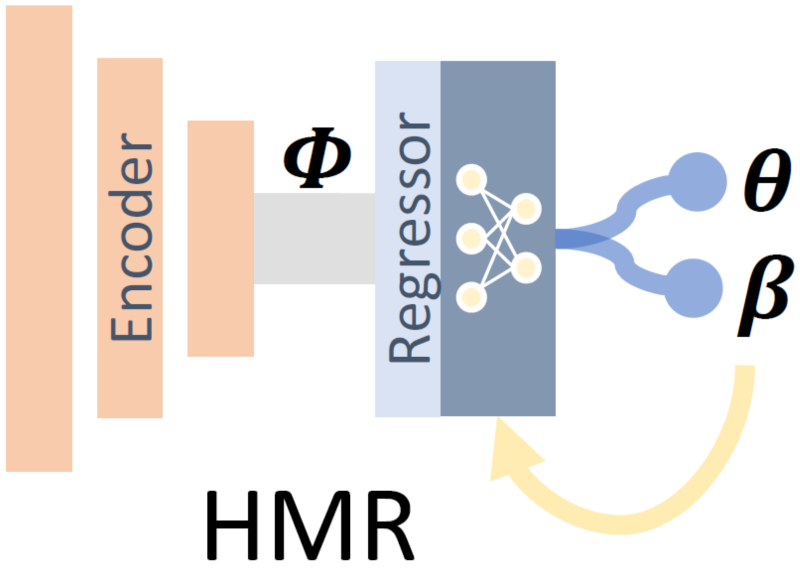

Jana Kosecka


Ziyan Wu

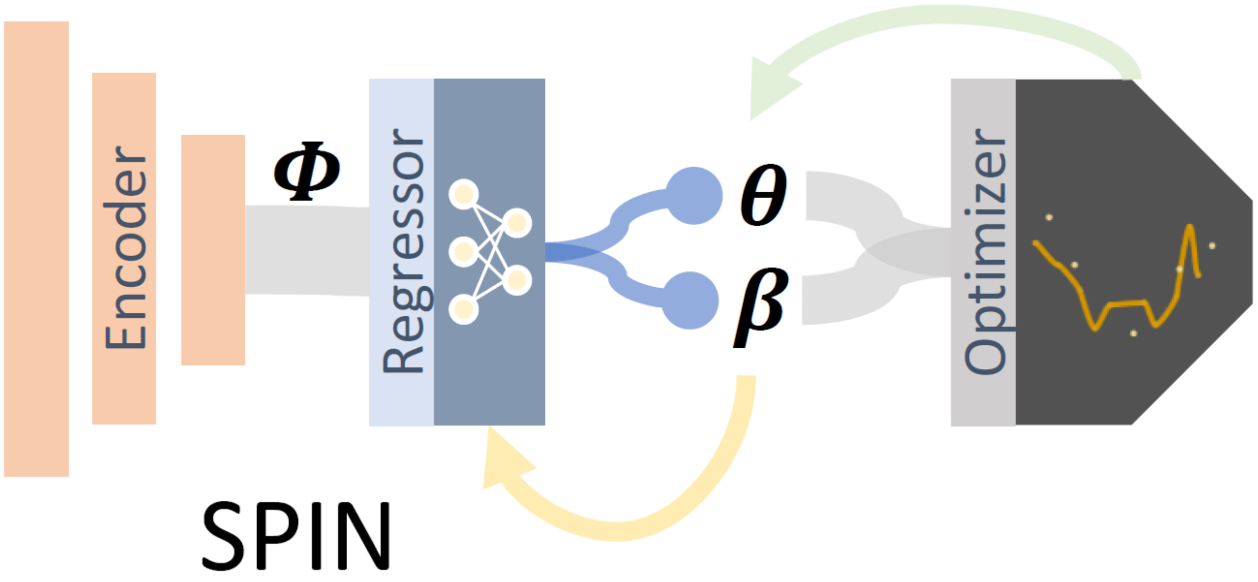
*equal contribution, work during internship with United Imaging Intelligence

Current Works

[Kanazawa et al. 2018]

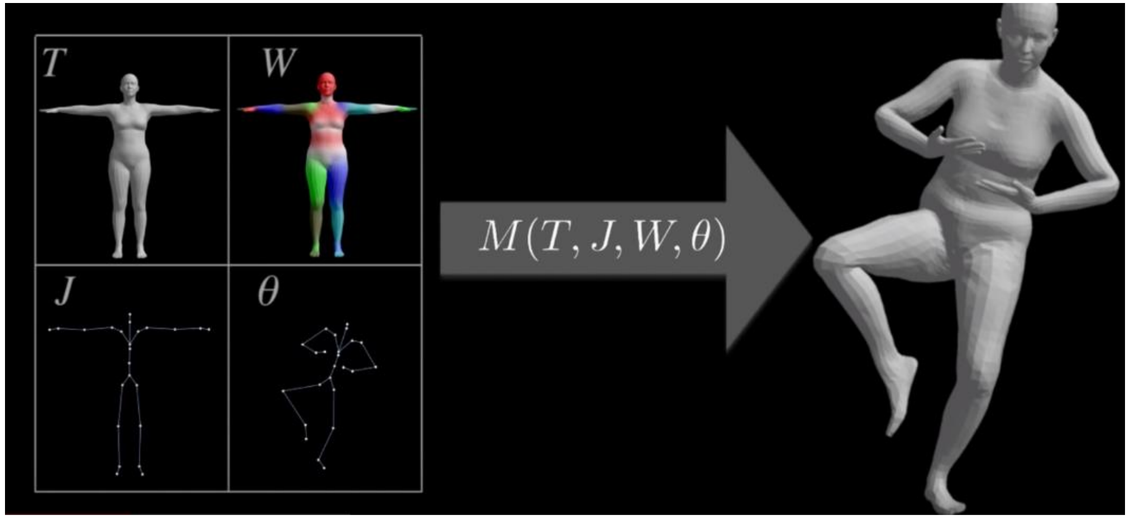


[Kolotouros et al. 2019]

