Video Based Reconstruction of 3D People Models

Thiemo Alldieck, Marcus Magnor , **Weipeng Xu** , Christian Theobalt , Gerard Pons-Moll

Technische Universität Braunschweig

max planck institut informatik

2018 CVPR SALT LAKE CITY • JUNE 18-22

Motivation

Applications of 3D Human Model

© Fitnect via youtube.com

VR/AR

© Mario Botsch et al

Medicine and self-perception

© Oculus VR

Telepresence

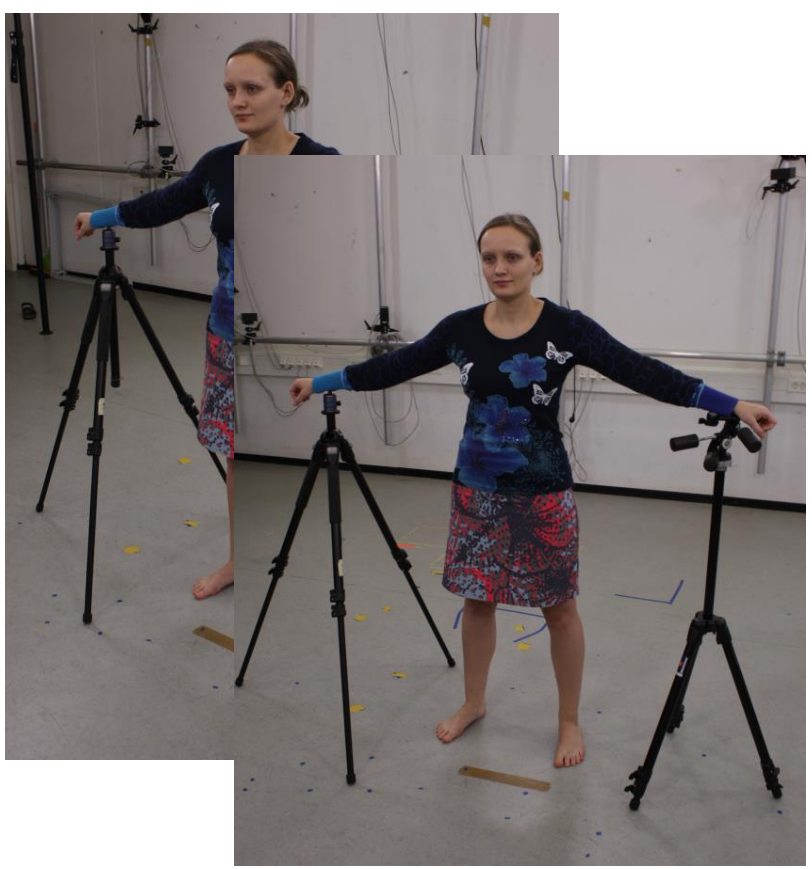
Previous works

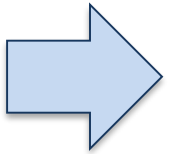
- 3D body scanner

© Twindom

Previous works

- Monocular 3D reconstruction





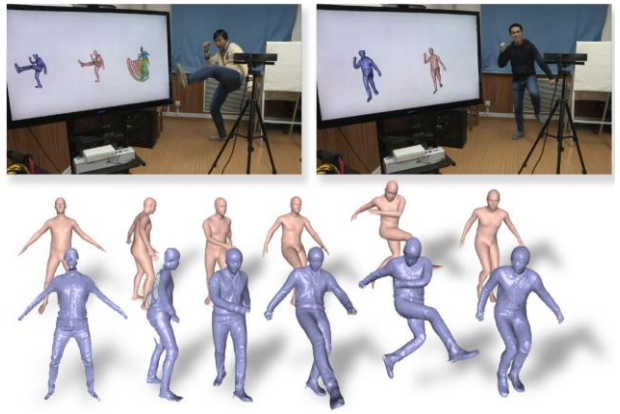
Human body is not rigid!

