Fraunhofer Heinrich Hertz Institute
Computer Vision & Graphics
Image Processing Department

Visual Computing
Department of Computer Science
Humboldt-Universität zu Berlin

KINEMATIC ICP FOR ARTICULATED TEMPLATE FITTING

Philipp Fechteler, Anna Hilsmann, and Peter Eisert

Introduction

We present a robust and efficient optimization to recover the 3D pose of a human from a point cloud (single scan, visual hull, z-cam, Kinect …) :

- the well-known ICP algorithm is adapted to fit a generic articulated template model to the given 3D points
- each iteration jointly refines the parameters for rigid alignment, uniform scale as well as all joint angles, and optionally with consideration of Linear Blend Skinning weights
- sample based learning of optimal template model in terms of
  - joint rotation centers and
  - Linear Blend Skinning Weights

Kinematic Template Fitting

The objective function is modeled according to kinematic dependencies:

$$d_0 = \|\boldsymbol{S}_0 - \boldsymbol{t}\boldsymbol{1}^T - s\boldsymbol{R}\boldsymbol{T}_0\|$$
$$d_1 = \|\boldsymbol{S}_1 - \boldsymbol{t}\boldsymbol{1}^T - s\boldsymbol{R}[\boldsymbol{R}_1(\boldsymbol{T}_1 - \boldsymbol{t}_1\boldsymbol{1}^T) + \boldsymbol{t}_1\boldsymbol{1}^T\ ]\ \|$$
$$d_2 = \|\boldsymbol{S}_2 - \boldsymbol{t}\boldsymbol{1}^T - s\boldsymbol{R}[\boldsymbol{R}_1(\boldsymbol{R}_2(\boldsymbol{T}_2 - \boldsymbol{t}_2\boldsymbol{1}^T) + \boldsymbol{t}_2\boldsymbol{1}^T - \boldsymbol{t}_1\boldsymbol{1}^T) + \boldsymbol{t}_1\boldsymbol{1}^T\ ]\ \|$$

with $\mathbf{S}_0$, $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{T}_0$, $\mathbf{T}_1$, $\mathbf{T}_2$ being matrices of scan resp. template vertices with 0, 1 and 2 joints to the root limb, s, $\mathbf{R}$, $\mathbf{t}$ being the uniform similarity parameters, $\mathbf{R}_1$, $\mathbf{R}_2$ being the rotation matrices of 1st and 2nd joint in the kinematic chain $\mathbf{t}_1$, $\mathbf{t}_2$ being the joint rotation centers and $\mathbf{1}$ being a vector of 1's.

Introducing changes in parameters $\Delta[s\ \boldsymbol{t}\ \boldsymbol{r}\ \boldsymbol{r}_1\boldsymbol{r}_2 \ldots]$ and using the linearization of rotations via skew-symetric cross-product matrices: $\boldsymbol{R}\cdot\boldsymbol{p} \approx \boldsymbol{p} + [\boldsymbol{p}]_\times \cdot \boldsymbol{r}$, we can reformulate the upper equation system to

