

Modern Techniques and Applications for Real-Time Non-Rigid Registration

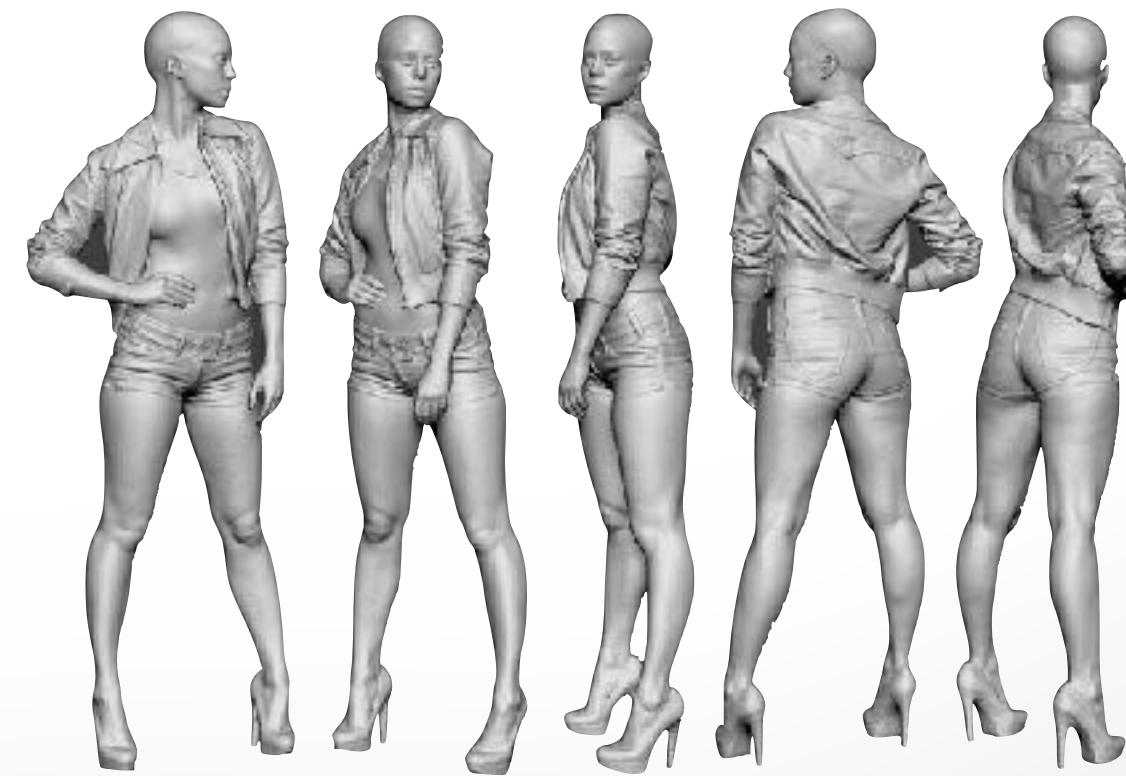
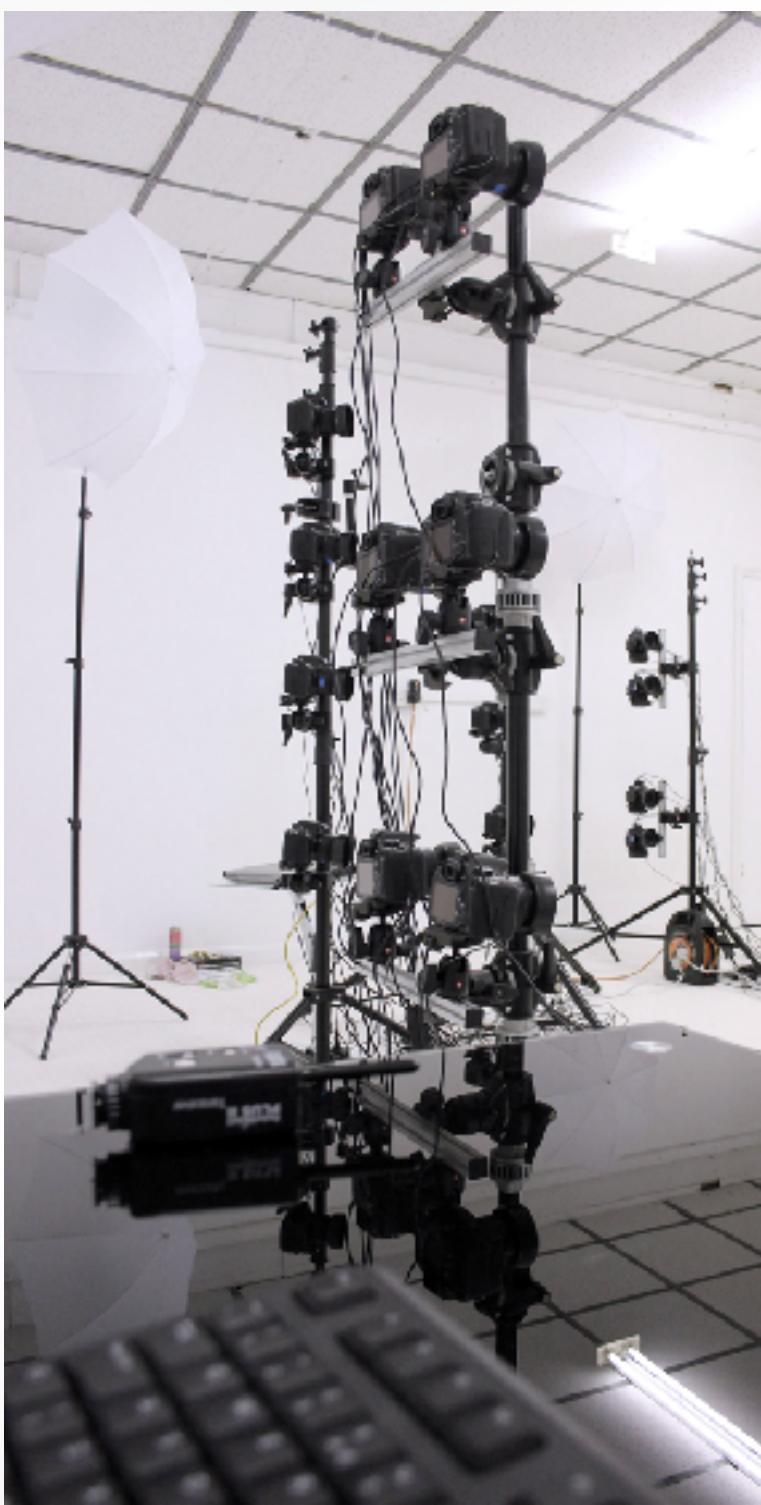
Andrea Tagliasacchi

Hao Li



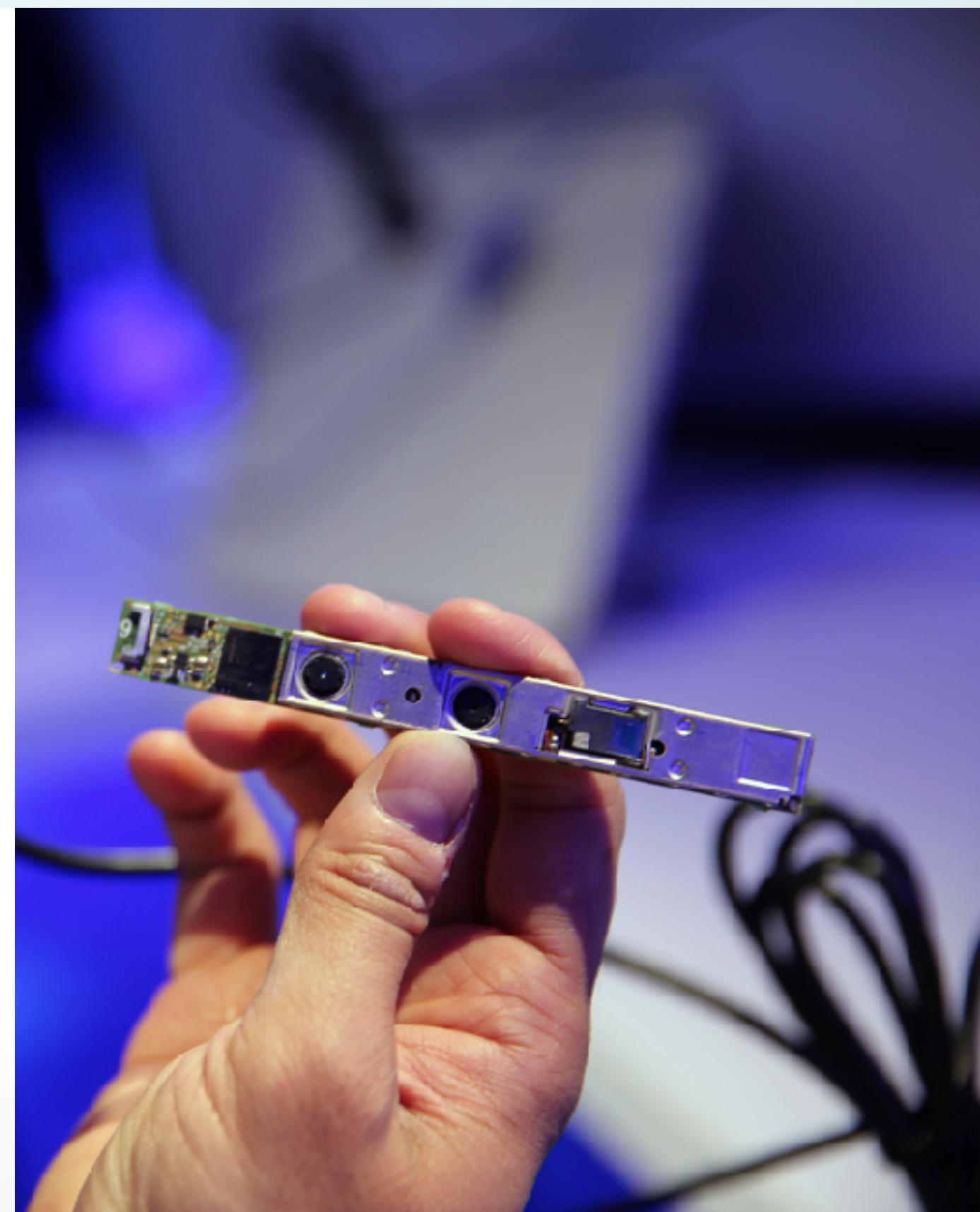
Non-Rigid Registration

3D Scanning



Democratization

Microsoft, Intel

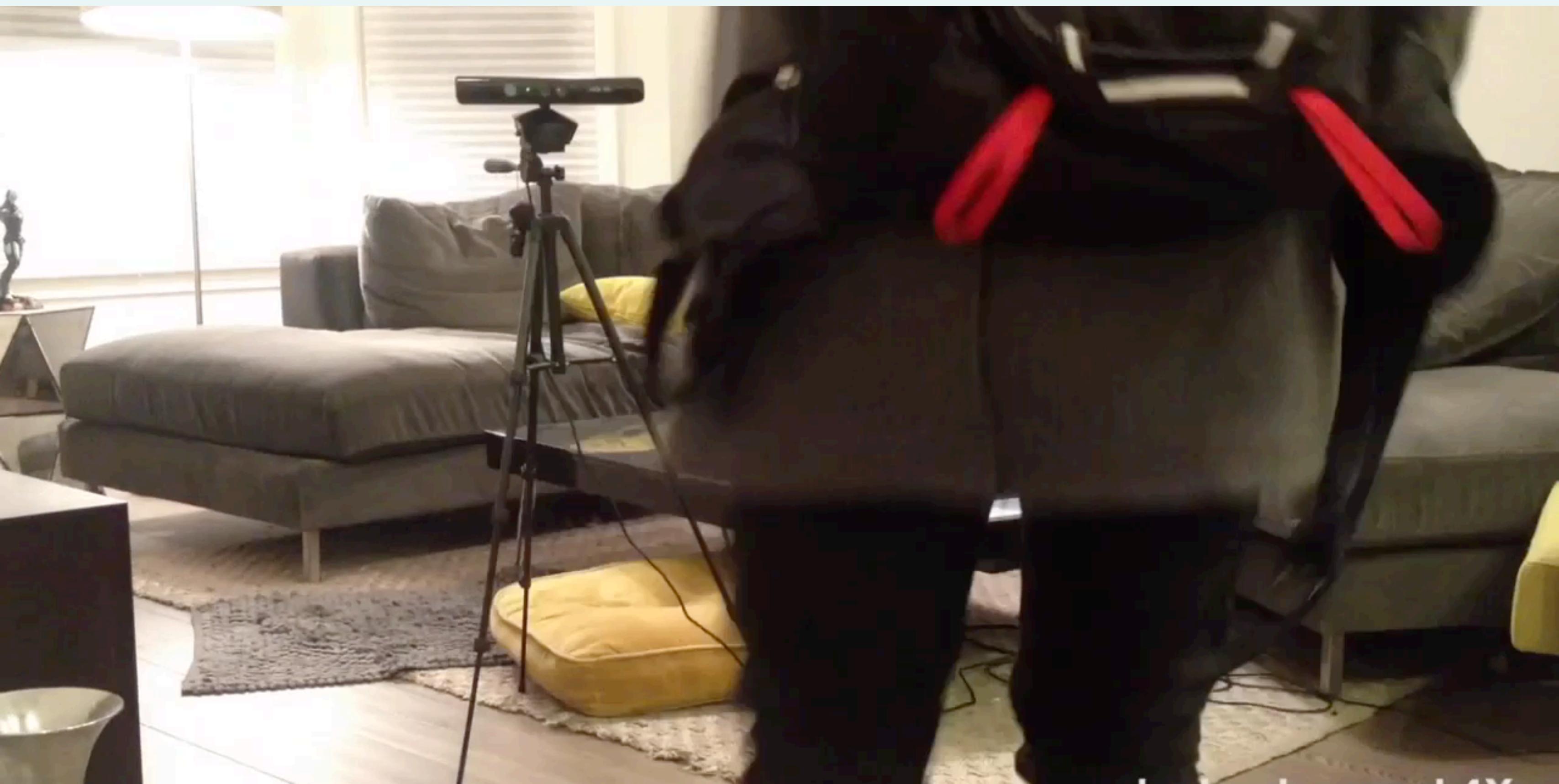


3D Scanning Process



3D Self-Portraits

Shapify.me

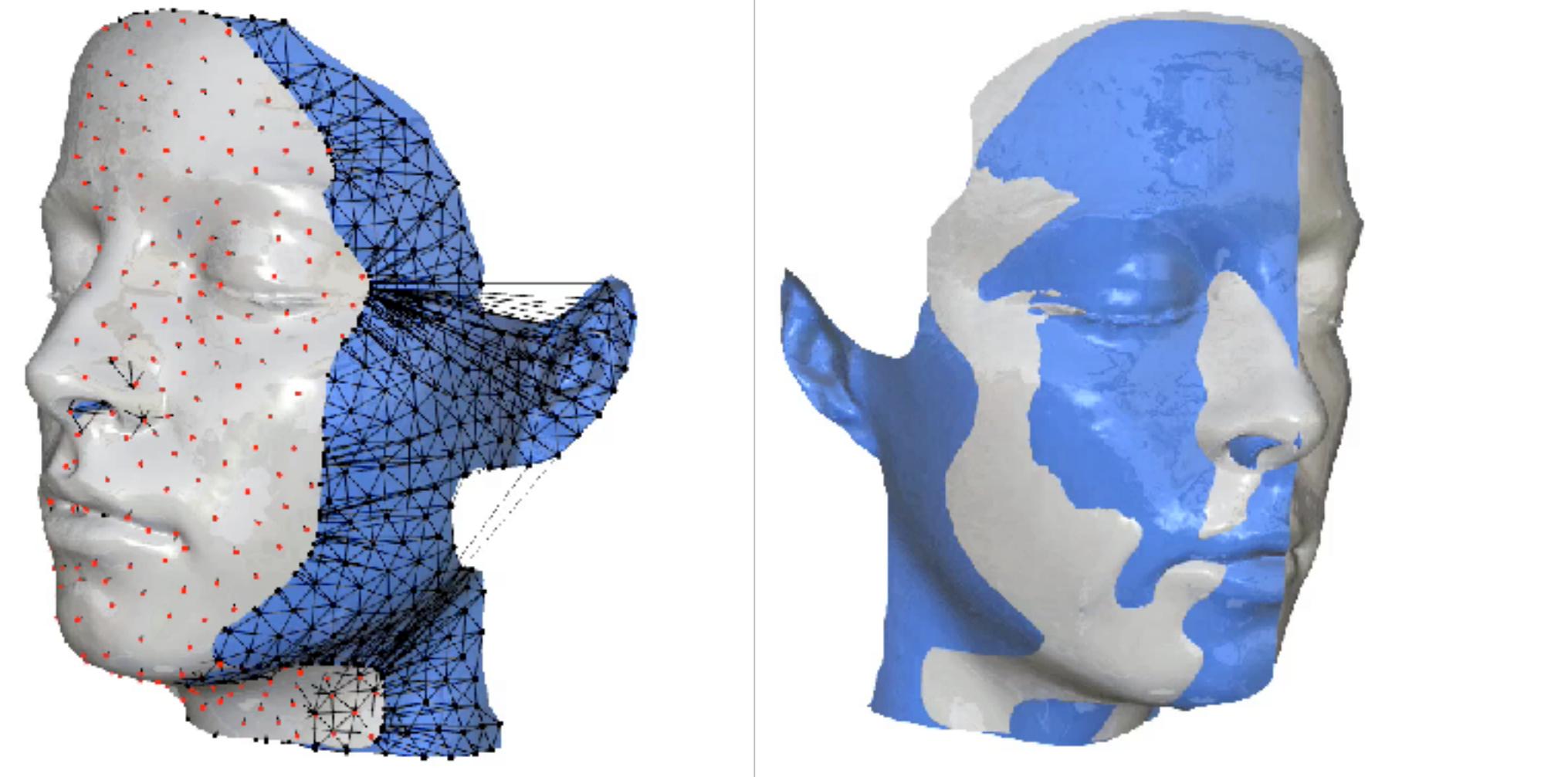


3D Self-Portraits

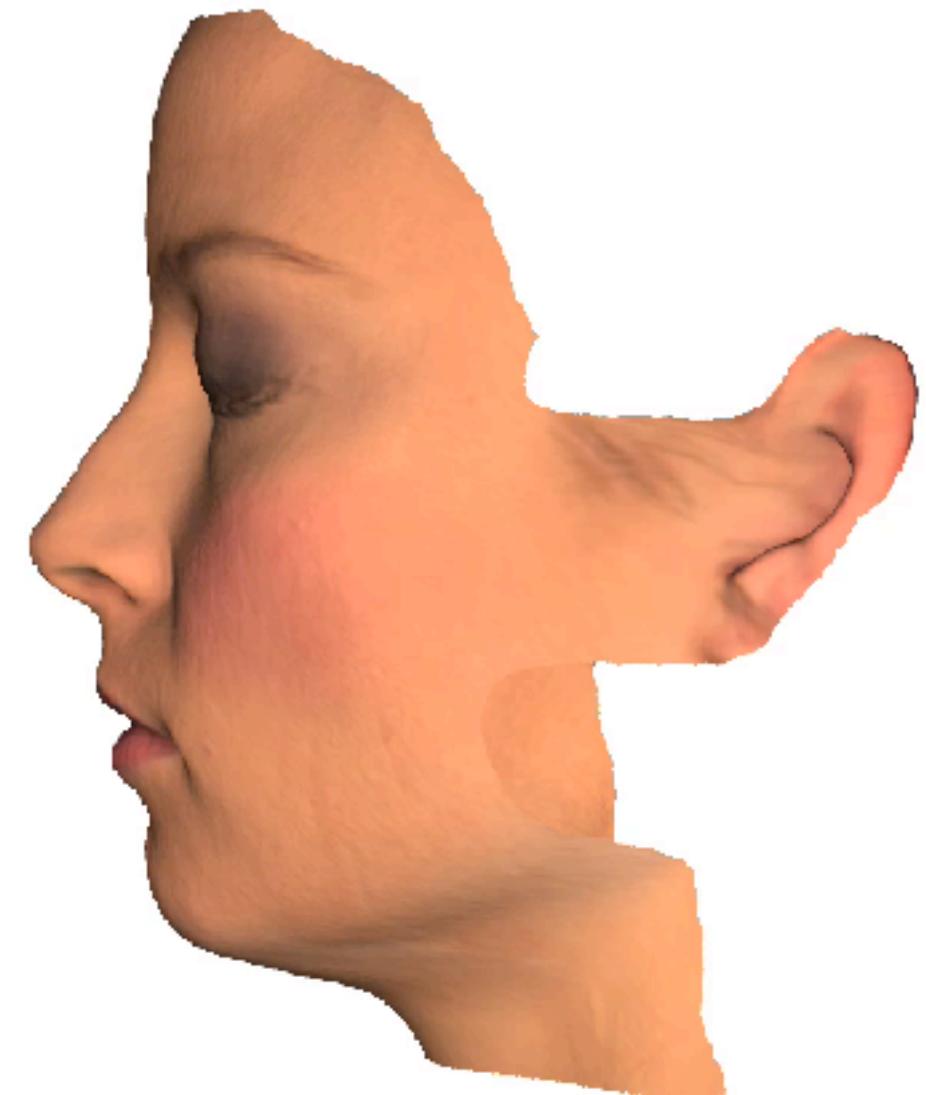
Shapify.me



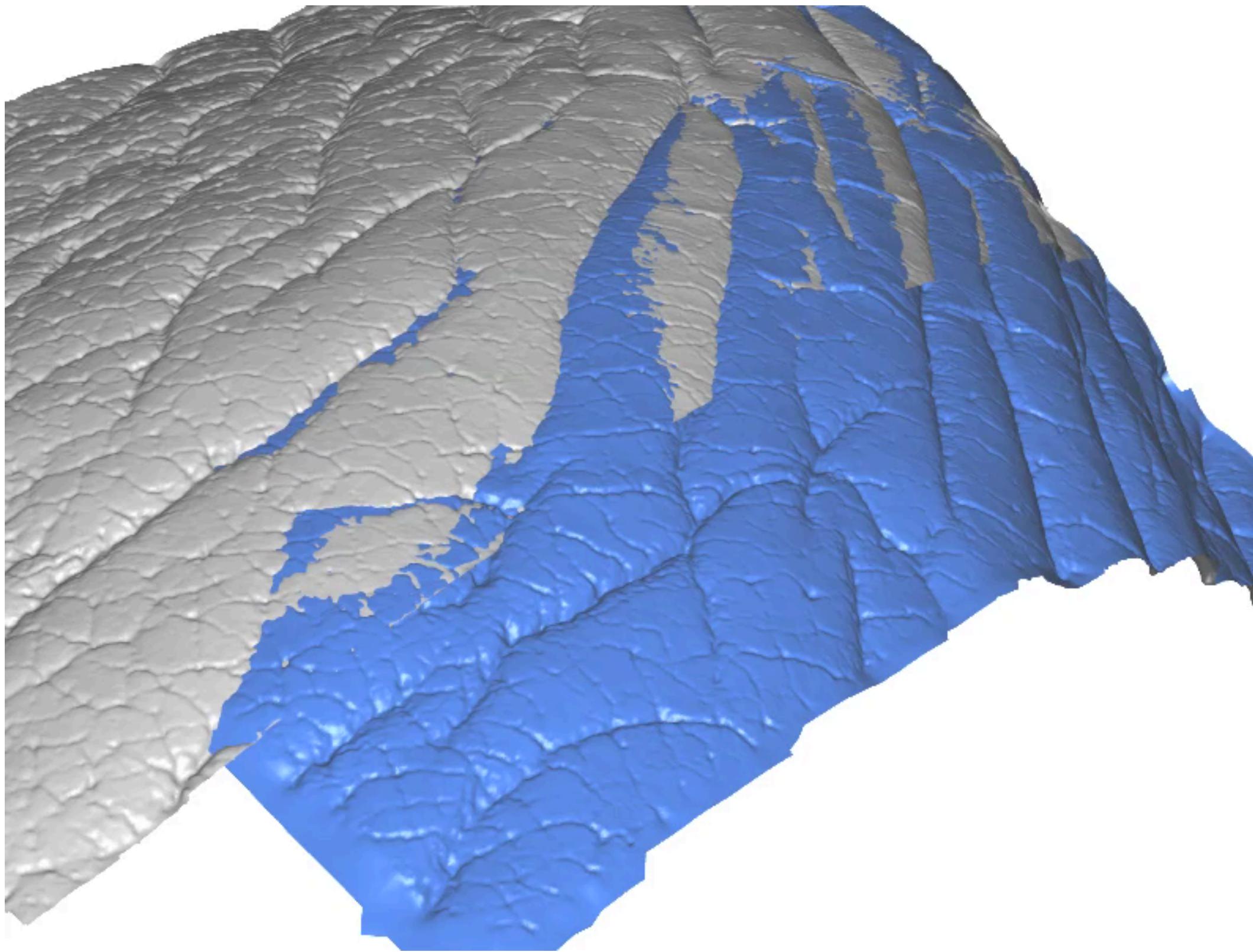
Shape Completion



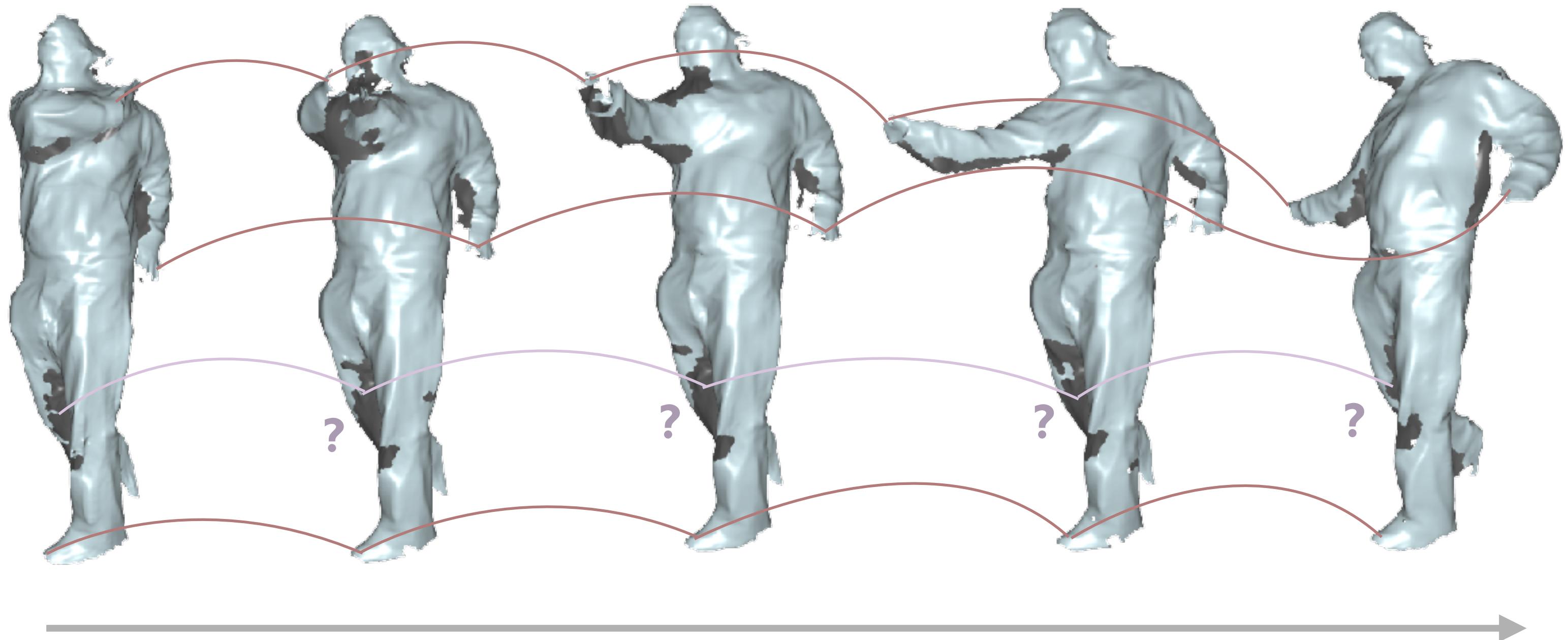
Full Reconstruction



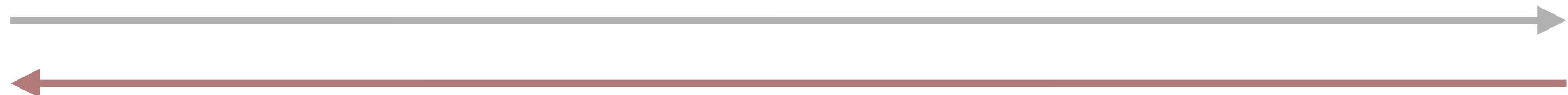
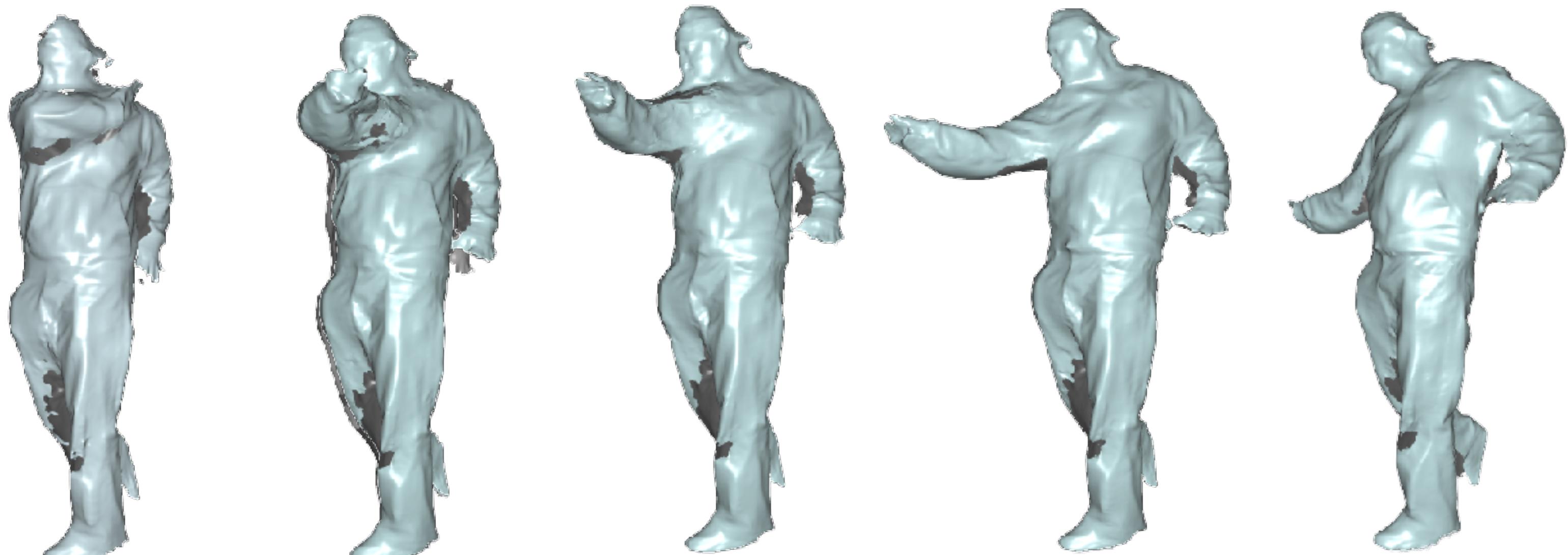
Life Sciences



Inter-Frame Correspondences



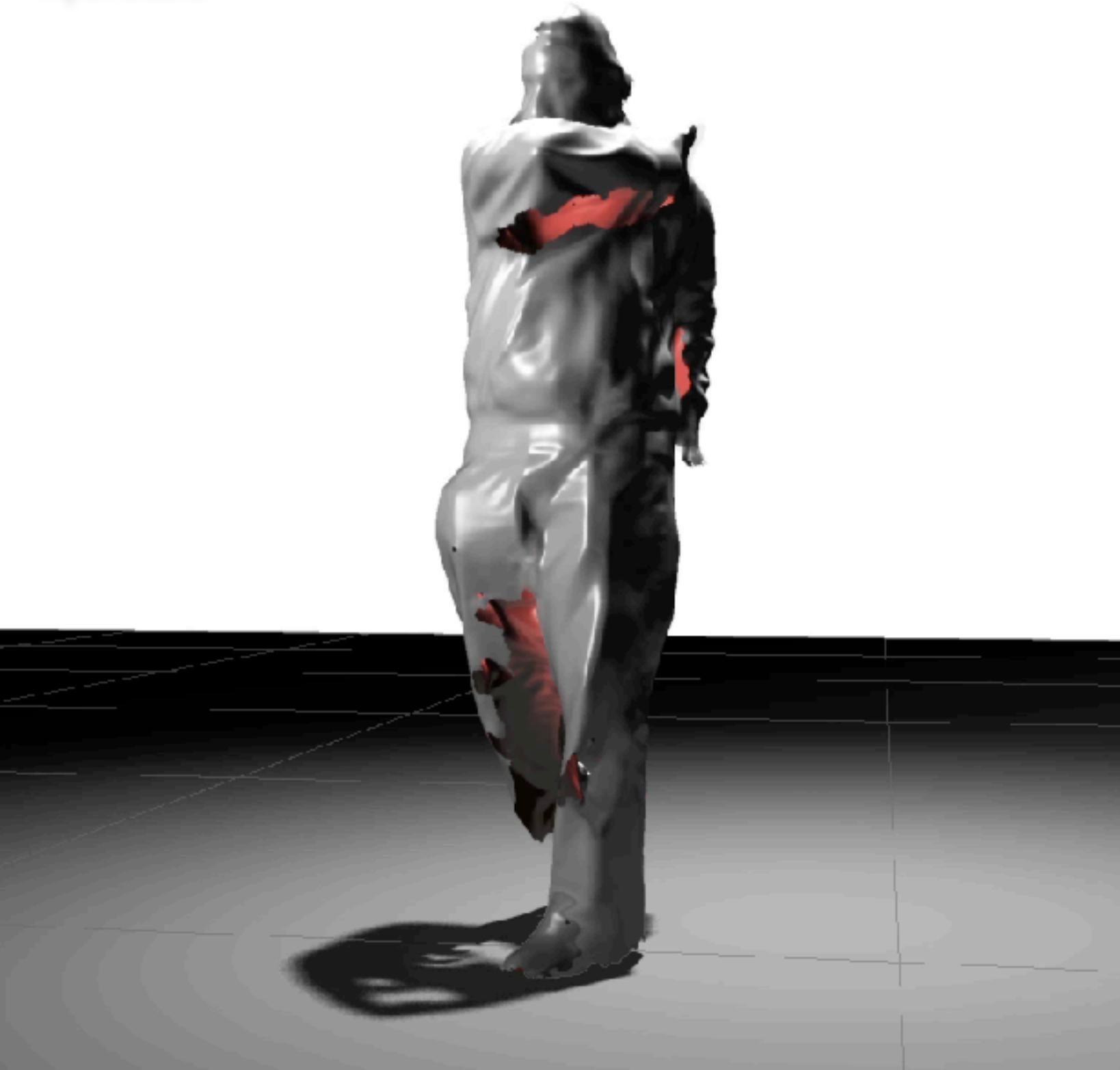
Propagating Correspondence through Holes