Previous works

- Depth based

© 3D Self-Portraits, Li et al

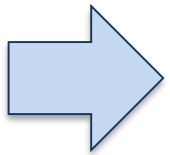
© DynamicFusion, Newcombe et al



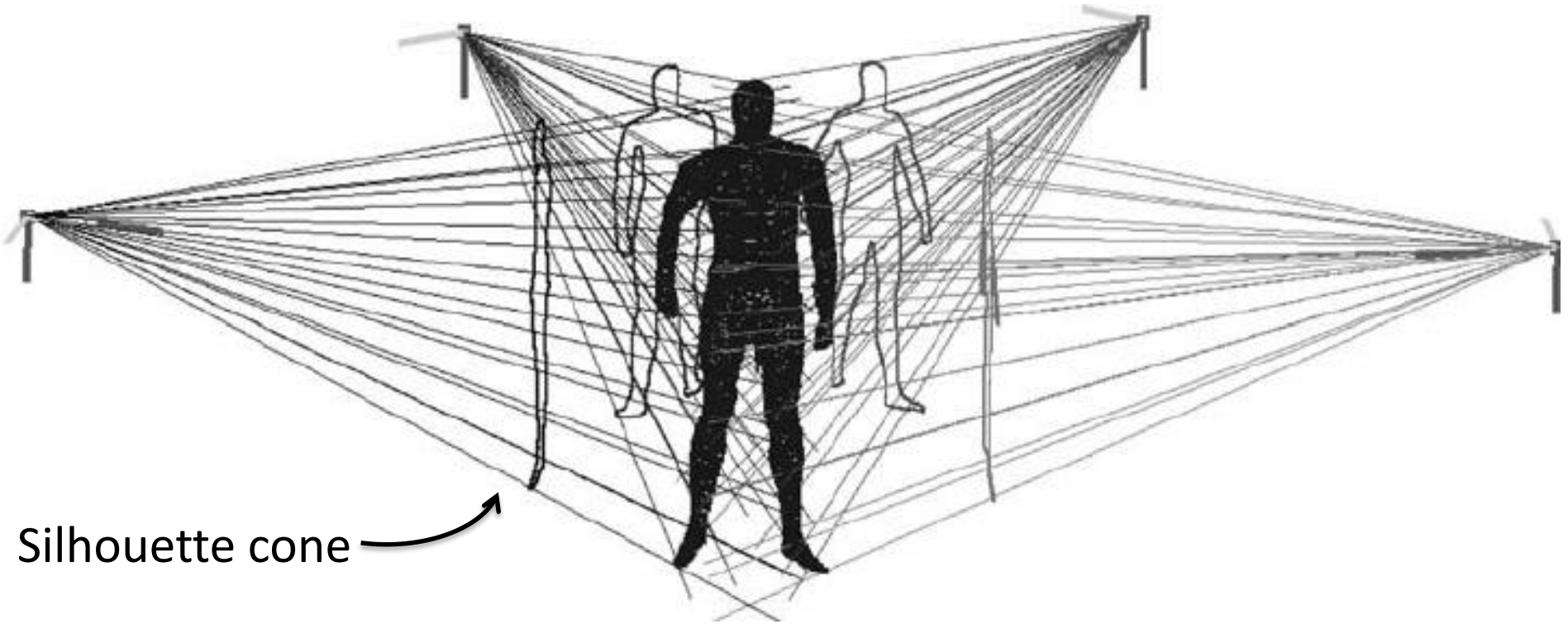
© DoubleFusion, Yu et al

Goal:

3D Reconstruction from a **Single RGB Video** of **Moving** People **in Clothes**



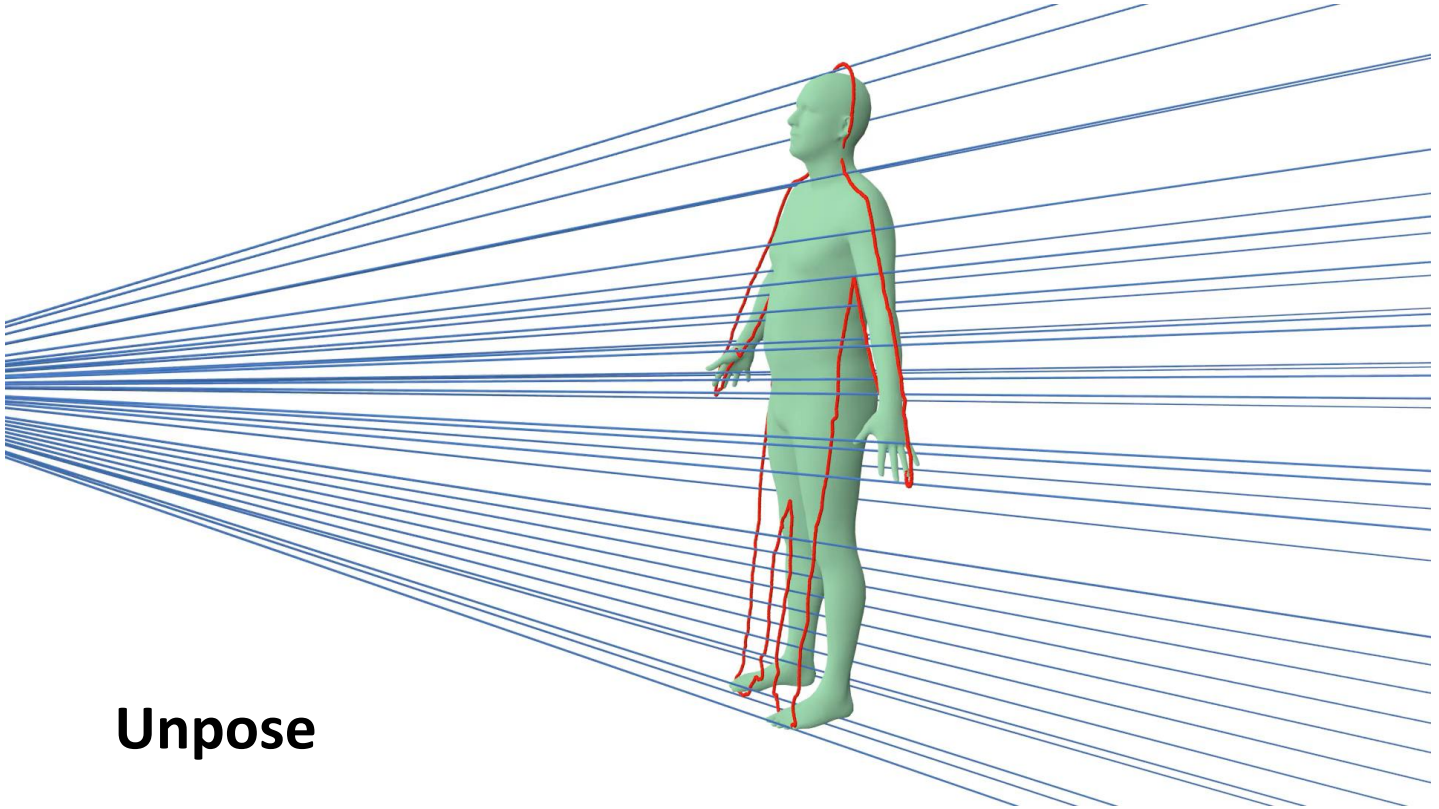
Key idea: Extend visual hulls to dynamic human motion



© Stefano Corazza

Problem: standard visual hull requires a **static** object captured by multiple views