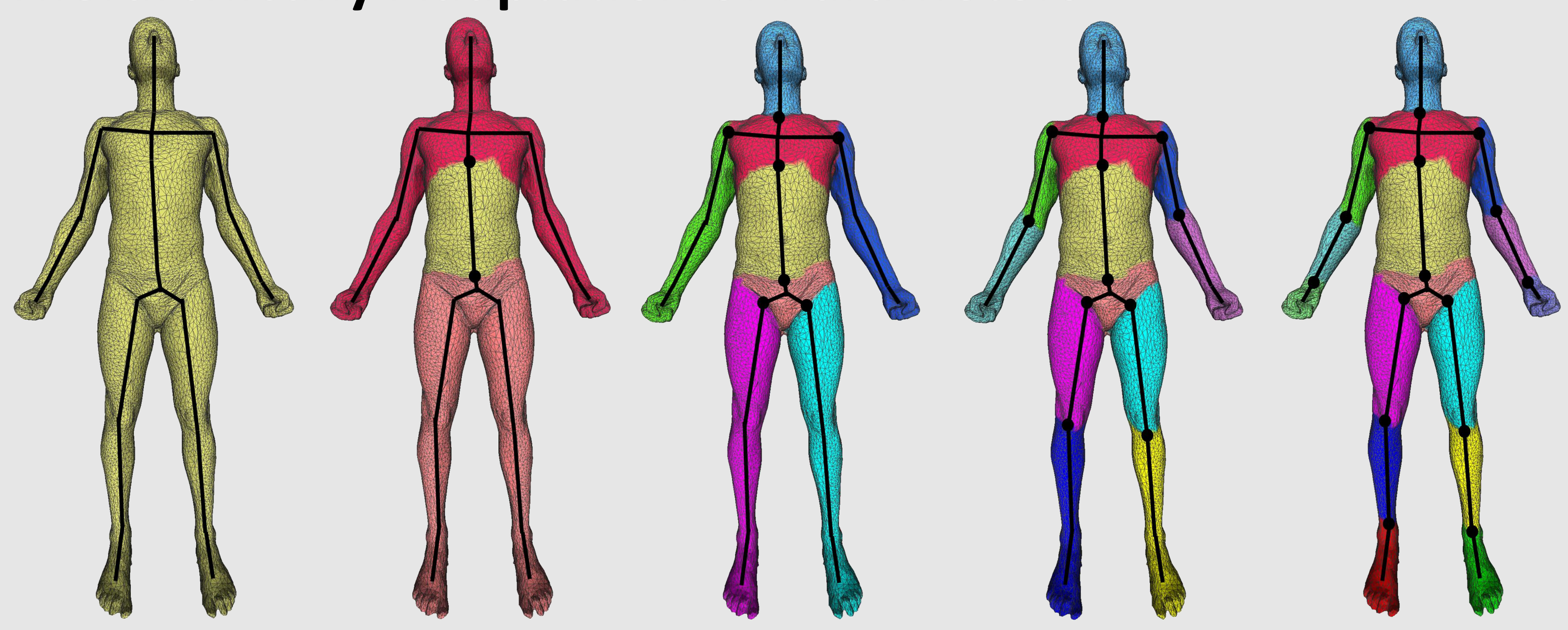
$$\Delta[s\ \boldsymbol{t}\ \boldsymbol{r}\ \boldsymbol{r}_1\ \boldsymbol{r}_2\ \ldots] = (\boldsymbol{M}^T\boldsymbol{M})^{-1}(\boldsymbol{M}^T\boldsymbol{N})$$

with $\boldsymbol{M}$ being a $3n \times p$ matrix, and $\boldsymbol{N}$ being a $3n$-vector, which can be setup in parallel and solved efficiently.

Since this approach is based on ICP, all ICP refinements can be introduced, e.g. by weighting each correspondence with...

- the reciprocal of its distance, we achieve robustness against outliers,
- the scalar product of the vertex normals, we omit correspondences of too different orientations,
- Linear Blend Skinning weights, we achieve smooth deformations

