

SEMANTIC PARAMETRIC BODY SHAPE ESTIMATION FROM NOISY DEPTH SEQUENCES

Alexandru E. Ichim



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Federico Tombari



TECHNISCHE
UNIVERSITÄT
MÜNCHEN



SOME OF MY OTHER WORK I



Static

8 MP, ~100 images



Dynamic

HD, ~2 min

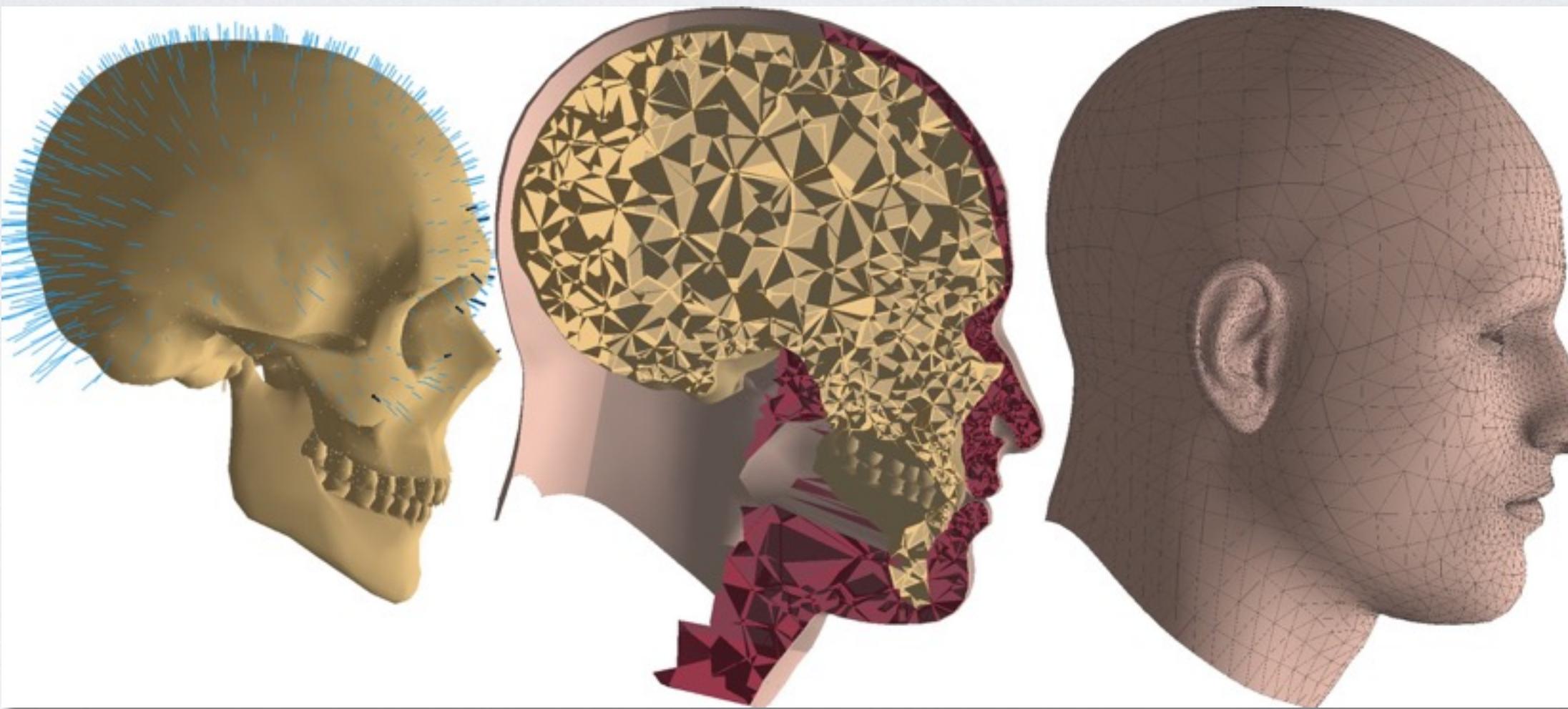


Dynamic 3D Avatar Creation From Hand-held Video Input