Key idea: Extend visual hulls to dynamic human motion

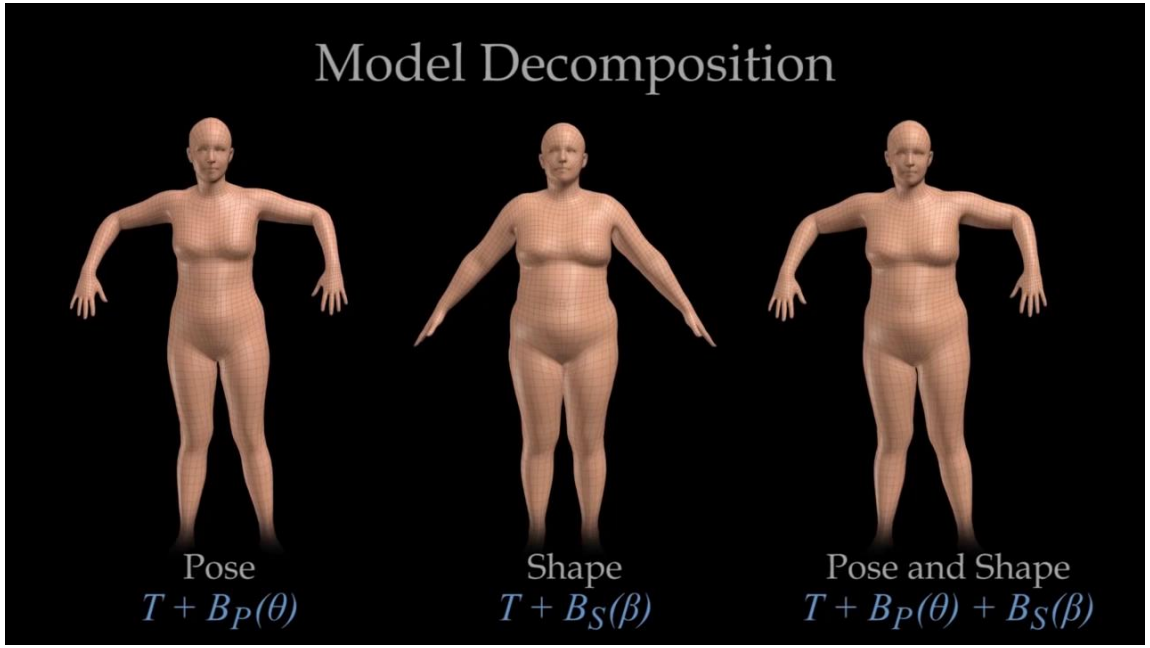


Unpose

Transform the silhouette cones according to the inverse of non-rigid motion

Method

- Shape representative



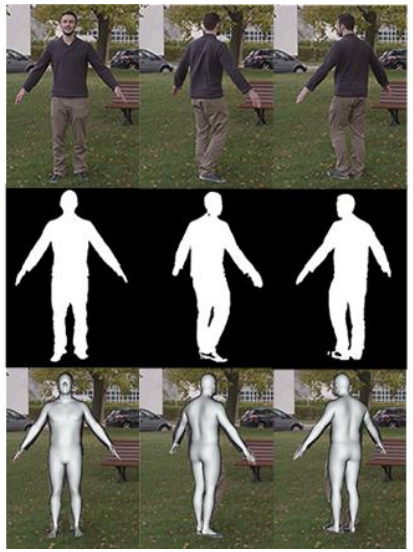
SMPL [Loper et al, Siggraph Asia '15]

3D mesh:

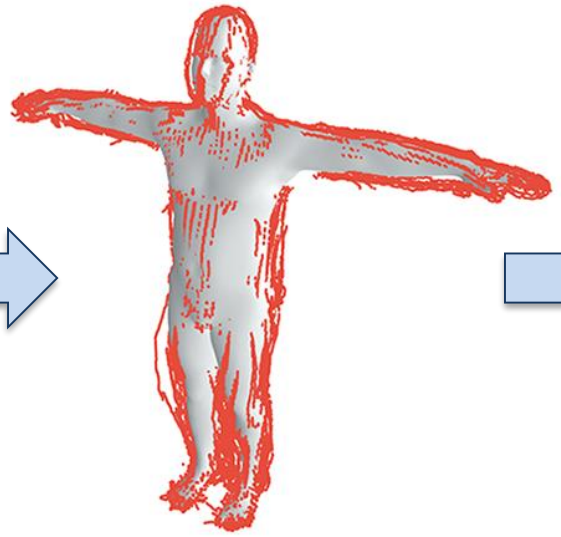
$$\mathbf{T}_\mu + B_s(\boldsymbol{\beta}) + B_p(\boldsymbol{\theta}) + \mathbf{D}$$