Results Abhijeet (textured)

[Chuang et al. '09]

input scans



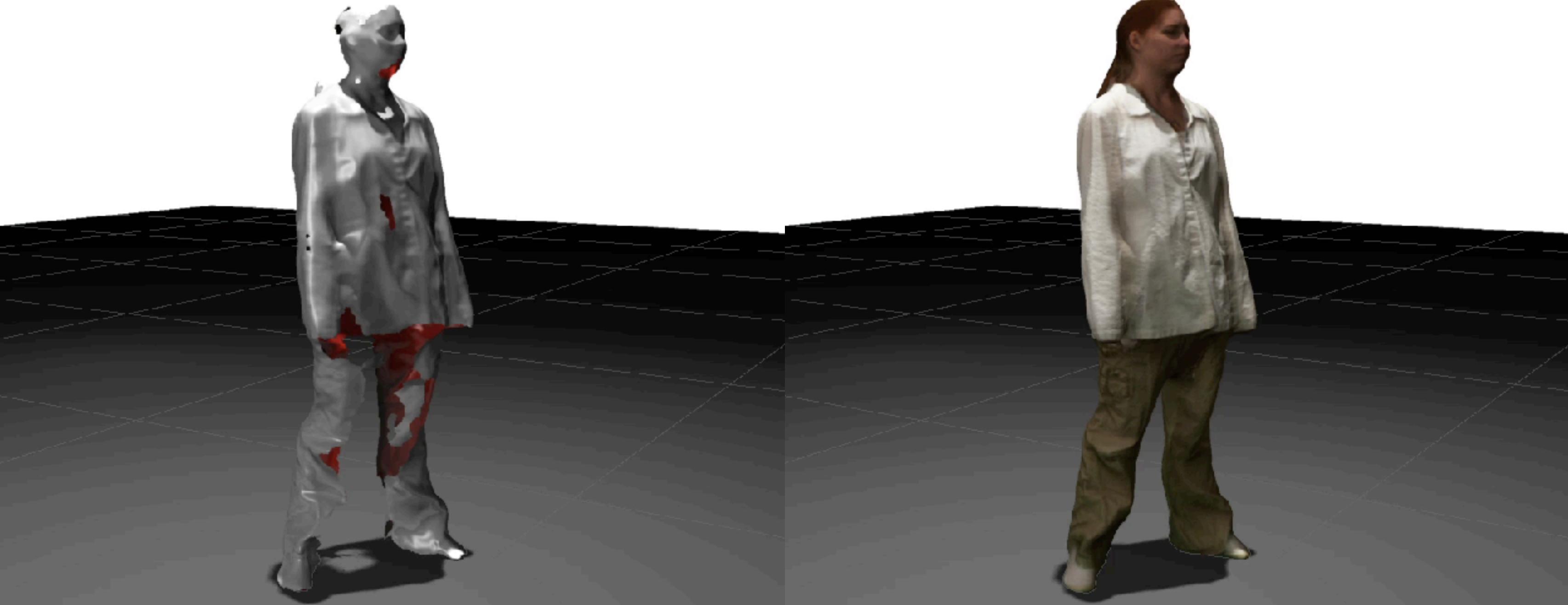
shape completion



Results Jay (Textured)

[Chuang et al. '09]

input scans



shape completion

Template-Based Reconstruction

input data

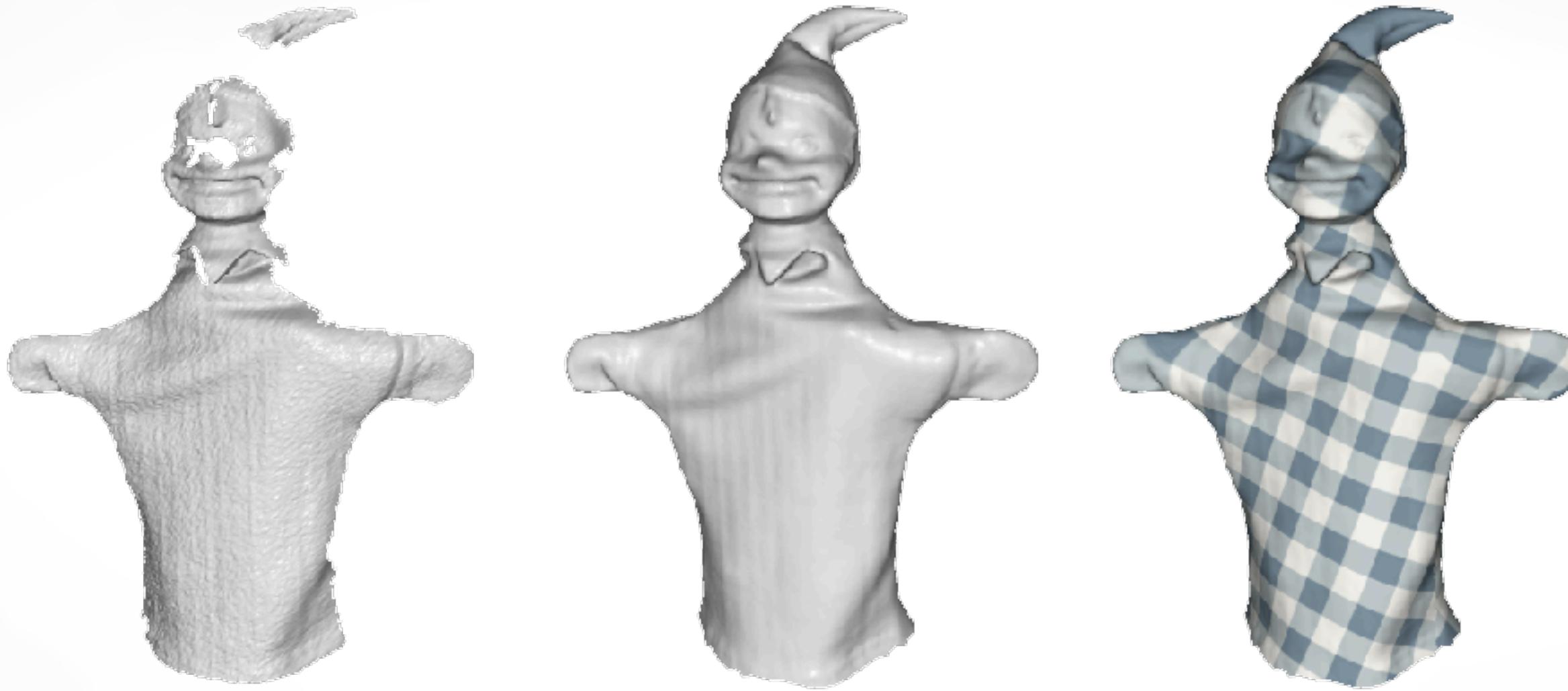


template fitting



data provided by Stanford and MPI Saarbrücken

Template-Based Reconstruction



Input Scans

Reconstruction

Textured Reconstruction

Template-Based Reconstruction



Input Scans

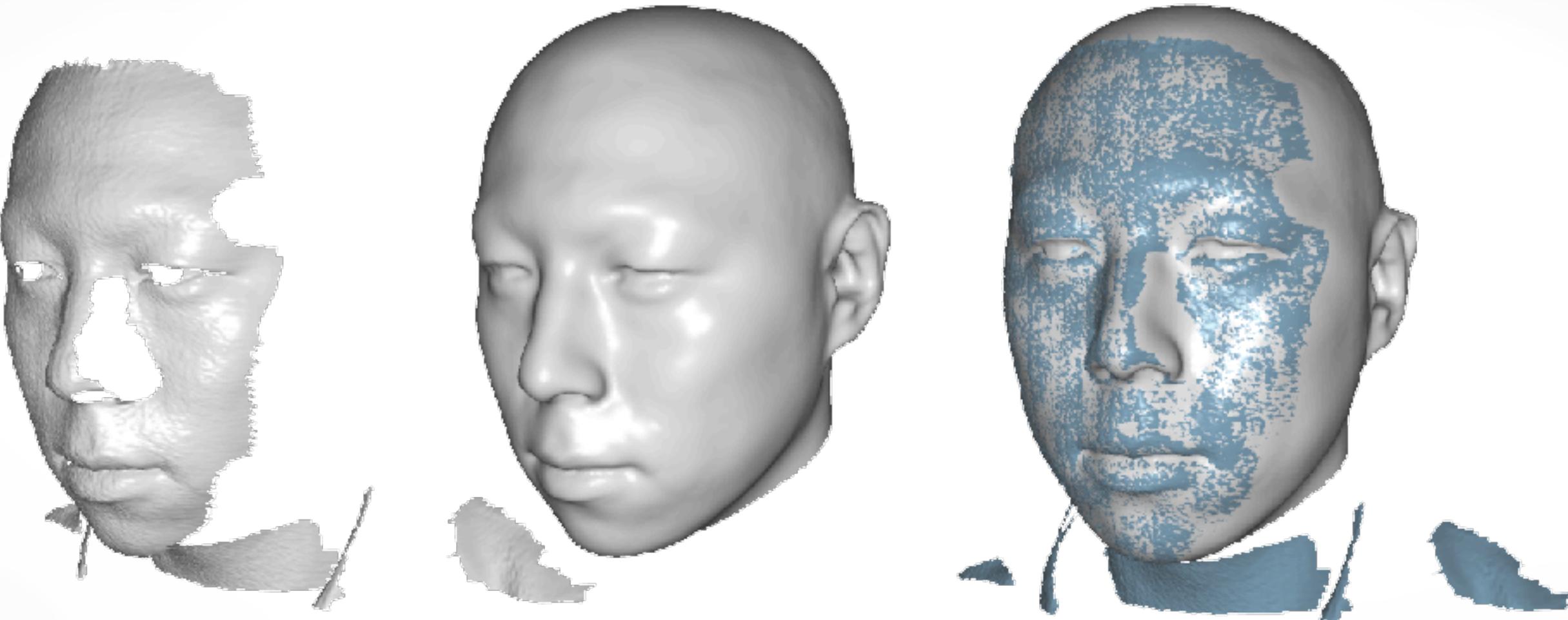


Reconstruction



Textured Reconstruction

Template-Based Reconstruction

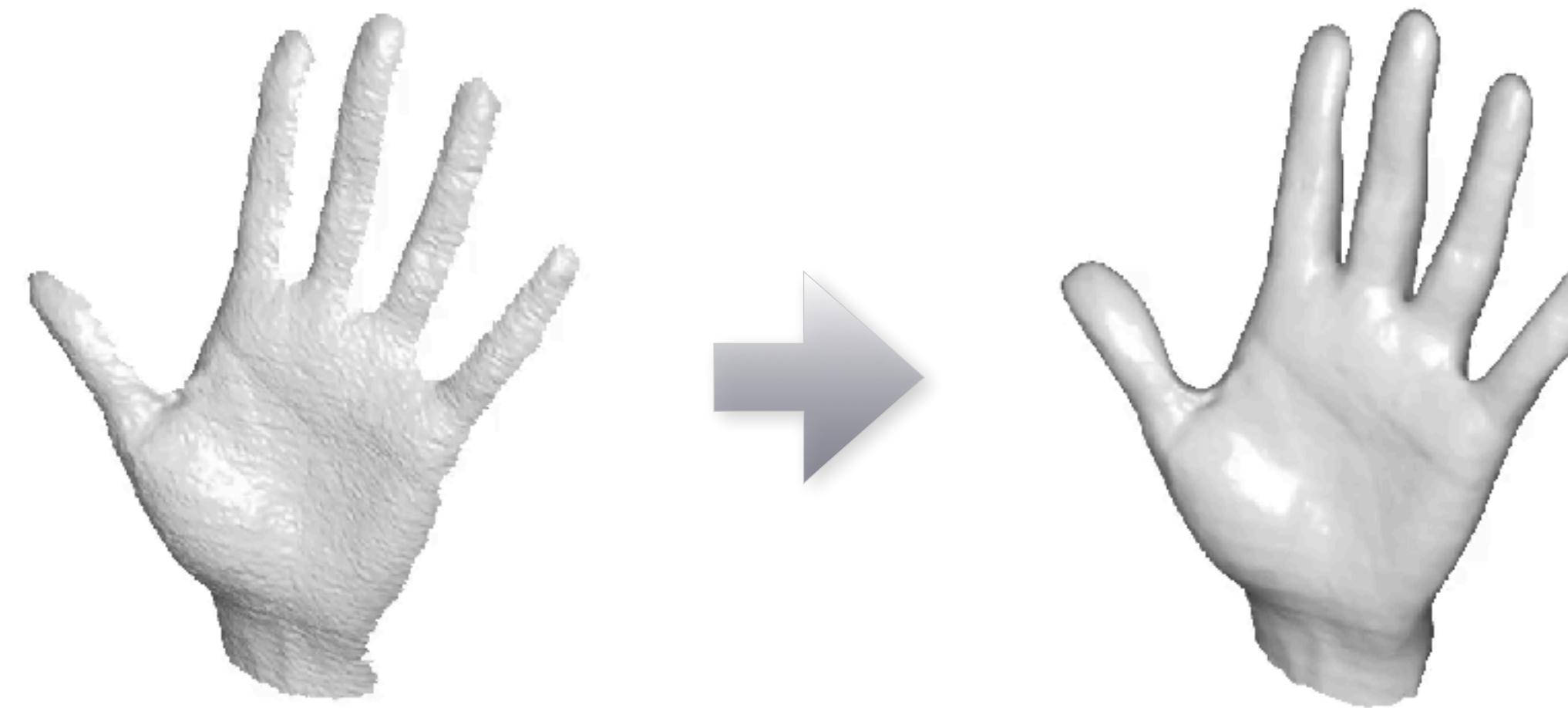


Input Scans

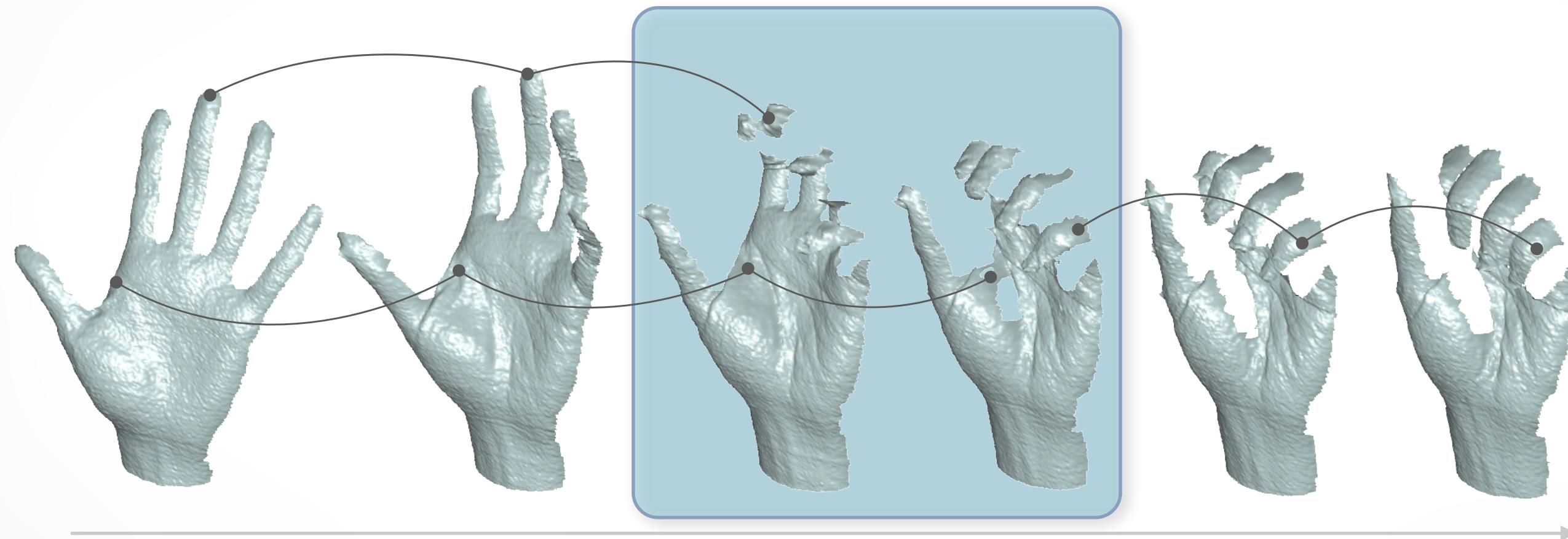
Reconstruction

Overlaid Scans

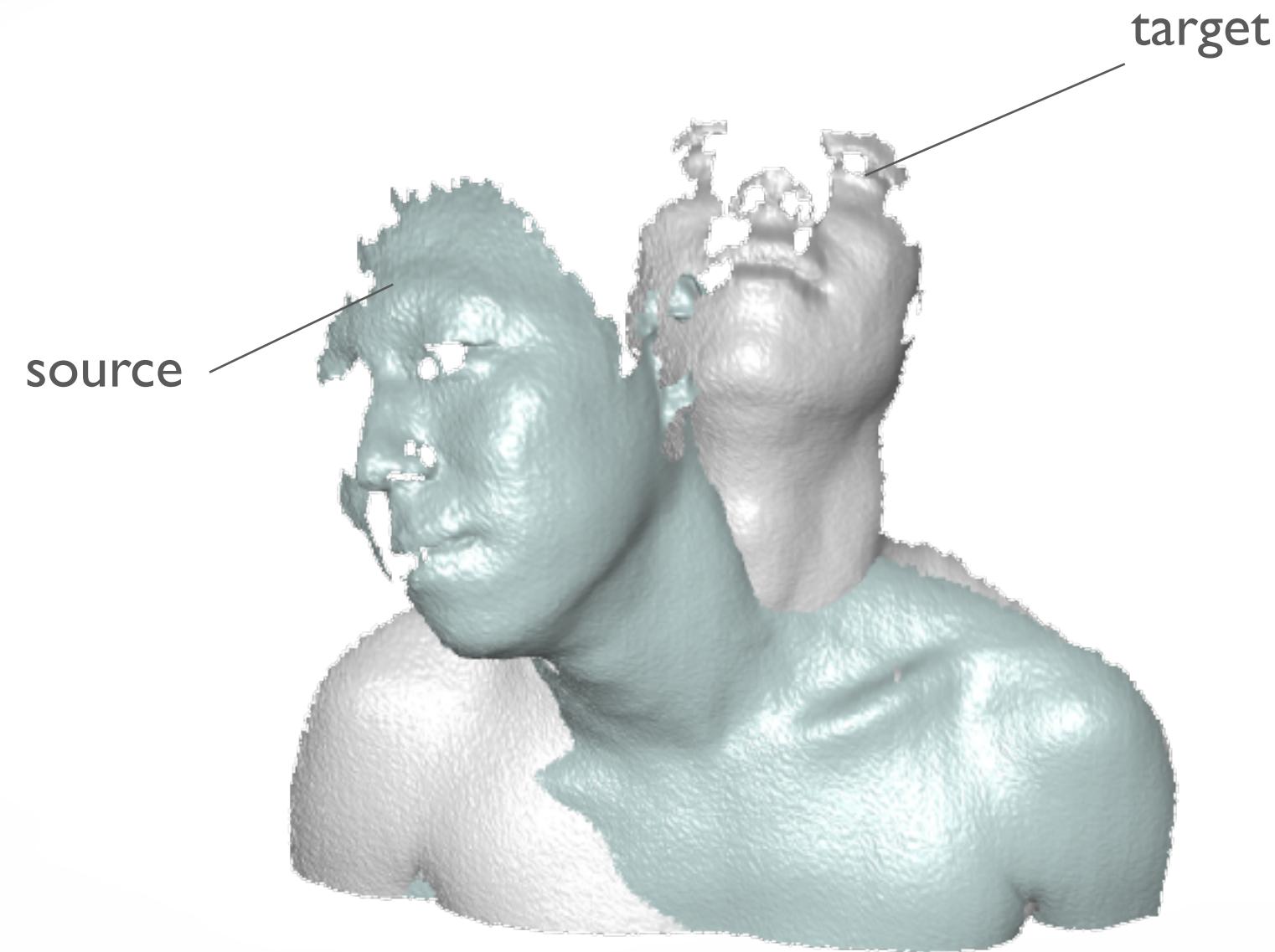
Template-Based Reconstruction



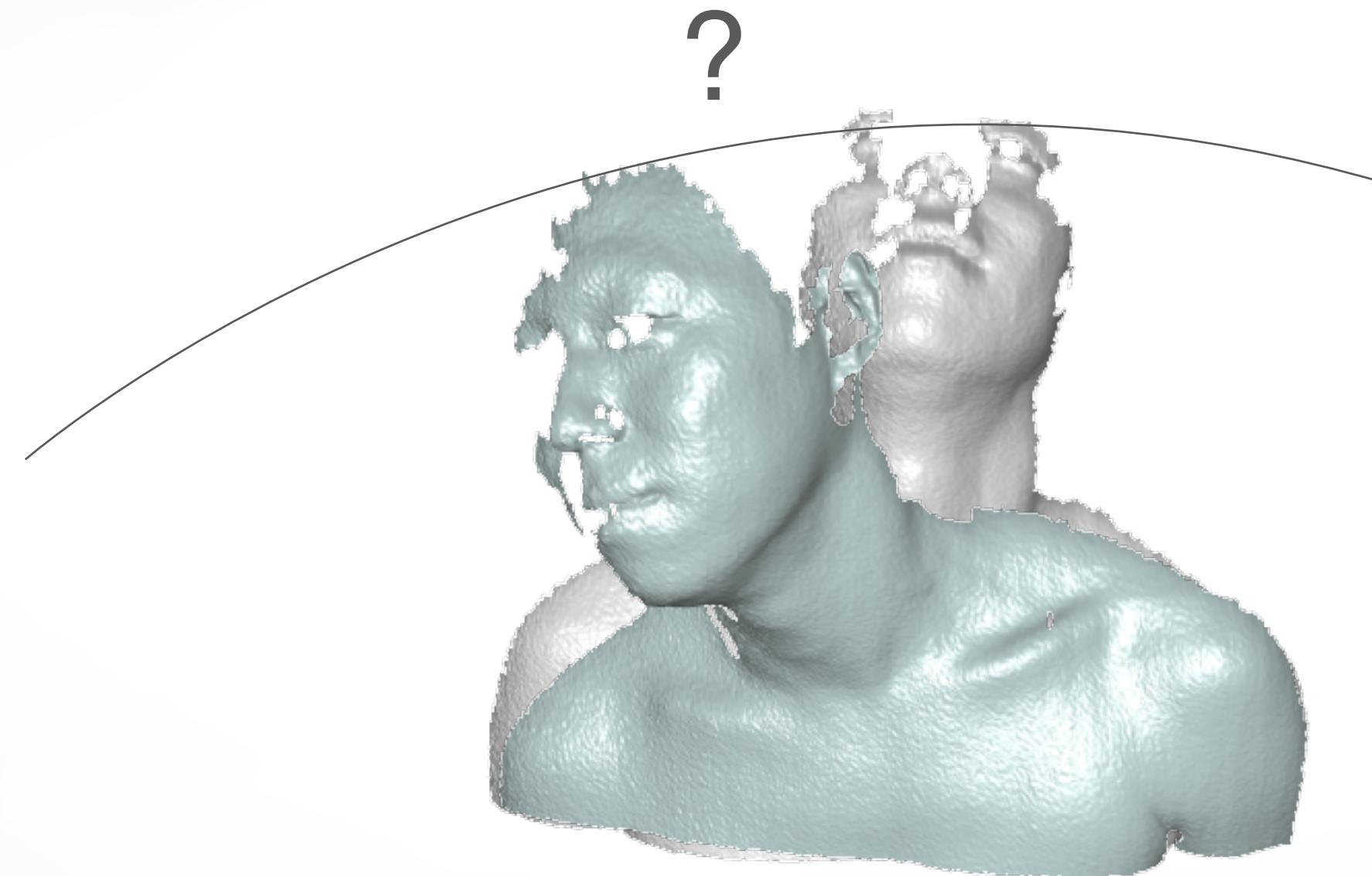
Correspondence Problem



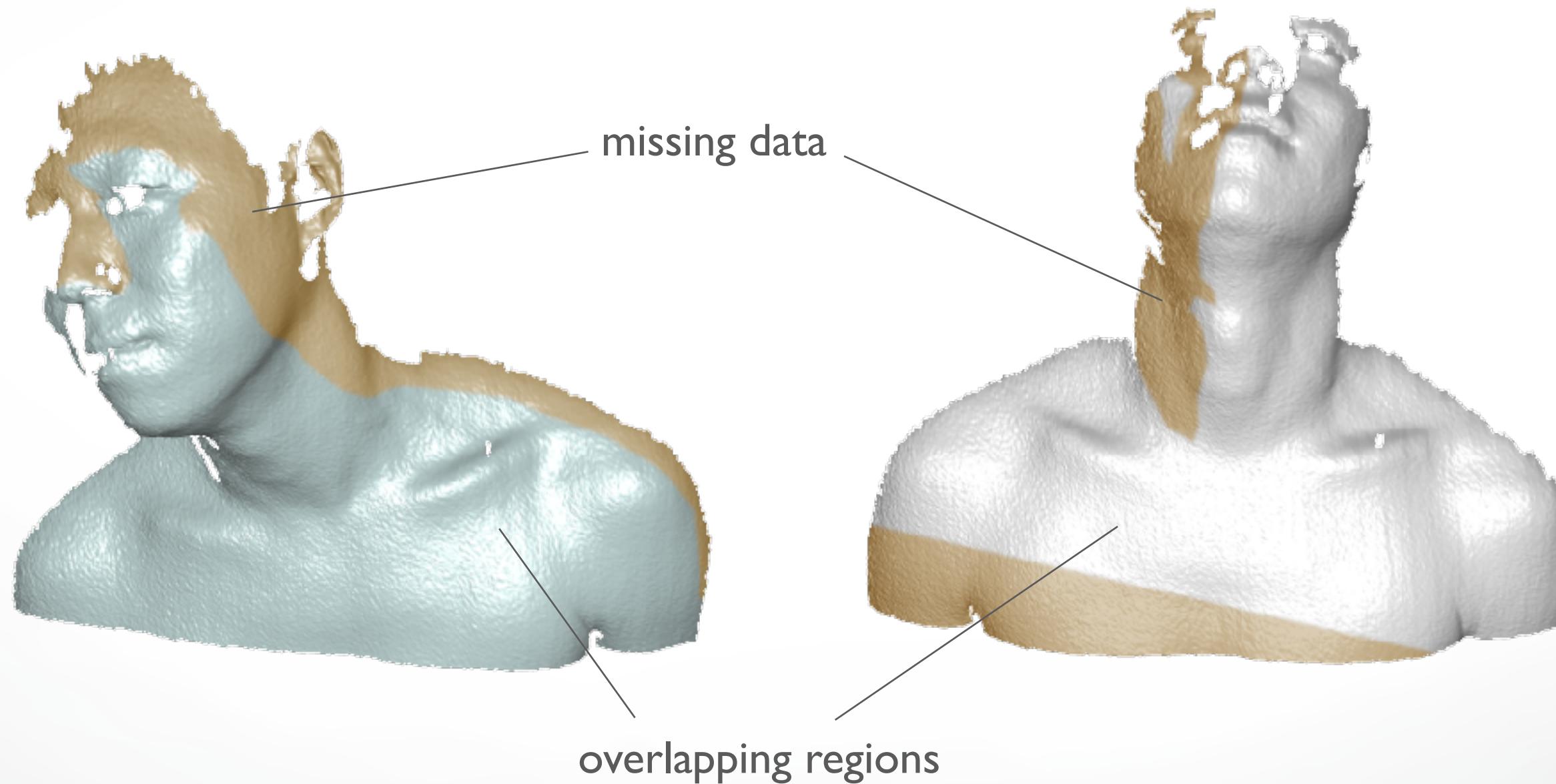
Pair of 3D Scans



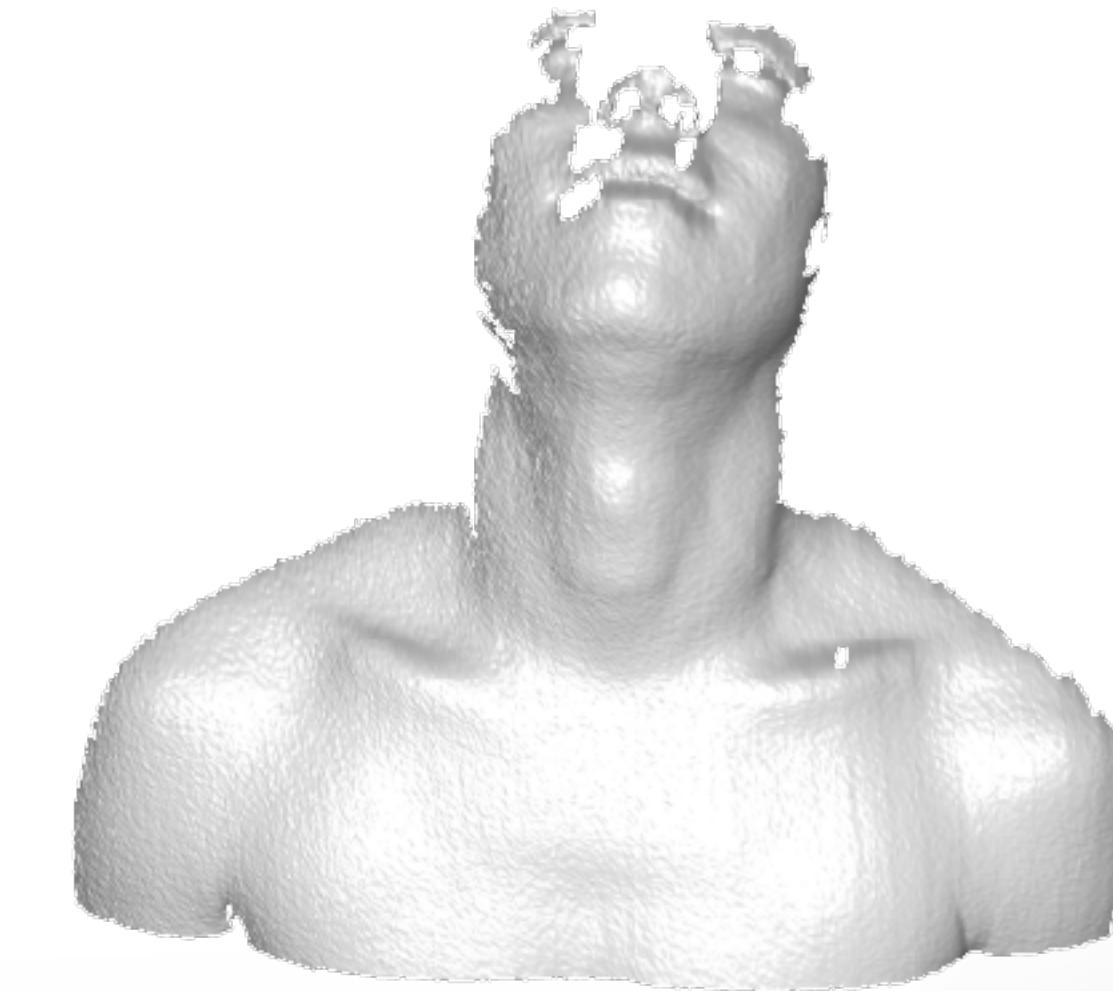
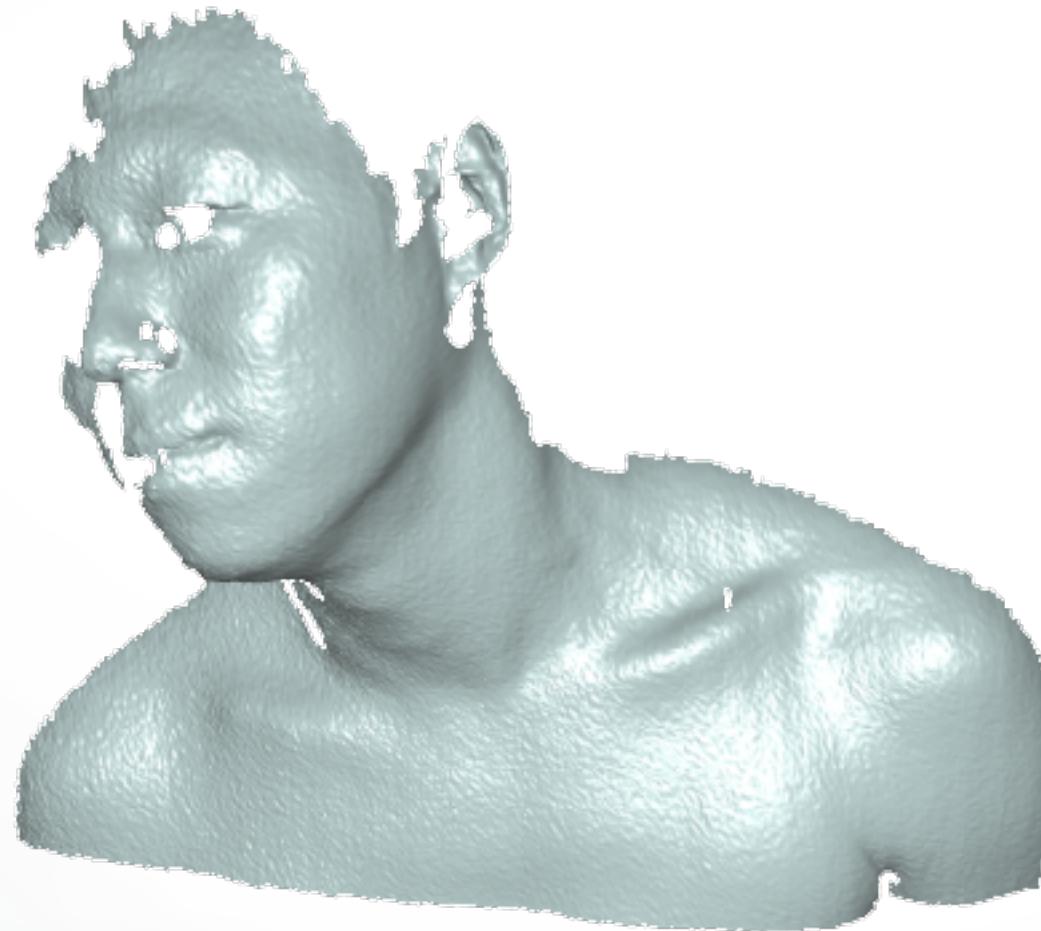
Correspondences are Lost