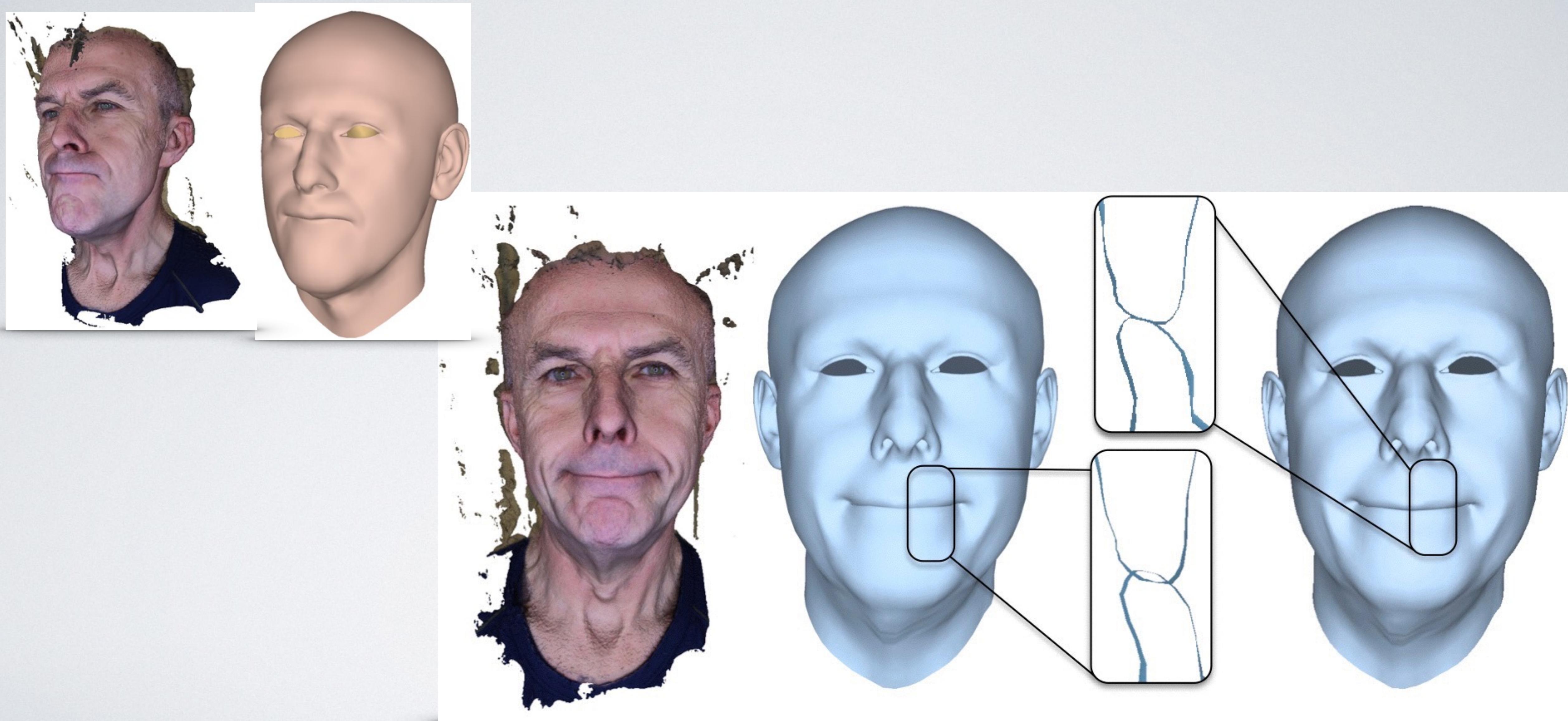
SOME OF MY OTHER WORK I



SOME OF MY OTHER WORK 2



SOME OF MY OTHER WORK 2



SOME OF MY OTHER WORK 2

Sequence: chewing_0



Linear blendshapes

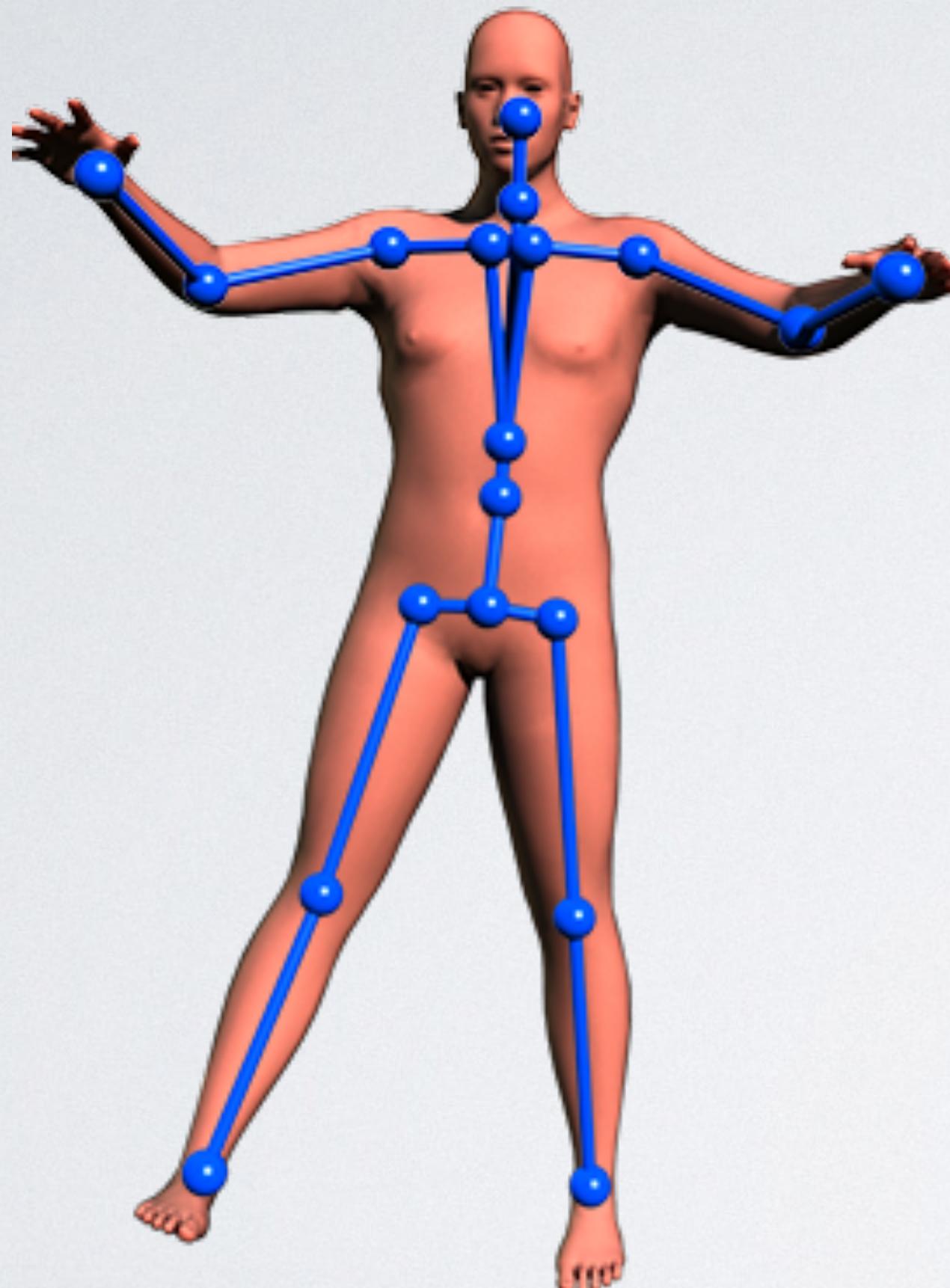


Our method



MOTIVATION

Generic framework for tracking and modeling articulated bodies



in our experiments:

human bodies

RGB-D input

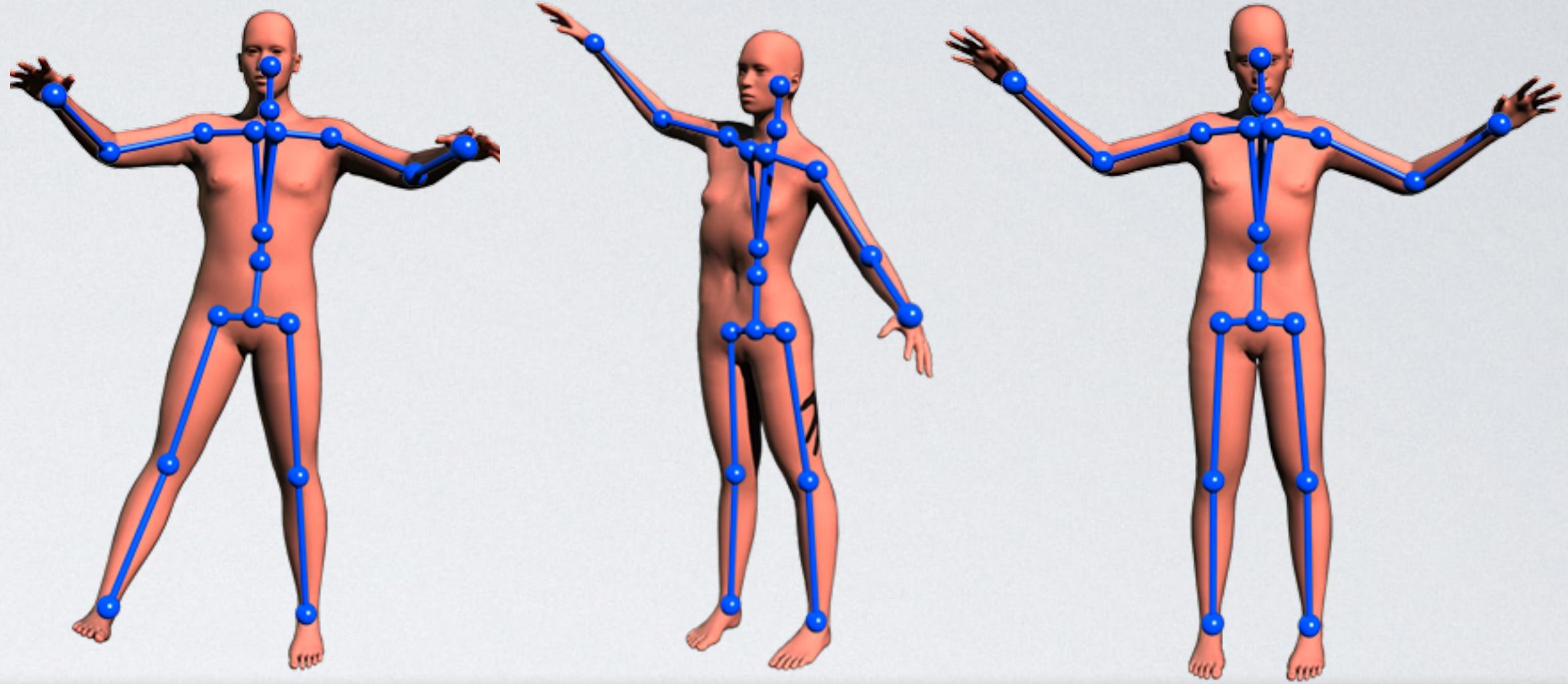
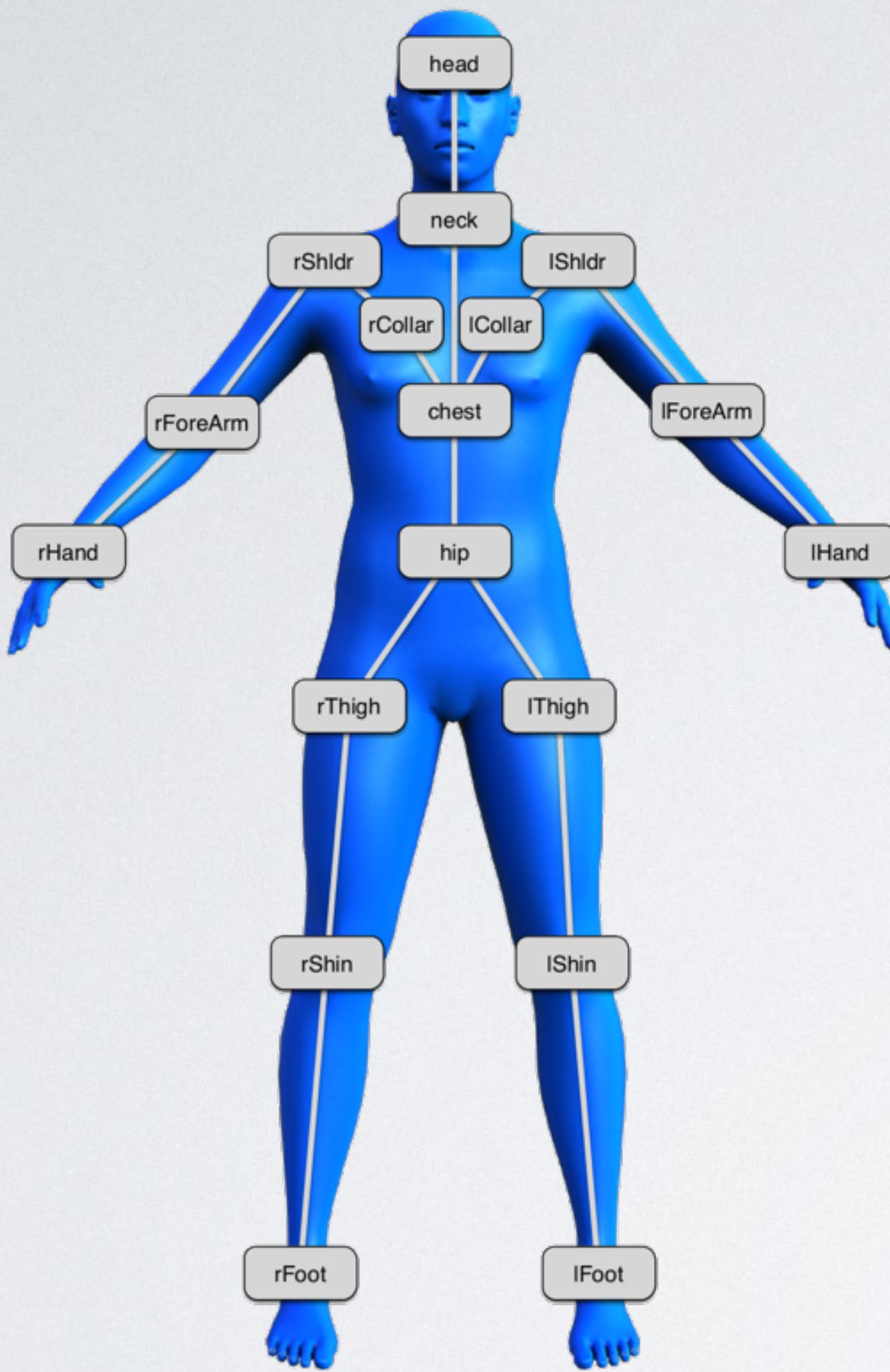
can be applied to:

hand tracking

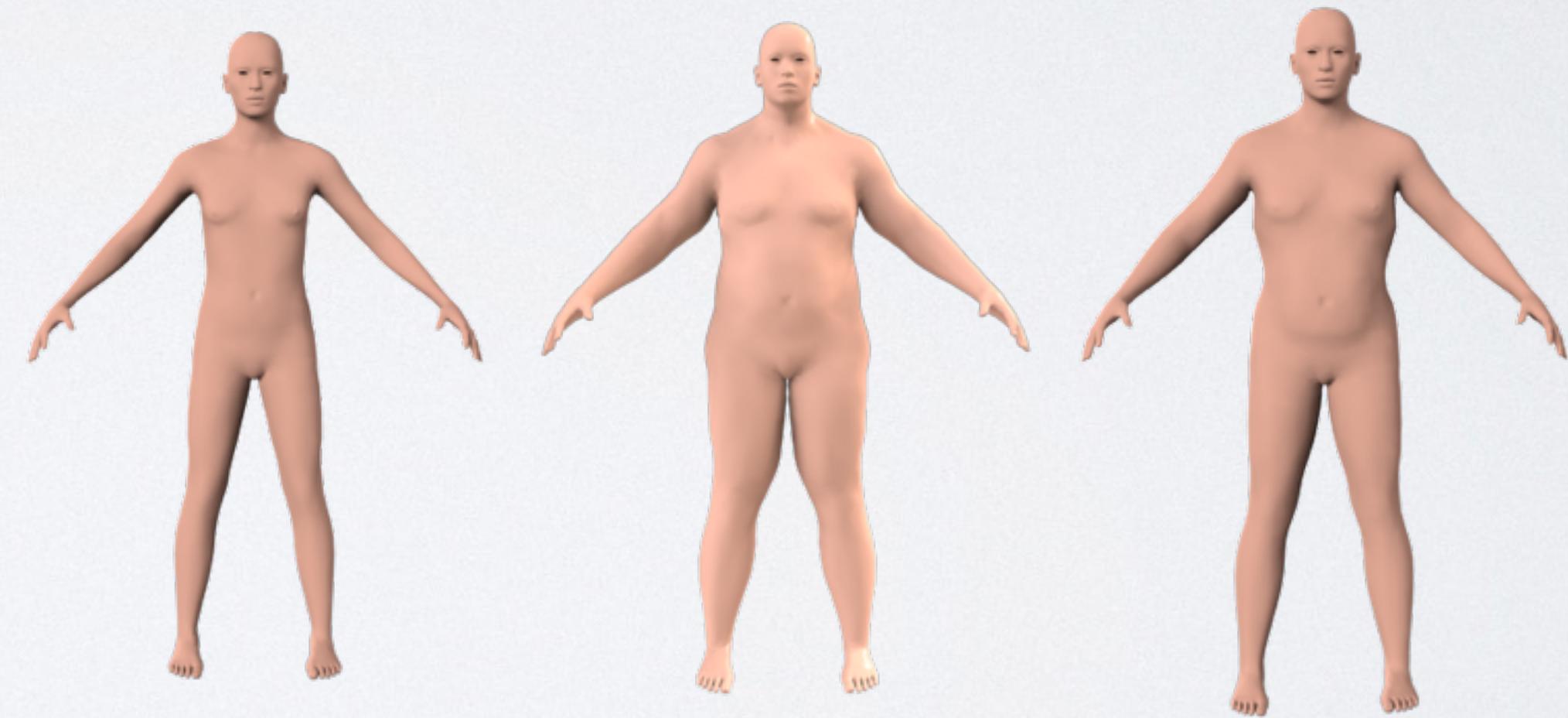
animal tracking



BODY REPRESENTATION



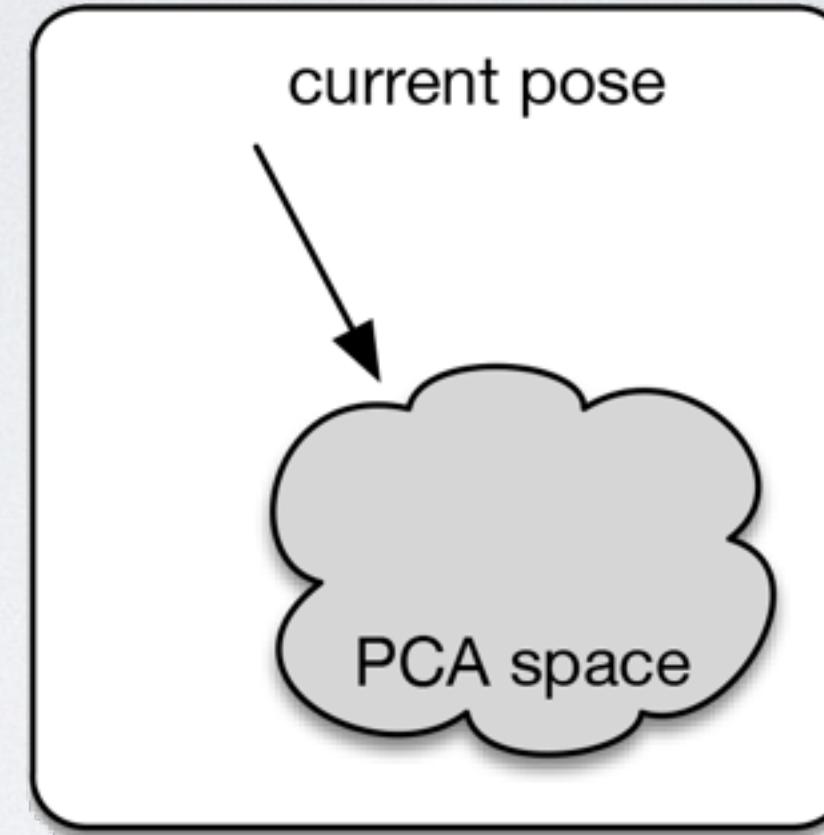
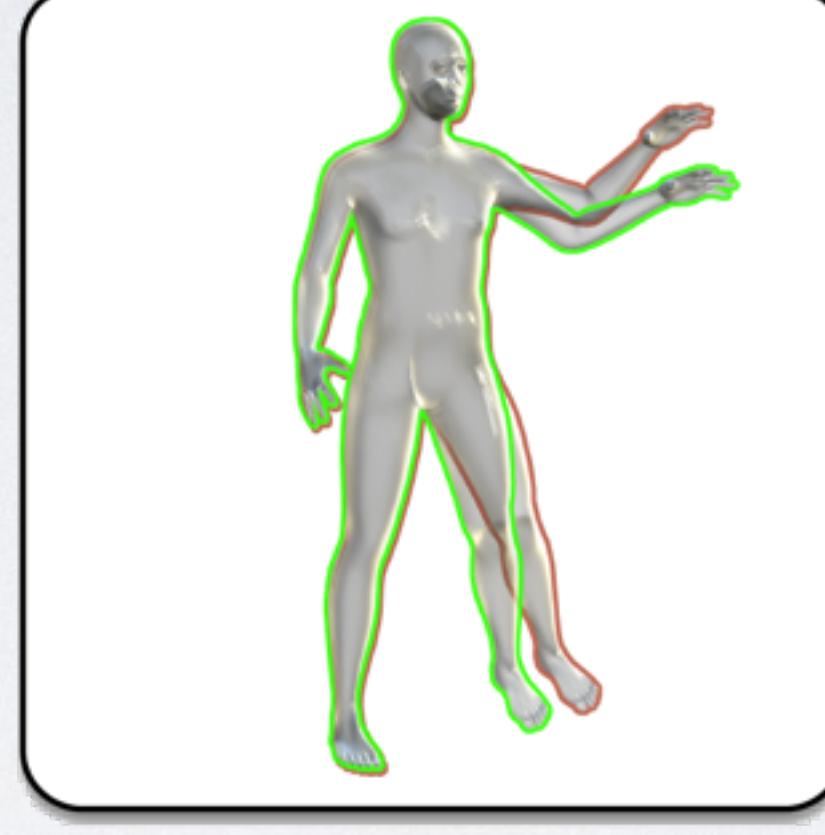
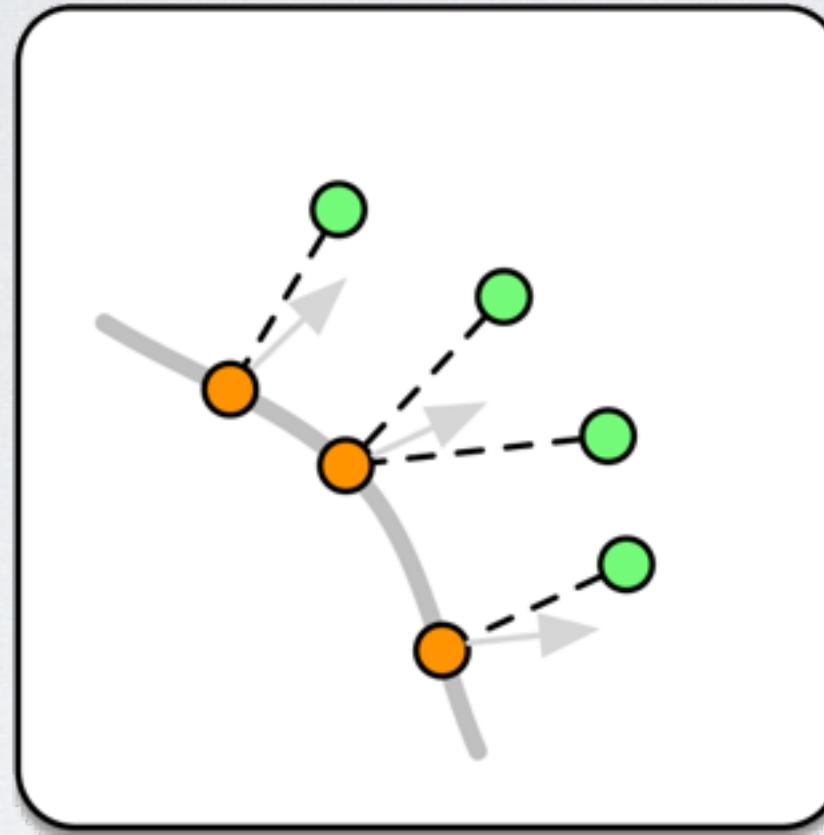
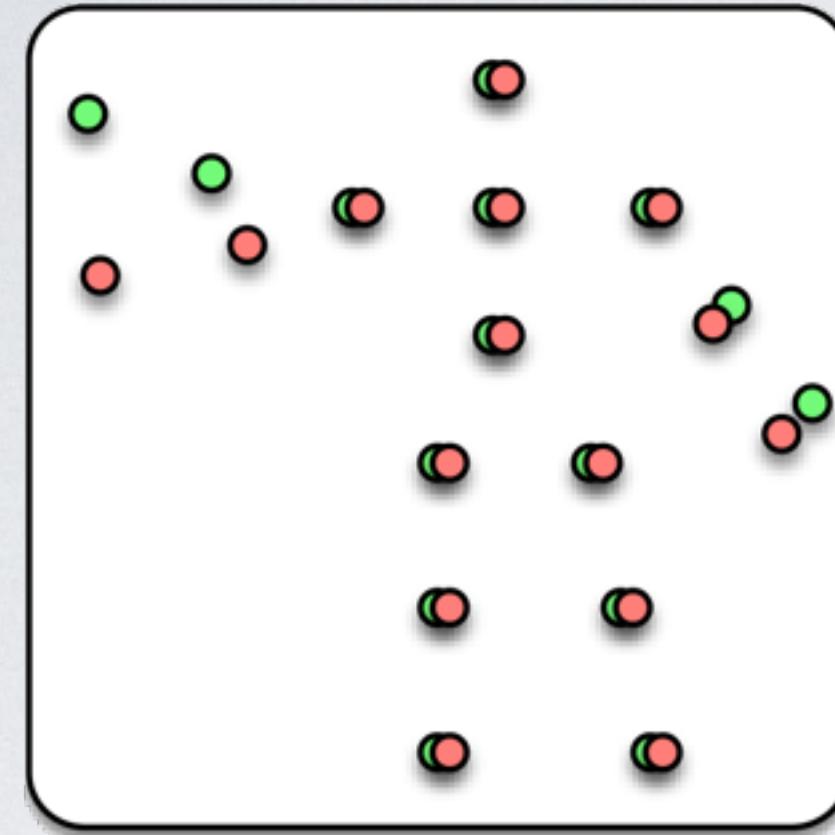
Pose variations - skeleton + linear blend skinning