mean shape shape pose offsets

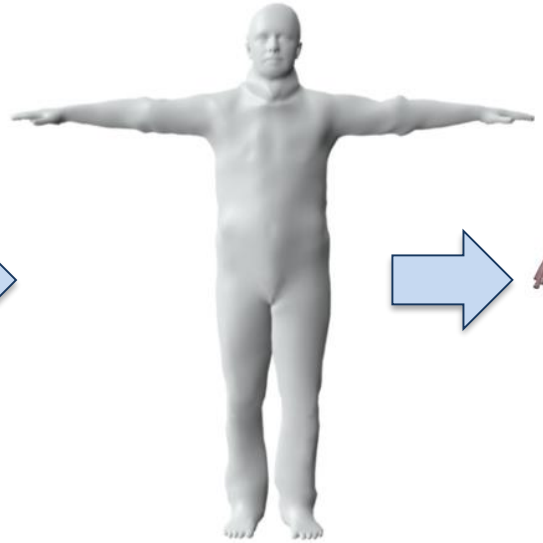
Method



Silhouette estimation
& Pose tracking



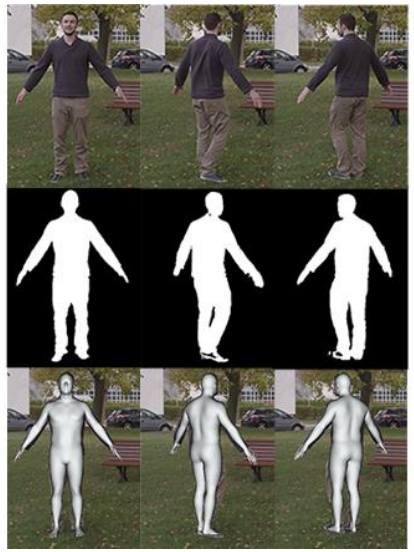
Silhouette cone unposing



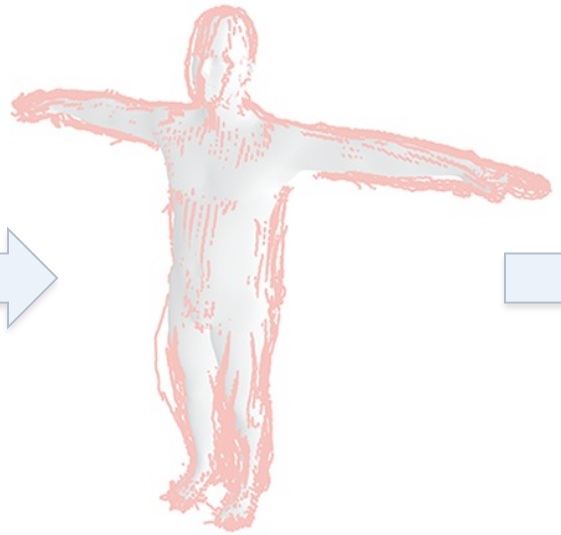
Consensus Shape
reconstruction

Texture generation

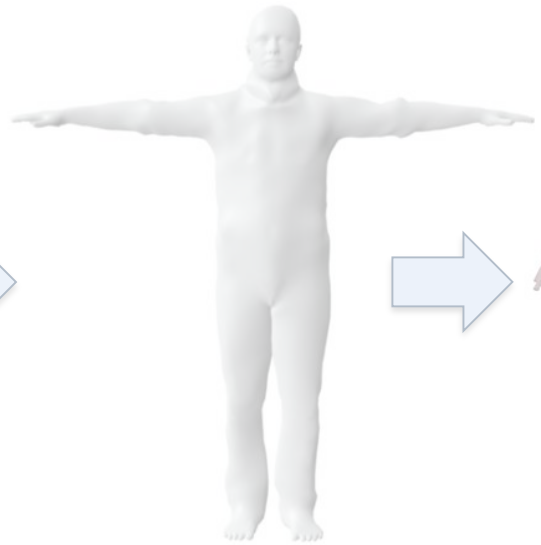
Method



Silhouette estimation
& Pose tracking



Silhouette cone unposing

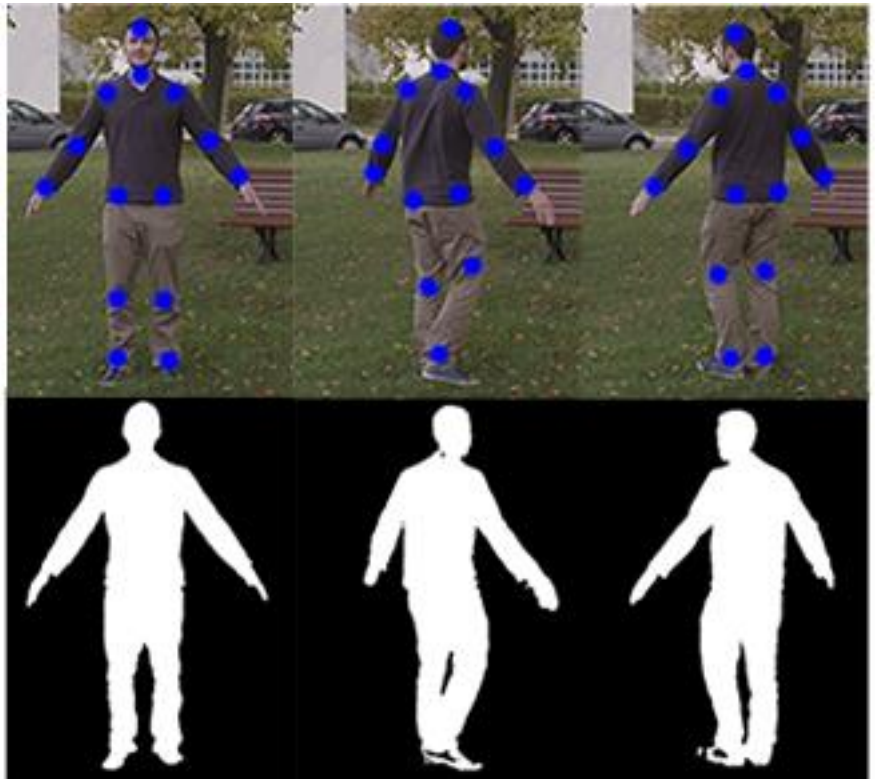


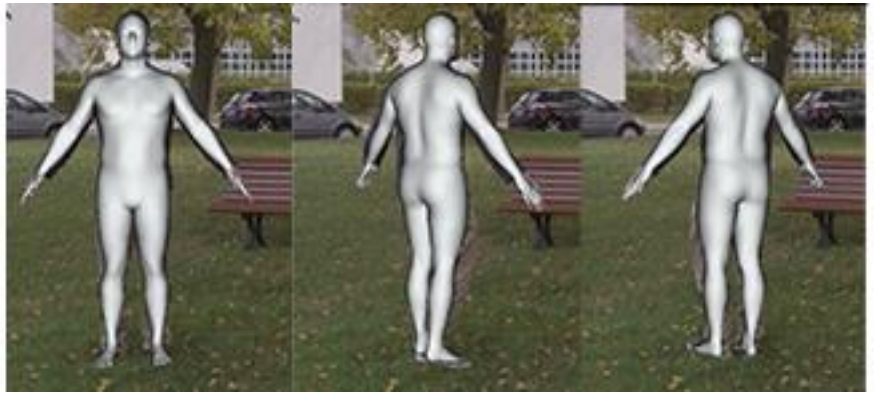
Consensus Shape
reconstruction

Texture generation

Method

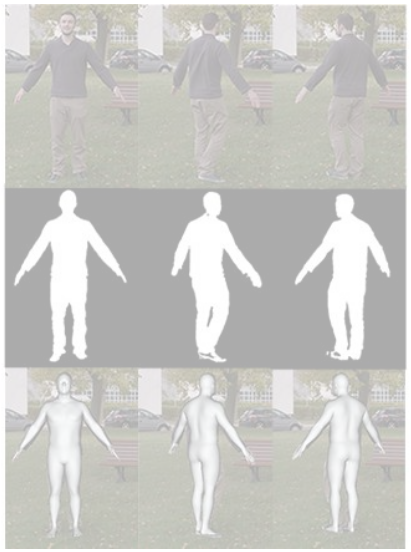