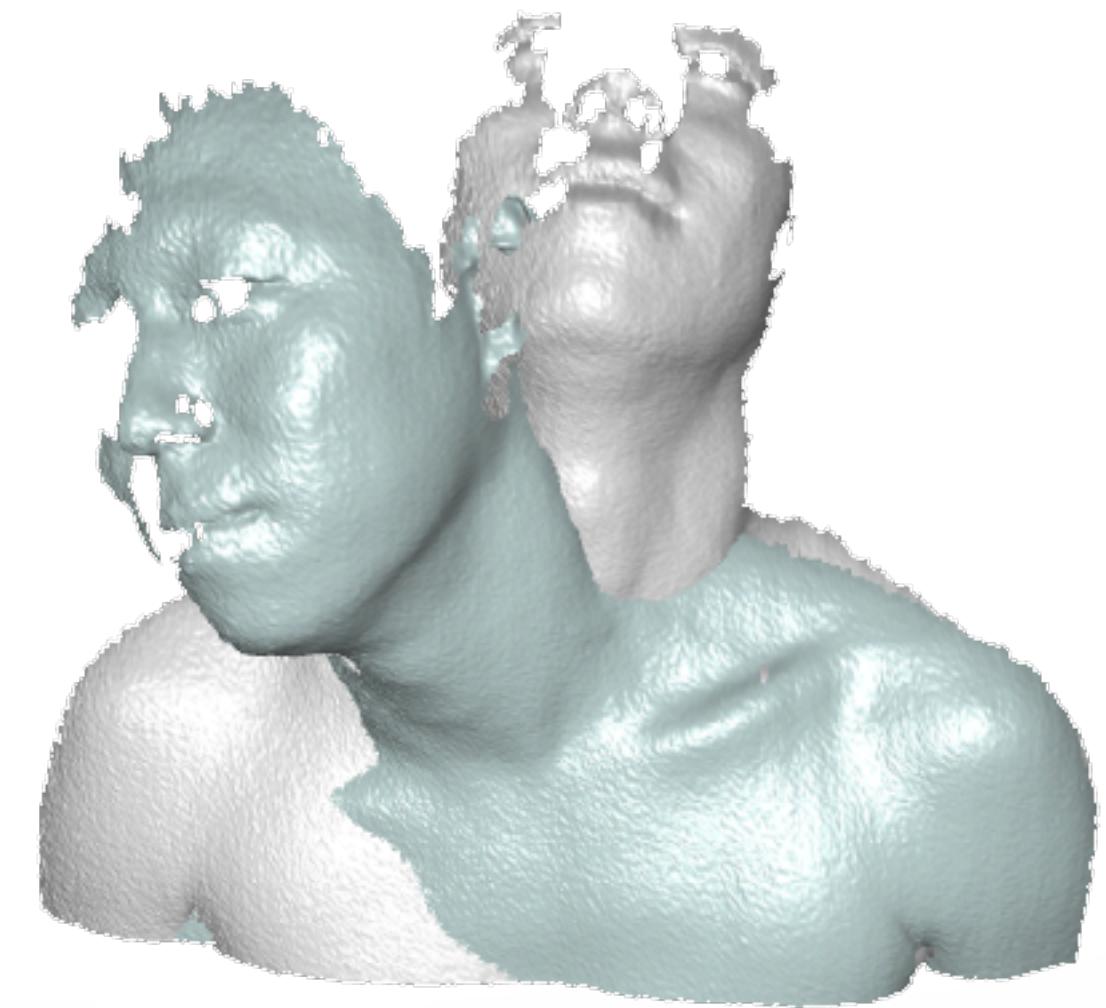
Overlapping Regions are Lost



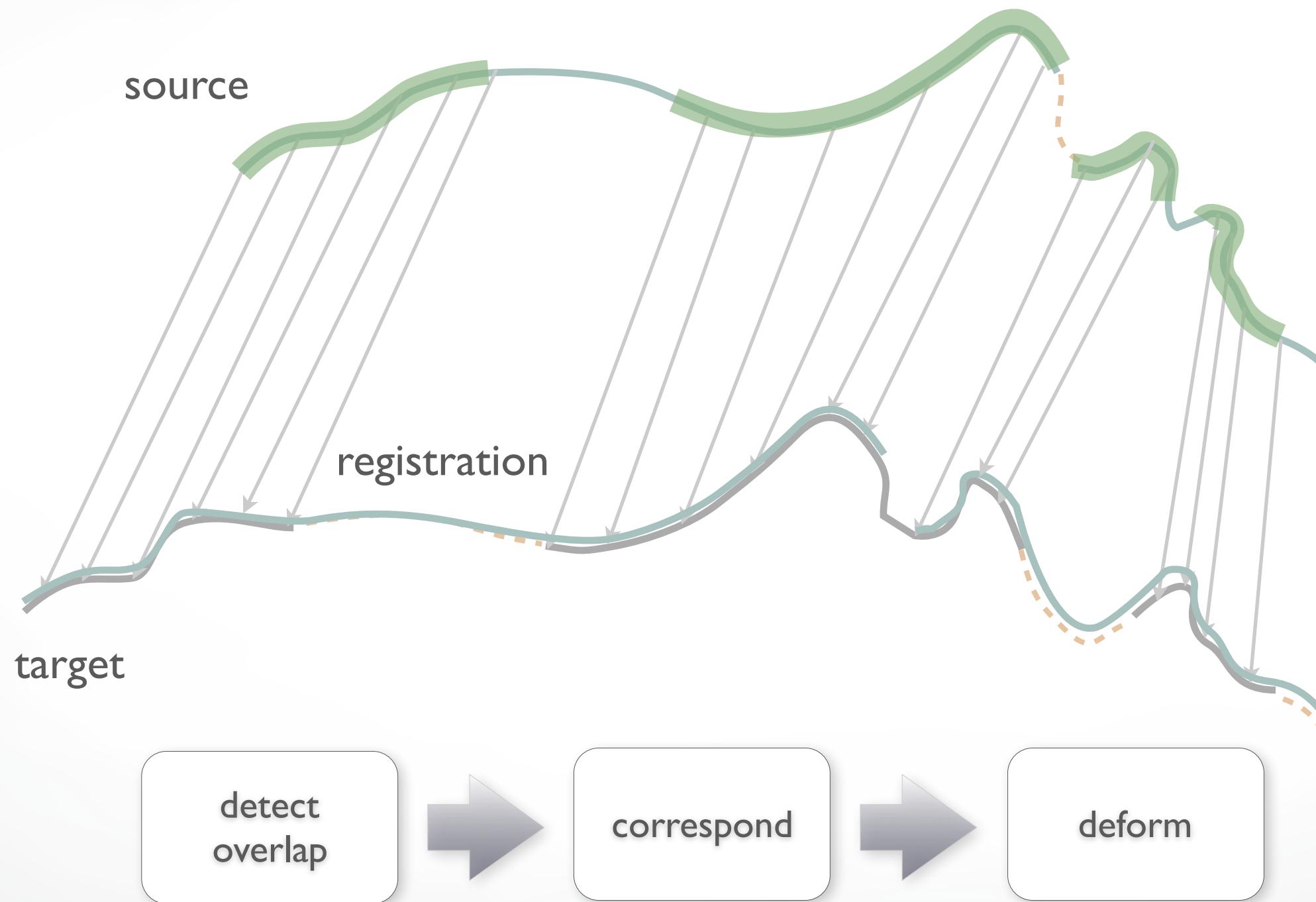
Overlapping Regions are Lost



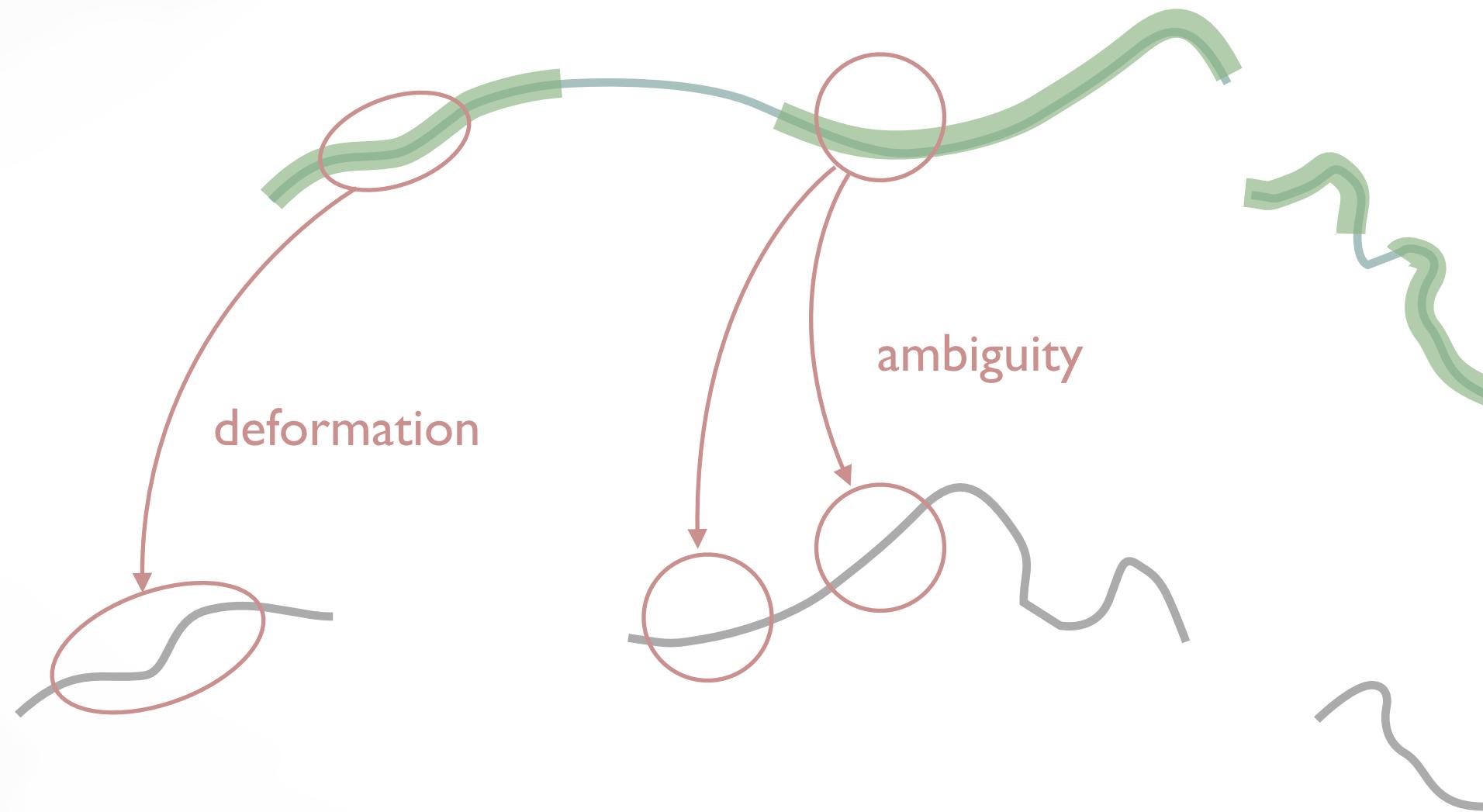
Non-Rigid Registration



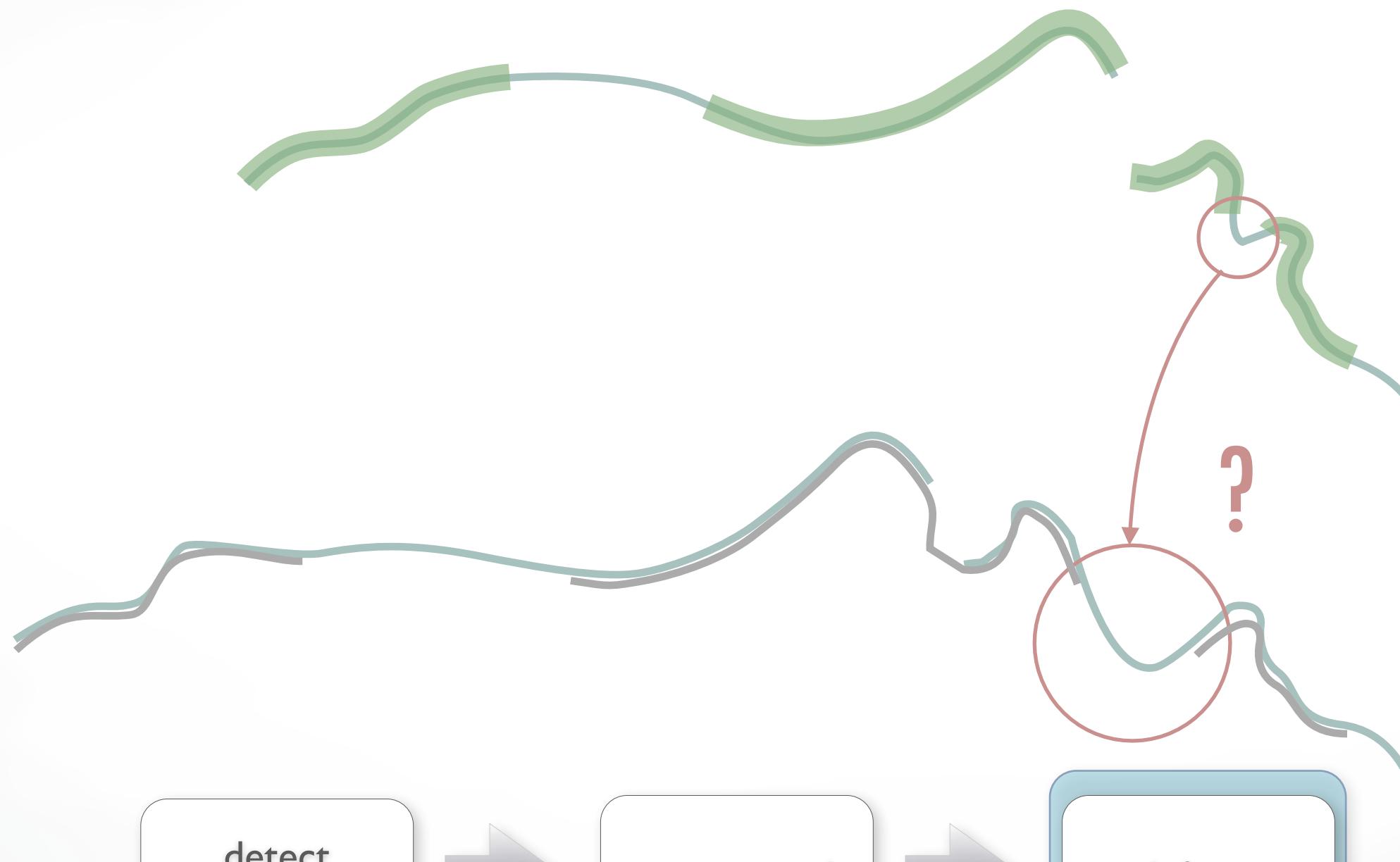
Recipe



Challenge



Challenge



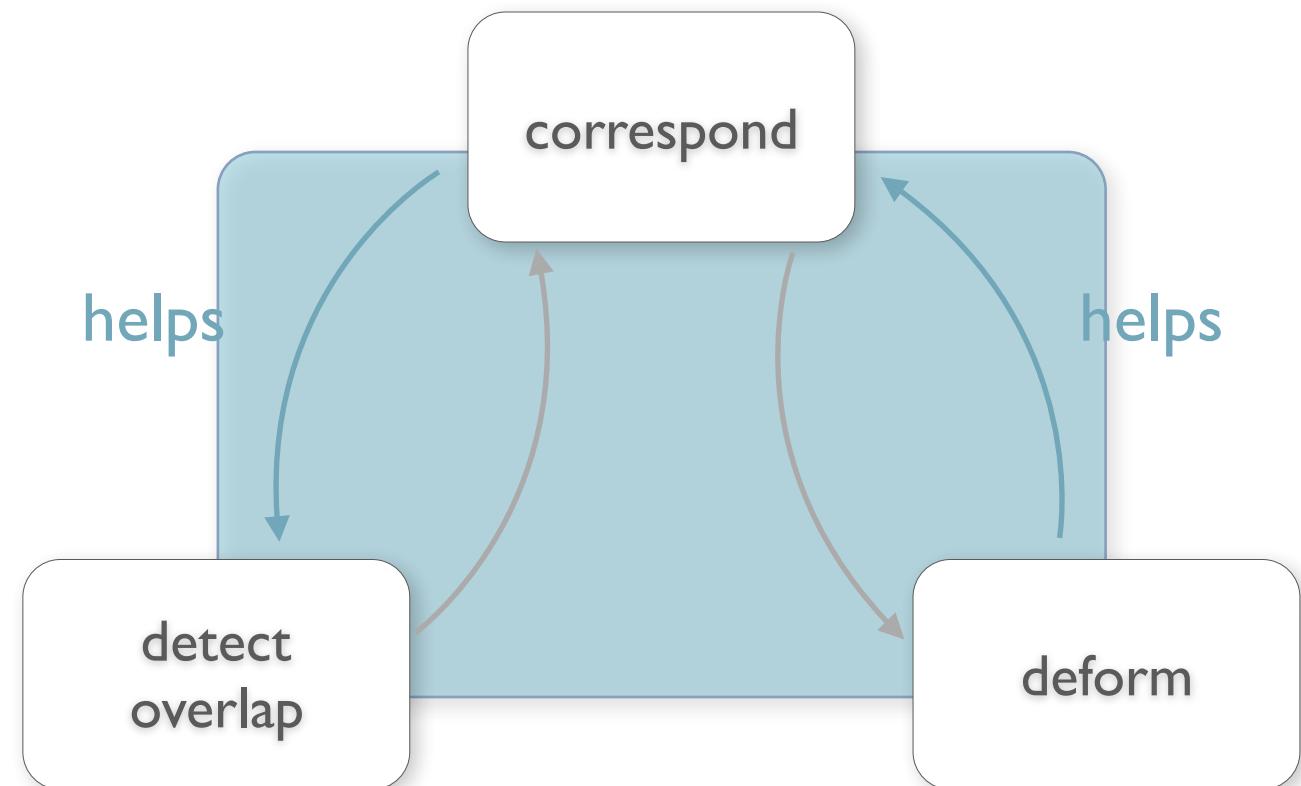
Challenge

detect
overlap

correspond

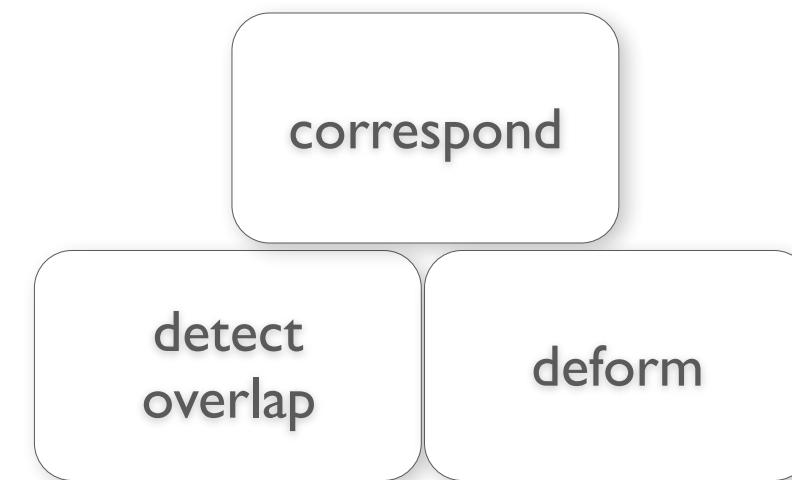
deform

Observation

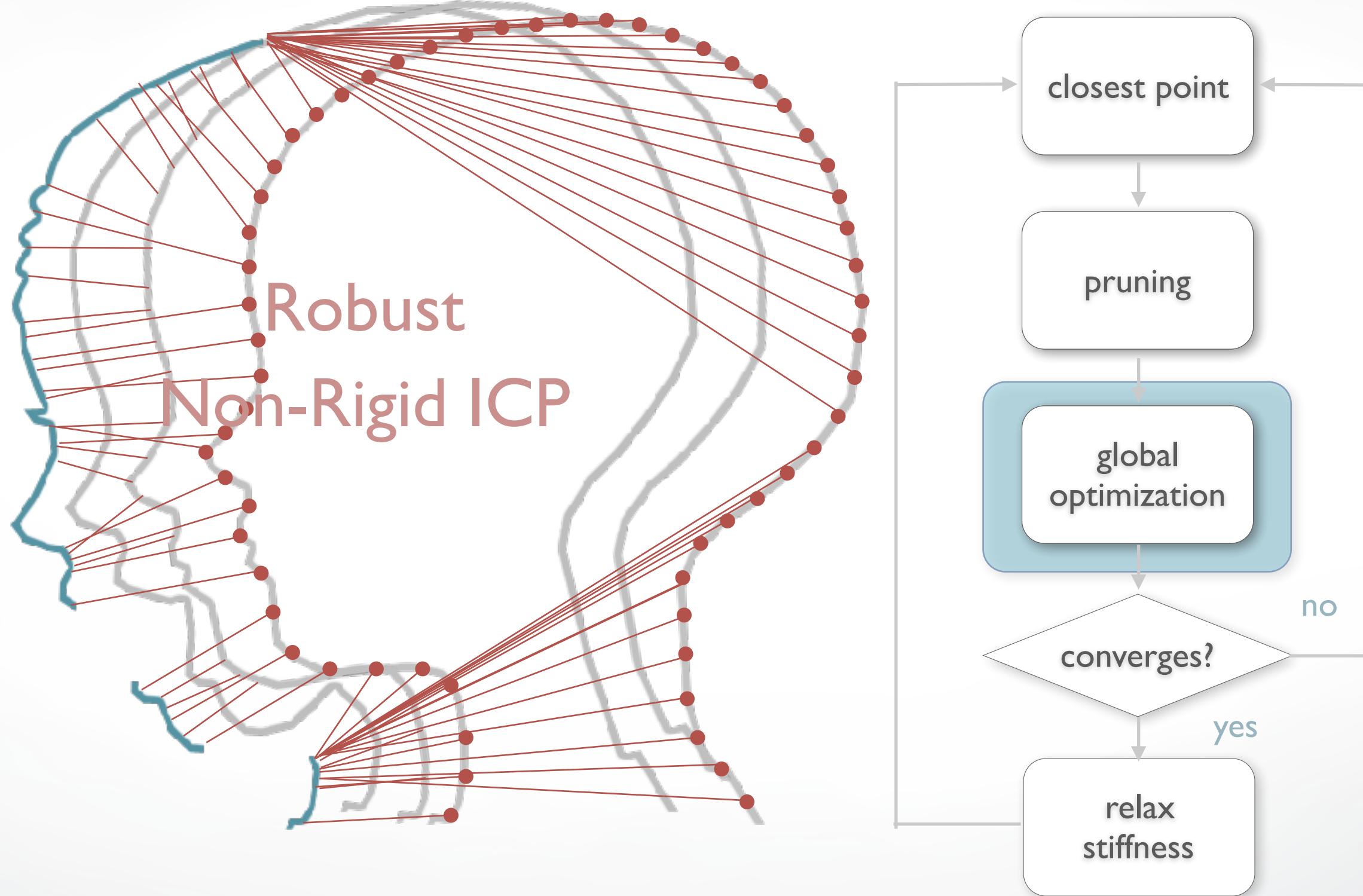


global optimization via local refinement

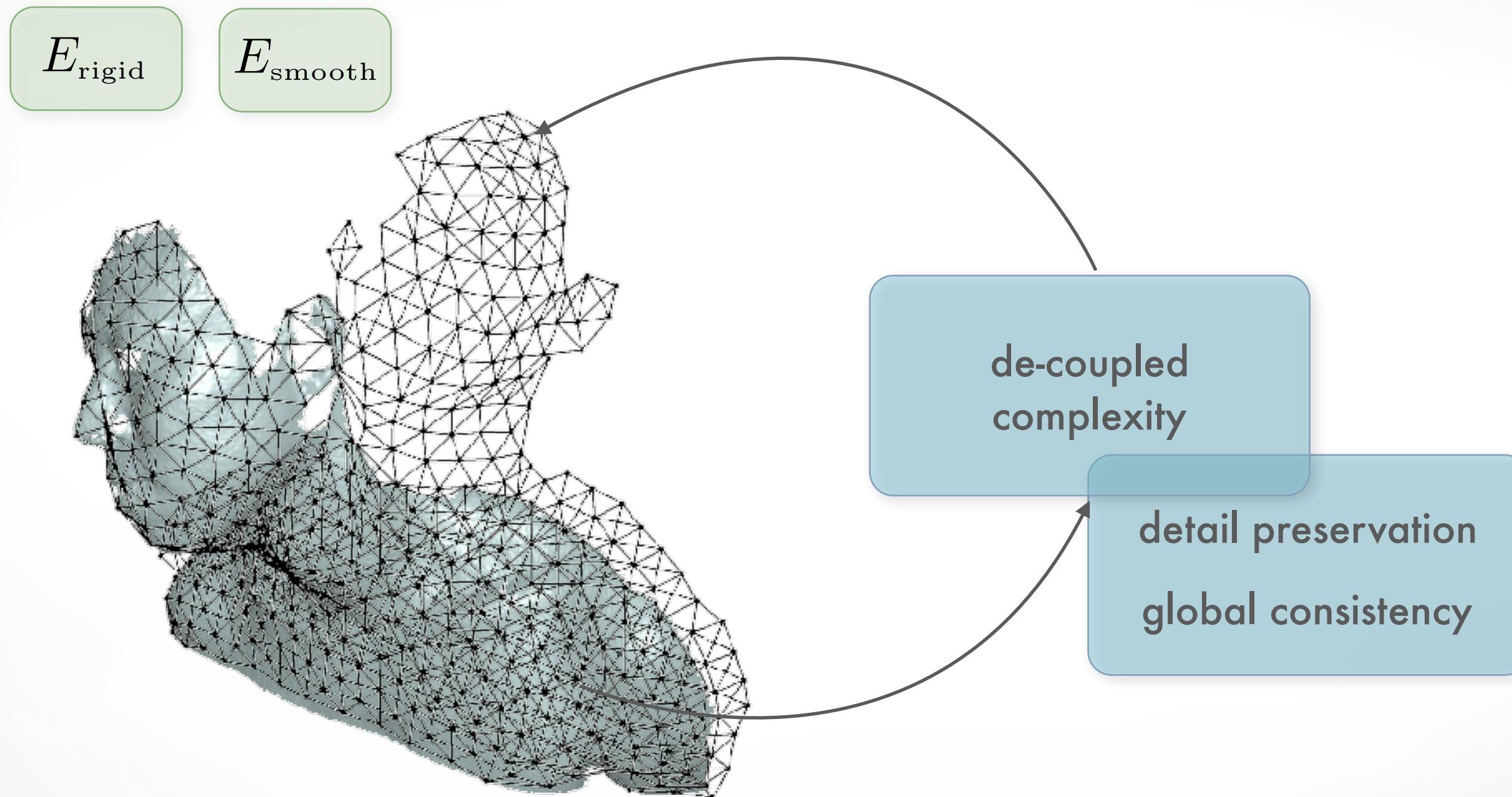
Iterative Optimization



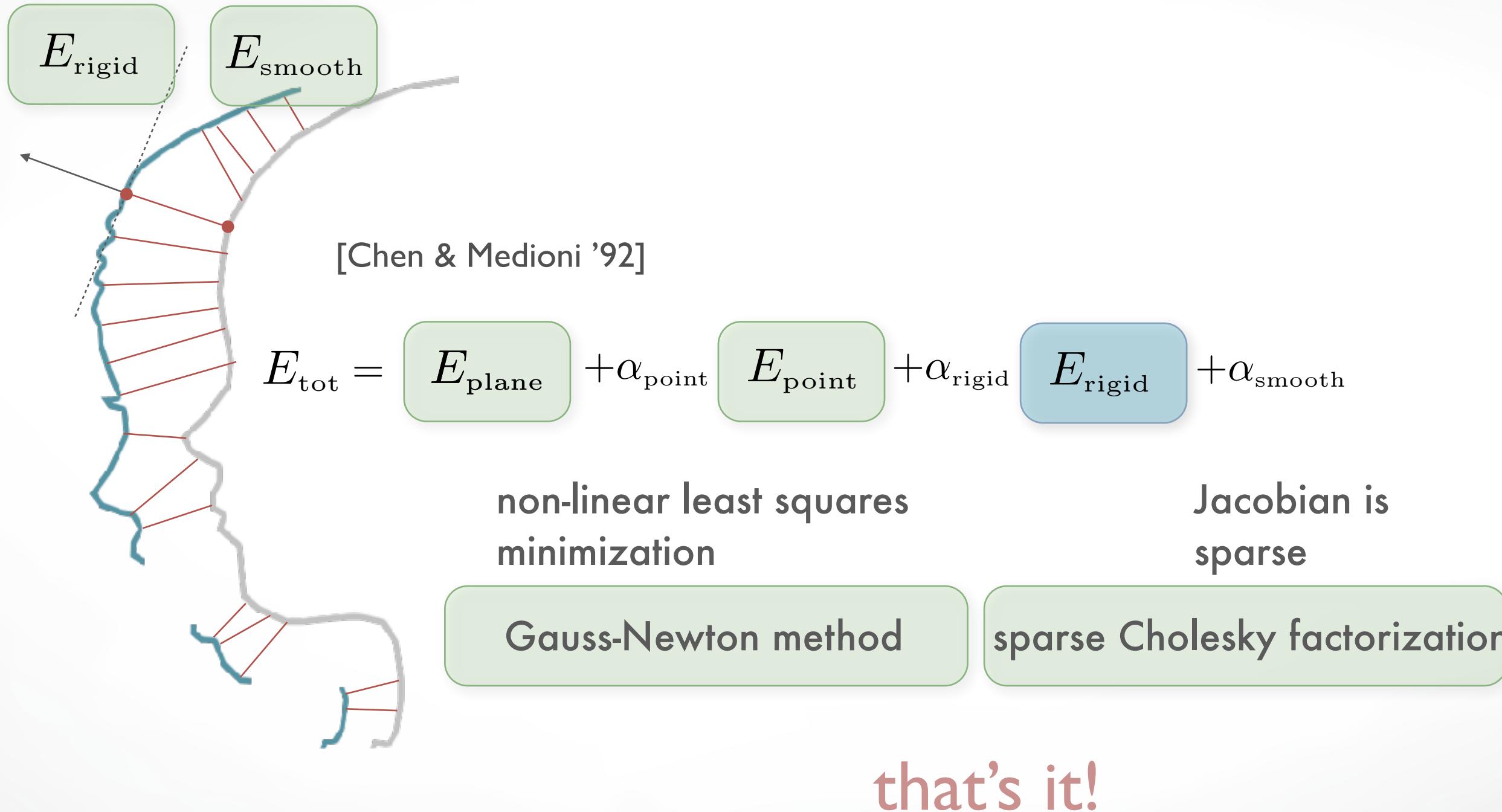
Iterative Optimization



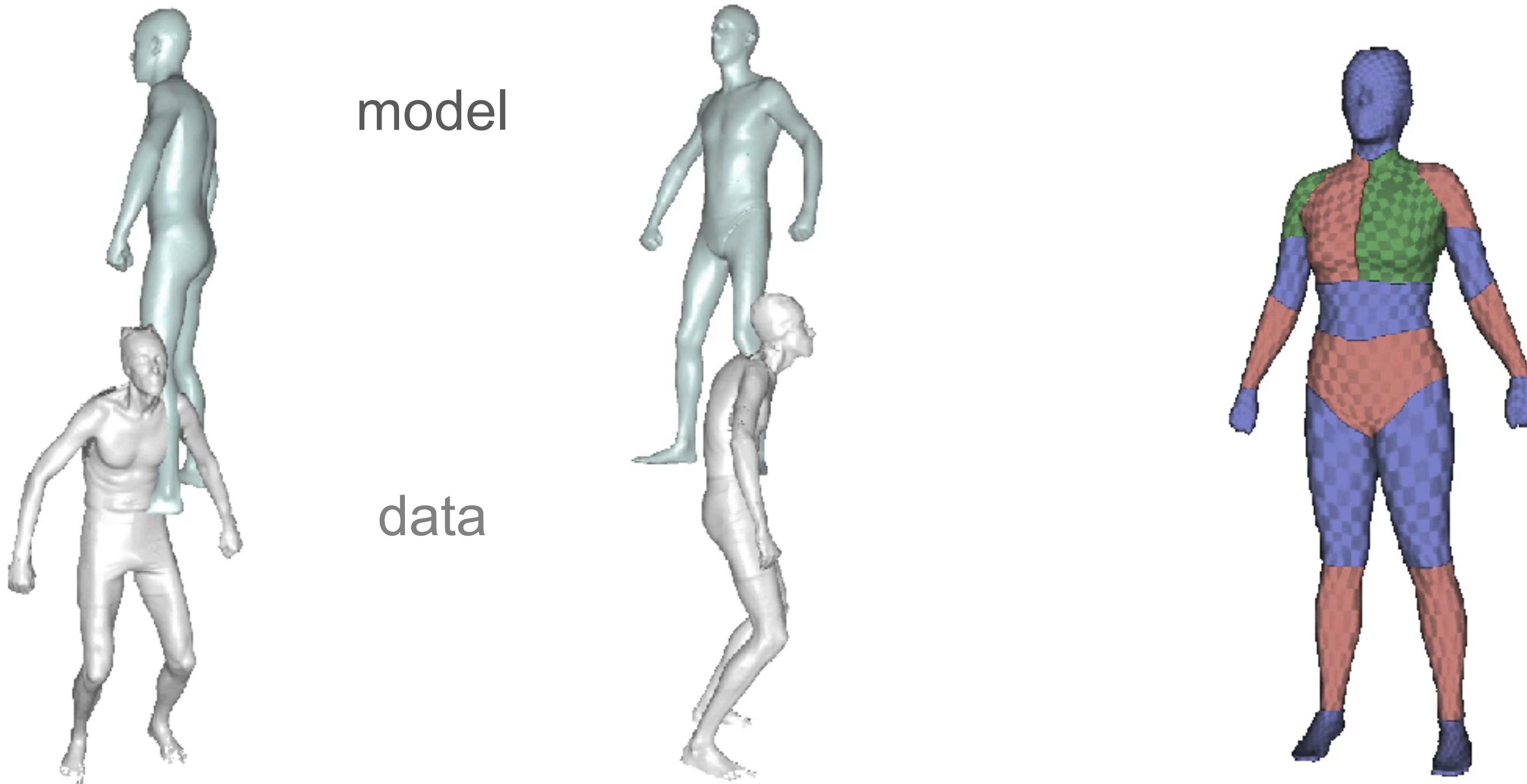
Deformation Model



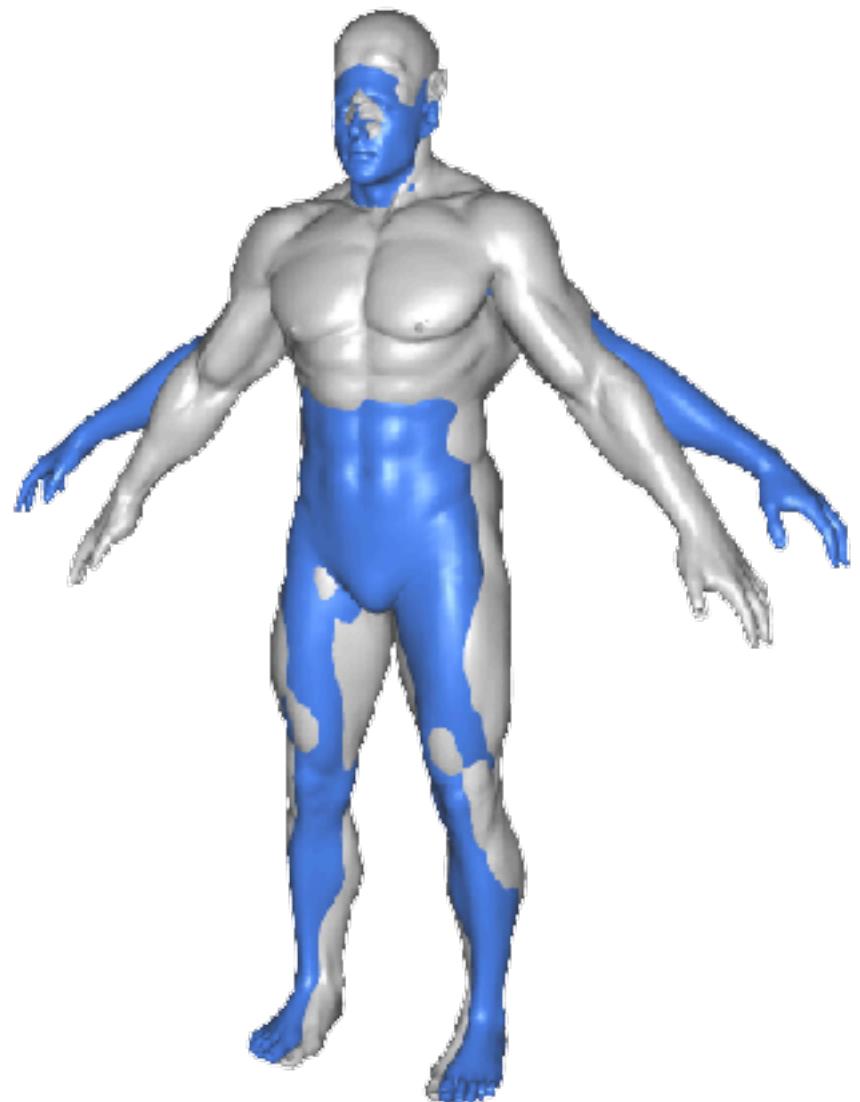
Non-Linear Energy Minimization



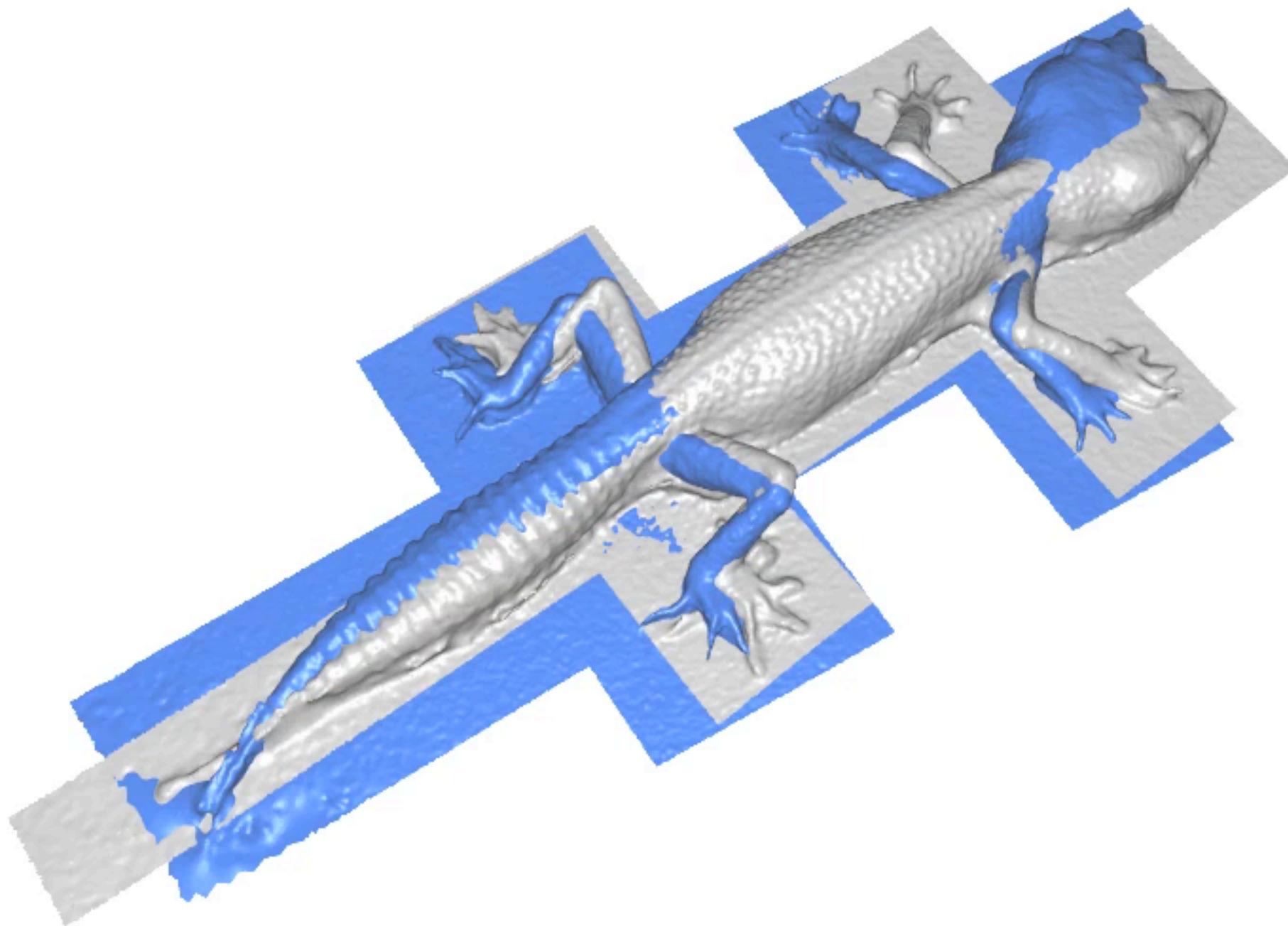
Shape Analysis



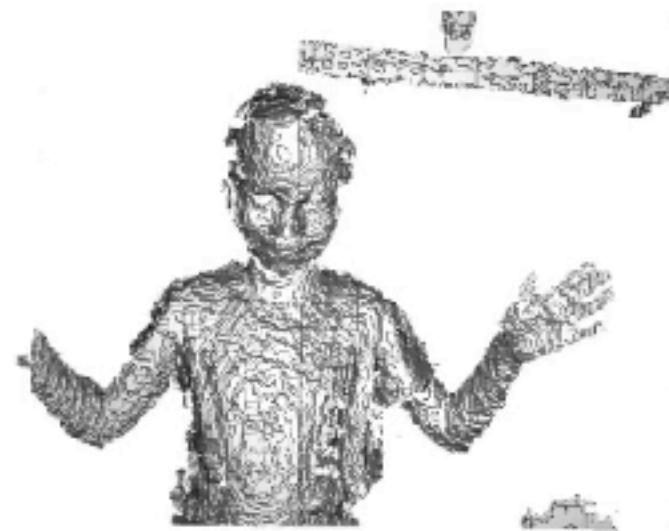
Re-topology