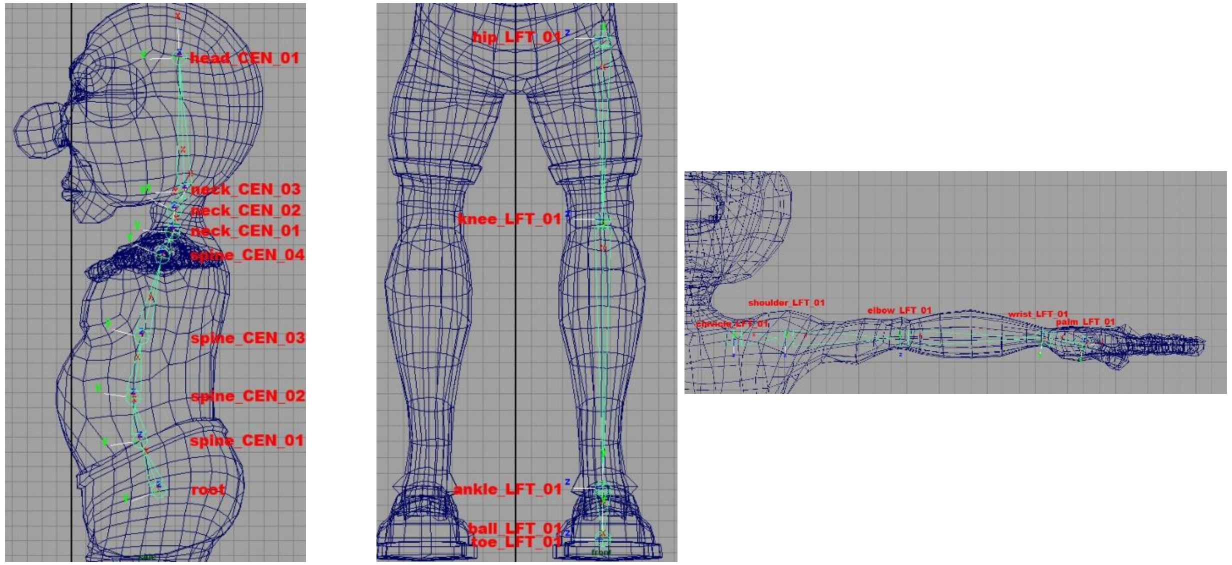


SMPL Body Model

- Differentiable generative model parameterized by pose θ and shape β.
- Hierarchical structure inspired by the standard skeletal rig.