Rest pose variations - blendshapes + bone lengths



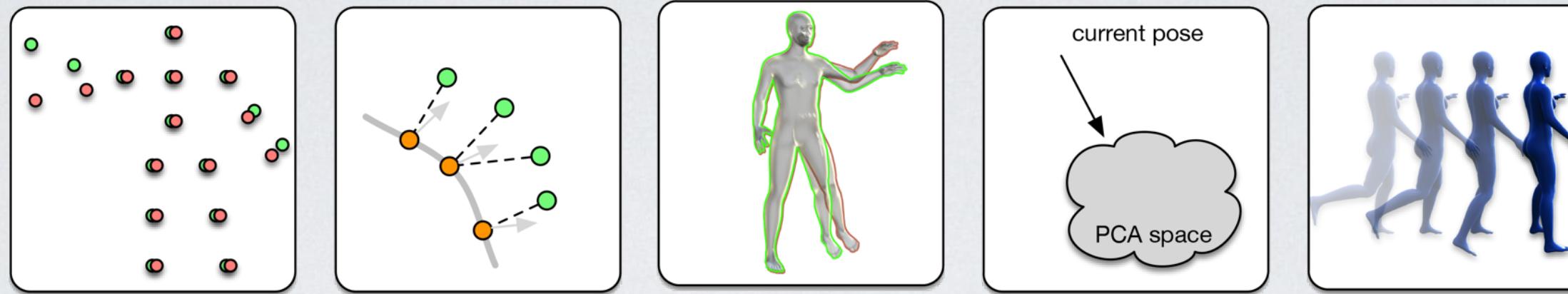
REGISTRATION ENERGY



$$E_{\text{total}} = \lambda_1 E_{\text{features}} + \lambda_2 E_{\text{surface}} + \lambda_3 E_{\text{contour}} + E_{\text{prior}} + \lambda_4 E_{\text{smoothness}}$$