Skin Texture Analysis



Dynamic Fusion: Real-Time Non-Rigid Registration



Live Input Depth Map



Live Model Output



Live RGB Image (unused)

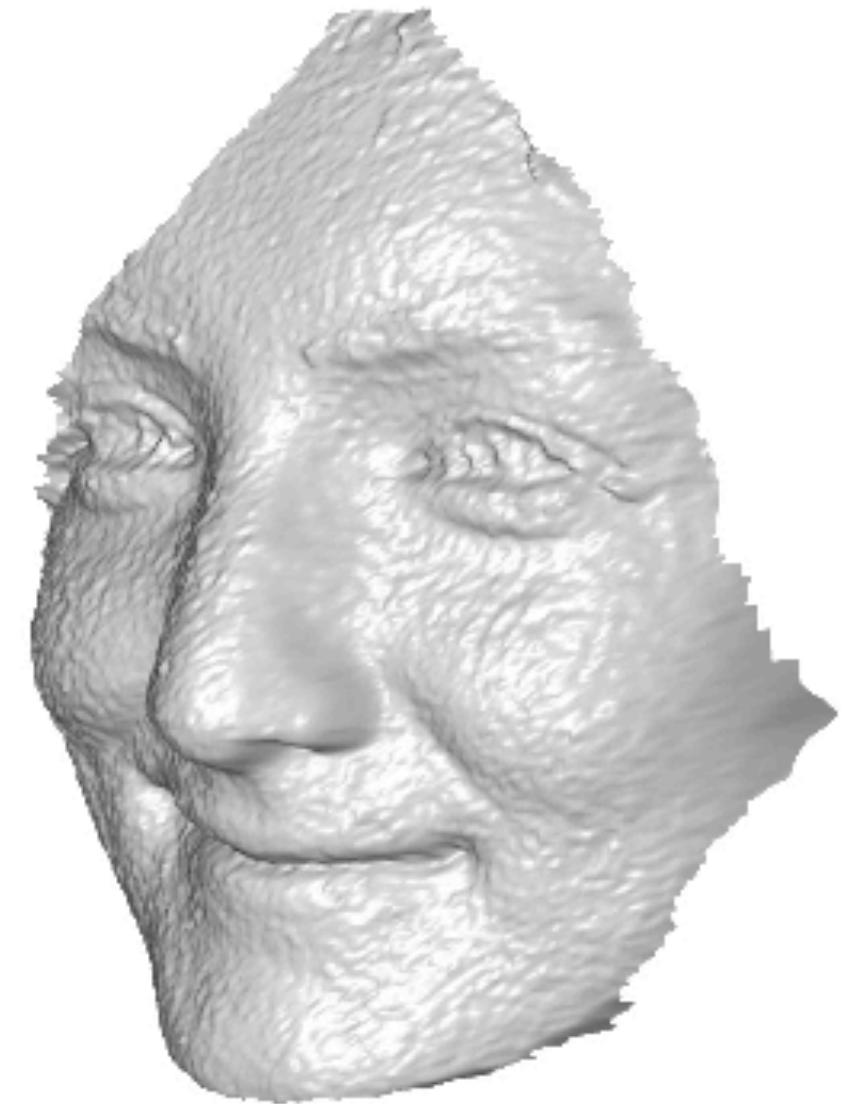


Canonical Model Reconstruction



Warped Model

Using a Linear PCA Model



Facial Performance Capture

Universal Capture



tracked template model



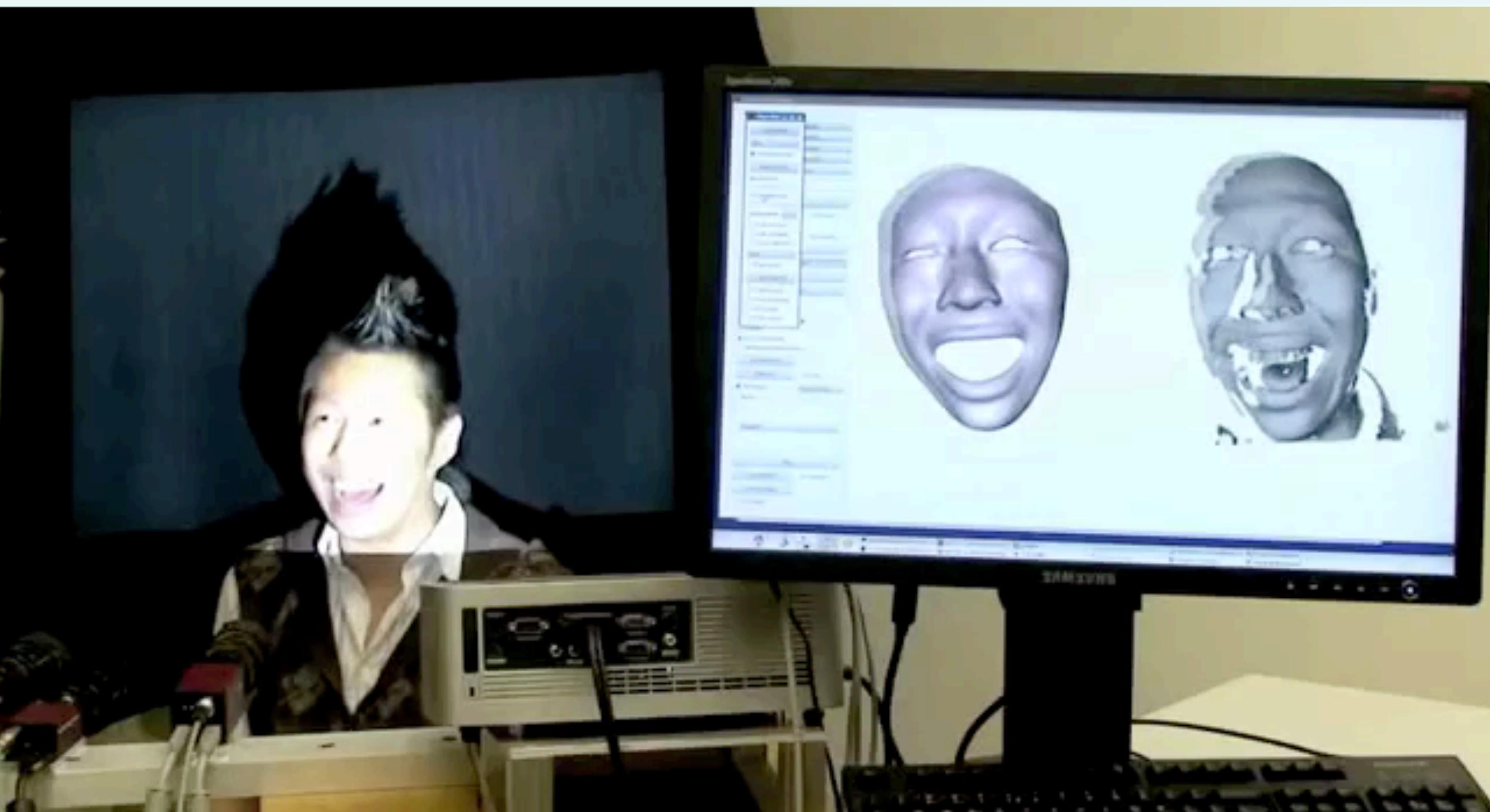
shape completion in prior model

universal capture

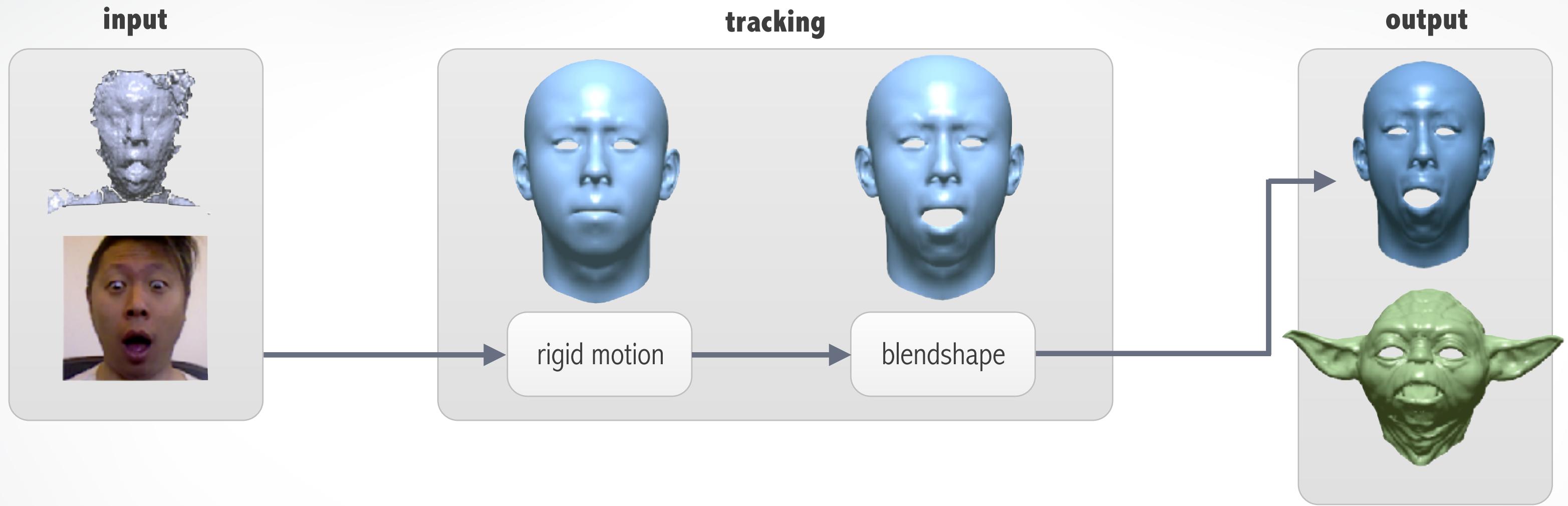


Realtime 3D Scanning

Weise et al. 2009



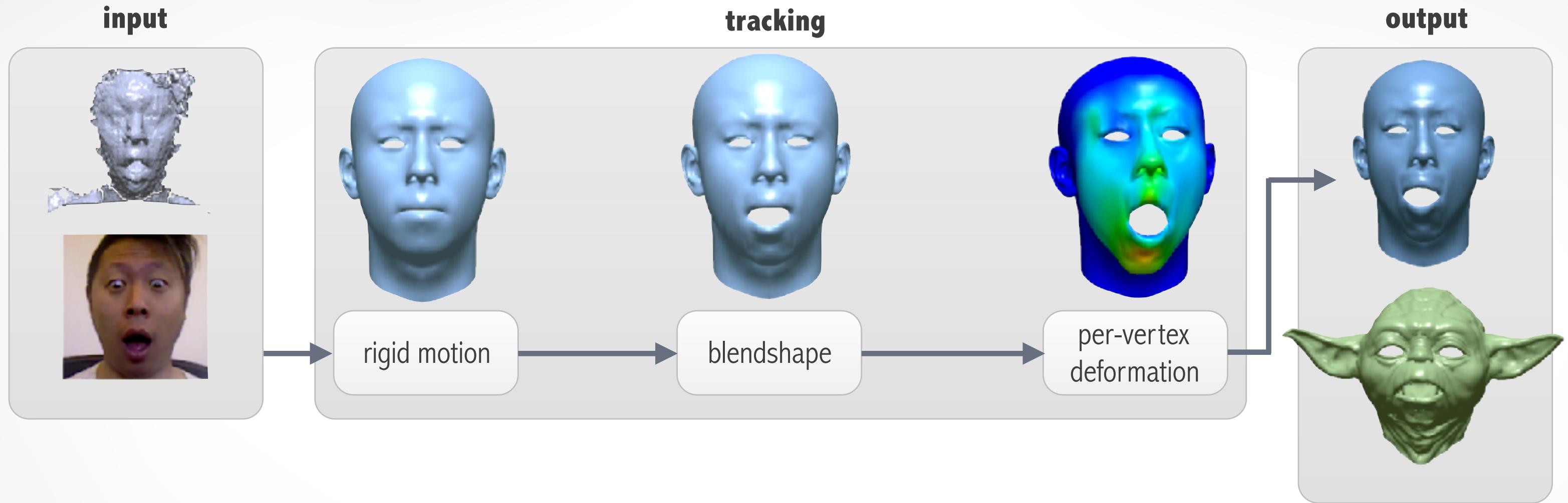
Pipeline Overview



$$\mathbf{v}_i(\mathbf{x}) = \mathbf{v}_i^{(0)} + \sum_l \mathbf{v}_i^{(l)} x_l$$

$$x_l \in [0, 1]$$

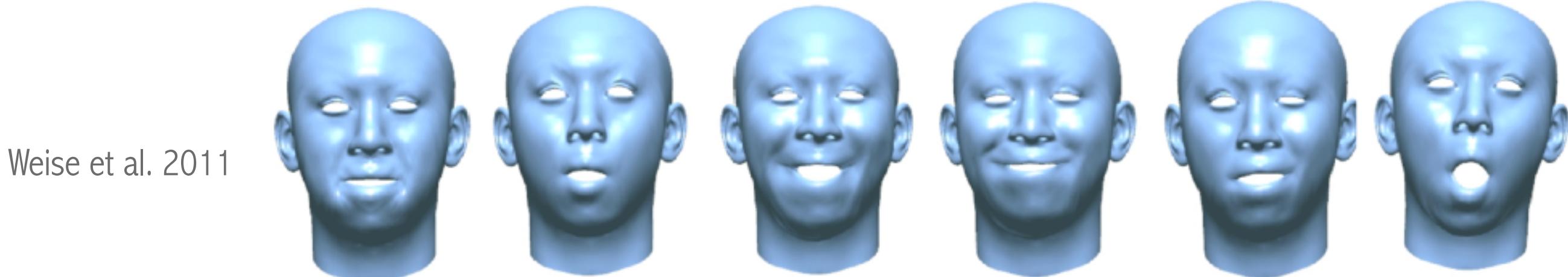
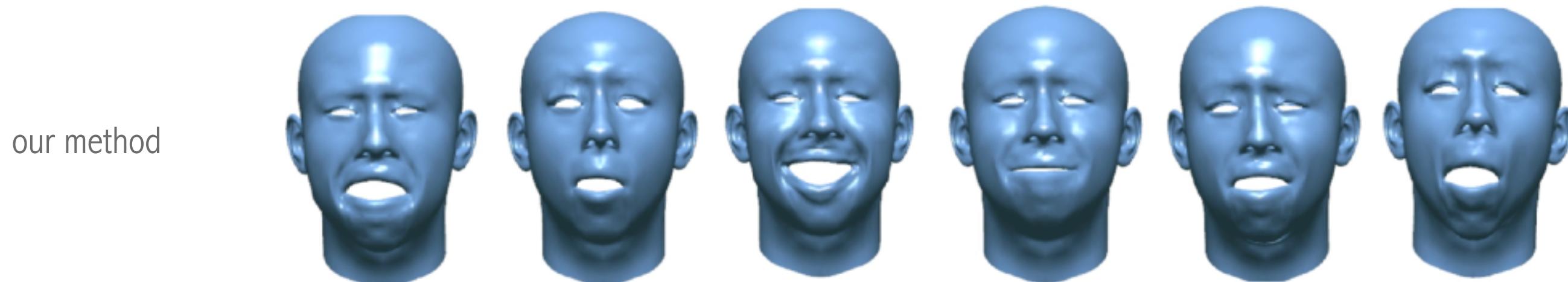
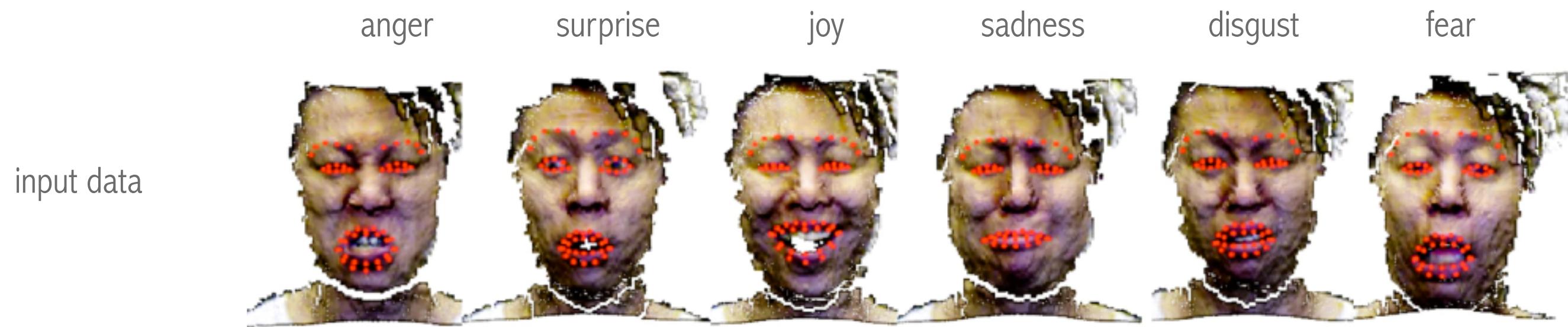
Pipeline Overview



$$\tilde{\mathbf{v}}_i(\Delta \mathbf{v}_i) = \mathbf{v}_i + \Delta \mathbf{v}_i$$