Forward process of the SMPL model

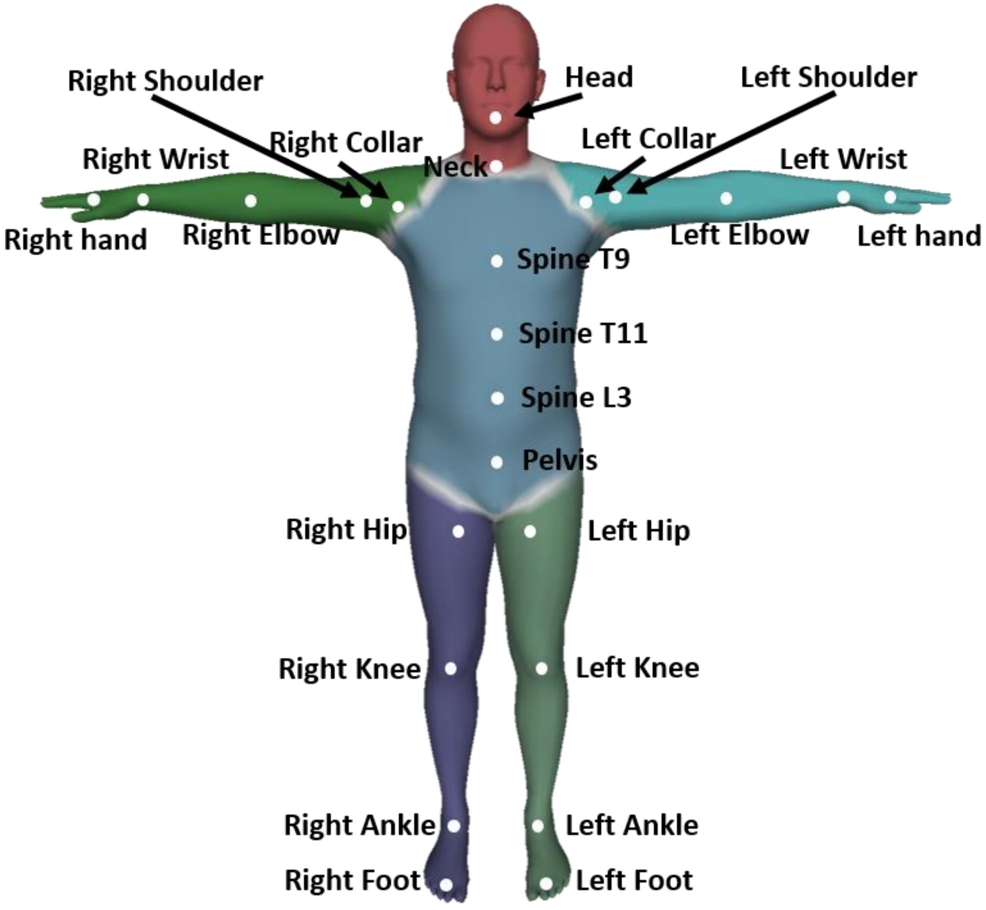


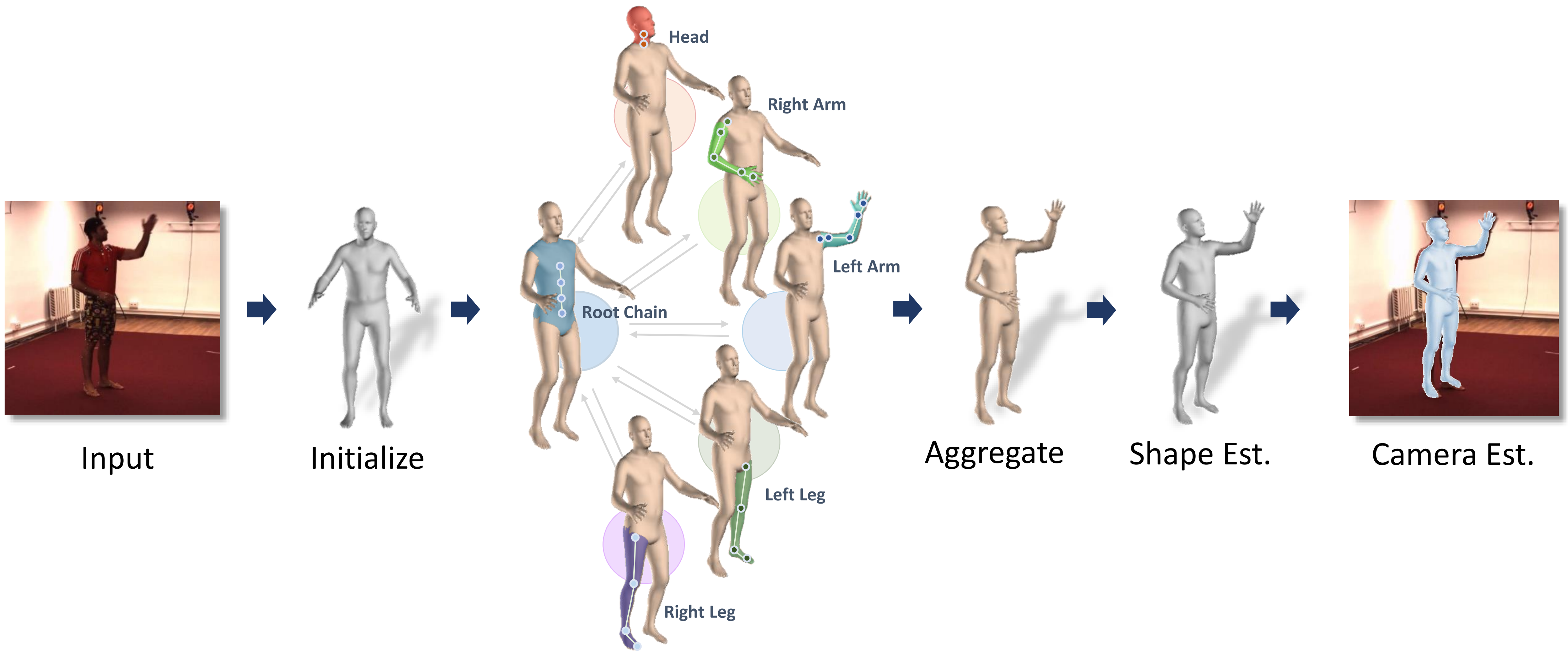
The standard skeletal rig: root first followed by other joints

Regressing 3D Rotations

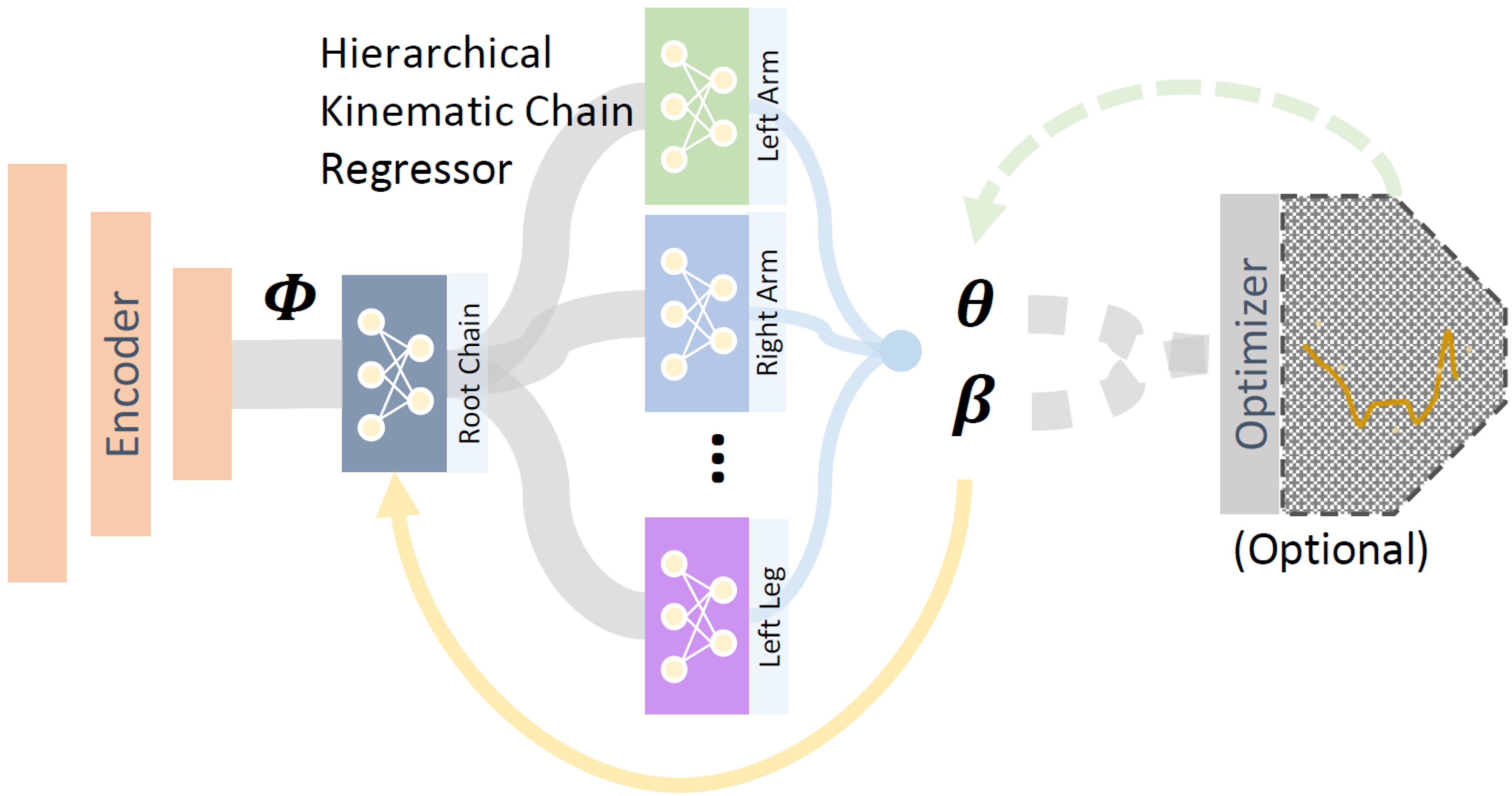
- Direct regression of rotation parameters is very challenging [Kendall et al. CVPR 2017]
  - Euler angles wrap around 2π radians.
  - Rotation matrices are overparametrized.
  - More pronounced in occlusions.