- Silhouette estimation & Pose tracking



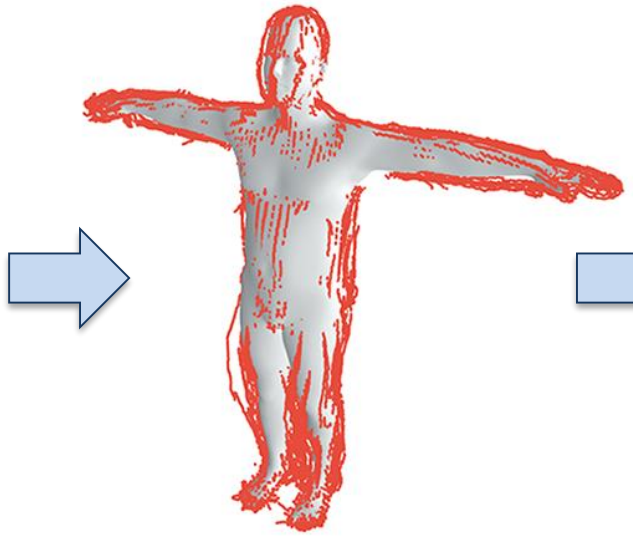


$$E_{Joint}(\theta,\beta) + E_{Silh}(\theta) + E_{Reg}(\theta,\beta)$$

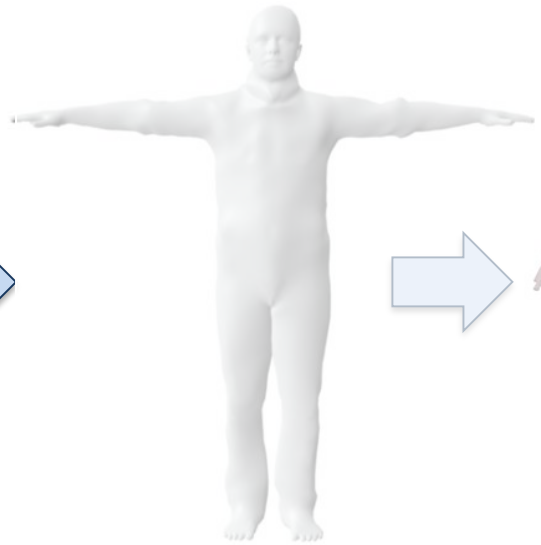
Method



Silhouette estimation
& Pose tracking



Silhouette cone unposing

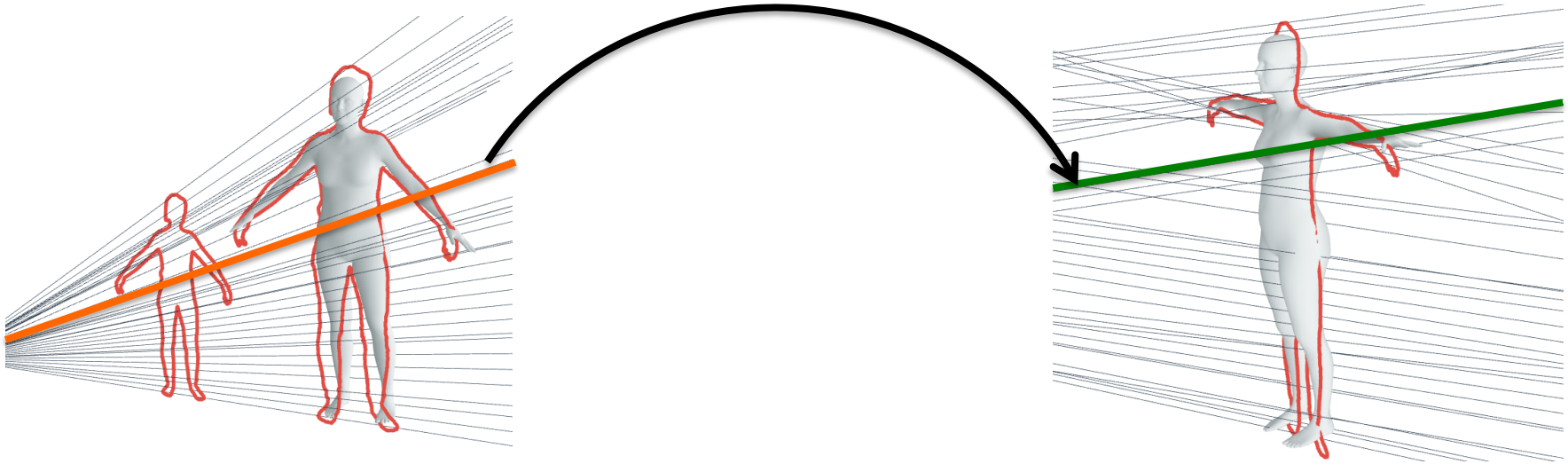


Consensus Shape
reconstruction

Texture generation

Method

- Silhouette cone **unposing**



Method

- Silhouette cone **unposing**

$$\mathbf{v}_i' = \sum_{k=1}^{K} w_{k,i} G_k(\boldsymbol{\theta}, J(\boldsymbol{\beta}))(\mathbf{v}_i + b_{s,i}(\boldsymbol{\beta}) + b_{P,i}(\boldsymbol{\theta}))$$