Tracking **Basic Emotions**

Li et al. SIGGRAPH 2013



Facial Performance Capture

Li et al. SIGGRAPH 2013



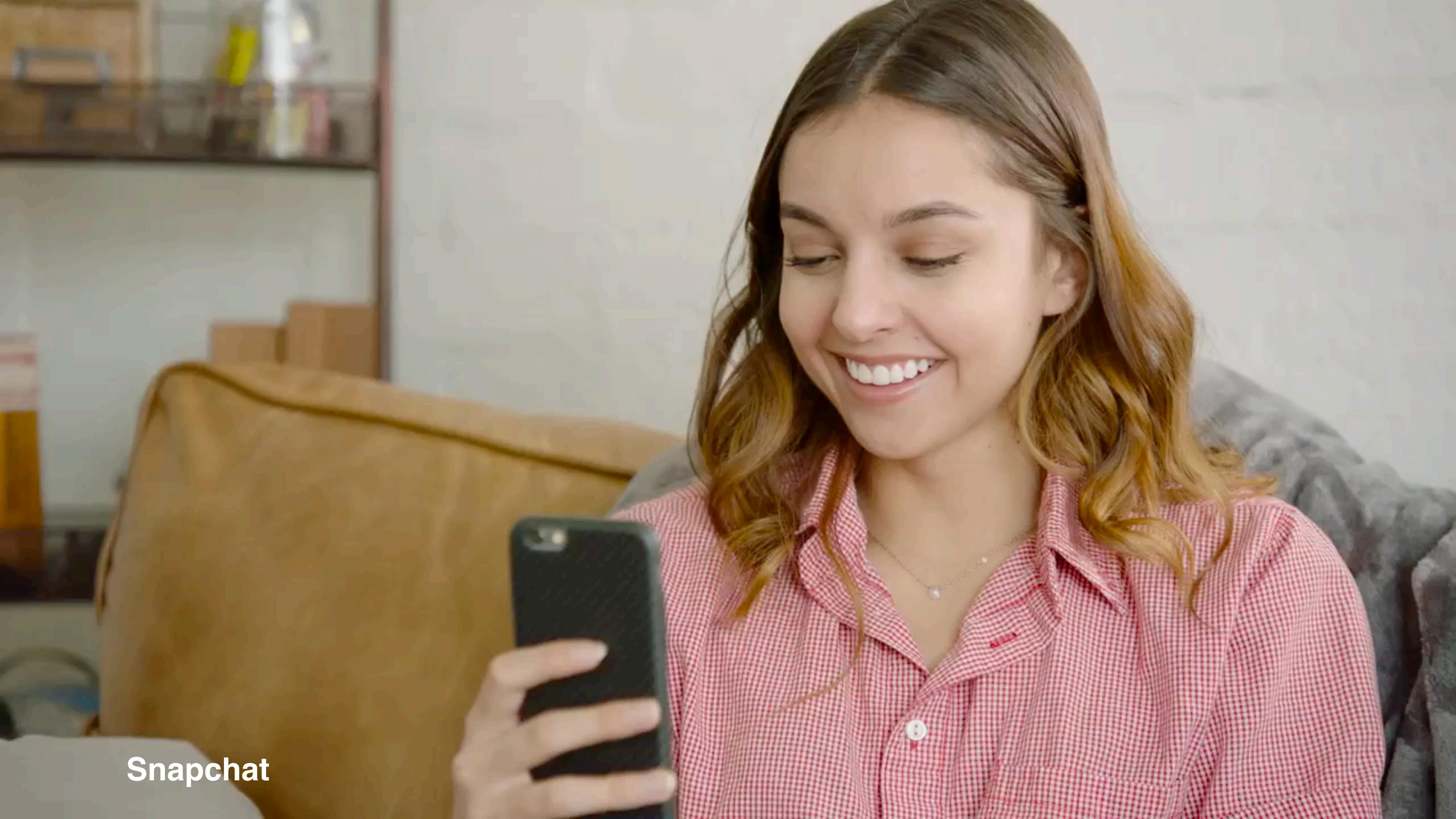
FaceX Kids



input video



face segmentation



Snapchat



Facebook / MSQRD

live demo

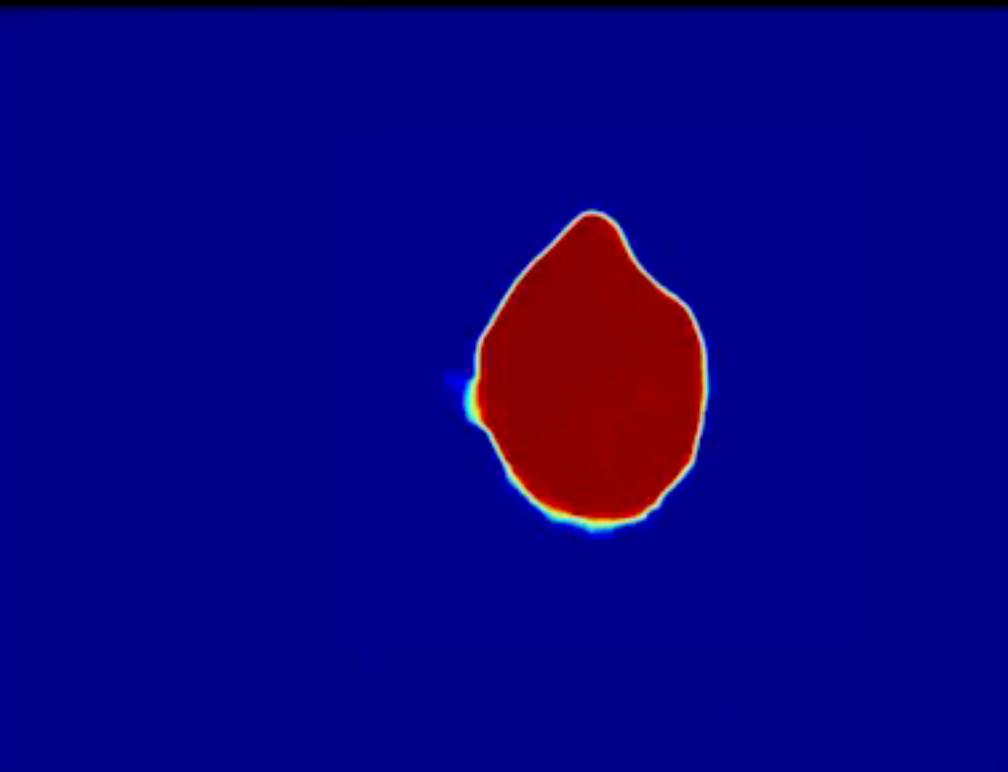




input video



facial segmentation/tracking



probability map



composed result



face data



occlusion / cropping



hand over face compositing



negative hand data

Preliminary Findings: Segmentation

submitted to SIGGRAPH 2016

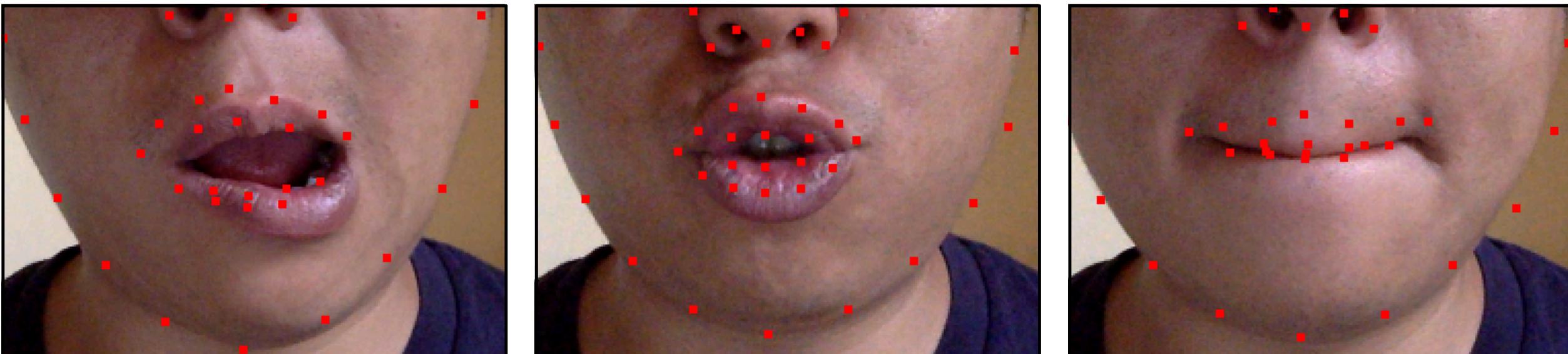


Occlusions

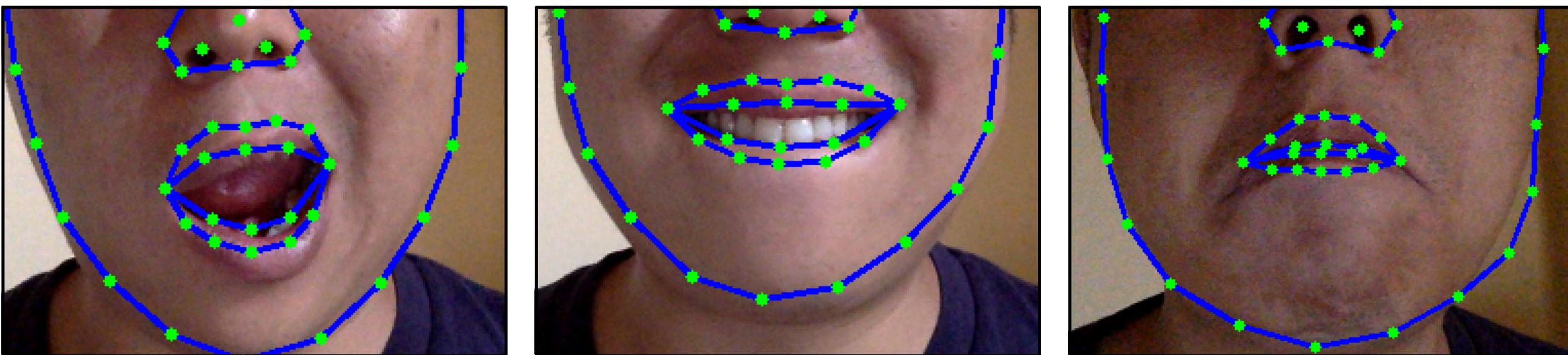




Feature-Based Tracking

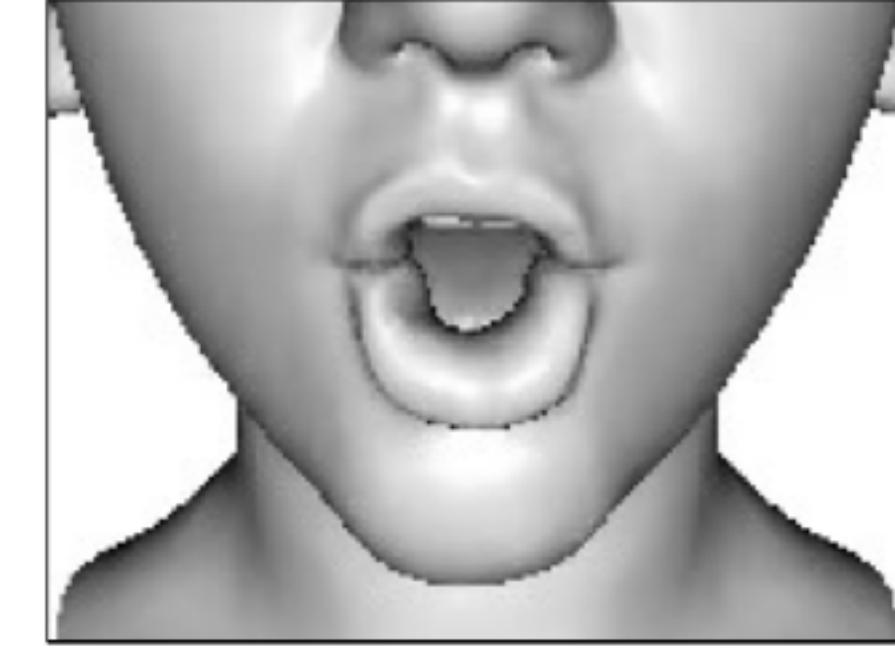
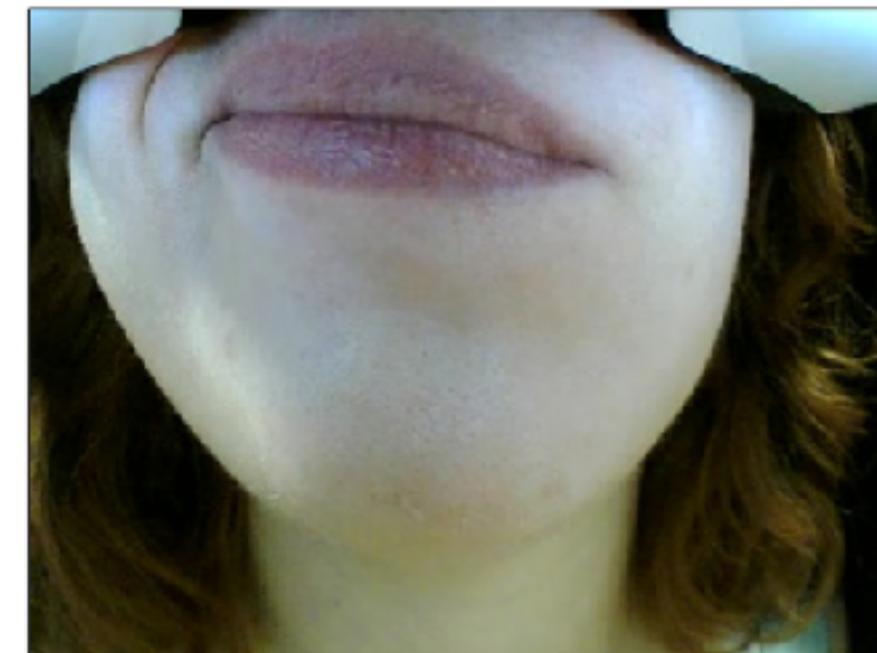
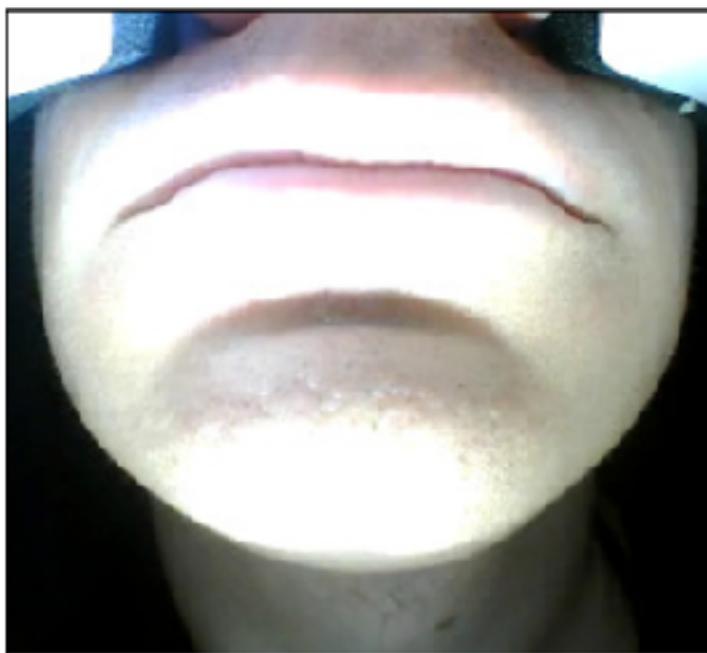
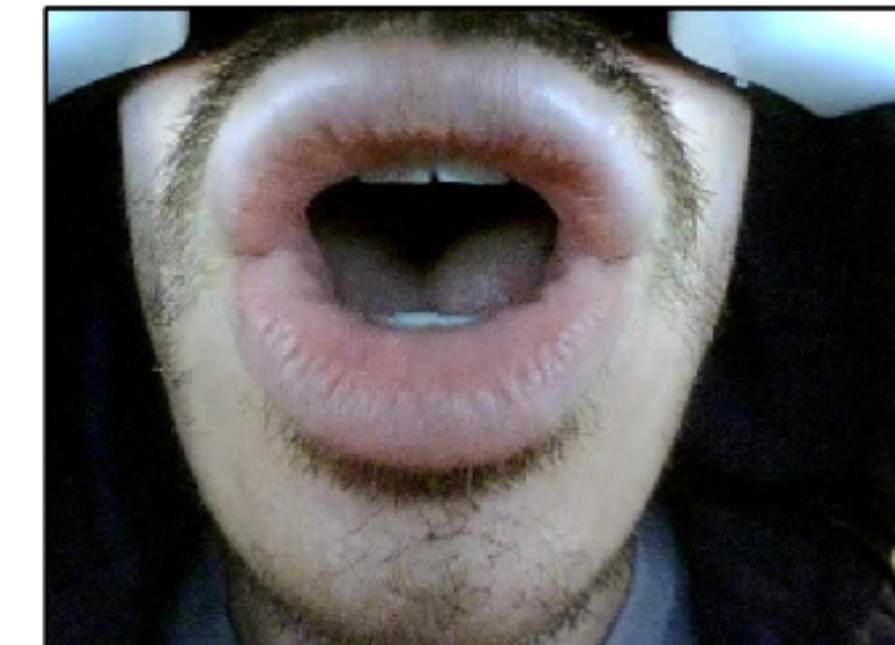
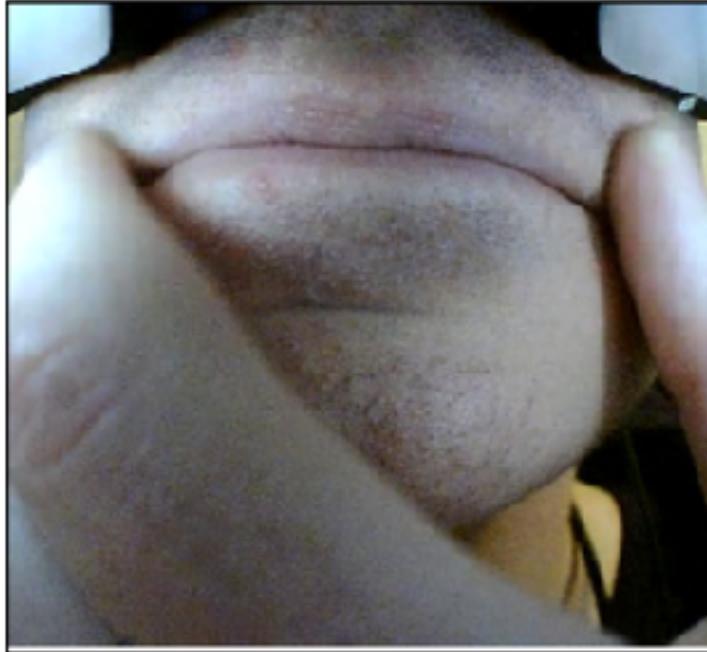


Kazemi and Sullivan, CVPR 2014



Cao et al., SIGGRAPH 2014

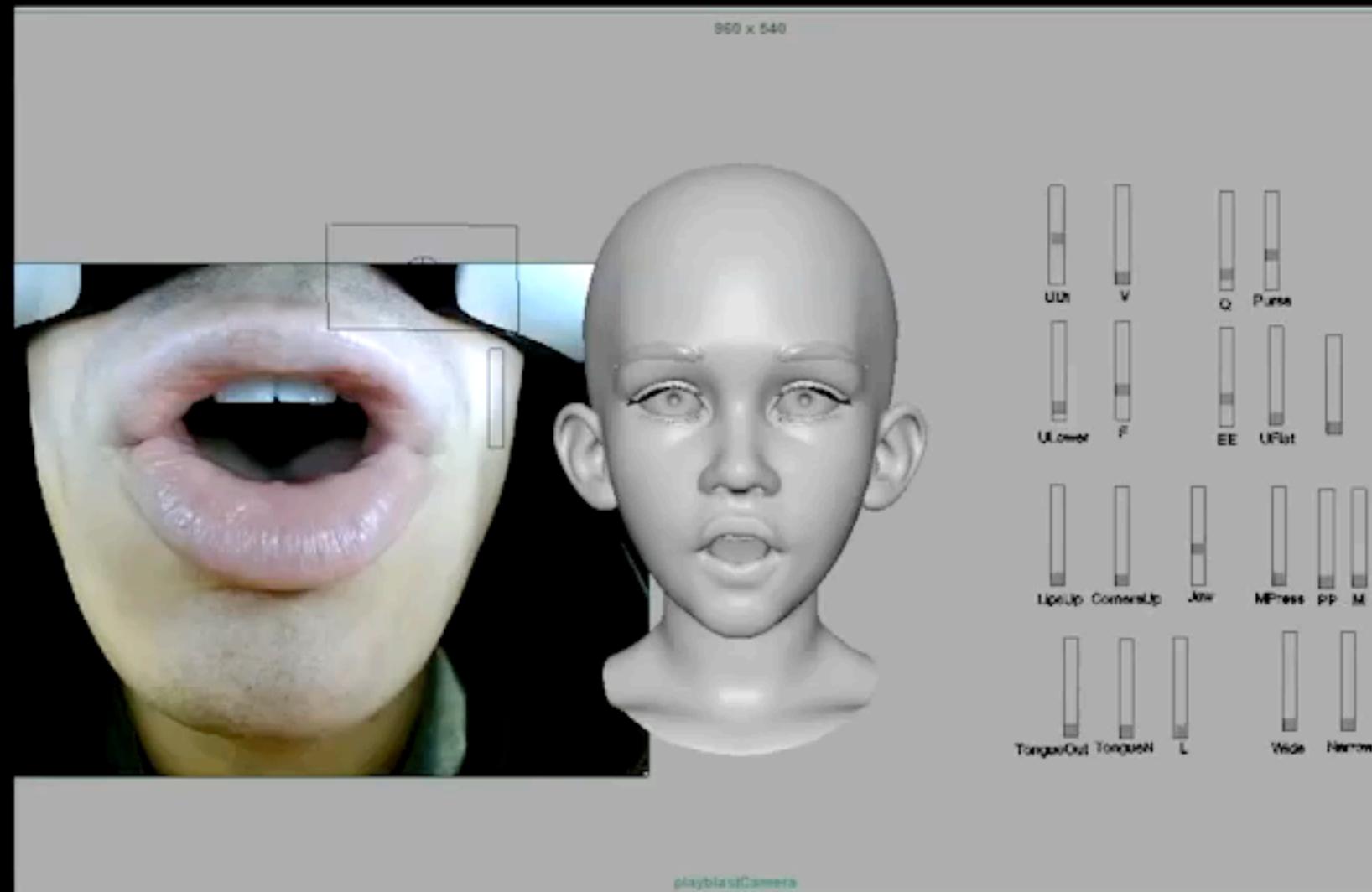
High Dimensionality & Non-Linearity



occlusions, lighting, expressions, identity

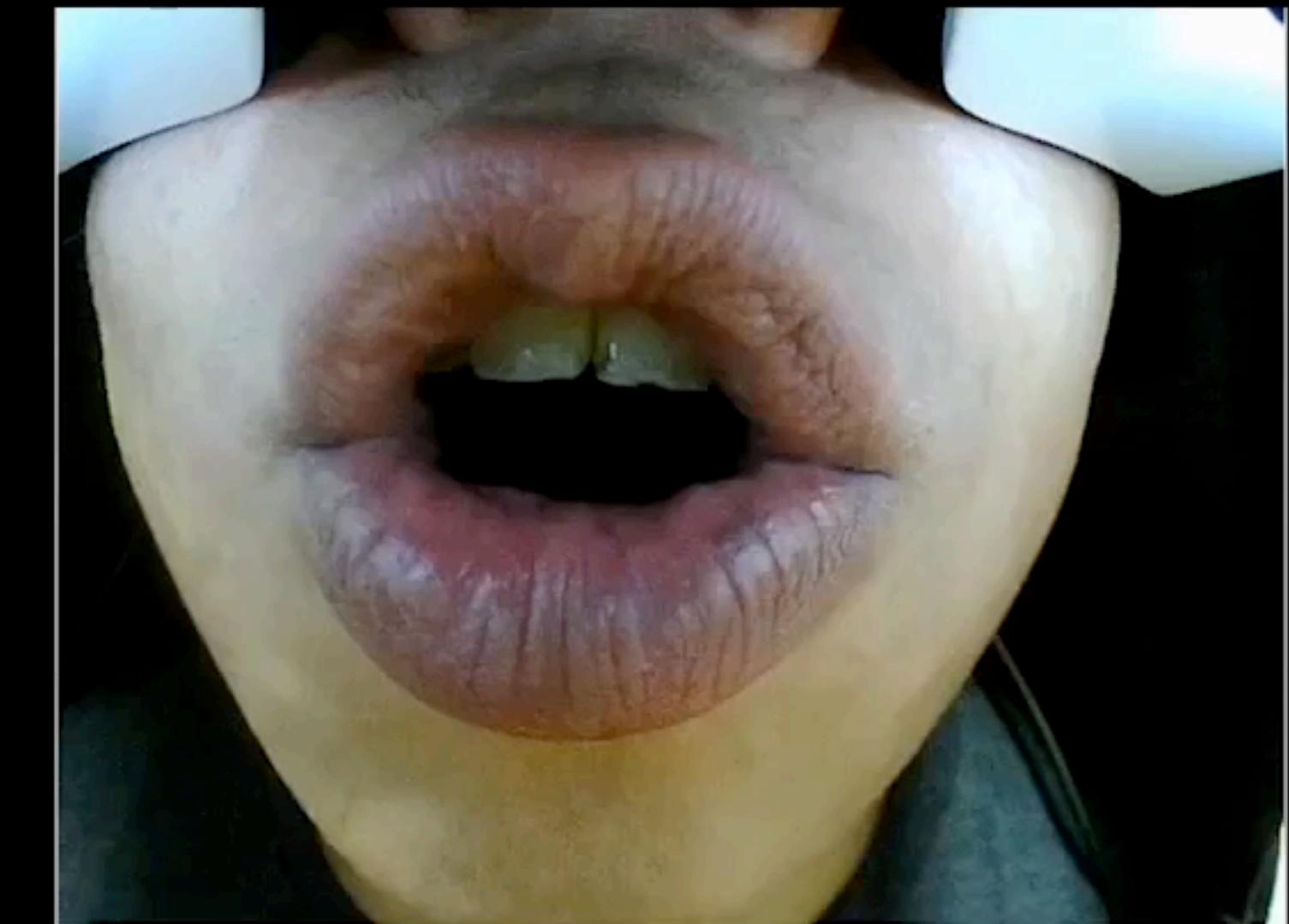
sticky lips, biting, visemes, ...

Label Transfer via Audio Alignment

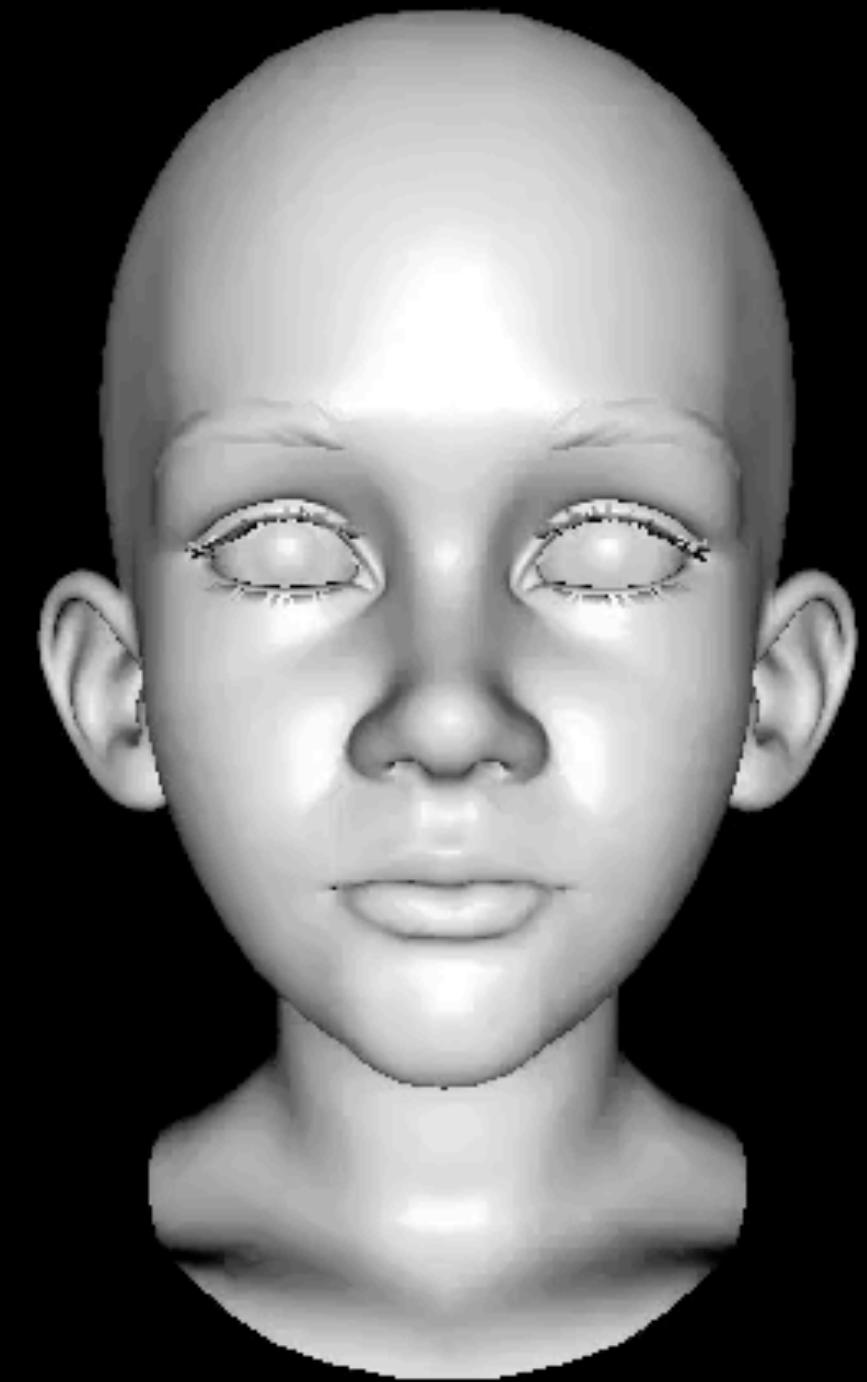


reference
data

reference
animation



dynamic time warped
training data



Correspondences

with Convolutional Neural Network

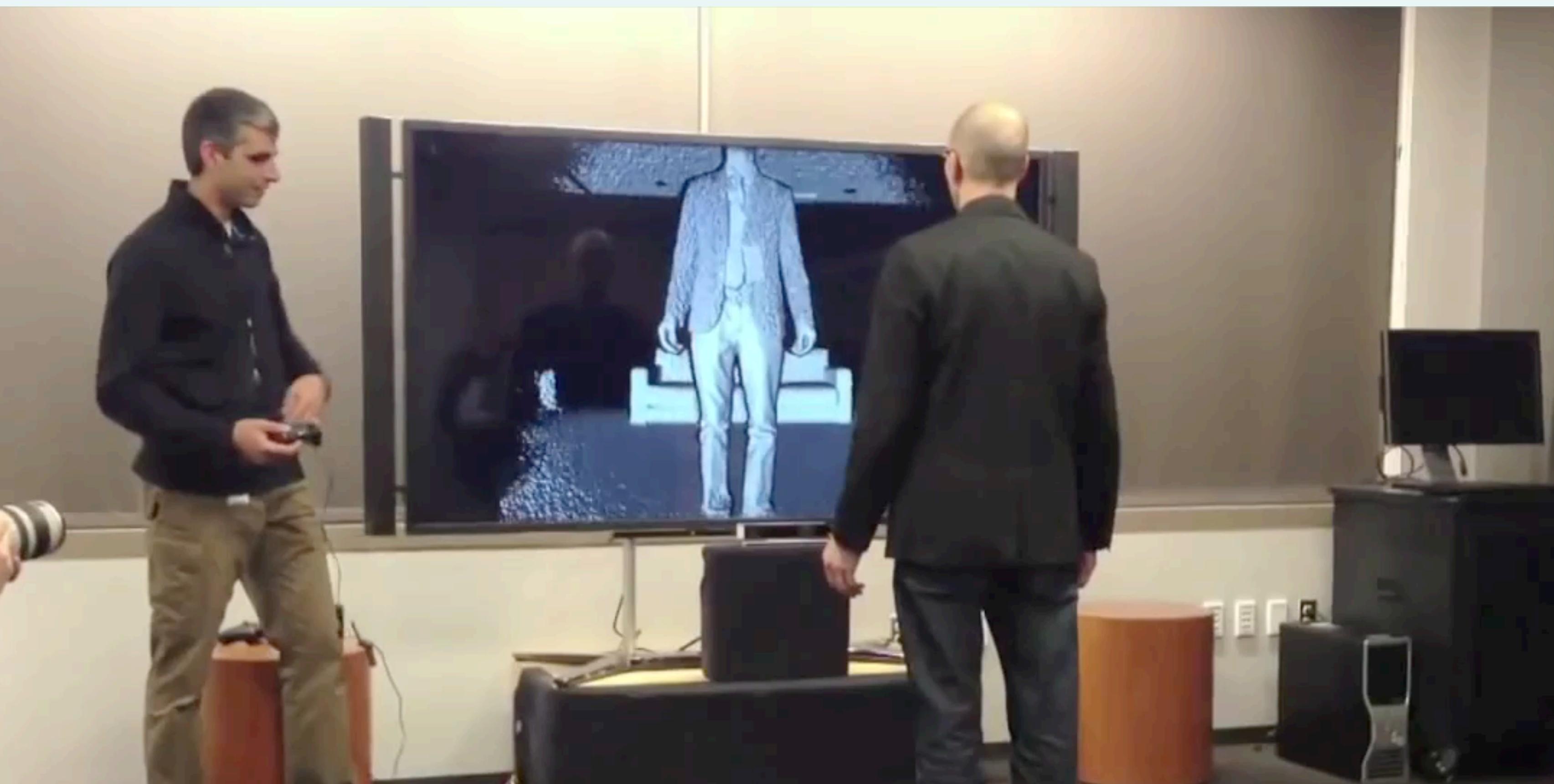
Credits

Lingyu Wei Qixing Huang Duygu Ceylan Etienne Vouga
Ruizhe Wang Gerard Medioni



3D Human Capture

Microsoft 2013



3D Human Capture

[Dou et al. '16]