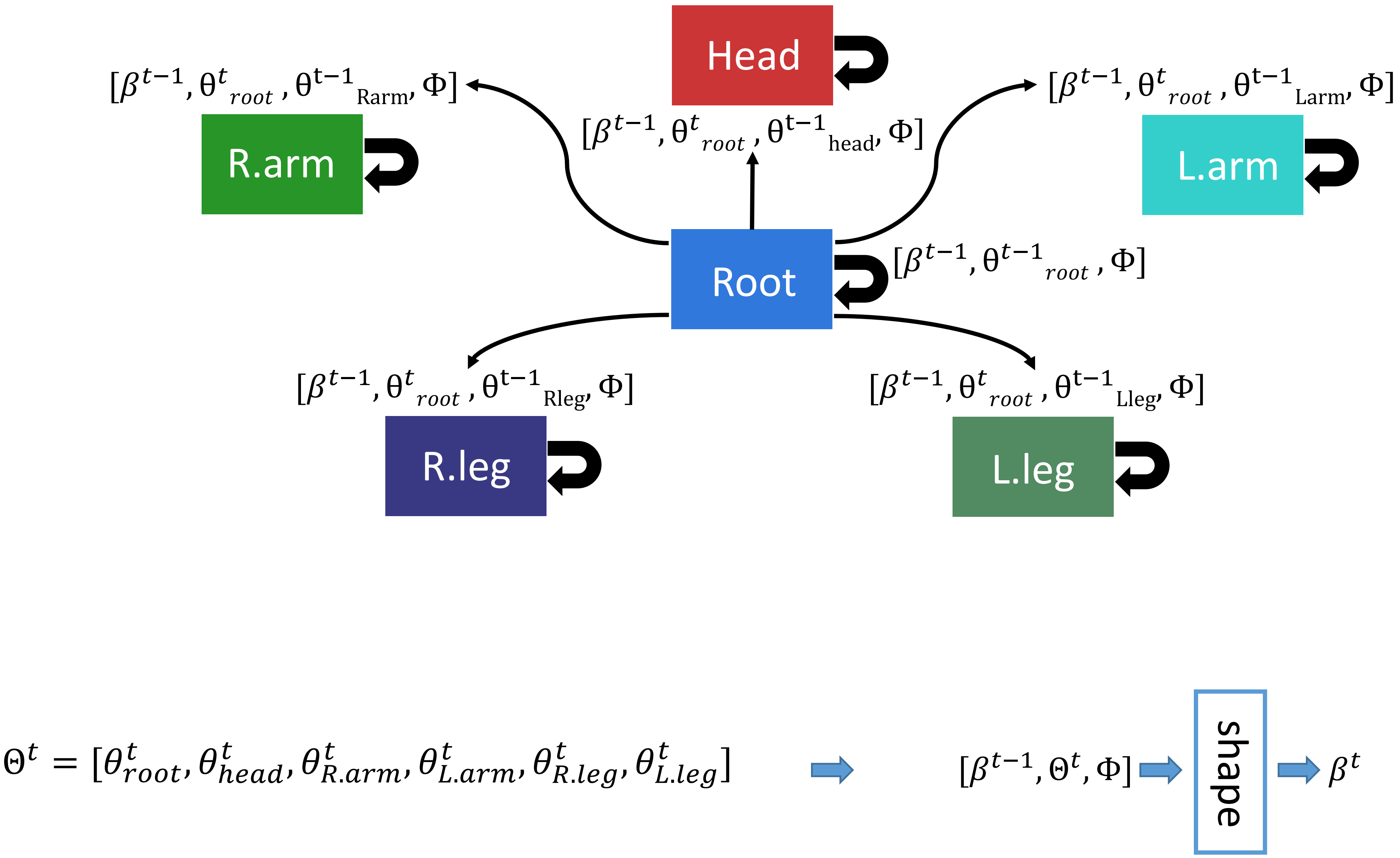
- Considering geometry when designing the regressor.
  - Geometry of human body model ➔ Modeling the interdependencies of the limbs and the joints.
  - Can help infer occluded joints.





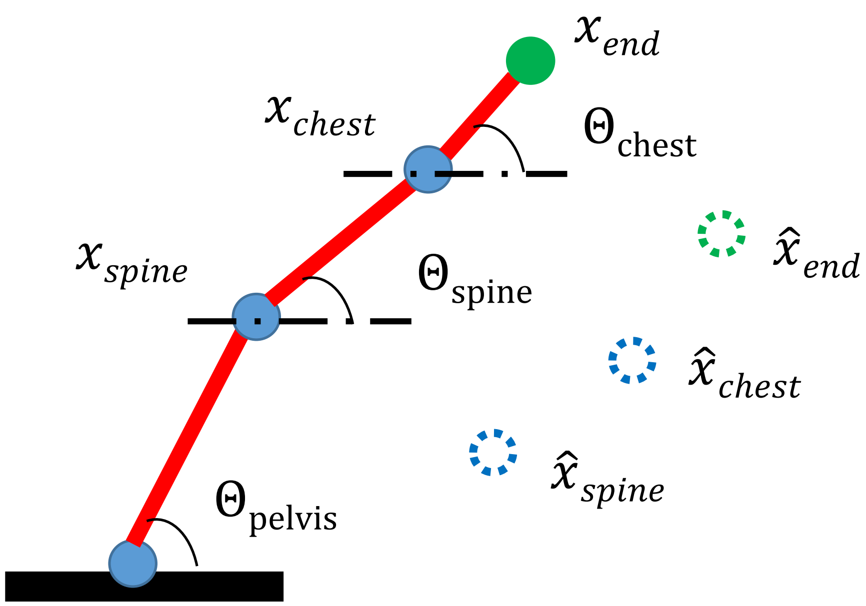
Input
Initialize
Head
Right Arm
Left Arm
Root Chain
Left Leg
Right Leg
Aggregate
Shape Est.
Camera Est.