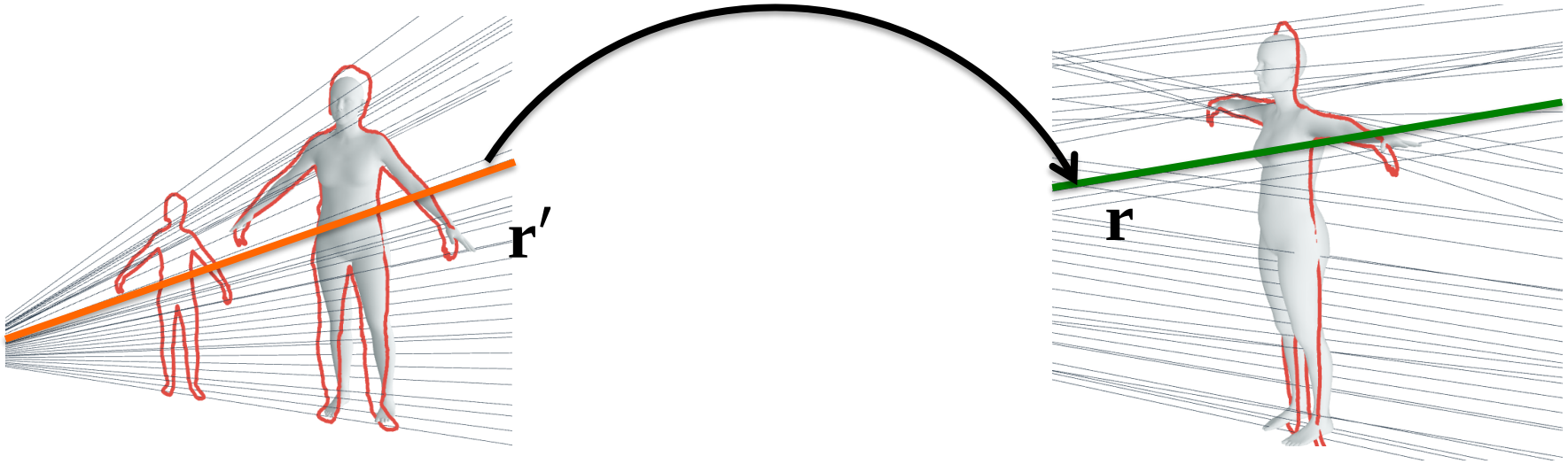
posed

Articulated Motion

unposed

Method

- Silhouette cone **unposing**



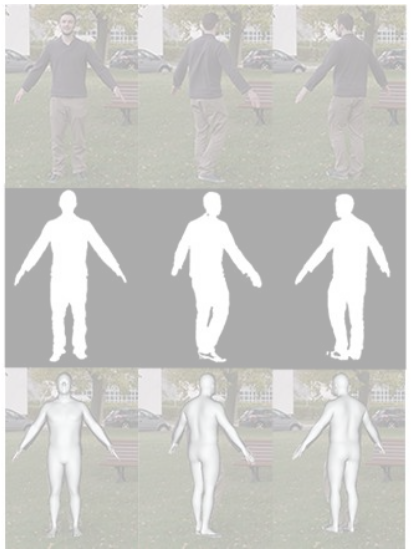
$$\mathbf{r} = \left(\sum_{k=1}^{K} w_{k,i} G_k(\boldsymbol{\theta}, \boldsymbol{J}_{\boldsymbol{\beta}}) \right)^{-1} \mathbf{r}' - b_{P,i}(\boldsymbol{\theta}).$$