REGISTRATION ENERGY



$$E_{\text{total}} = \lambda_1 E_{\text{features}} + \lambda_2 E_{\text{surface}} + \lambda_3 E_{\text{contour}} + E_{\text{prior}} + \lambda_4 E_{\text{smoothness}}$$

$$E_{\text{features}} = \sum_j \left\| t_j^{\text{pose}} - \chi_j \right\|_2^2$$

$$E_{\text{surface}} = \sum_i \left\| \mathbf{n}_i^T (\mathbf{x}_i - \mathbf{v}_i) \right\|_2^2$$

Features

Point-to-plane dense closeness

NiTE landmarks - replaceable

ICP iterations



REGISTRATION ENERGY



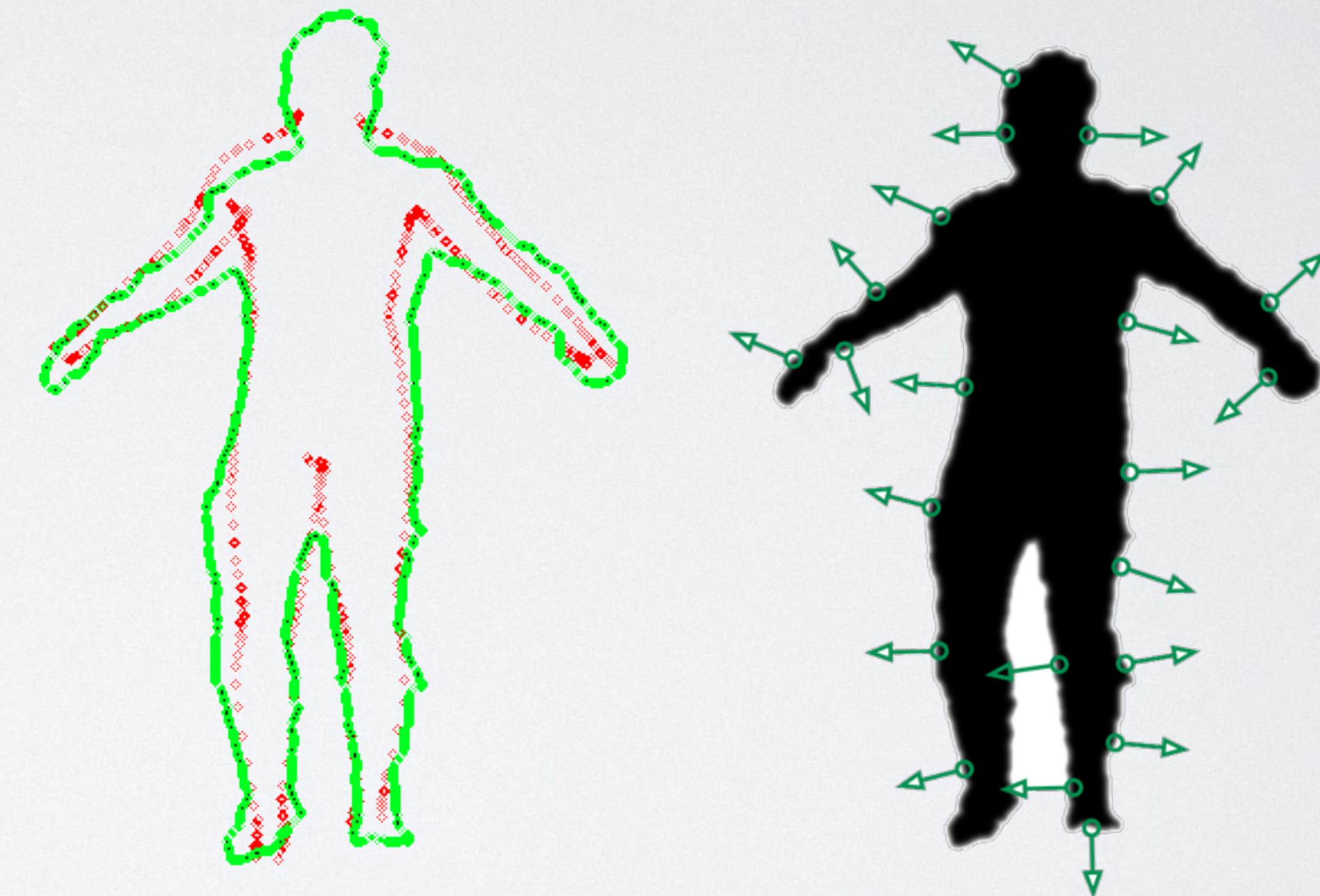
$$E_{\text{total}} = \lambda_1 E_{\text{features}} + \lambda_2 E_{\text{surface}} + \lambda_3 E_{\text{contour}} + E_{\text{prior}} + \lambda_4 E_{\text{smoothness}}$$

$$\mathbf{n} = \frac{(\nabla_x \bar{\mathbf{I}}_{\text{contour}}, \nabla_y \bar{\mathbf{I}}_{\text{contour}}, 0)^T}{\|(\nabla_x \bar{\mathbf{I}}_{\text{contour}}, \nabla_y \bar{\mathbf{I}}_{\text{contour}}, 0)\|}$$

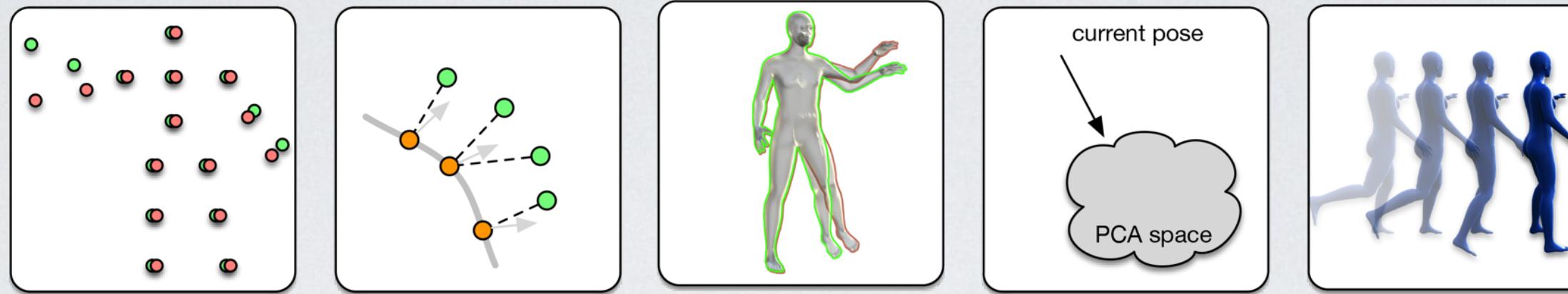
Contour

extracted from the 2D depth image

can be extended to RGB-segmentation



REGISTRATION ENERGY



$$E_{\text{total}} = \lambda_1 E_{\text{features}} + \lambda_2 E_{\text{surface}} + \lambda_3 E_{\text{contour}} + E_{\text{prior}} + \lambda_4 E_{\text{smoothness}}$$

$$E_{\text{prior}} = \lambda_5 E_{\text{prior-proj}} + \lambda_6 E_{\text{prior-dev}}$$