Hierarchical
Kinematic Chain
Regressor
Encoder
Φ
Root Chain
Left Arm
Right Arm
Left Leg
θ
β
Optimizer
(Optional)


Head
R.arm
L.arm
Root
R.leg
L.leg
$[\beta^{t-1}, \theta^{t}{}_{root}, \theta^{t-1}{}_{Rarm}, \Phi]$
$[\beta^{t-1}, \theta^{t}{}_{root}, \theta^{t-1}{}_{head}, \Phi]$
$[\beta^{t-1}, \theta^{t}{}_{root}, \theta^{t-1}{}_{Larm}, \Phi]$
$[\beta^{t-1}, \theta^{t-1}{}_{root}, \Phi]$
$[\beta^{t-1}, \theta^{t}{}_{root}, \theta^{t-1}{}_{Rleg}, \Phi]$
$[\beta^{t-1}, \theta^{t}{}_{root}, \theta^{t-1}{}_{Lleg}, \Phi]$
$\Theta^t = [\theta^t_{root}, \theta^t_{head}, \theta^t_{R.arm}, \theta^t_{L.arm}, \theta^t_{R.leg}, \theta^t_{L.leg}]$
$[\beta^{t-1}, \Theta^t, \Phi]$
shape
β^t

Drawing Inspiration from Inverse Kinematics

- Estimate how joint angles change together to reach certain pose.
- Typically solved in iterative fashion.



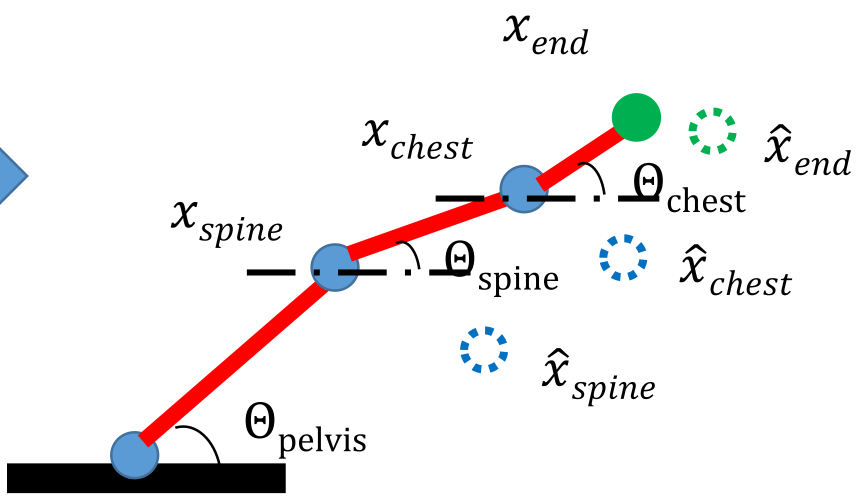
Iter t-1

Solver

$$\Delta\hat{x} = J(\theta)\Delta\theta$$

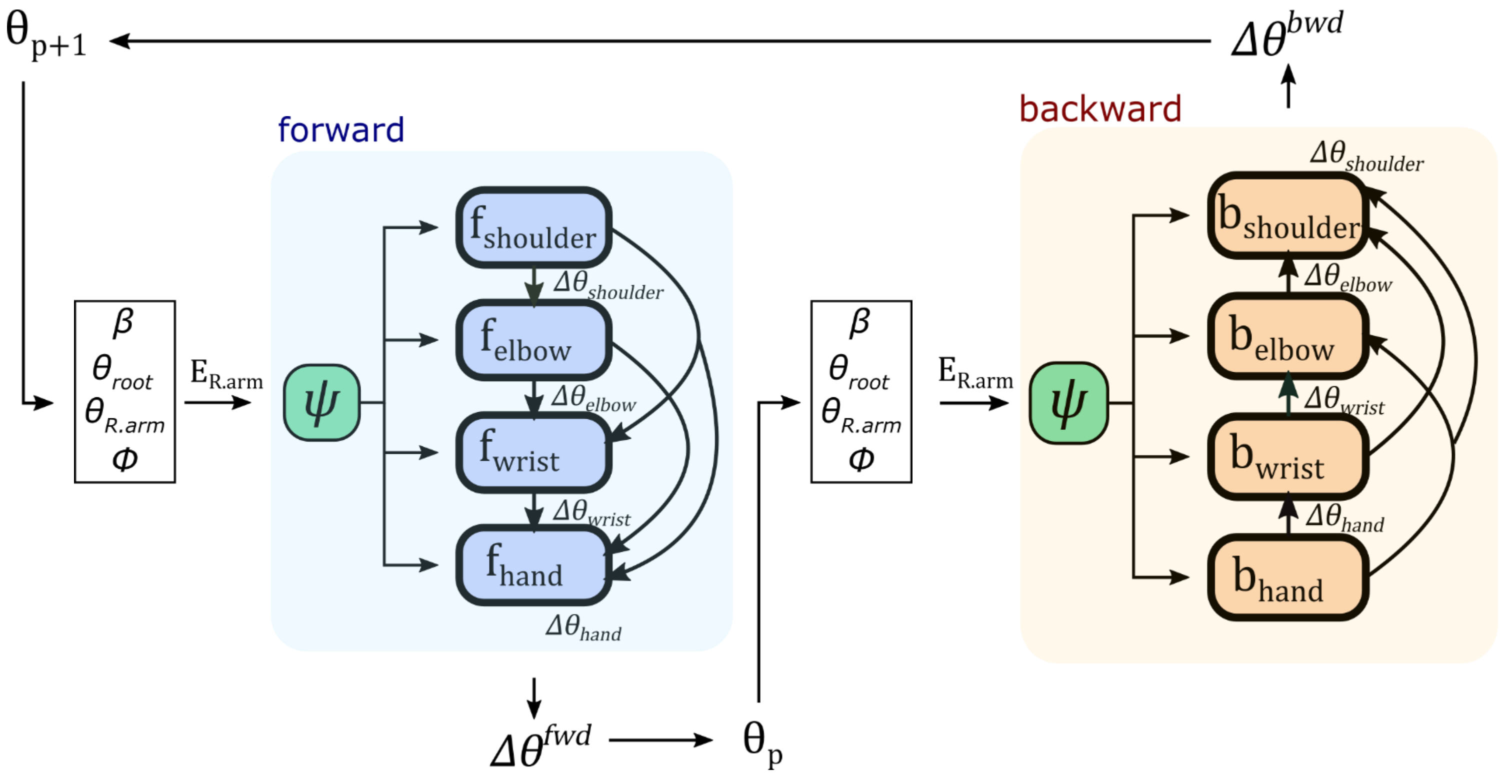
$$\Delta\theta = J^{+}(\theta)\Delta\hat{x}$$