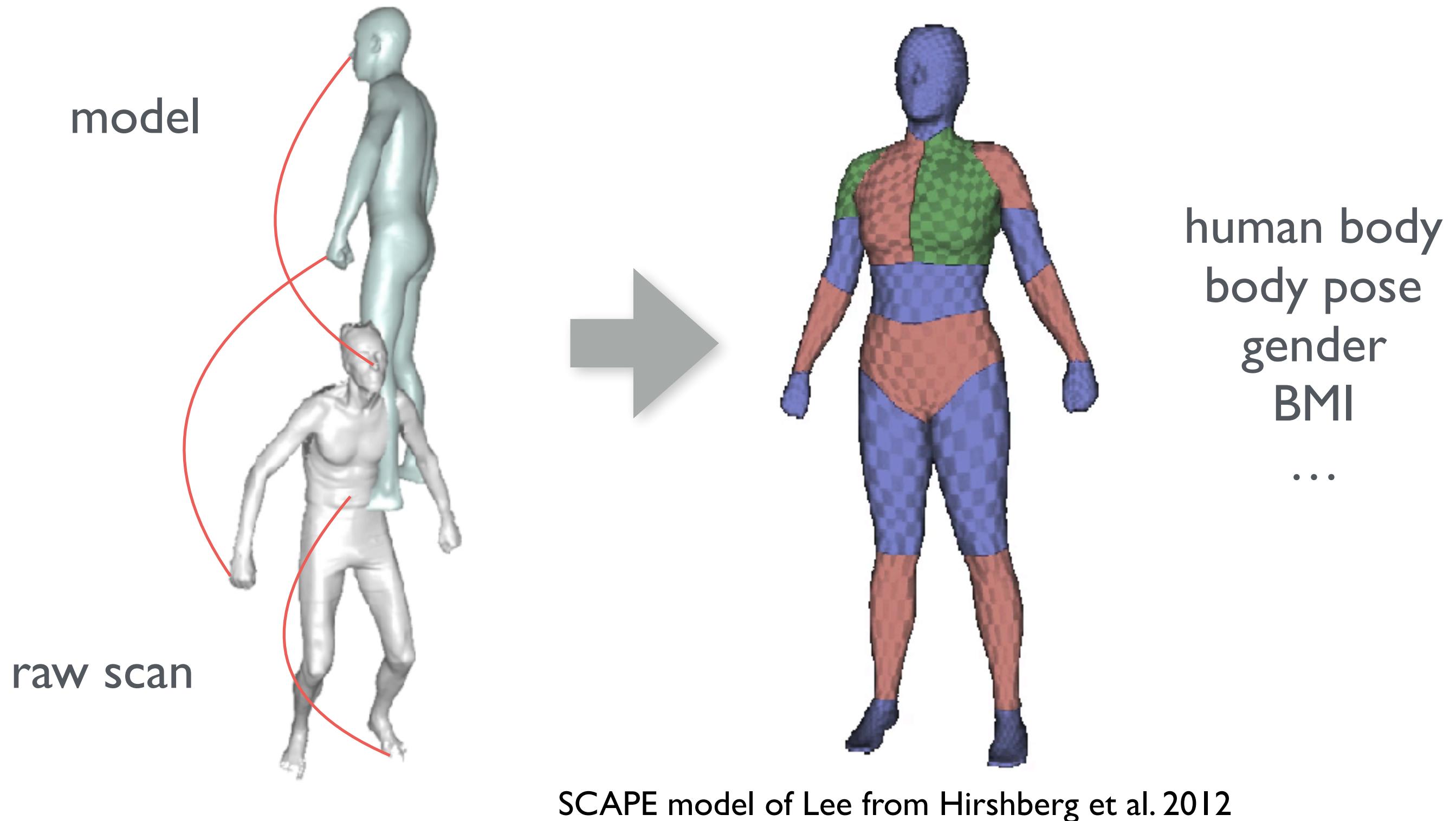


Analysis & Reasoning

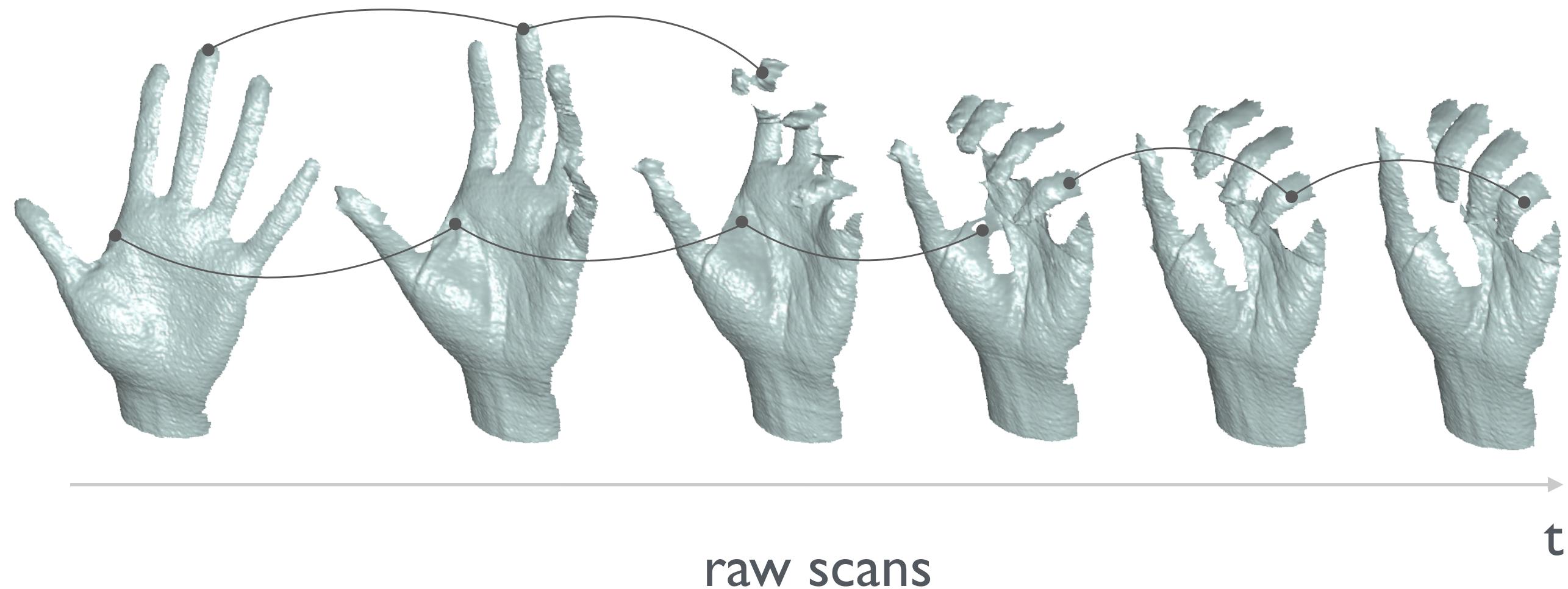
Correspondences on Clothed Human Bodies



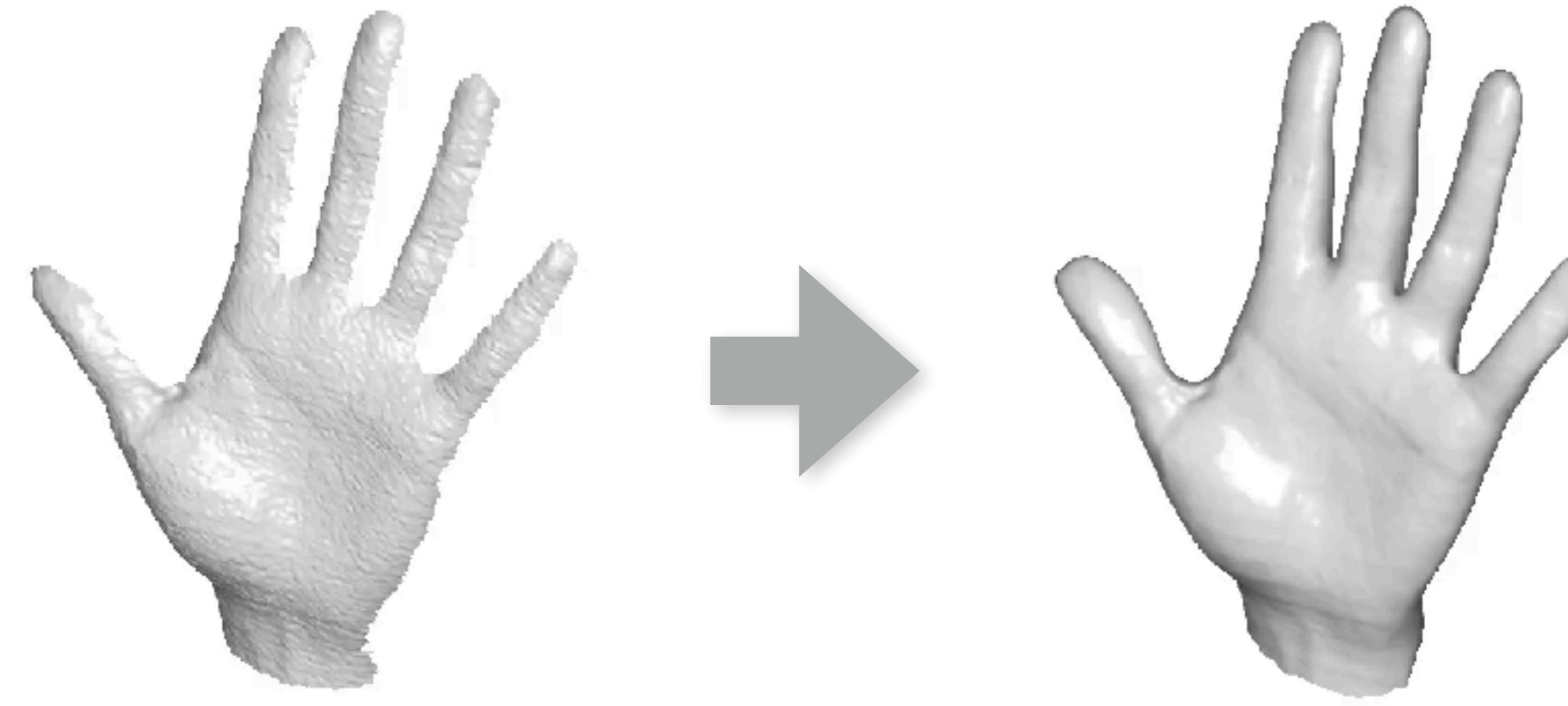
Shape Analysis



Motion Understanding



Motion Understanding

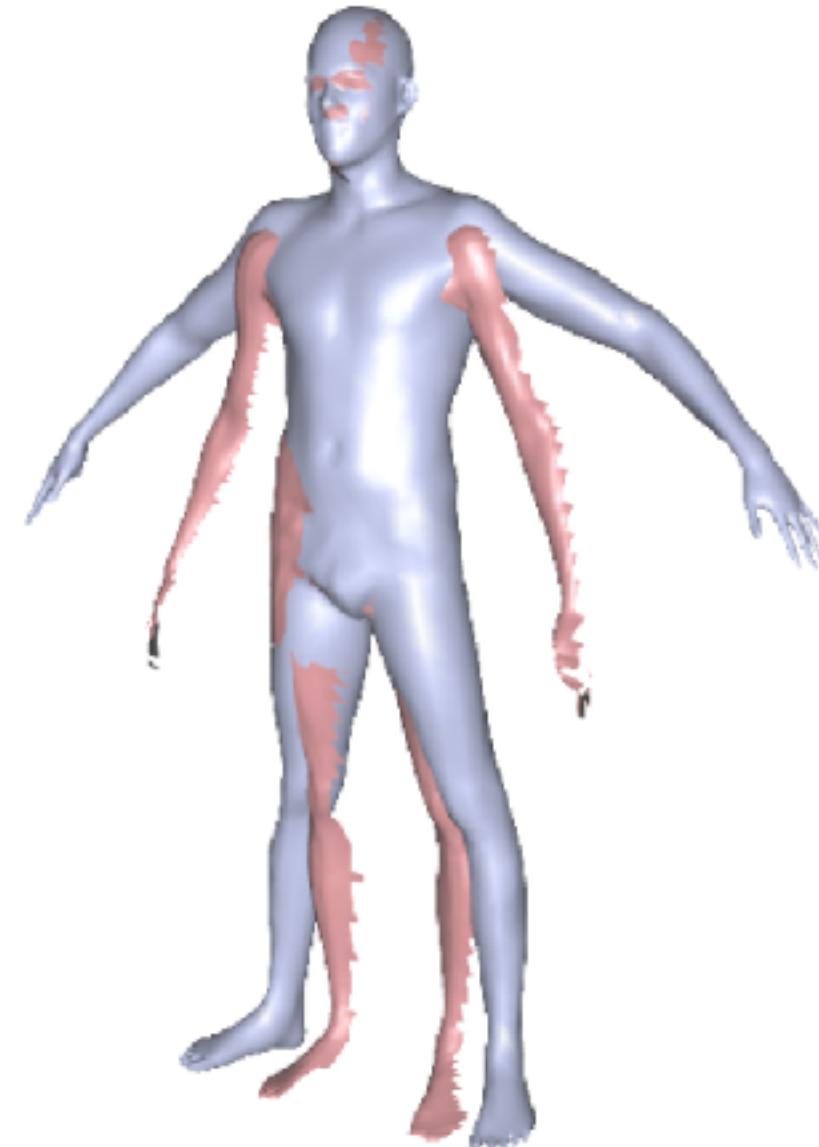


raw scans

“grasping”

Correspondences?

Large Pose Changes



source & target

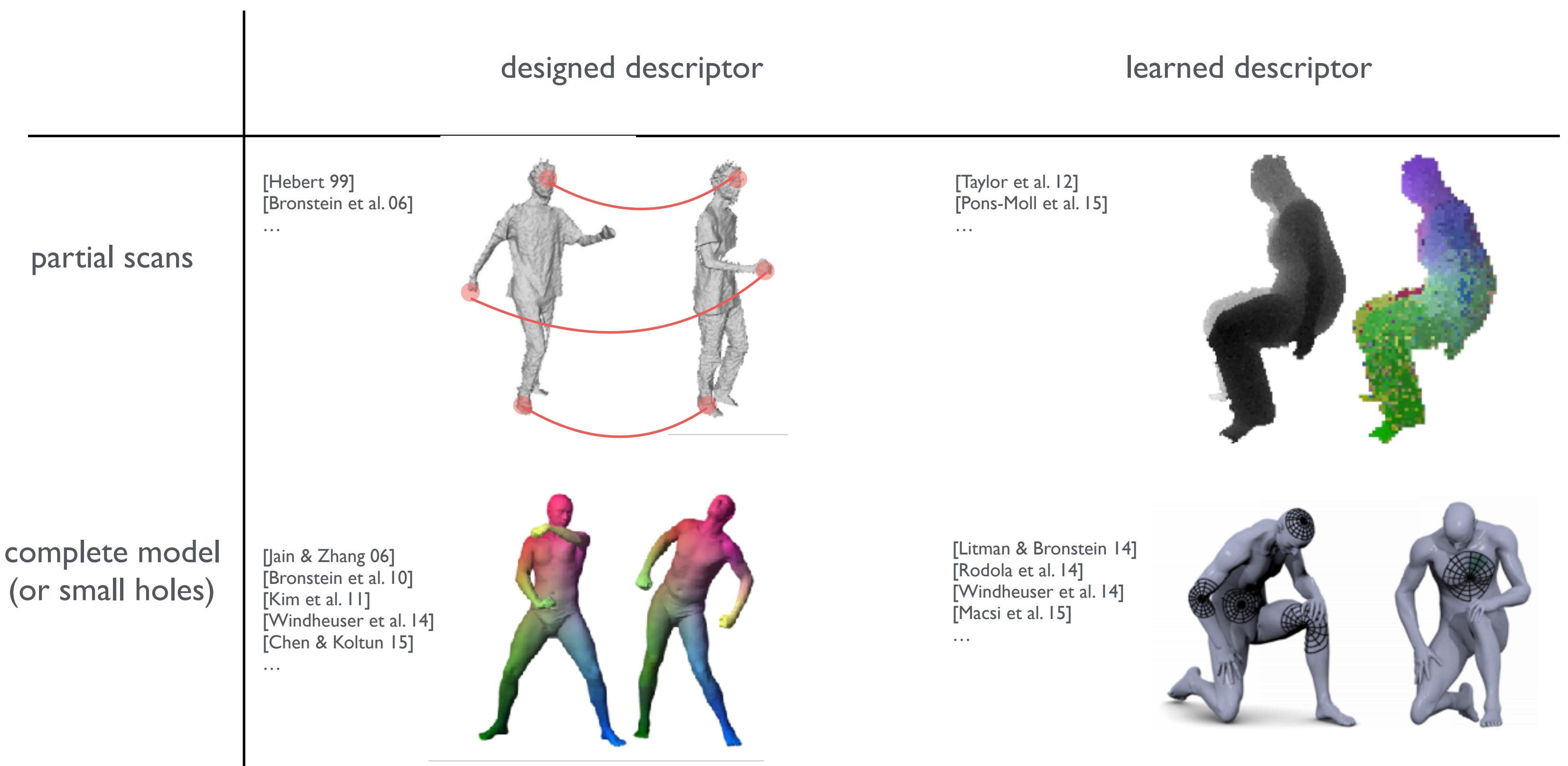


[Li et al. 09]

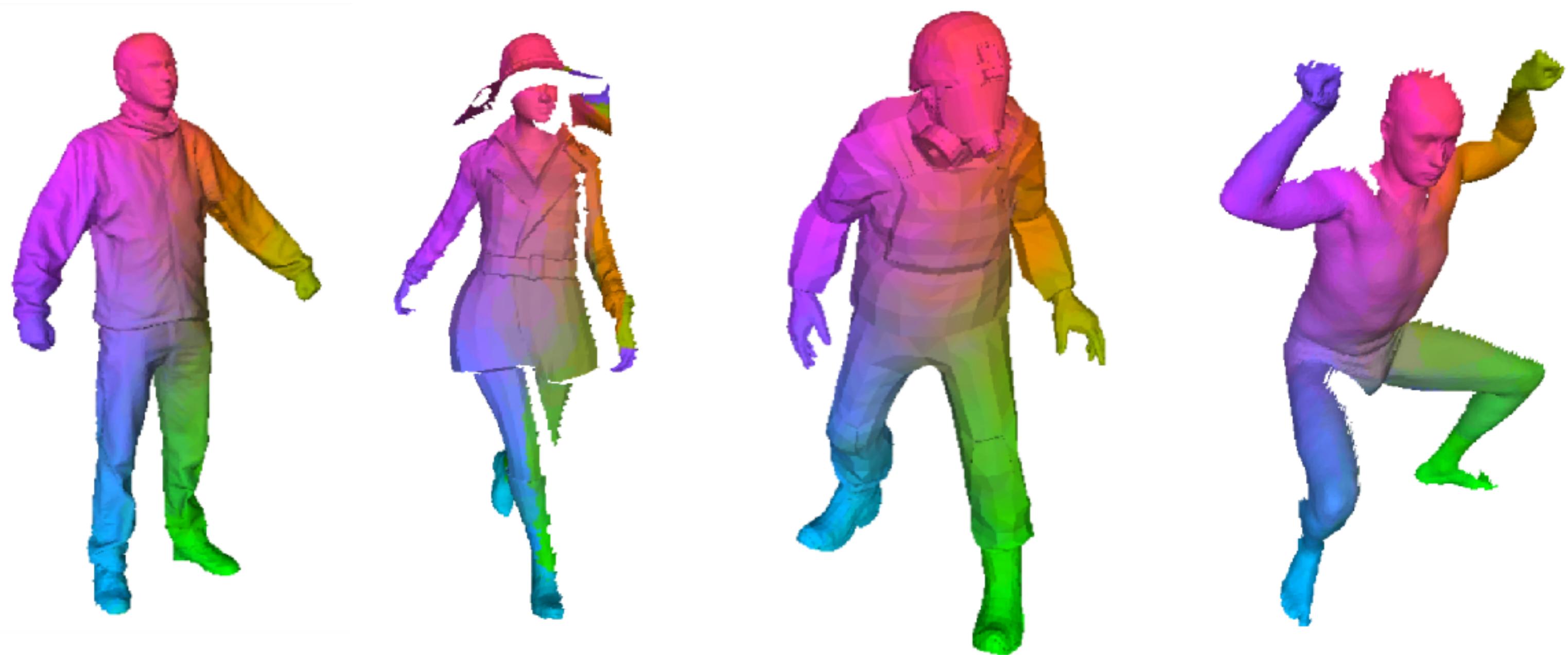


[Huang et al. 08]

Descriptors



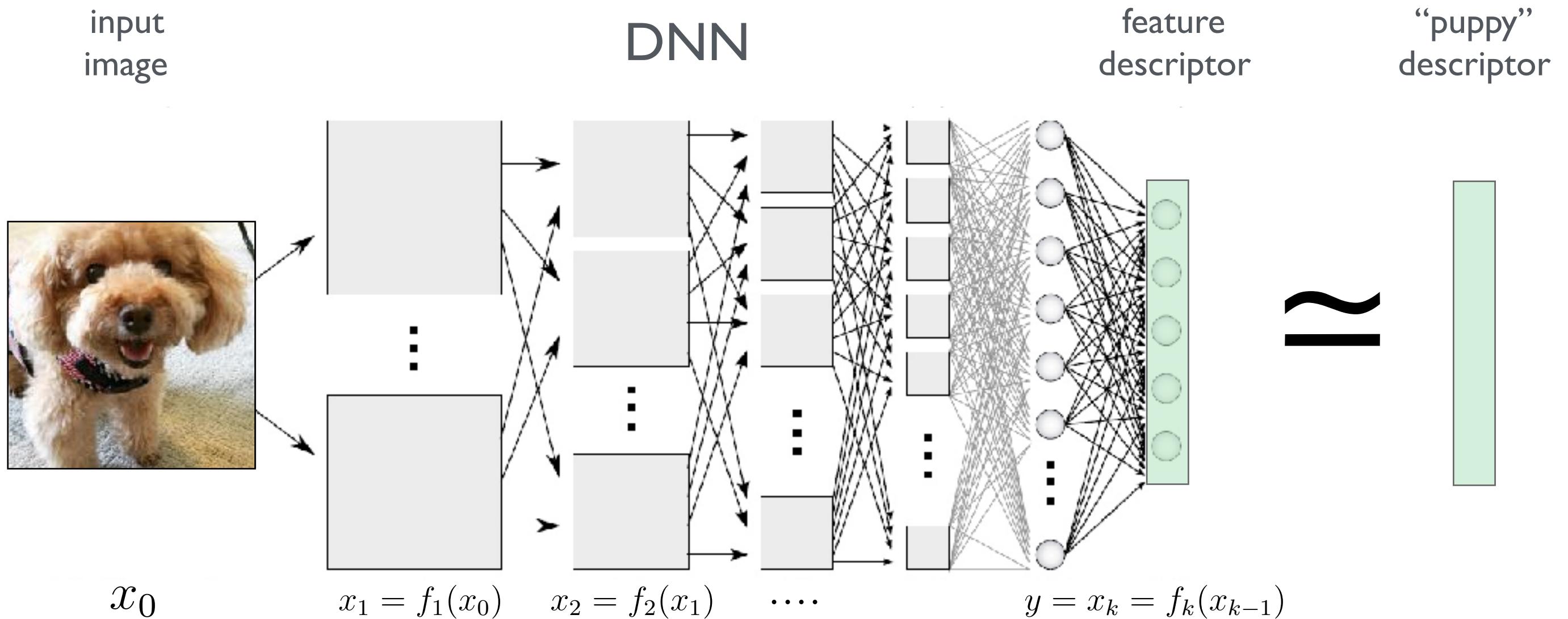
Clothed and Partial Data



immense space of variations

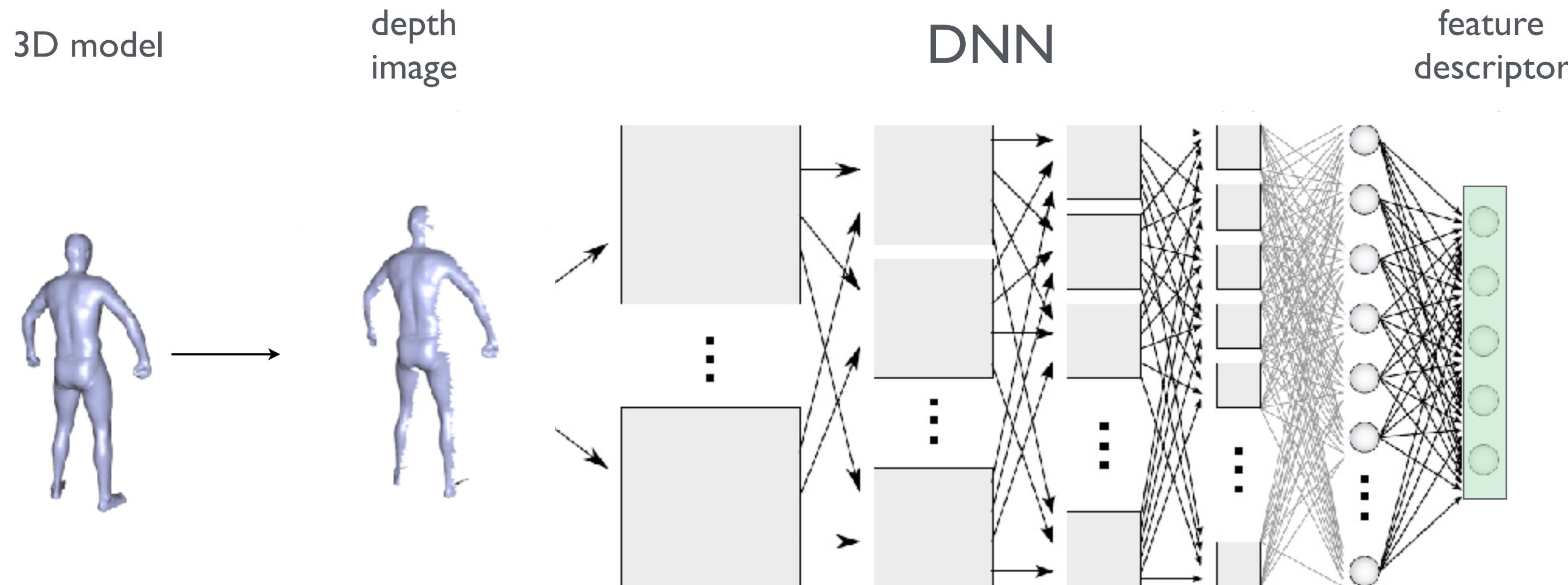
Classification Networks

Deep Convolutional Neural Network

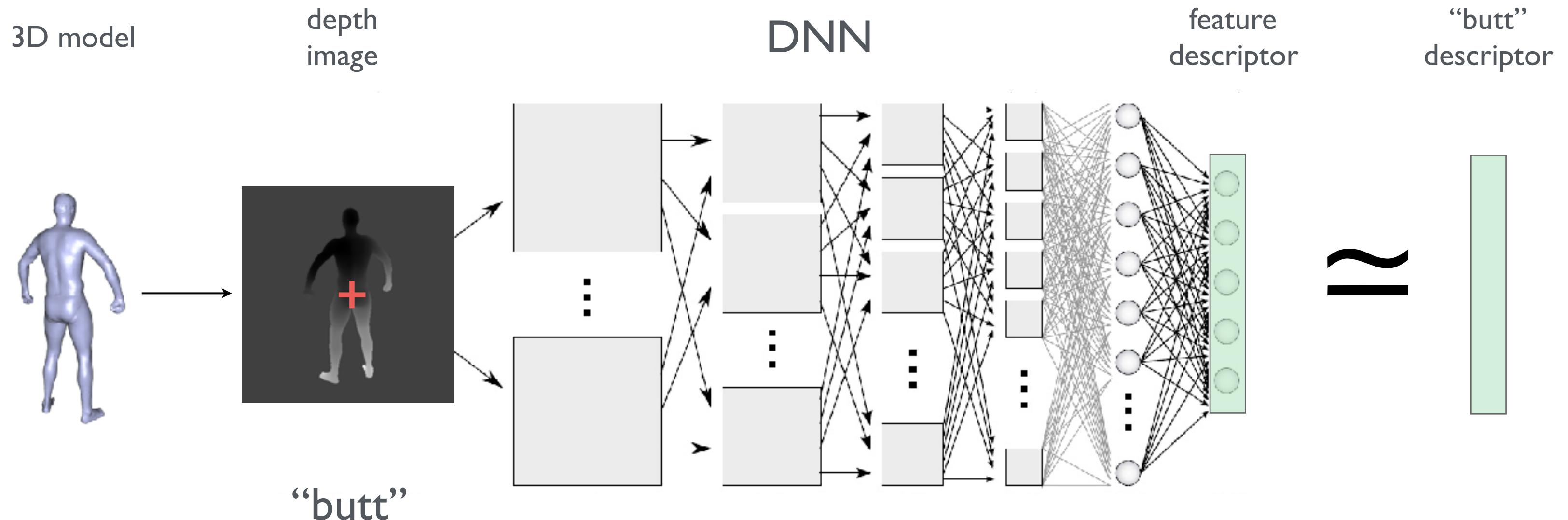


classification network, e.g. AlexNet [Krizhevsky et al. 2012]

Deep Convolutional Neural Network



Deep Convolutional Neural Network

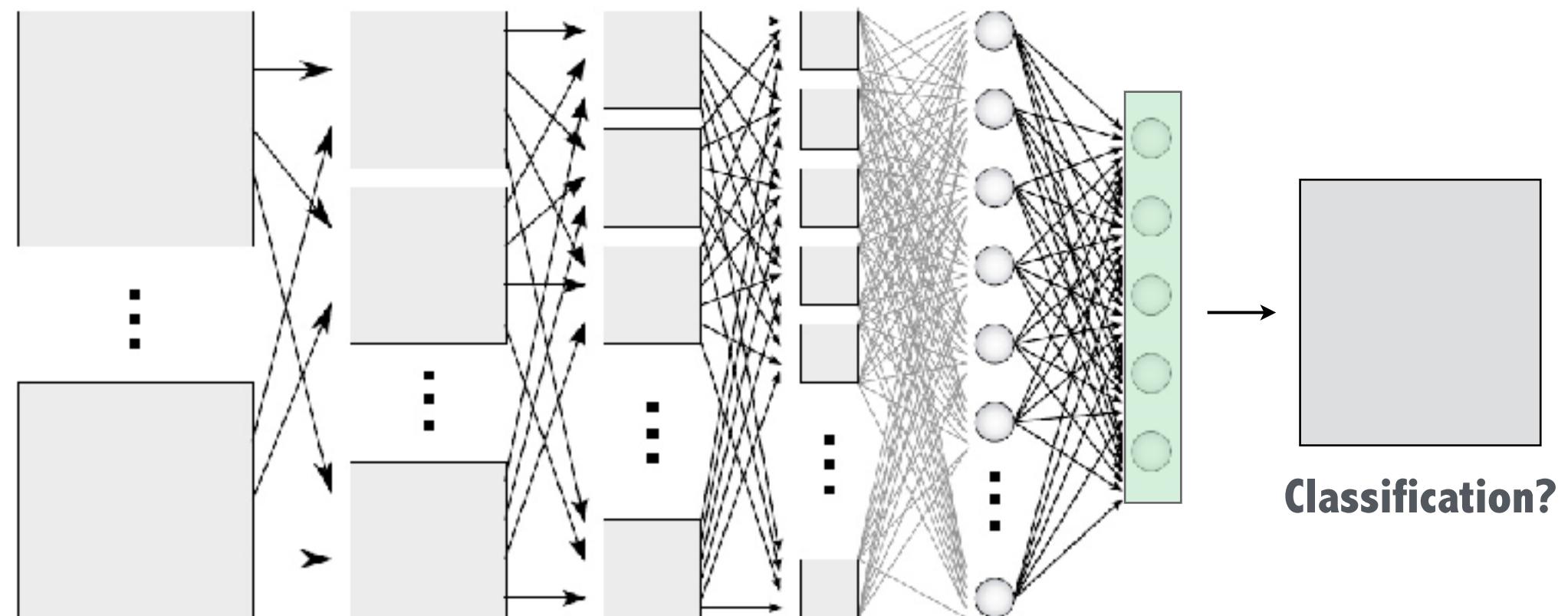


Loss Function

Training Data



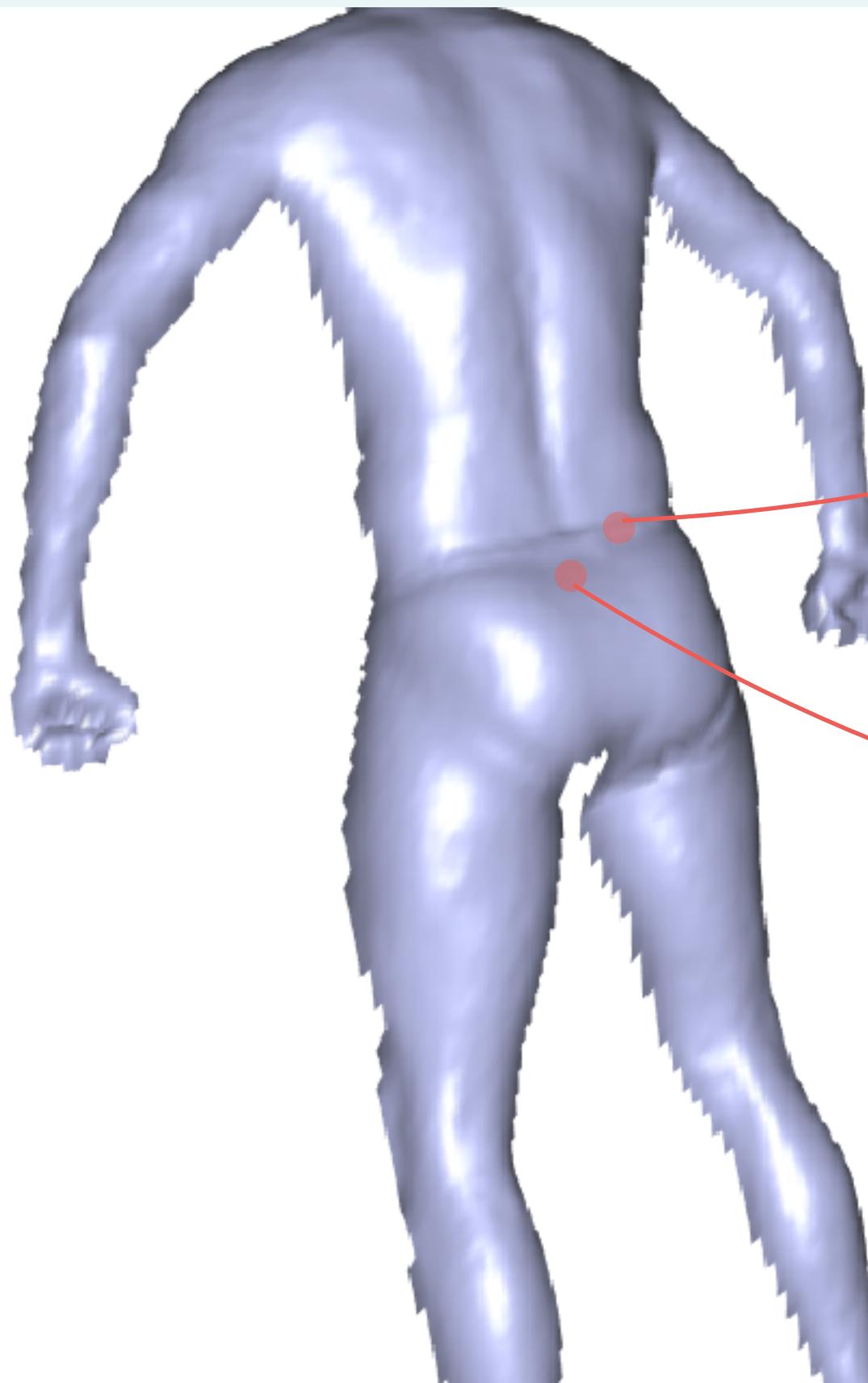
DNN



Loss Function

Classification?

Classification Task



descriptors are
far apart

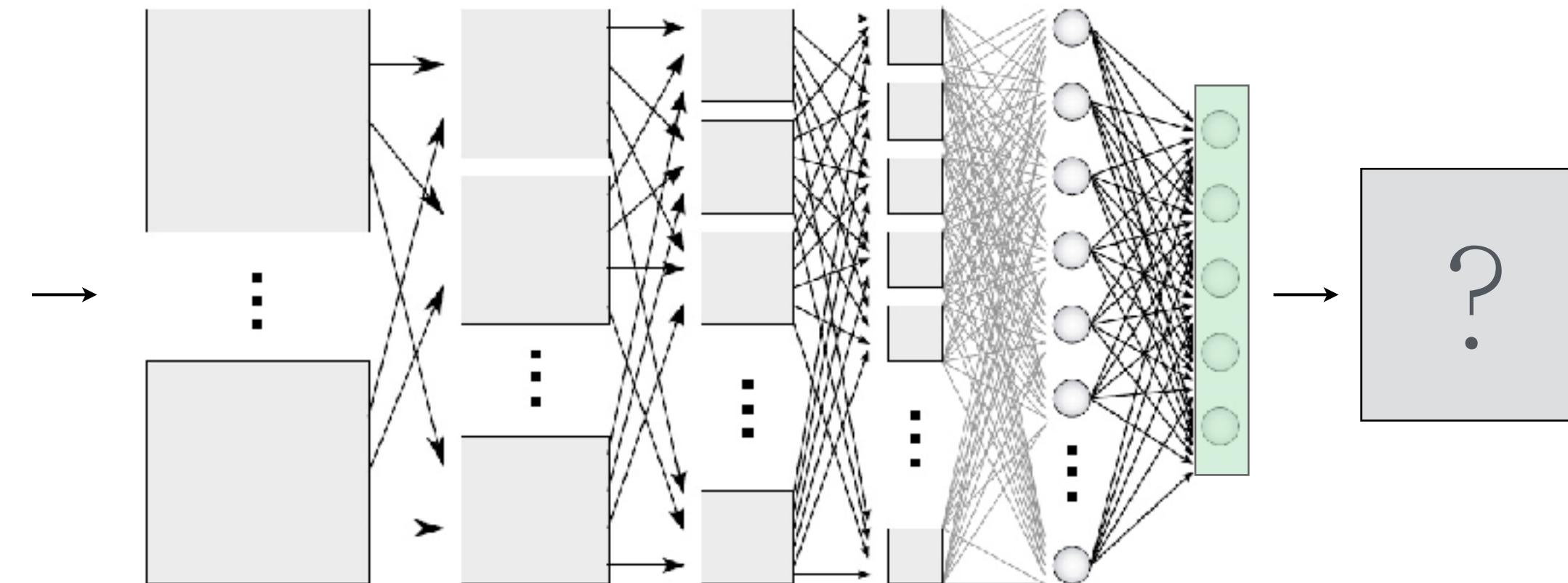
How to preserve distances?