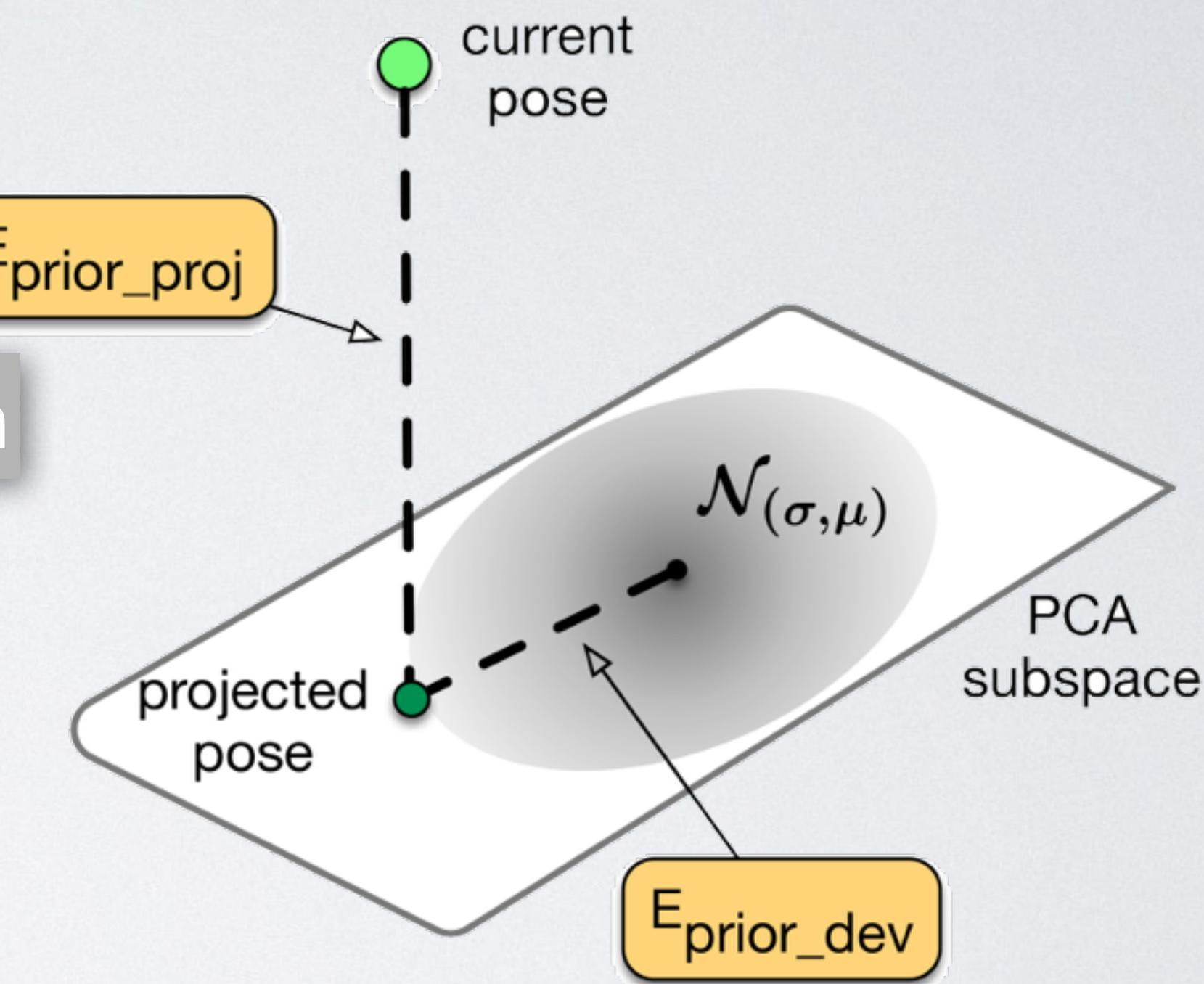
$$E_{\text{prior-proj}} = \left\| (\boldsymbol{\beta} - \boldsymbol{\mu}) - \mathbf{M} \mathbf{M}^T (\boldsymbol{\beta} - \boldsymbol{\mu}) \right\|_2^2$$

distance to projection

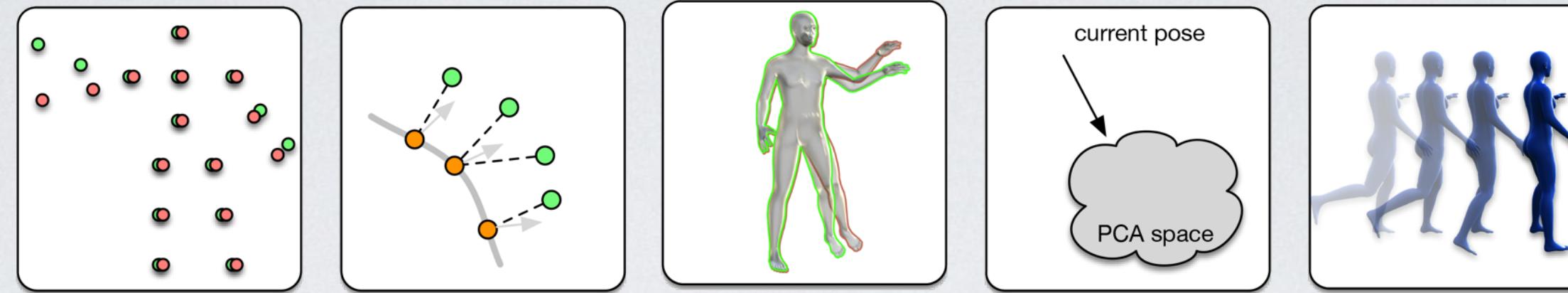
$$E_{\text{prior-dev}} = \left\| \boldsymbol{\Sigma}^{-1/2} \mathbf{M}^T (\boldsymbol{\beta} - \boldsymbol{\mu}) \right\|_2^2$$

std-dev-weighting
of the projection

Statistical Pose Prior



REGISTRATION ENERGY



$$E_{\text{total}} = \lambda_1 E_{\text{features}} + \lambda_2 E_{\text{surface}} + \lambda_3 E_{\text{contour}} + E_{\text{prior}} + \lambda_4 E_{\text{smoothness}}$$

$$E_{\text{smoothness}} = \left\| \boldsymbol{\beta}_{(t)} - \boldsymbol{\beta}_{(t-1)} \right\|_2^2$$