$$\theta \leftarrow \theta + \alpha\Delta\theta$$



Iter t

Inner Chain Iterations



Learning Objective

- 3D Joints: $\sum_{i=1}^{N} \left\| \hat{X}_i^t - X_i \right\|_1$
- 2D Joints: $\sum_{i=1}^{N} \left\| \hat{x}_i^t - x_i \right\|_1$
- SMPL parameters: $\left\| [\hat{\Theta}^t, \hat{\beta}^t] - [\Theta, \beta] \right\|_2^2$
- Pose prior: $KL\left(Z_{\hat{\Theta}^t} \| \mathcal{N}(0, \mathcal{I})\right)$ [Pavlakos et al. 2019]

We perform a single backward pass by adding all losses over T outer iterations.

Experiments

LSP

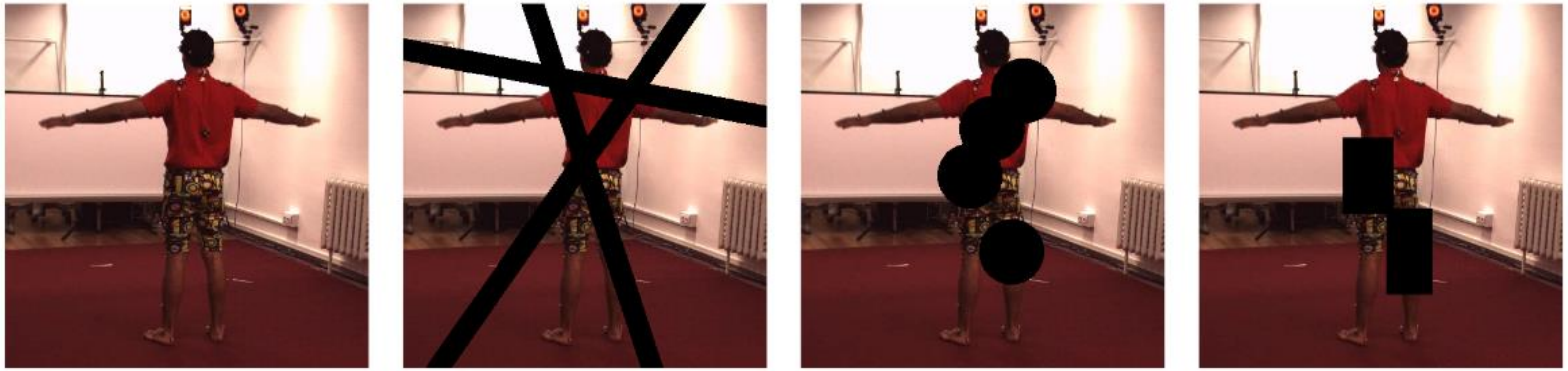
Human3.6M

MPII

COCO

- Evaluation metric: Mean per joint position error (MPJPE)