Ray in Canonical Frame

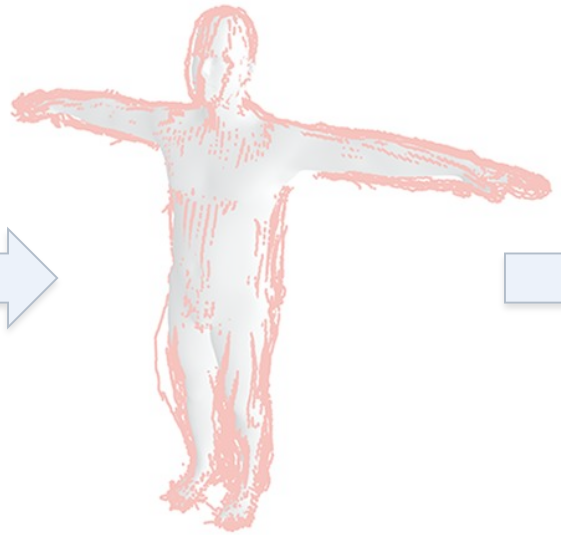
Inverse of Articulated Motion

Ray

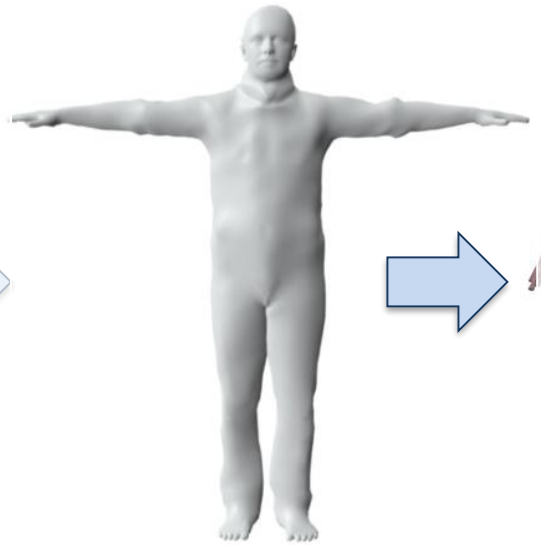
Method



Silhouette estimation
& Pose tracking



Silhouette cone unposing

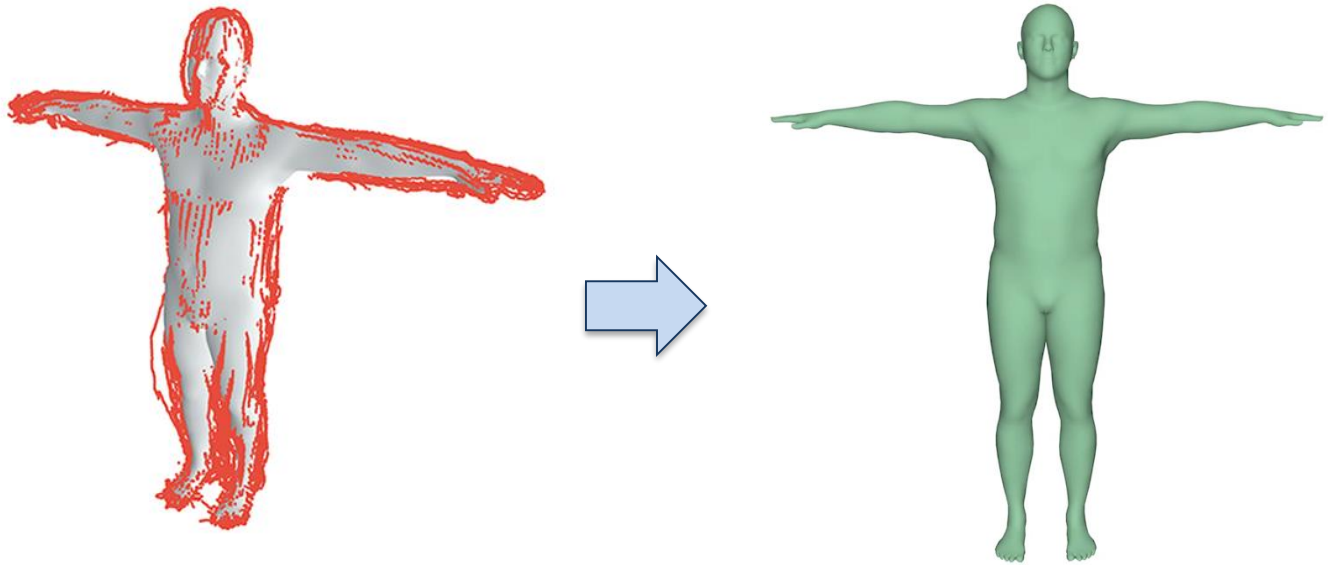


Consensus Shape
reconstruction

Texture generation

Method

- **Consensus shape** estimation
- Optimize a single shape to fit **unposed** silhouette cones from all frames



$$E_{\text{cons}} = E_{\text{data}} + w_{\text{lp}} E_{\text{lp}} + w_{\text{var}} E_{\text{var}} + w_{\text{sym}} E_{\text{sym}}$$

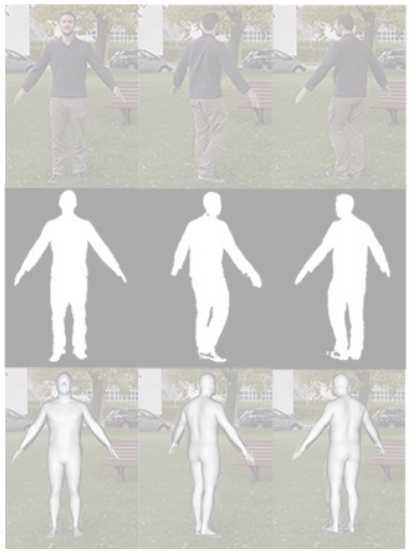
$$\sum_{(\mathbf{v},\mathbf{r}) \in \mathcal{M}} \rho(\mathbf{v} \times \mathbf{r}_n - \mathbf{r}_m)$$