Hierarchically Adaptation of Parameters



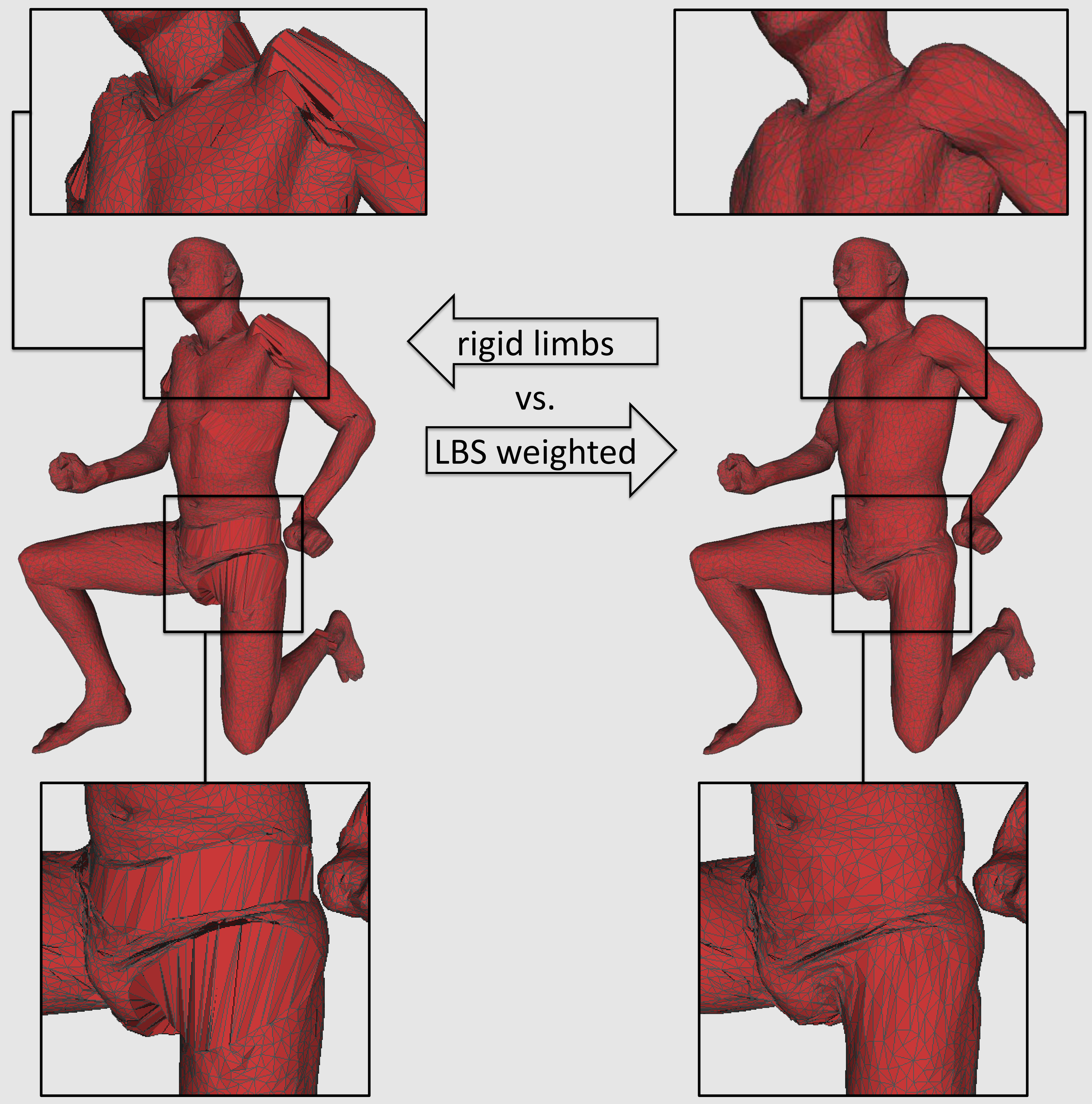
Optimization of Articulated Template Model

Joint rotation centers calculated from samples to better reflect realistic human surfaces:

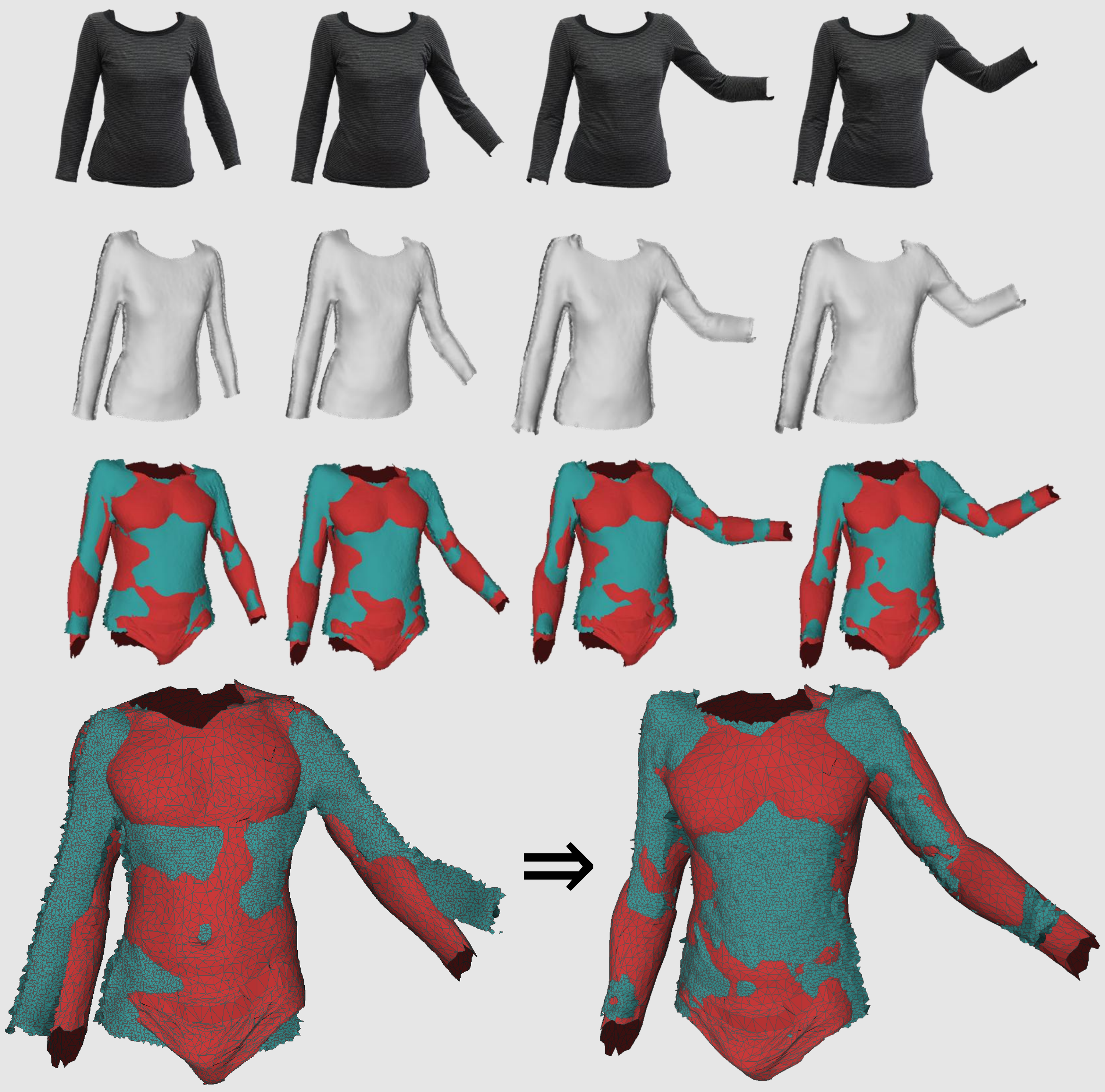
- Introduction of Δt_i's into the equation system above
- Optimization against all sample poses simultaneously, with one common set of rotation centers and separate sets of rotation matrices for different target poses

Smooth shape adaptations via Linear Blend Skinning techniques:

- Sample based optimization of weights to optimally reflect true human shape deformation (one weight per vertex and per bodypart)
- Solved via linear least squares optimization with equality and non-linear inequality constraints ($0 \le w_i \le 1$ and $\sum w_i = 1$)