Smoothness term



MODELING ENERGY

$$E_{modeling} = \gamma_1 E_{surface} + \gamma_2 E_{contour} + \gamma_3 E_{bs_reg}$$

$$E_{bs_reg_L2} = \|\mathbf{s}\|_2^2$$

$$E_{bs_reg_L1} = \|\mathbf{s}\|_1$$

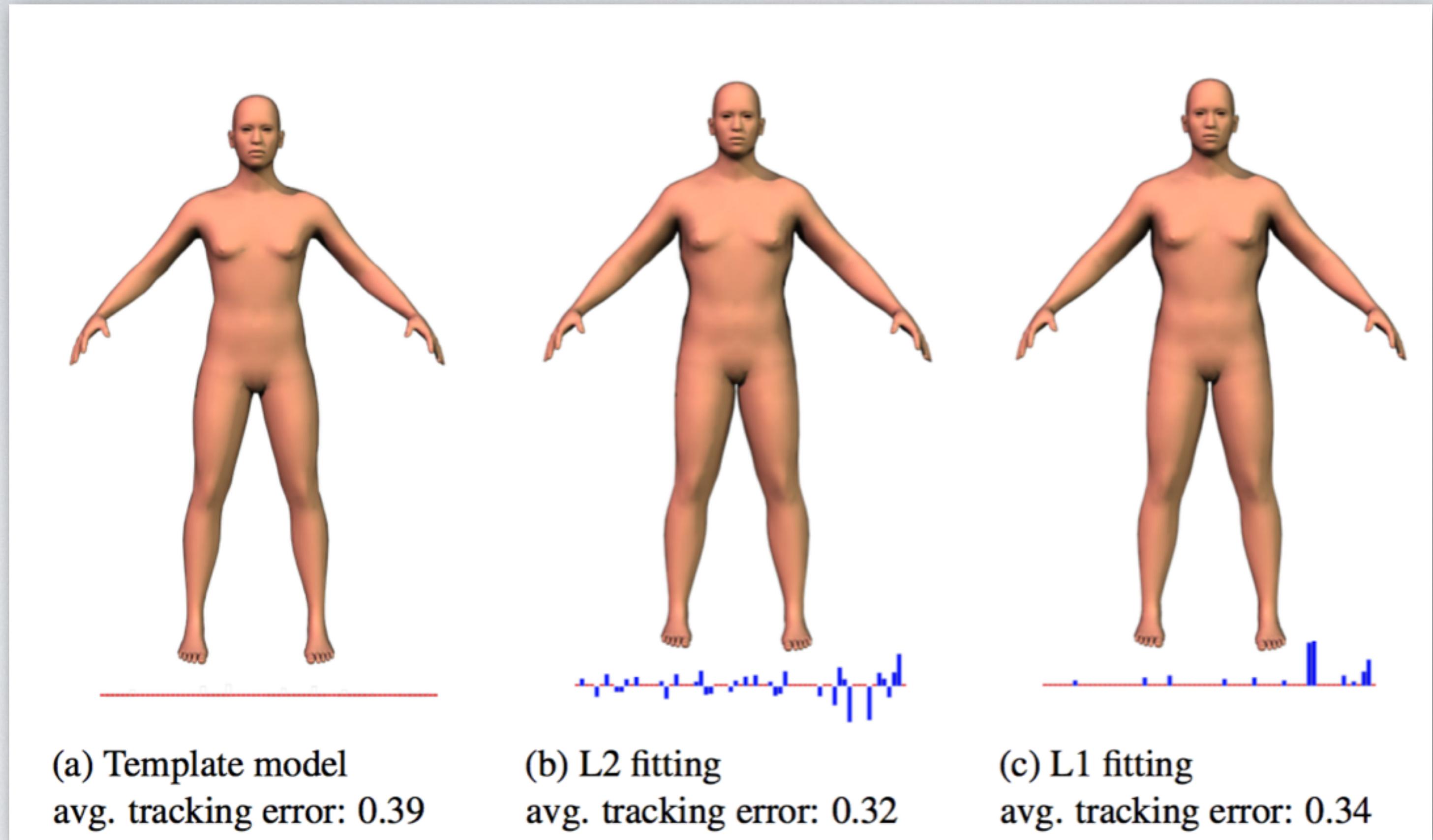
Regularization terms

accumulate constraints over windows of frames -
performance trade-off

solve and update the body template shape



MODELING ENERGY



L2 regularization

-decr r-upperleg-skinny
l-lowerleg-muscle-decr r-leg-genu-valgus l-lowerleg-skinny
stomach-pregnant-incr bulge-decr buttocks-volume-decr
bulge-incr universal-female-young-maxmuscle-minweight
r-upperarm-muscle-decr l-upperarm-skinny l-upperarm-muscle-decr stomach-pregnant-decr
stomach-tone-incr universal-female-young-averagemuscle-minweight
r-upperarm-shoulder-muscle-incr stomach-navel-up l-upperarm-fat
l-lowerarm-fat torso-vshape-more l-lowerleg-fat r-upperarm-muscle-incr
r-lowerleg-fat r-leg-genu-varun r-lowerleg-skinny stomach-navel-down
r-upperarm-fat r-lowerarm-skinny l-upperleg-muscle-decr torso-vshape-less
r-lowerarm-fat l-lowerleg-muscle-incr l-upperleg-muscle-incr
pelvis-tone-decr universal-female-young-maxmuscle-maxweight
l-lowerarm-muscle-incr buttocks-volume-incr r-upperarm-shoulder-muscle-decr
r-upperarm-skinny l-upperleg-fat pelvis-tone-incr
universal-female-young-minmuscle-minweight r-upperleg-muscle-incr l-upperleg-skinny l-leg-genu-valgus

L1 regularization

young-averagemuscle-minweight
universal-female-young-maxmuscle-maxweight
universal-female-young-minmuscle-minweight
r-upperleg-fat pelvis-tone-incr
universal-female-young-minmuscle-averageweight
torso-muscle-dorsi-incr
l-upperleg-fat torso-vshape-less
torso-muscle-pectoral-decr
l-leg-genu-varun
universal-female-young-minmuscle-maxweight
l-lowerarm-fat r-upperarm-fat
stomach-pregnant-decr

word cloud
visualizations



RESULTS

Semantic Parametric Body Shape Estimation from Noisy Depth Sequences

Alexandru E. Ichim

Federico Tombari

project webpage: http://lgg.epfl.ch/~ichim/bodies_ras_2015/

Supplementary video



Thank you!