  - Synthetically generated occlusion set

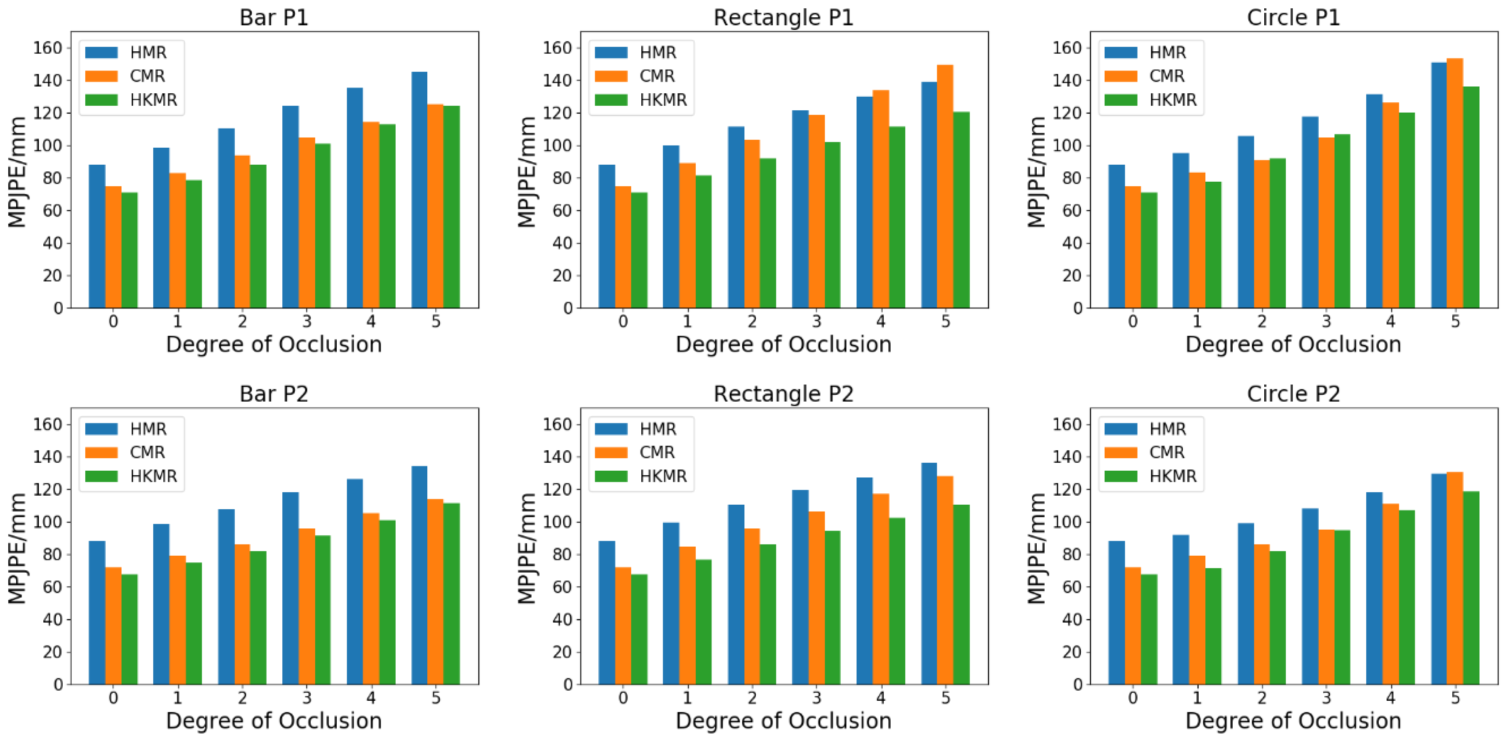


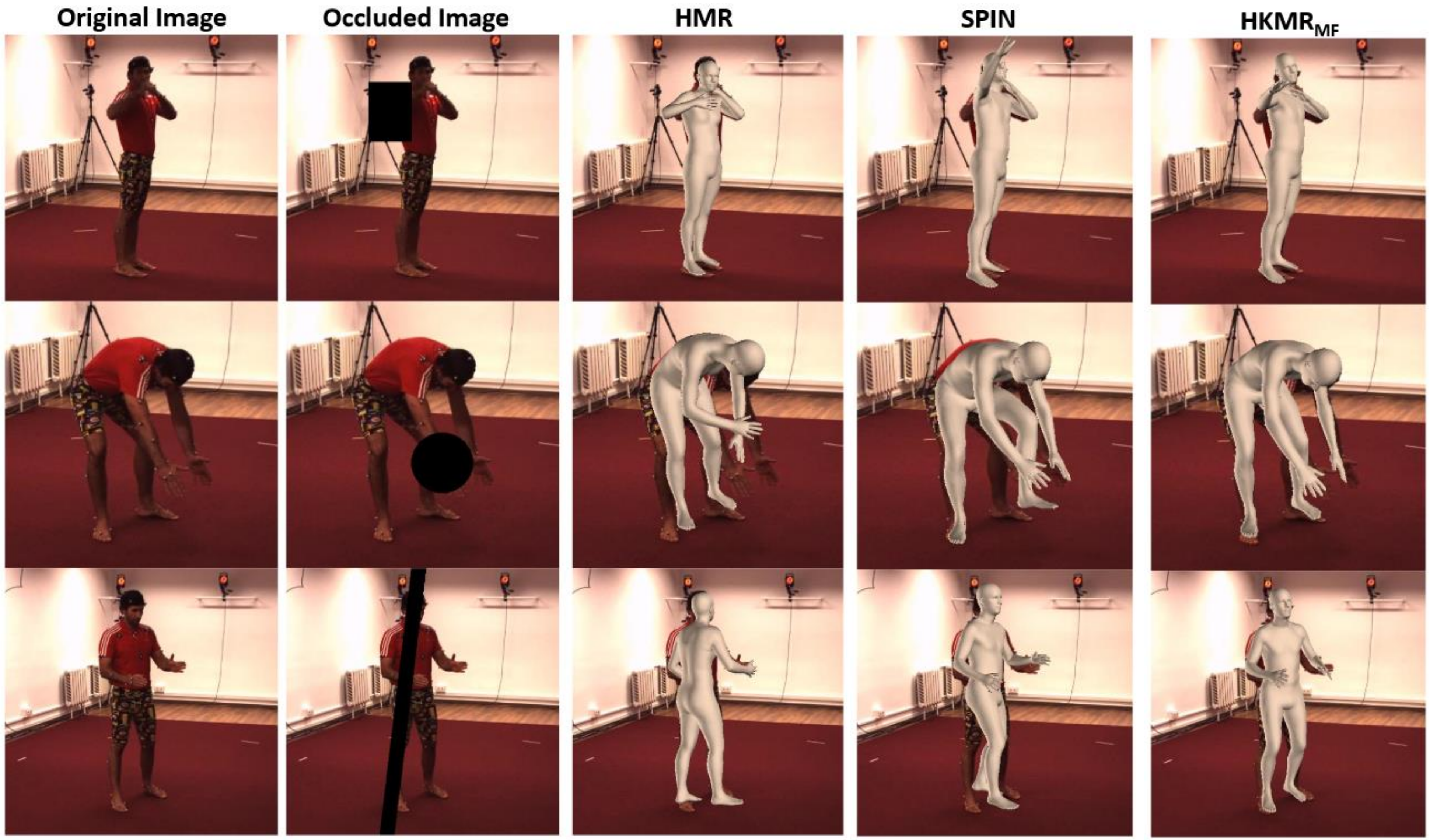
(a) Original (b) Oriented bar (c) Circle (d) Rectangle