Deep Convolutional Neural Network

Training Data



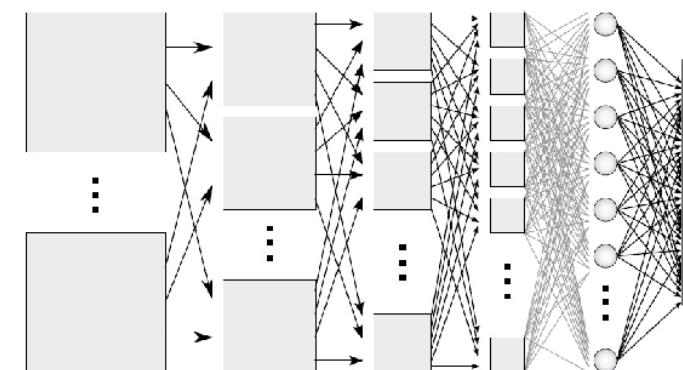
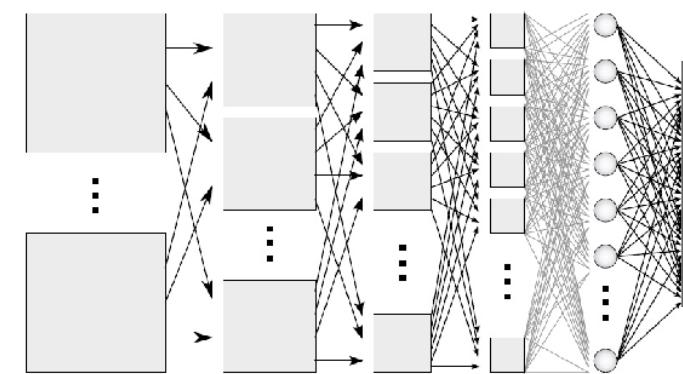
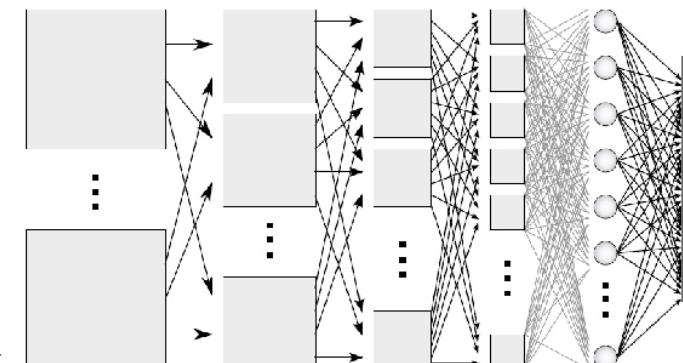
DNN



Loss Function

Loss Function

Training Data

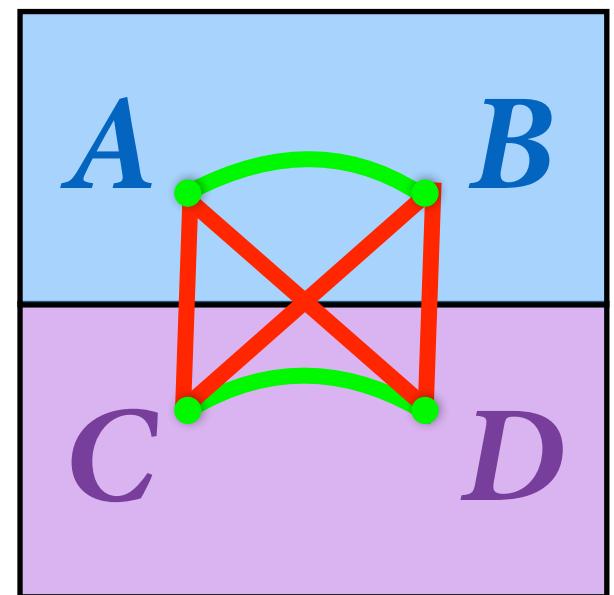
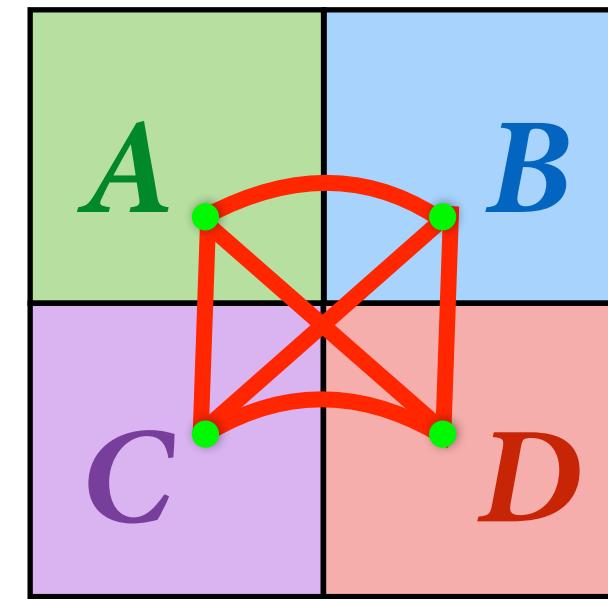
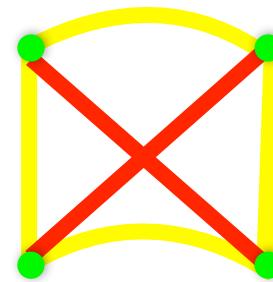
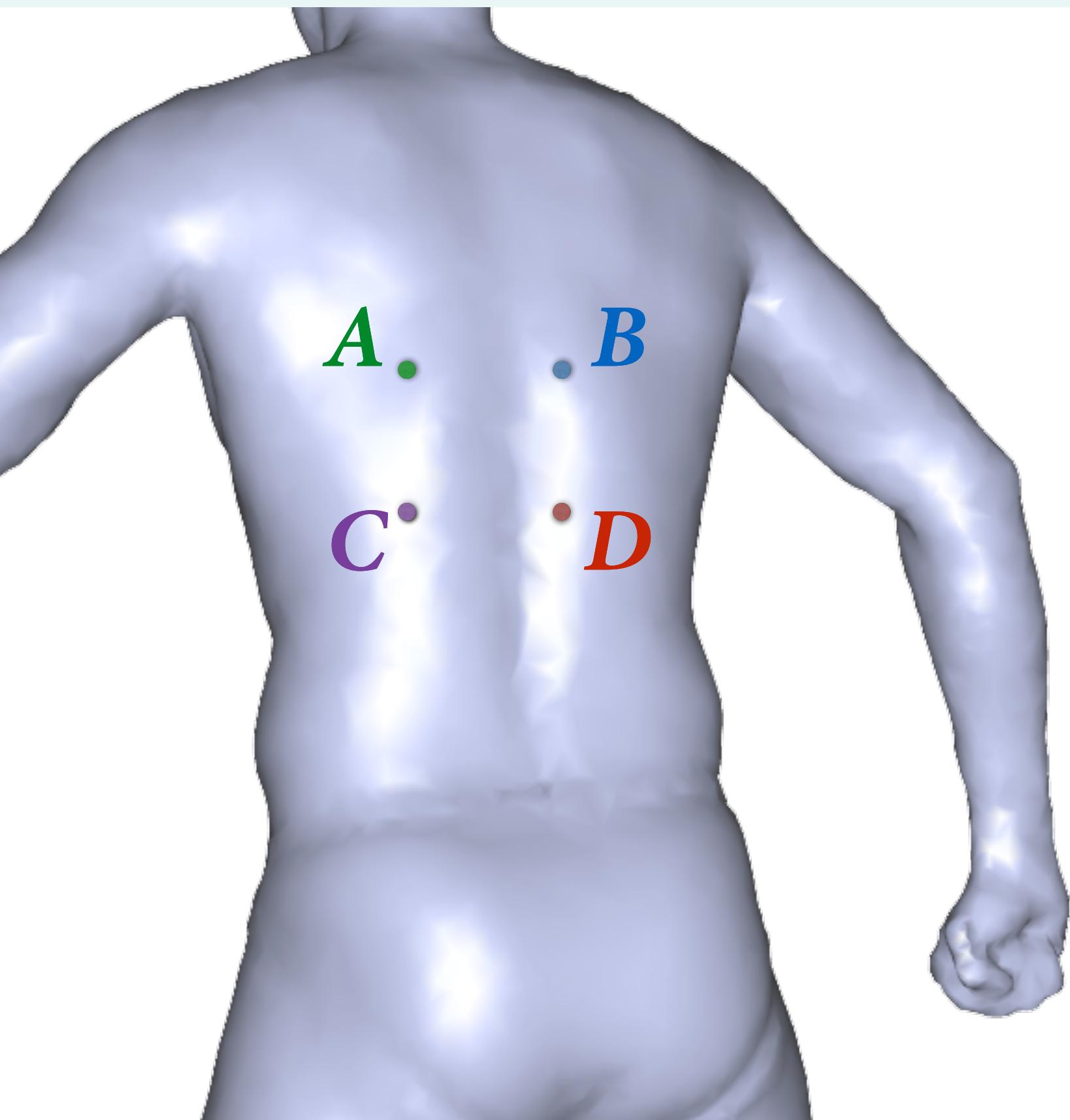


Loss Function

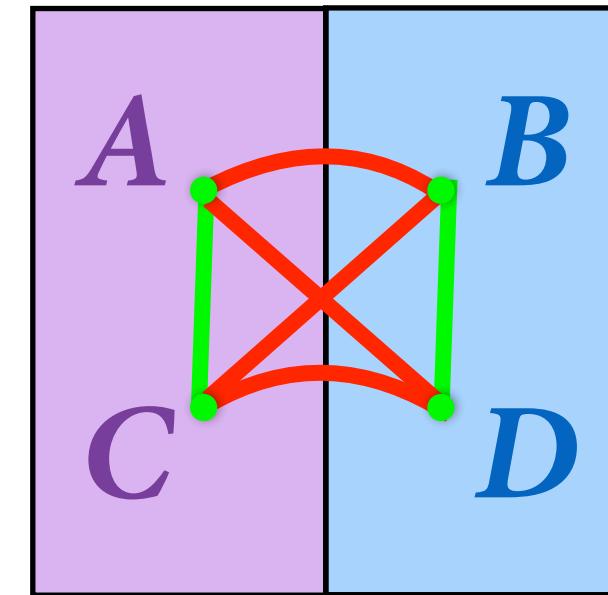
Triplet Loss

(Anchor,Positive,Negative)

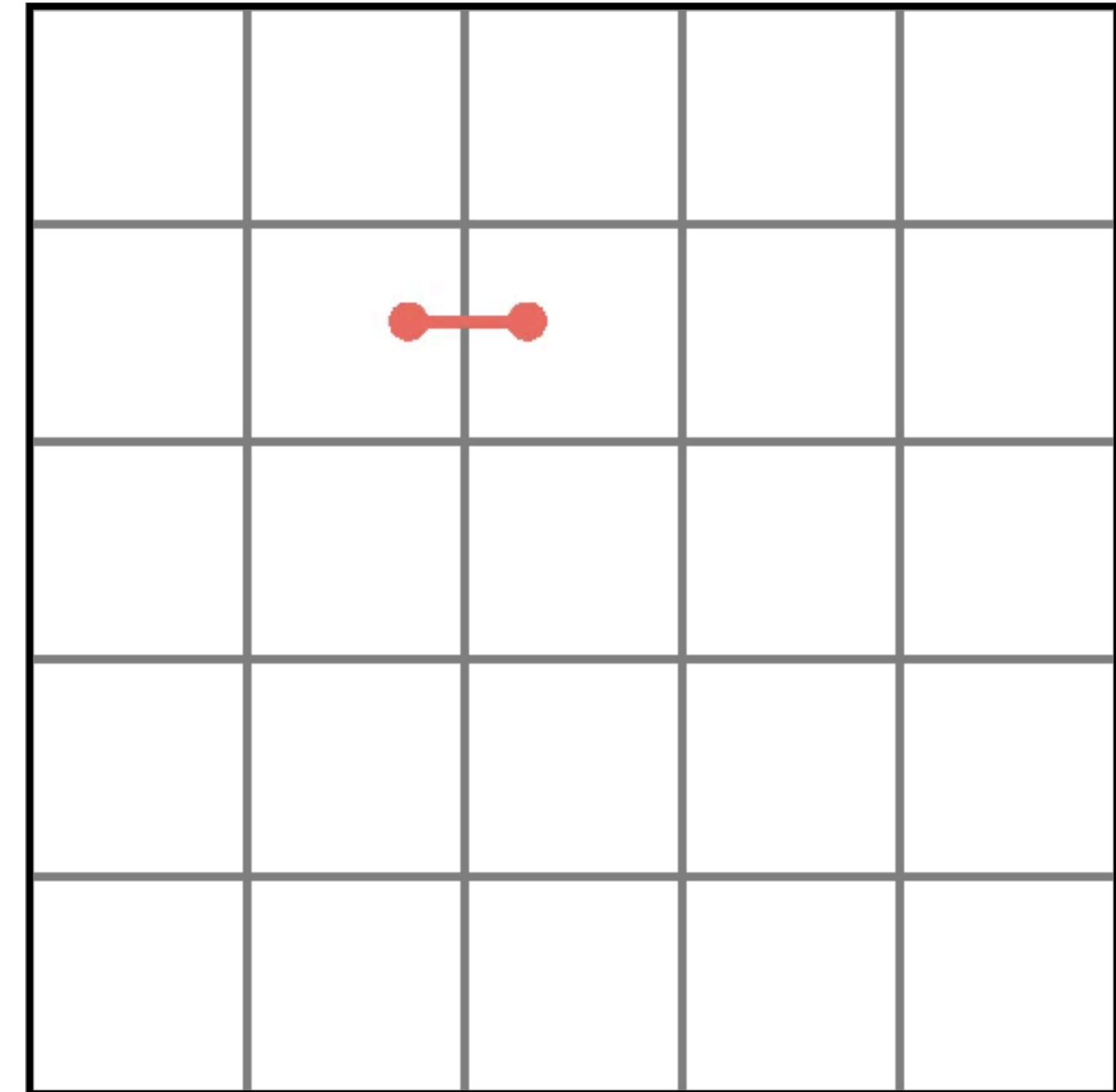
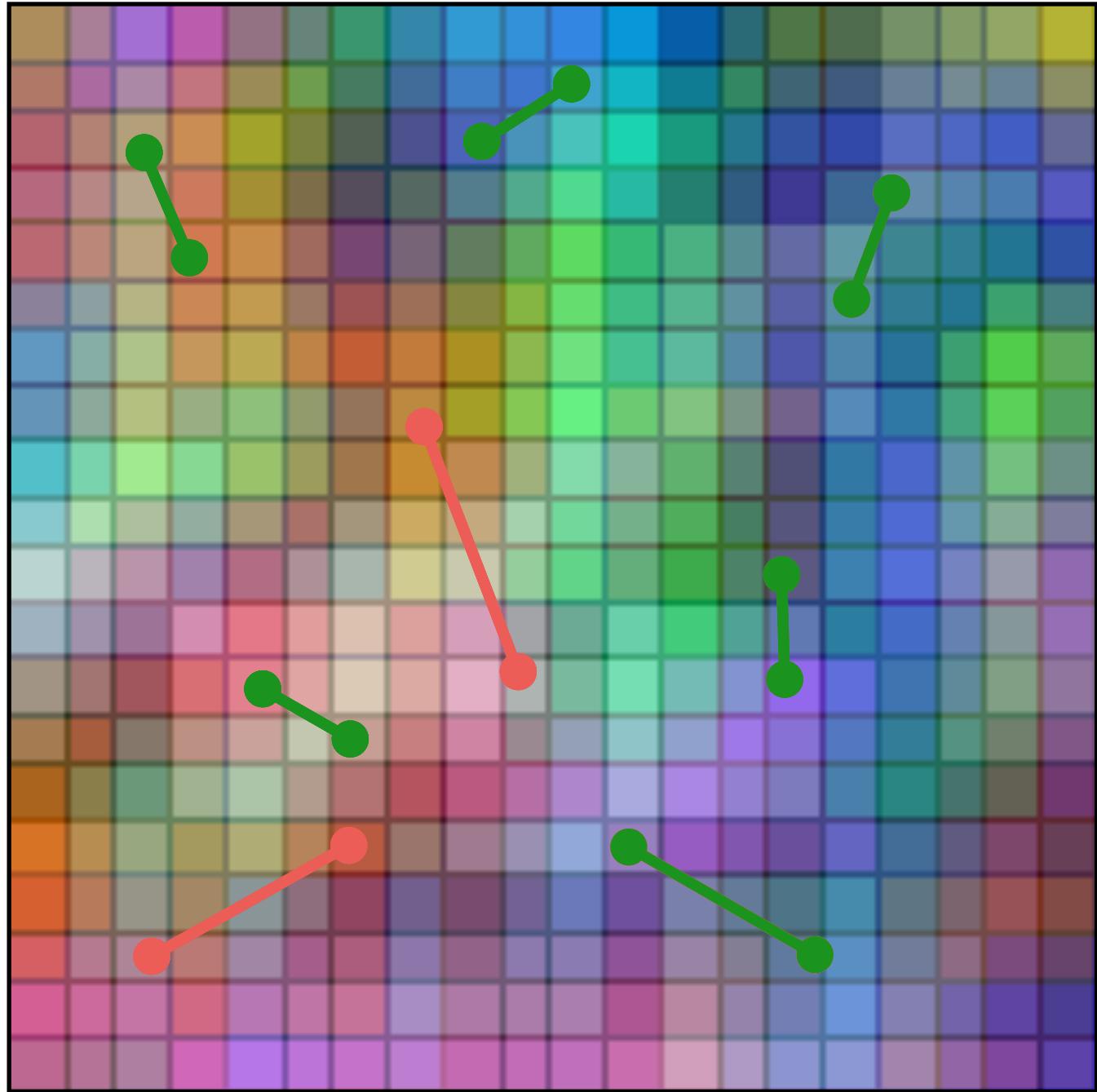
Multi-Segmentation



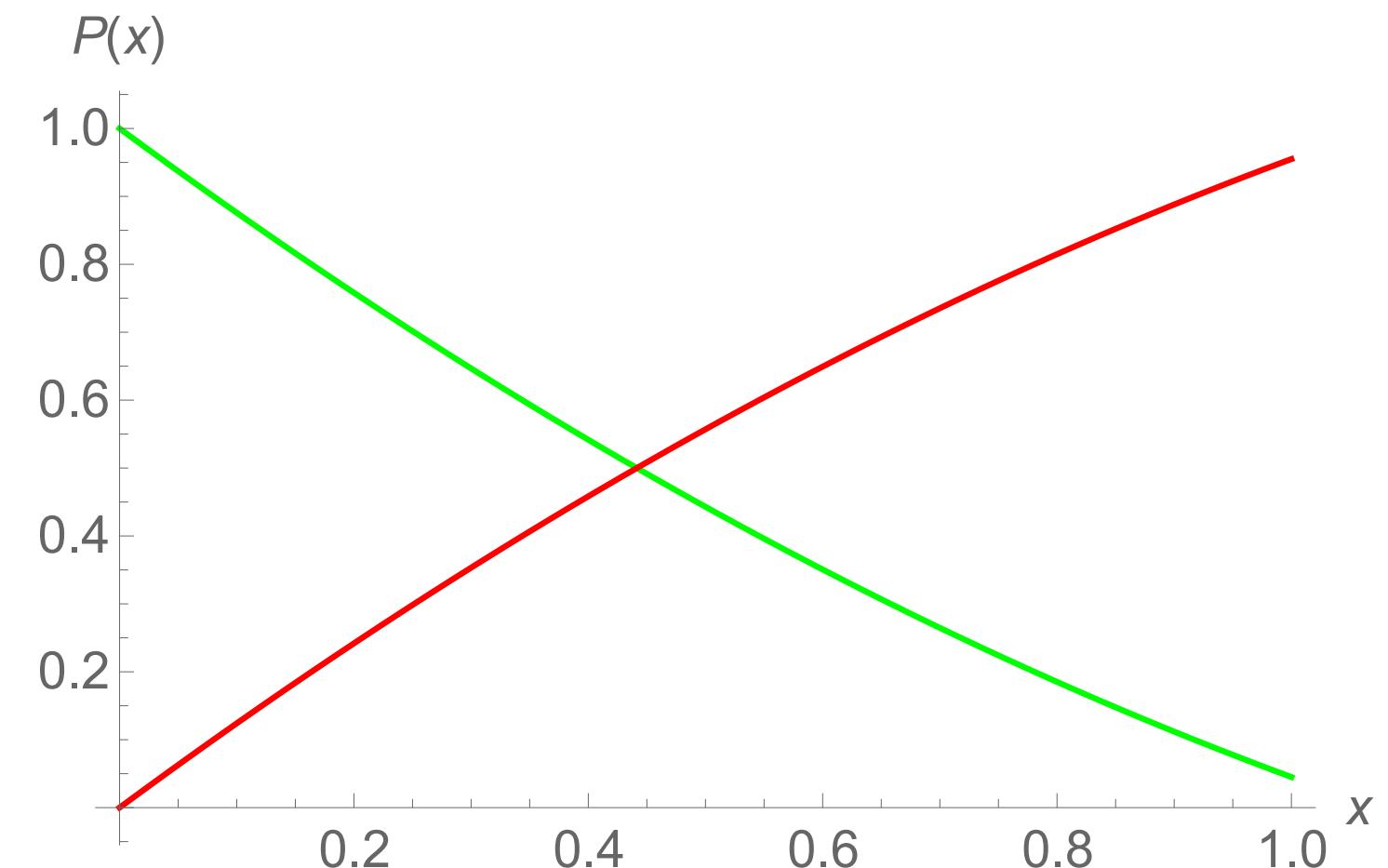
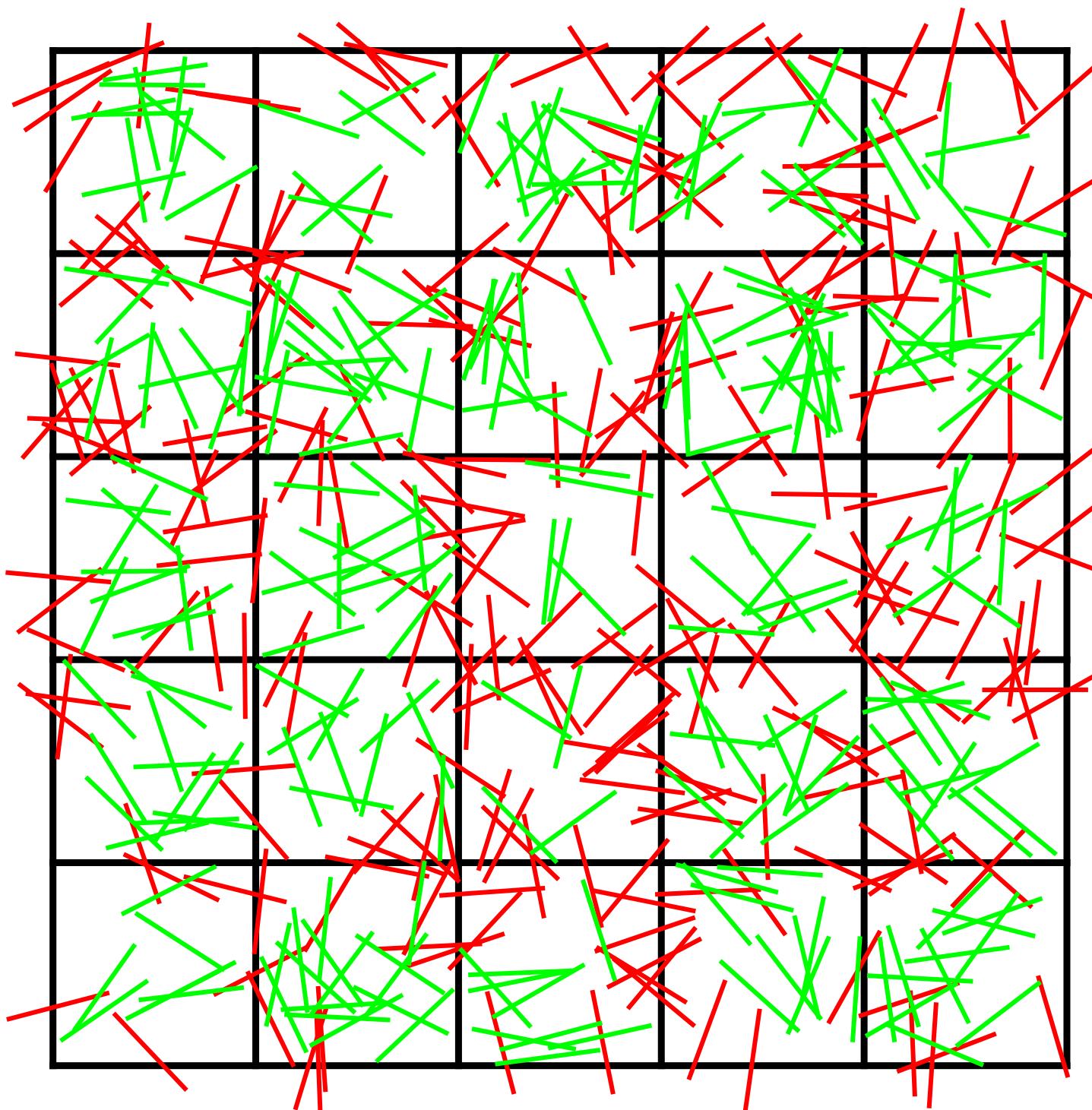
+



Multiple Segmentation

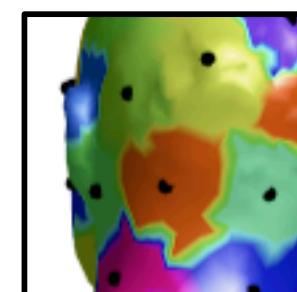
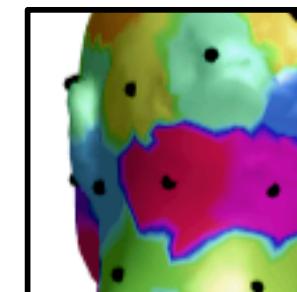
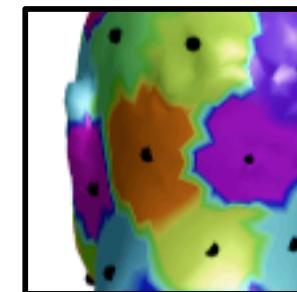


Buffon-Laplace Needle Problem (18th Century)



Distance Preserving Learning

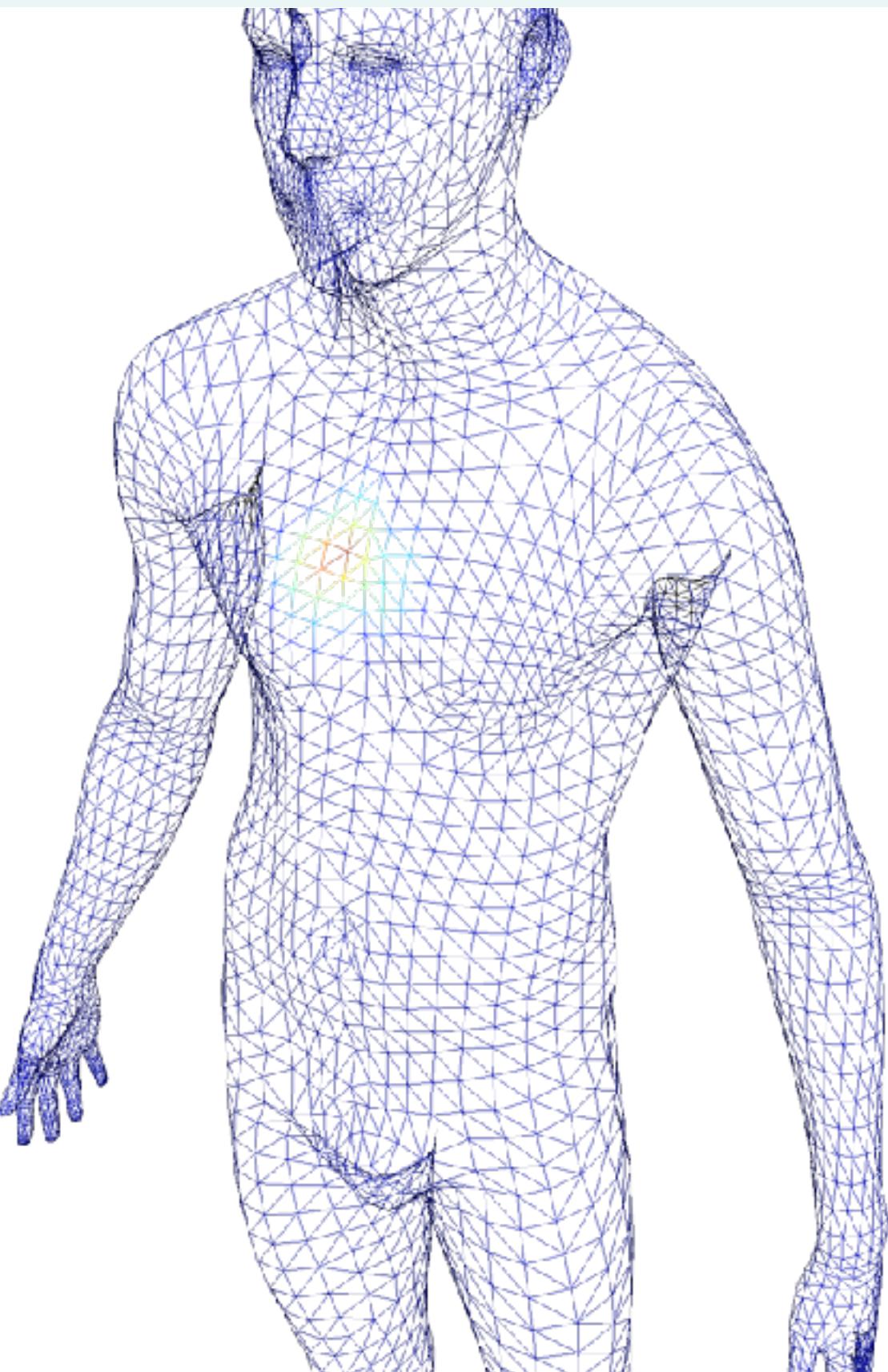
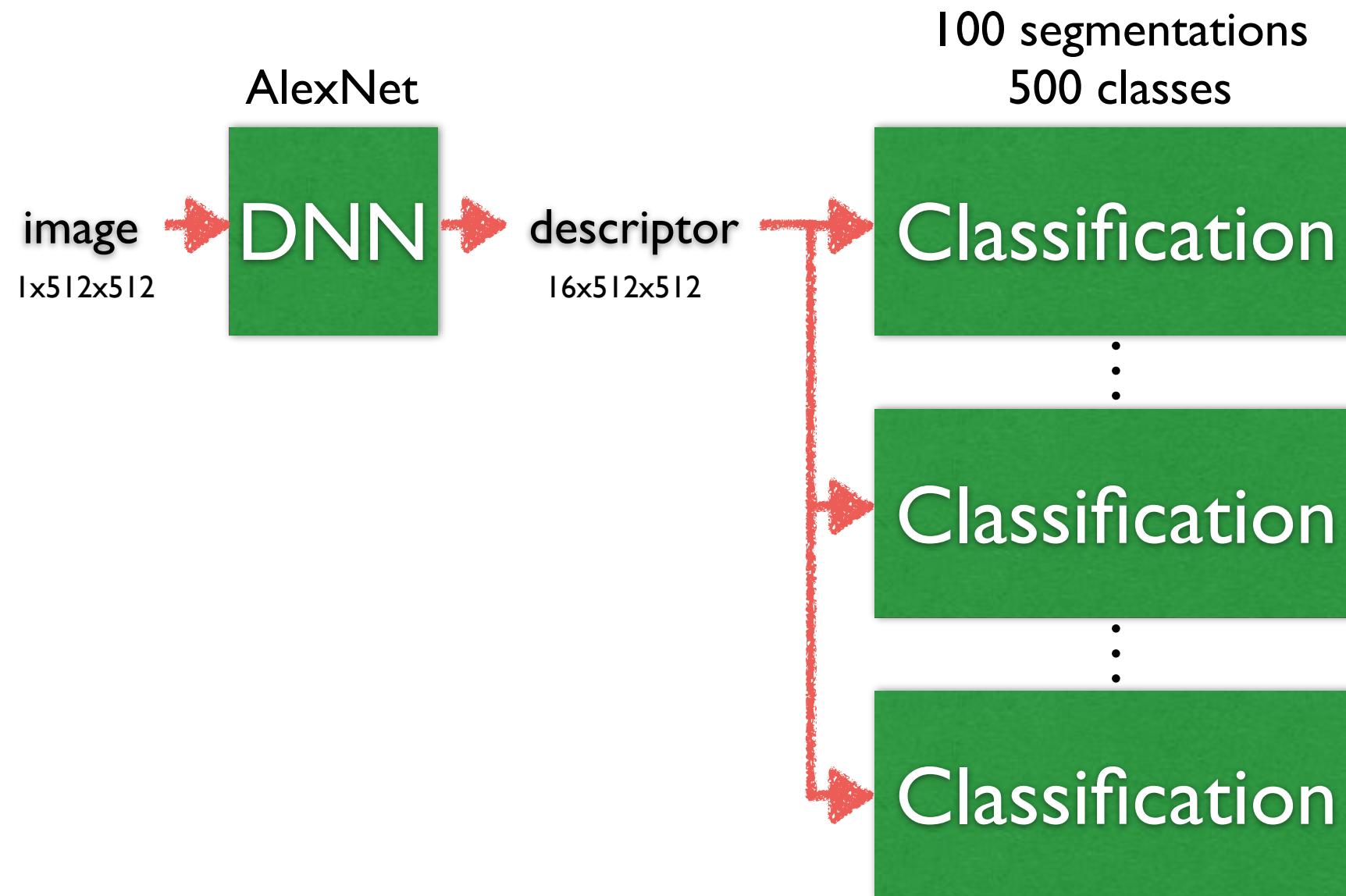
500 classes