Laplacian term

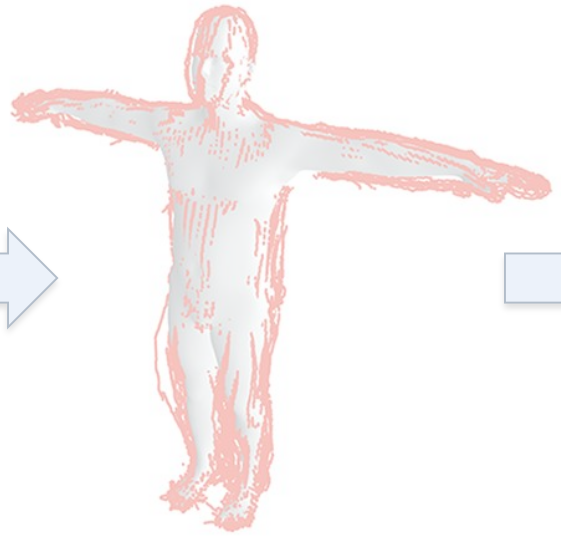
Body model term

Symmetry term

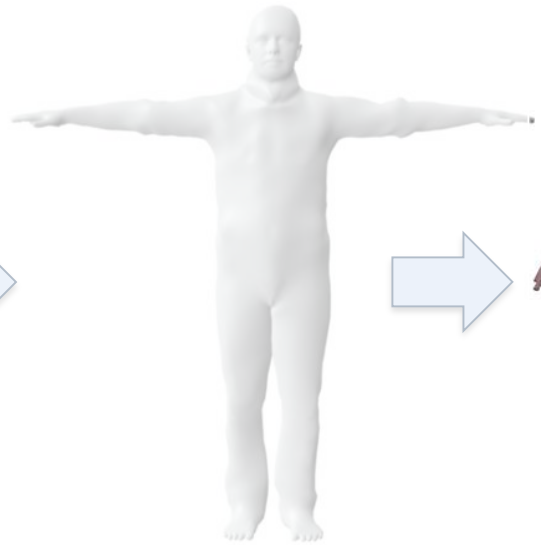
Method



Silhouette estimation
& Pose tracking



Silhouette cone unposing

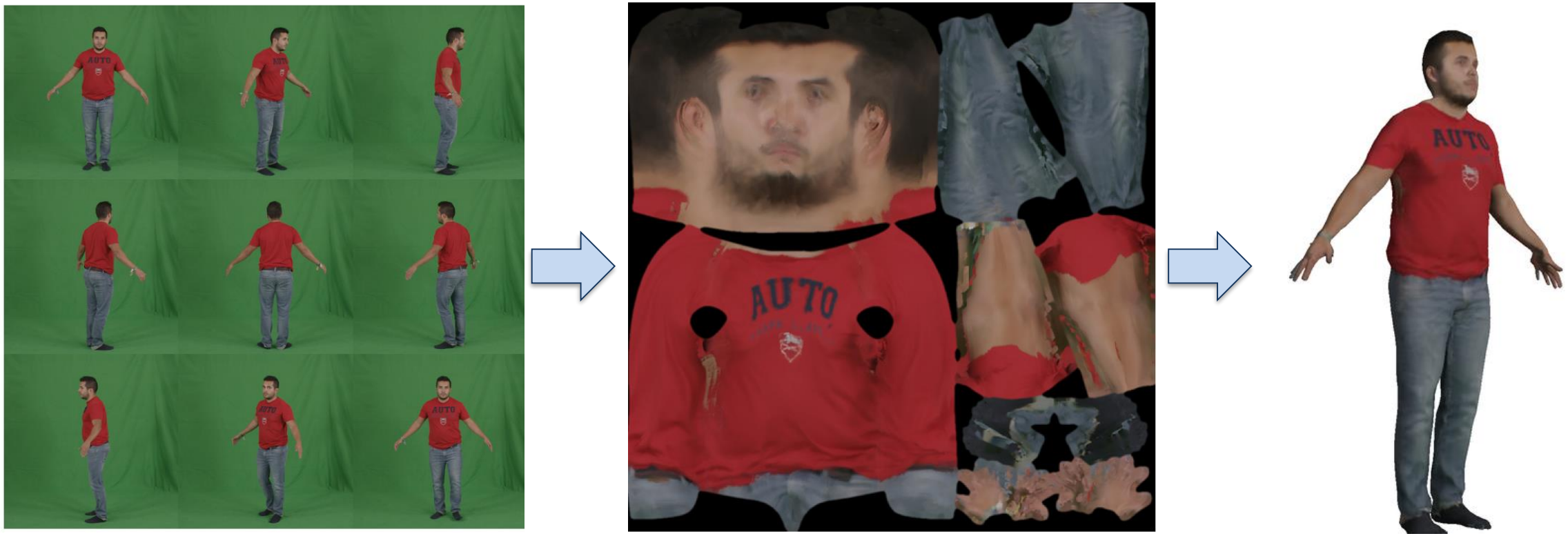


Consensus Shape
reconstruction

Texture generation

Method

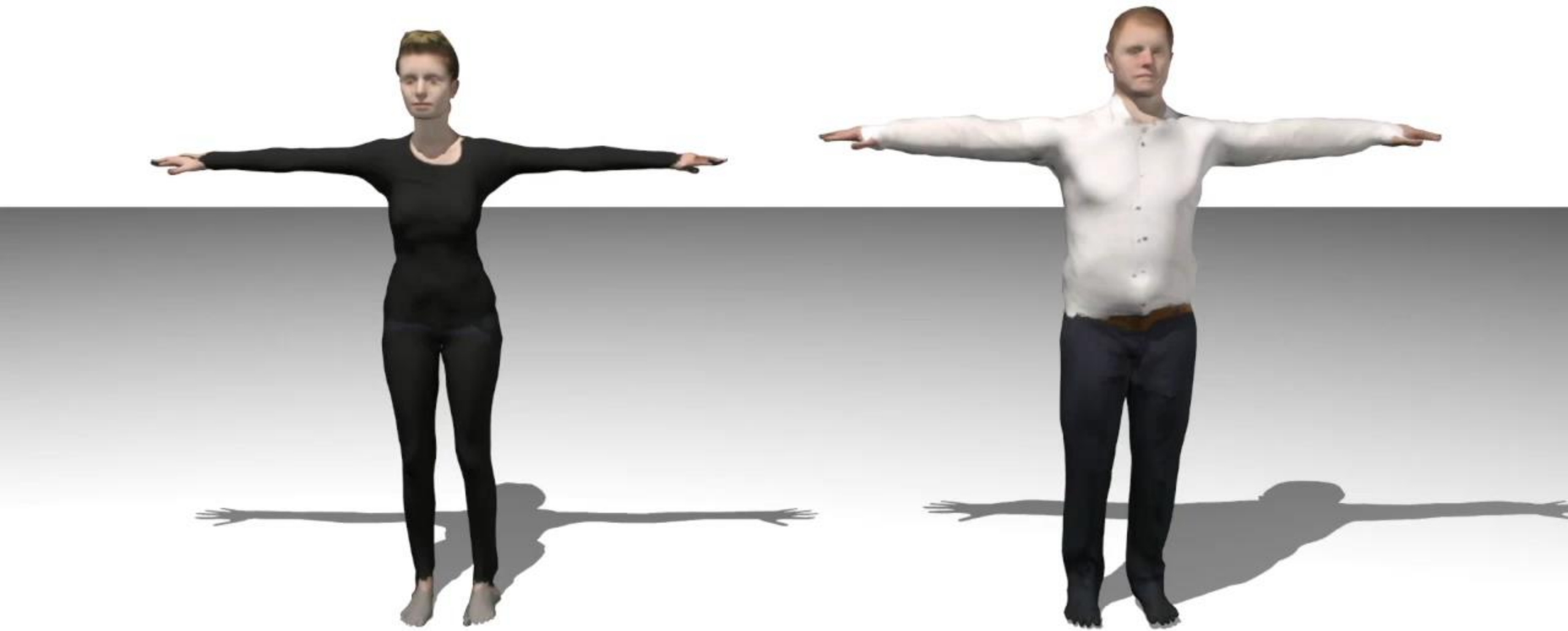
- Texture generation



Input

Input

AUTO



People-Snapshot dataset

* Dataset and code, publicly available: https://graphics.tu-bs.de/people-snapshot

Thank you!