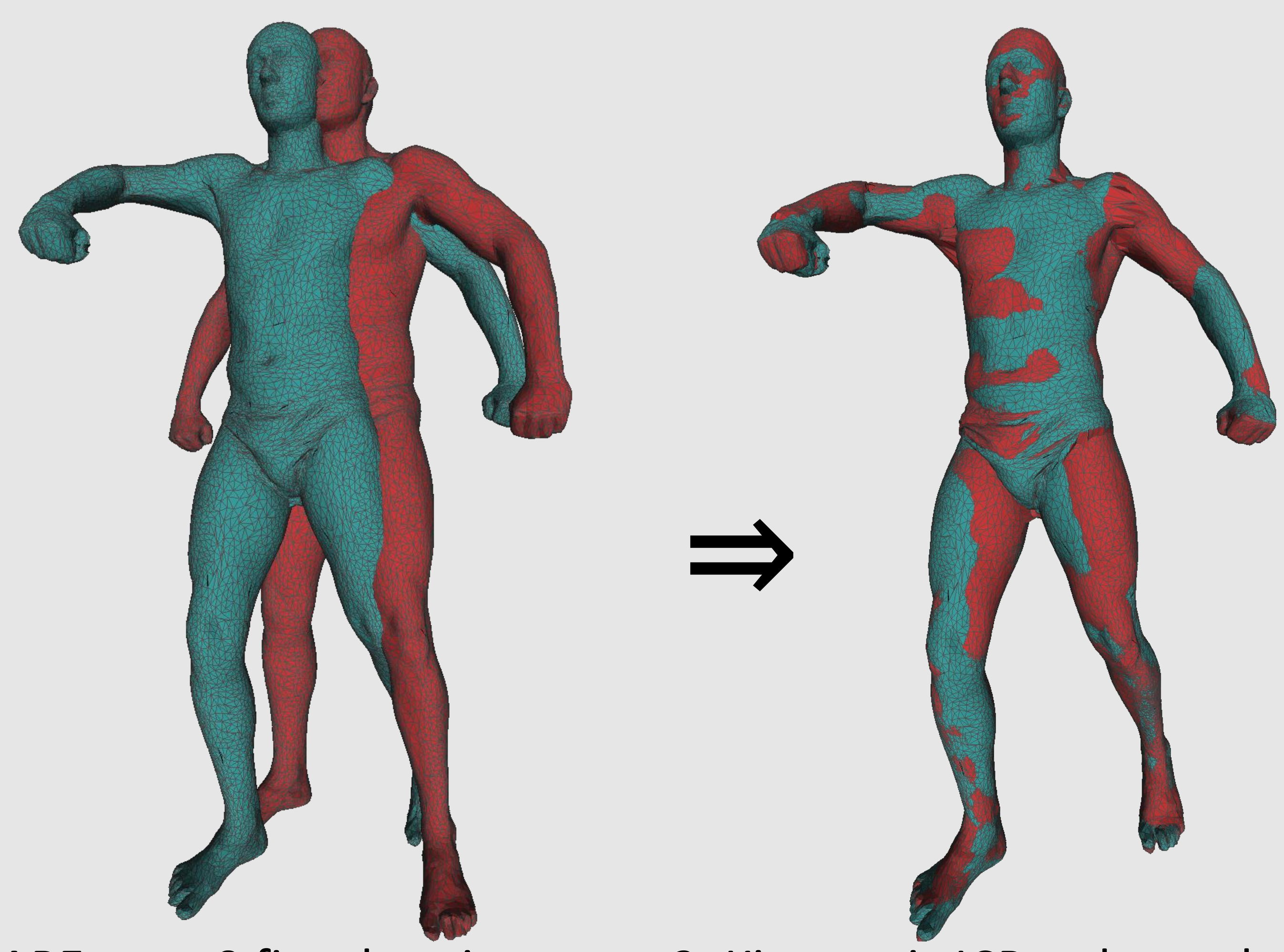


Application 1: Pose Tracking from Partial Depth Maps



Depth maps created with multi-view analysis. Kinematic ICP reduces the initial mean distance by a factor of 4.

Application 2: Pose Recovery from Full Body Mesh



SCAPE pose 0 fitted against pose 2. Kinematic ICP reduces the initial mean distance by a factor of 7 in 8 iterations.