Encoder-Regressor

	#Param	Standard		Bar		Circle		Rectangle	
		P1	P2	P1	P2	P1	P2	P1	P2
HMR [5]	26.8M	87.97	88.00	98.74	98.54	95.28	91.71	100.23	99.61
CMR [8]	42.7M	74.70	71.90	82.99	78.85	83.50	79.24	89.01	84.73
HKMR	26.2M	**71.08**	**67.74**	**78.34**	**74.91**	**77.60**	**71.38**	**81.33**	**76.79**

Robustness to varying degrees of occlusion

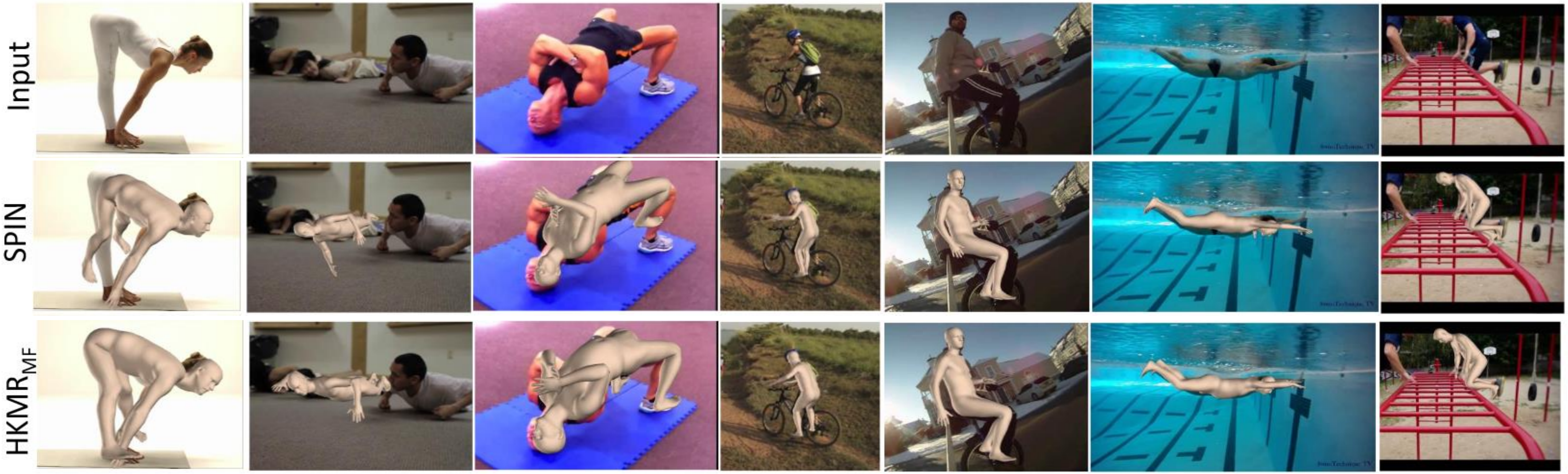



Original Image
Occluded Image
HMR
SPIN
HKMR$_{MF}$

Encoder-Regressor-Optimizer

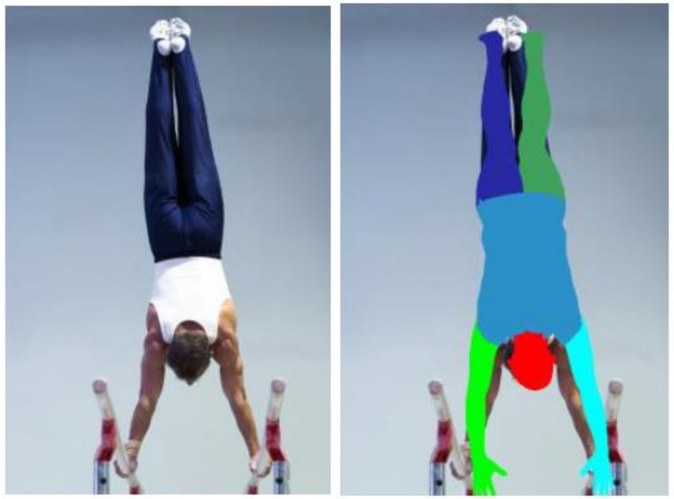
Human3.6M	Standard		Bar		Circle		Rectangle	
MPJPE (mm)↓	P1	P2	P1	P2	P1	P2	P1	P2
SPIN [6]	65.60	62.23	74.40	68.61	74.06	67.03	77.21	70.35
HKMR$_{MF}$	**64.02**	**59.62**	**70.10**	**64.91**	**69.60**	**63.22**	**70.10**	**64.91**

MPII Invisible	MPJPE (pixel)↓	PCK (%)↑
SPIN [6]	59.52	62.16
HKMR$_{MF}$	**55.56**	**66.24**



LSP	FB Seg.		Part Seg.	
	acc.	f1	acc.	f1
Oracle [3]	92.17	**0.88**	88.82	0.67
SMPLify [3]	91.89	**0.88**	87.71	0.64
SMPLify+[28]	92.17	**0.88**	88.24	0.64
HMR [5]	91.67	0.87	87.12	0.60
CMR [8]	91.46	0.87	88.69	0.66
TexturePose [21]	91.82	0.87	89.00	0.67
SPIN [6]	91.83	0.87	89.41	0.68
HKMR$_{MF}$	**92.23**	**0.88**	**89.59**	**0.69**

Human3.6M	P1	P2
HMR [5]	87.97	88.00
Arnab *et al.* [20]	-	77.80
HoloPose [16]	-	64.28
CMR [8]	74.70	71.90
DaNet [17]	-	61.50
DenseRaC [18]	76.80	-
VIBE [19]	-	65.60
SPIN [6]	65.60	62.23
HKMR$_{MF}$	**64.02**	**59.62**



Conclusion

- A new method for human mesh recovery that exploits the geometry of the parametric human model.
- Demonstrates robustness to occlusions and achieves state-of-art results on popular benchmarks.
- Compatible with existing model paradigms.