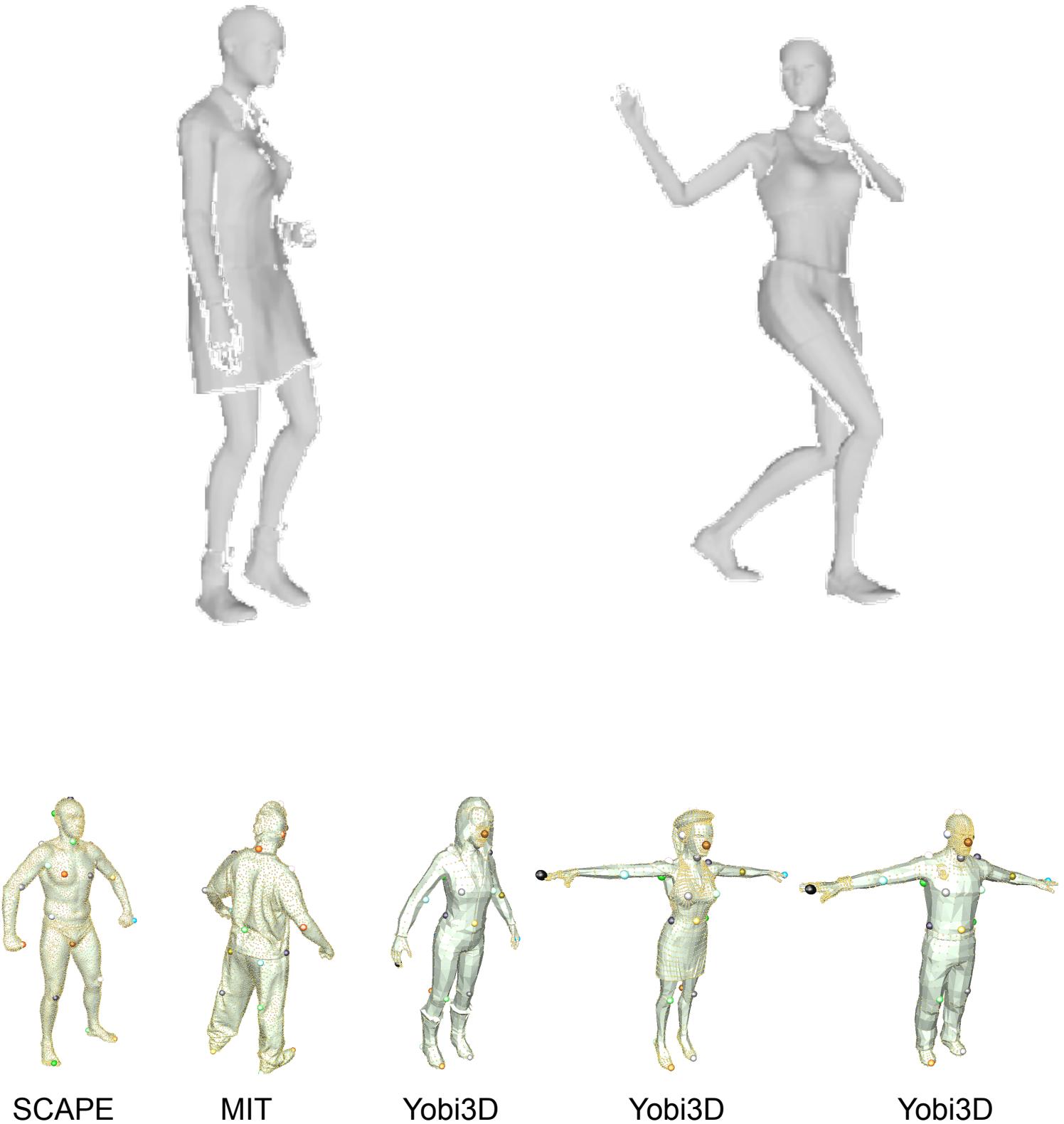
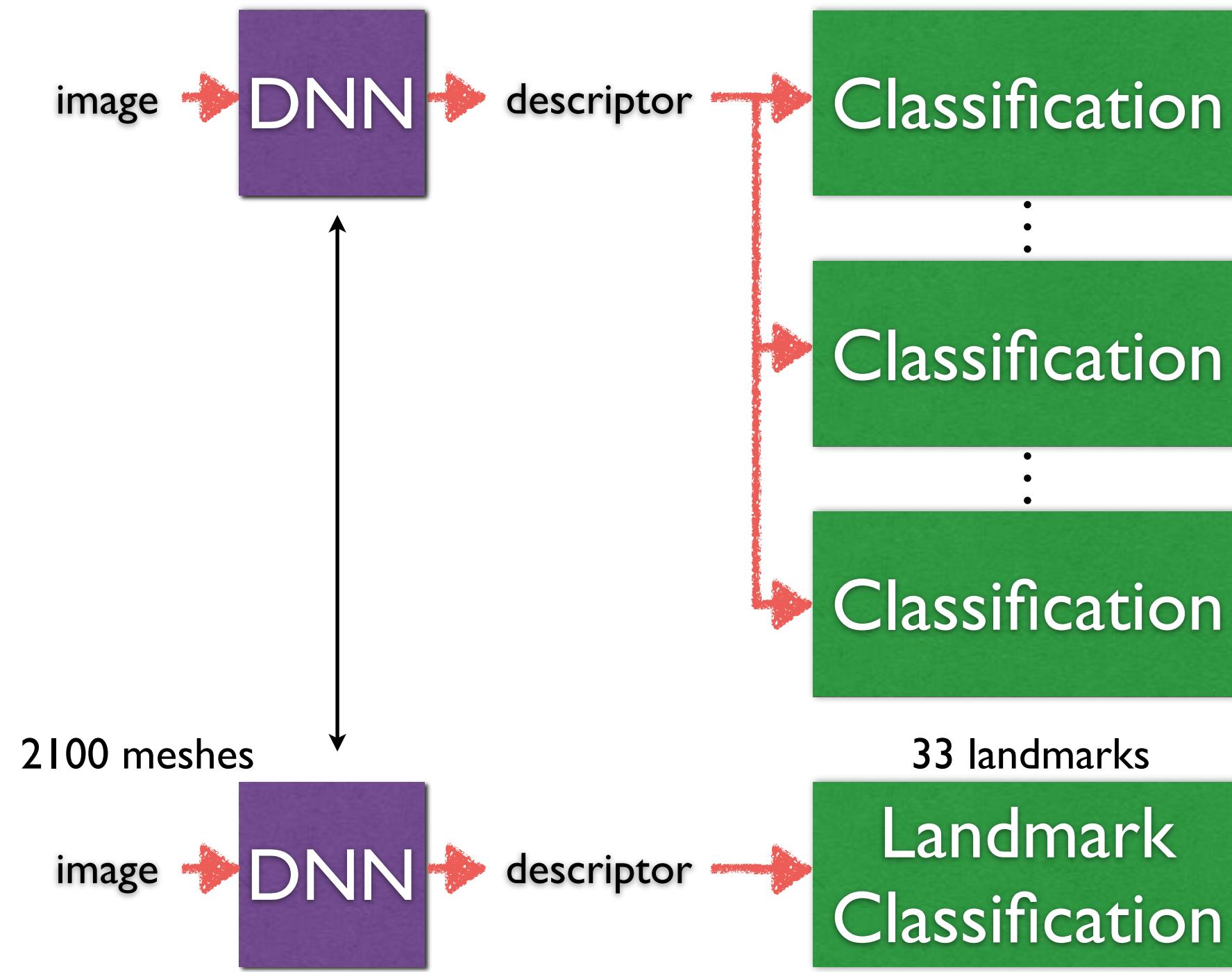
⋮

100 random
segmentations

Distance Preserving Learning

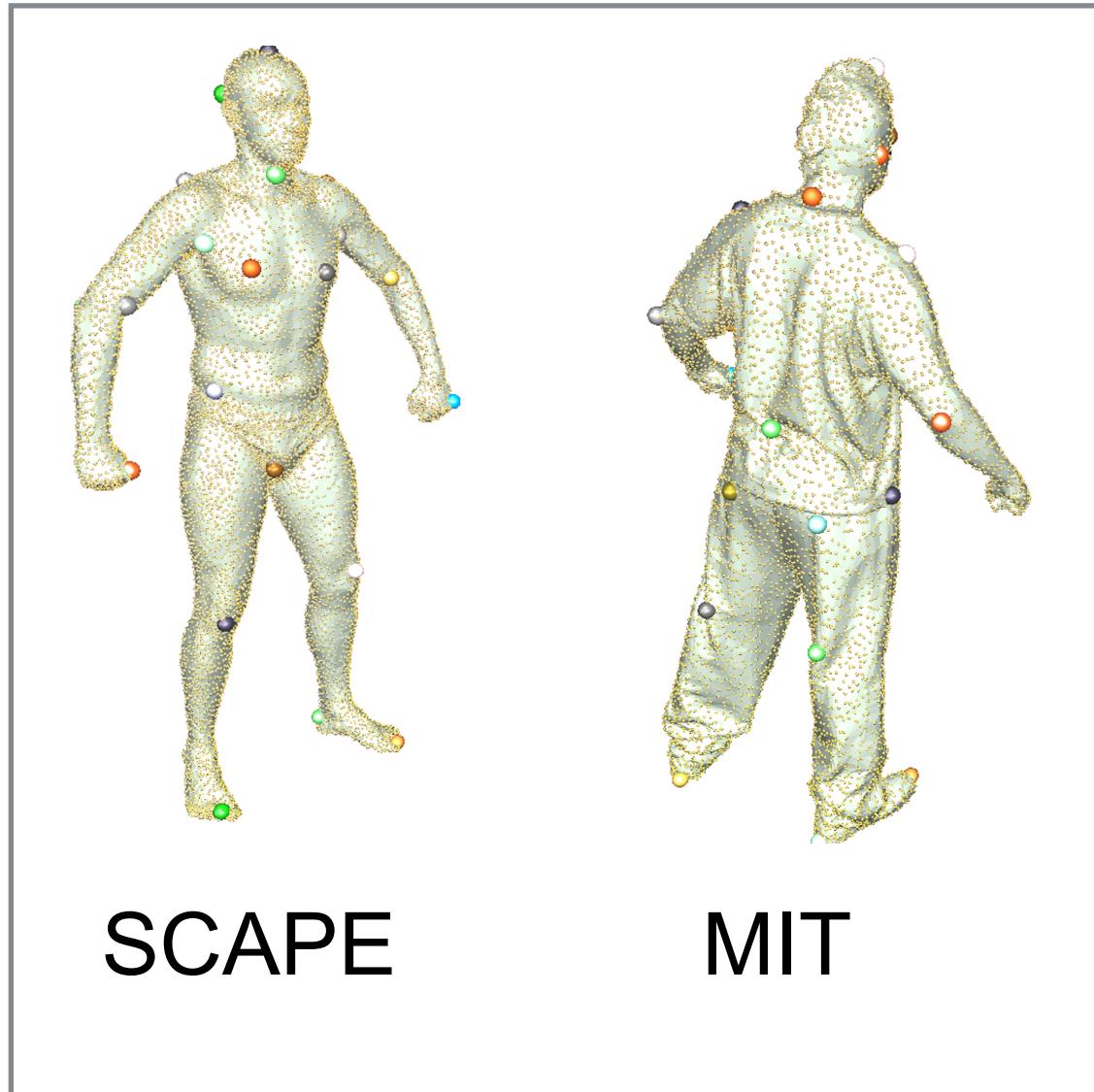


Variation on Clothing

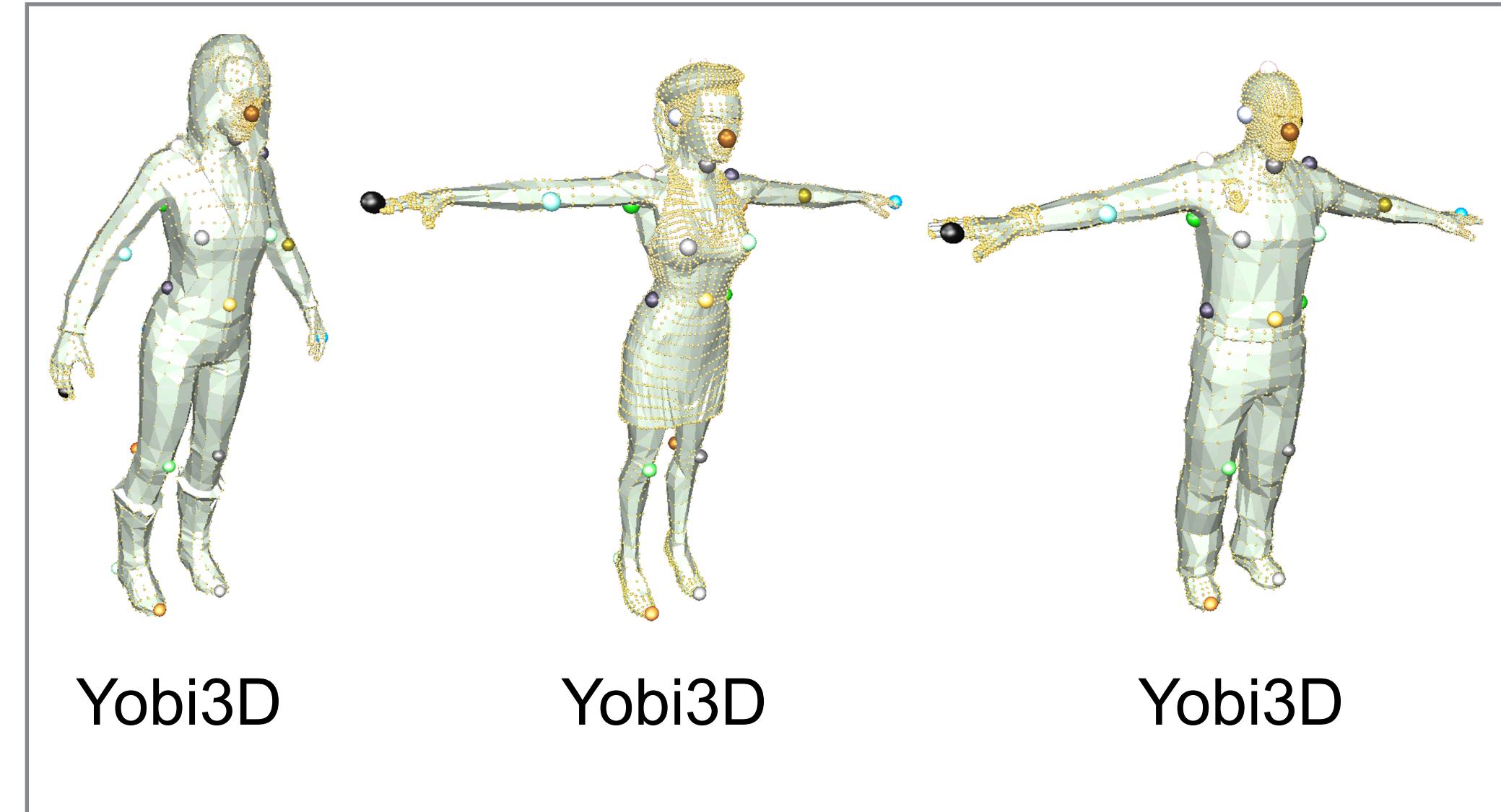


Training Data

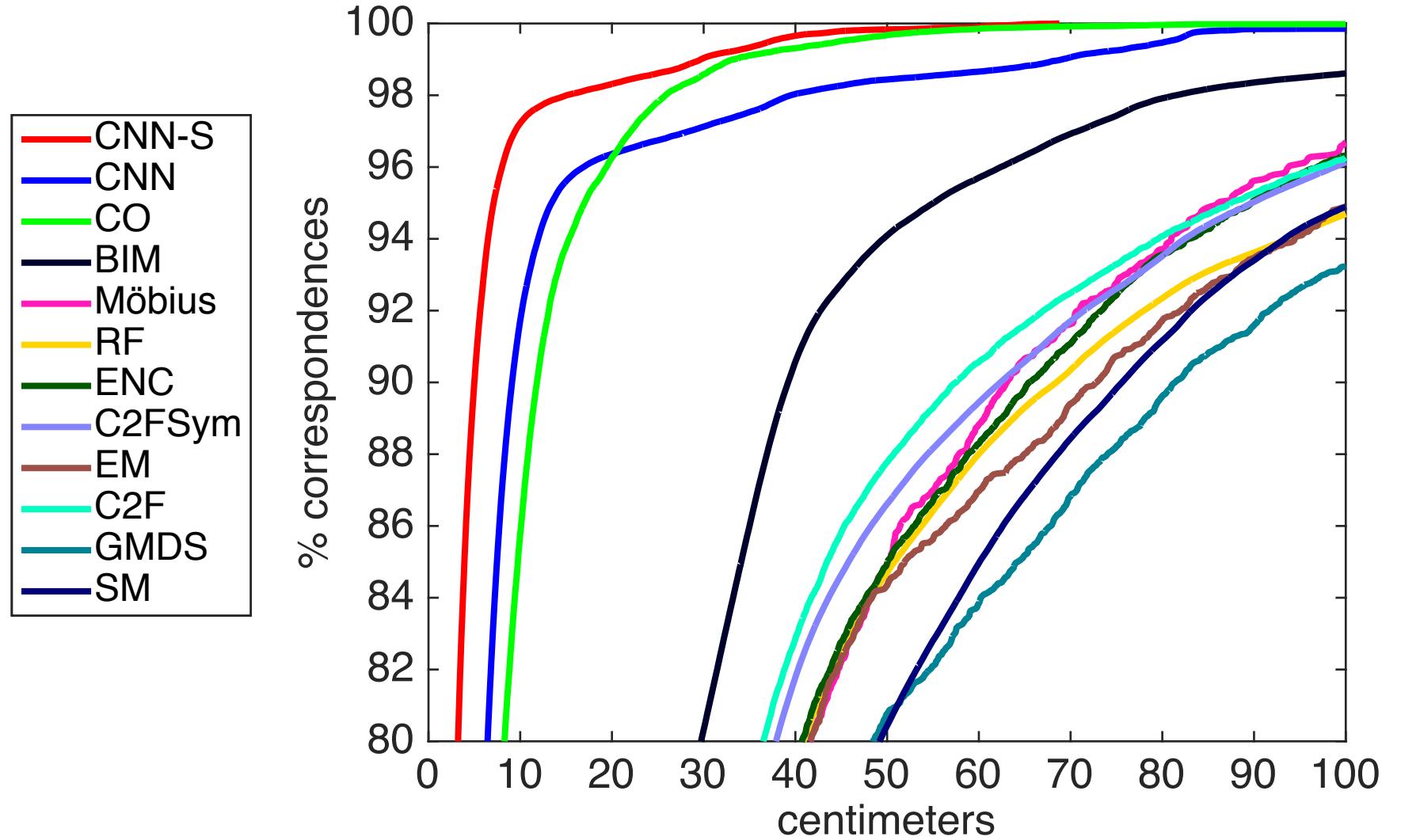
Shape & Pose



Clothing



Evaluation

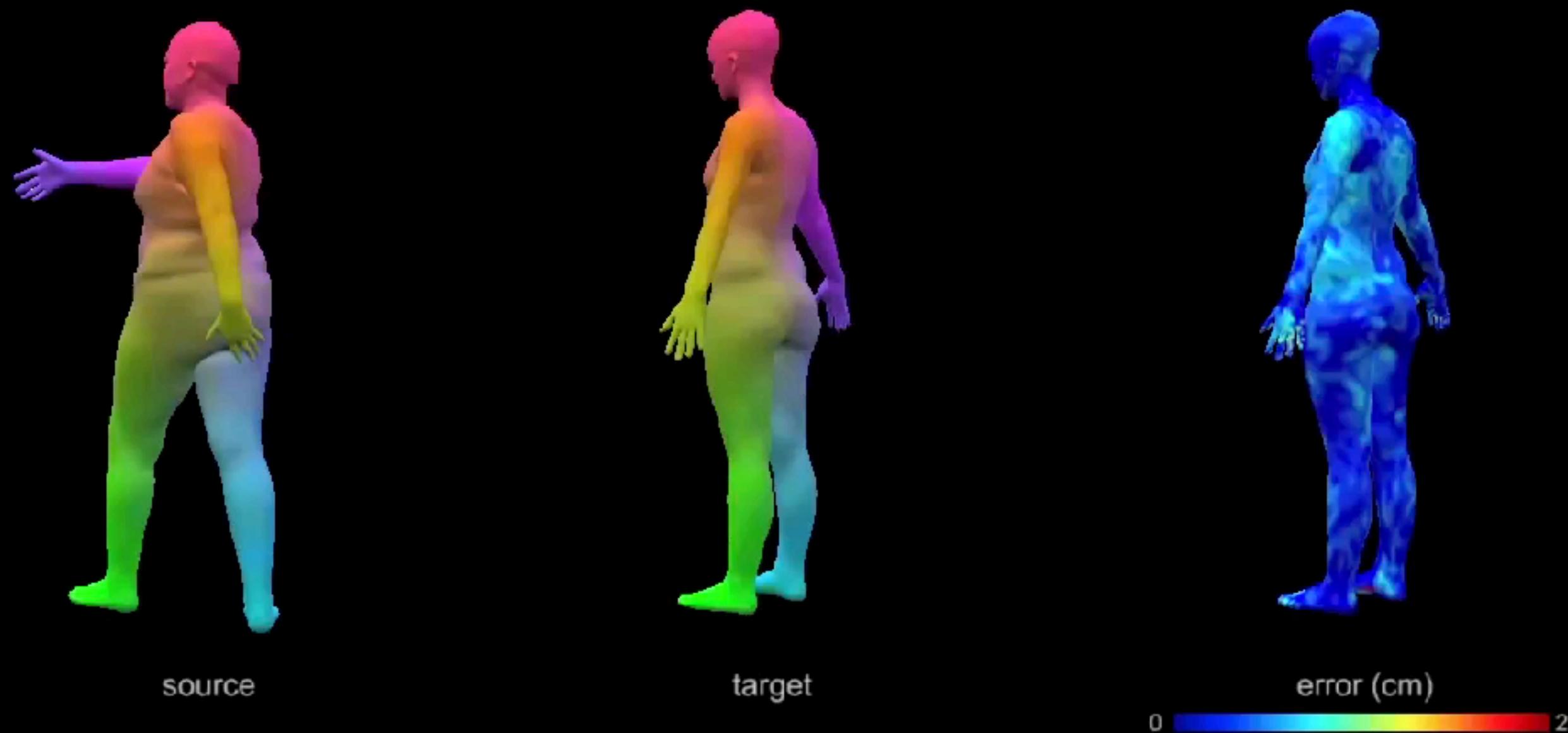


FAUST dataset

Results

Results: Static Shapes

full-to-full correspondences (synthetic data and naked)



Results: Static Shapes

full-to-partial correspondences (real data and clothed)



source



target

Results: Dynamic Shapes



input scans



correspondences (per frame)

Results: Dynamic Shape Reconstruction

dynamic correspondences (side view)



input scans

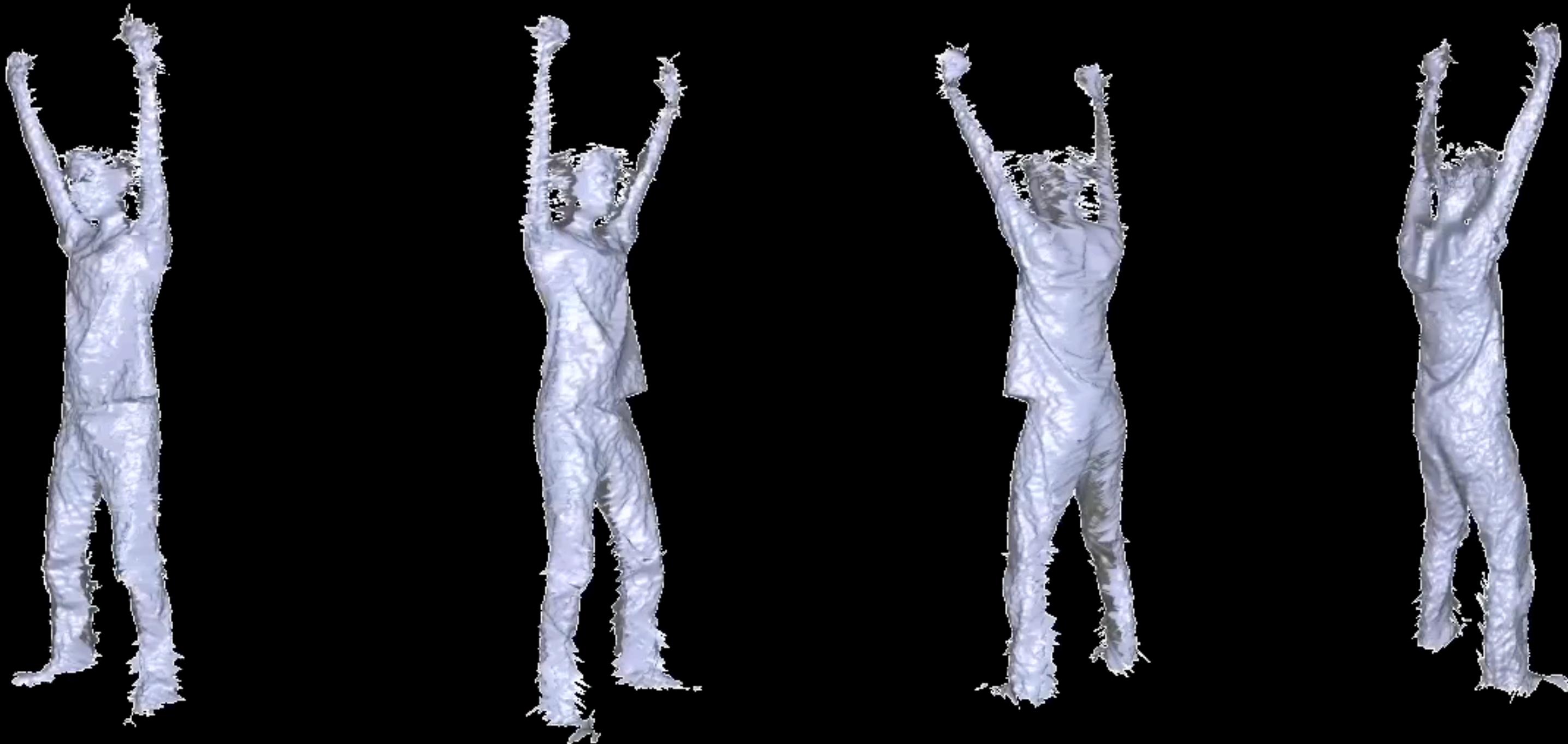


correspondences (per frame)

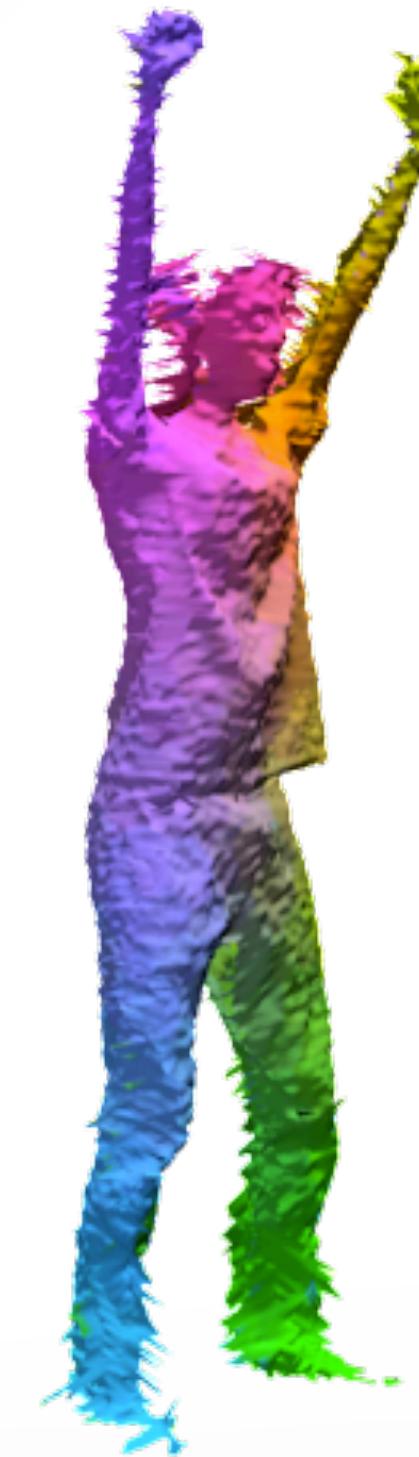


reconstruction and input scans

4 Stationary Kinects



Dense Correspondences



Applications

3D Human Capture

Microsoft 2015



Outside-in Capture

Microsoft



Low Cost Capture & Moving Target



Registration and Reconstruction



output scan alignment

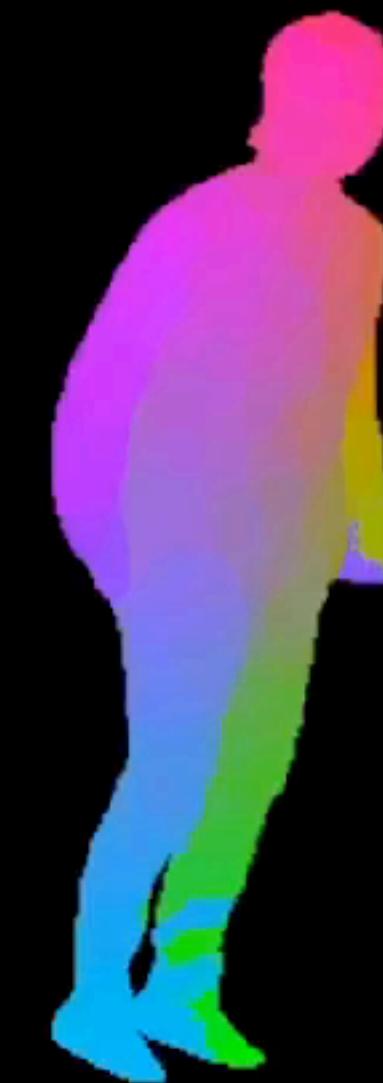


output textured reconstruction

Filtering and Texture Reconstruction



denoised
mesh

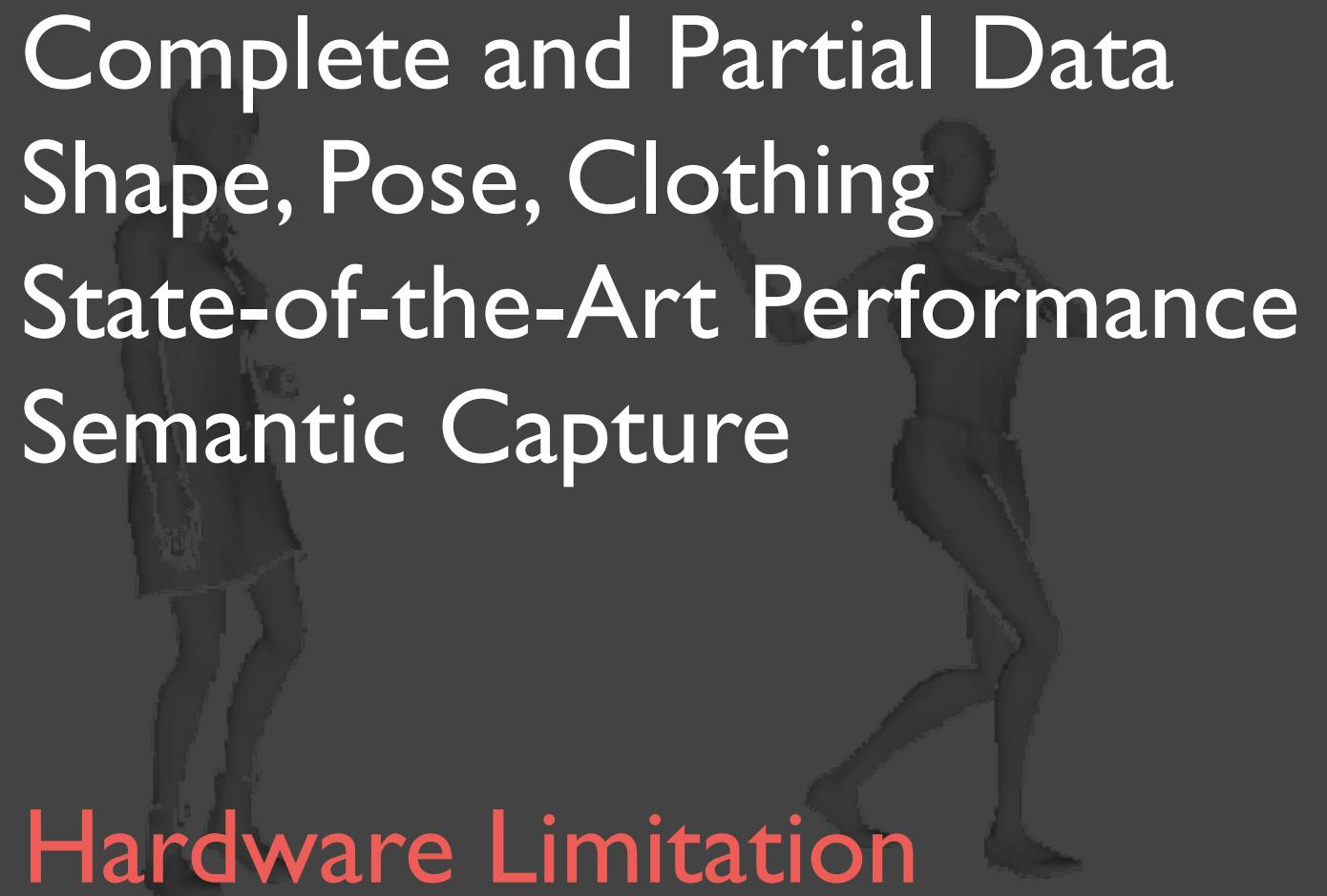


dense
correspondences



textured
mesh reconstruction

Discussion



Complete and Partial Data
Shape, Pose, Clothing
State-of-the-Art Performance
Semantic Capture

Hardware Limitation
Only Humans for Now

Thanks!