Vision, Modeling, and Visualization (2012)
M. Goesele, T. Grosch, B. Preim, H. Theisel, and K. Toennies (Eds.)

Poster

Kinematic ICP for Articulated Template Fitting

Philipp Fechteler, Anna Hilsmann and Peter Eisert

Fraunhofer HHI / Humboldt University, Berlin, Germany

Abstract

In this paper, we present an efficient optimization method to adapt an articulated 3D template model to a full or partial 3D mesh. The well-known ICP algorithm is enhanced to fit a generic template to a target mesh. Each iteration jointly refines the parameters for global rigid alignment, uniform scale as well as the rotation parameters of all joint angles. The articulated 3D template model is based on the publicly available SCAPE data set, enhanced with automatically learned rotation centers of the joints and Linear Blend Skinning weights for each vertex. In two example applications we demonstrate the effectiveness of this computationally efficient approach: pose recovery from full meshes and pose tracking from partial depth maps.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism —Animation I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Motion

Introduction In this paper, we address the problem of robustly and accurately fitting a generic articulated template model of a human to a 3D triangle mesh, point cloud or depth map, representing either a full object or only a partial view of it. The adaptation takes place in terms of global rigid alignment, uniform scaling, and pose adaptation (joint angles) where all parameters are optimized jointly such that the outer surface of the template model approximates the true body shape as close as possible. By adapting the Iterative Closest Point (ICP) algorithm to handle kinematic dependencies correctly, we calculate joint configurations, which concatenate body parts along a kinematic chain, each having 3 rotational degrees of freedom [FE11]. The template model is based on the publicly available SCAPE dataset [ASK*05]. To increase the accuracy, joint rotation centers and Linear Blend Skinning weights are calculated automatically from the set of SCAPE sample poses, similar to [MG03].

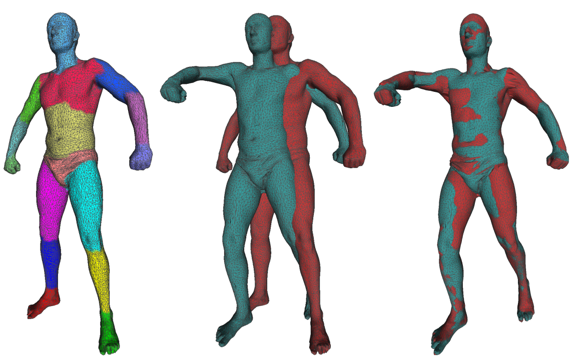

Figure 1: *Template in initial pose with colored body parts (left), target mesh in green superimposed onto template in red (middle), template model adapted to target model (right).*

Kinematic Template Fitting For preprocessing, we align the template with the target coarsely, by global rotation, scale and translation so that the principle axes and the covered regions on the dominant principle axis are aligned, followed by standard ICP. Kinematic ICP starts to insert the joints into the optimization loop in a hierarchical manner. During the first level of iterations, global similarity parameters and rotation parameters for the first set of innermost joints are optimized jointly. In following levels of iterations, further joints are added along the kinematic chain, until all joints are included. Note, that at each iteration each vertex influences the parameter adaptation of all joints along its kinematic chain up to the root body part (the abdomen in our case). The adaptation of global similarity alignment is always influenced by all vertex correspondences. To add joint parameters hierarchically, an additional loop is added around the base ICP.

In order to map the template vertices onto the target point cloud as close as possible, the ICP algorithm is modified to jointly adjust the joint angles besides refining the similar-

DOI: 10.2312/PE/VMV/VMV12/215-216

ity transformation parameters (totaling in $p = 52$ adjustable parameters in our case, 6 rigid, 1 scale and 45 kinematic parameters). For this purpose n correspondences are established. The objective function for minimization is modeled as sum of distances between all corresponding vertices. The template vertices, represented by the $3 \times n$ matrix $\mathbf{T}_i$, and corresponding target vertices $\mathbf{S}_i$ along each kinematic chain contribute distance costs d_i according to the kinematic chain equation (subscripts denote the number of joints from current to base body part):

$$\begin{aligned} d_0 &= ||\mathbf{S}_0 - s\mathbf{R}\mathbf{T}_0 - \mathbf{t}\mathbf{1}^T|| \\ d_1 &= ||\mathbf{S}_1 - \mathbf{R}(s\mathbf{R}_1\mathbf{T}_1 + (\mathbf{t}_1 - \mathbf{t})\mathbf{1}^T) - \mathbf{t}_1\mathbf{1}^T|| \\ &\dots \end{aligned}$$

with $\mathbf{1}$ being a n-vector of 1's, $\mathbf{R}$, $\mathbf{t}$, s being the similarity transformation parameters and $\mathbf{R}_1$ being the 3×3-rotation matrix for the first joint with rotation center $\mathbf{t}_1$. Introducing changes in parameters $\Delta[s\ \mathbf{t}\mathbf{R}\ \mathbf{R}_1 \dots]$ and assuming small Euler angle updates $\mathbf{r} = [r_x, r_y, r_z]^T$ in each iteration, the rotation (specified by $\mathbf{R}$) of a point $\mathbf{p}$ can be approximated by the skew-symmetric cross-product matrix $[\mathbf{r}]_\times$

$$\mathbf{R} \cdot \mathbf{p} \approx \mathbf{p} + [\mathbf{p}]_\times \cdot \mathbf{r} \text{ with } [\mathbf{p}]_\times = \begin{bmatrix} 0 & p_z & -p_y \\ -p_z & 0 & p_x \\ p_y & -p_x & 0 \end{bmatrix}$$

The upper equation system can be rearranged to yield the kinematic ICP update rule for all parameters of the form:

$$\Delta \left[s\ \mathbf{t}\ \mathbf{r}\ \mathbf{r_1}\ \dots \right]^T = (\mathbf{M}^T\mathbf{M})^{-1}(\mathbf{M}^T\mathbf{N})$$

with $\mathbf{M}$ being a $3n \times p$ matrix and $\mathbf{N}$ being a $3n$-vector. Each correspondence provides one row to $\mathbf{M}$, so that the $p \times p$ matrix $\mathbf{M}^T\mathbf{M}$ and the p-vector $\mathbf{M}^T\mathbf{N}$ can be setup efficiently by processing the correspondences in parallel.

Optimization of the Articulated Template Model In order to enable the template model to reflect realistic human surfaces as good as possible, the rotation centers of the joints are calculated from sample poses of the SCAPE dataset. Initially, the rotation centers are set to the mean of the cross section of neighboring body parts. After having calculated suitable joint rotations with them, we introduce $\Delta\mathbf{t}_i$'s and proceed the optimization with updates for each $\mathbf{t}_i$, now optimizing the template against all sample poses simultaneously, with one common set of rotation centers and separate sets of rotation matrices for the different target poses.

A further increase in accuracy is achieved by introducing Linear Blend Skinning techniques. Again, we use the sample poses of the SCAPE dataset to calculate optimal weights. In a linear least squares sense, we calculate one weight per vertex and per body part, which best reflects the influence of each body part movement onto this vertex.

Application 1: Pose Recovery A full 3D mesh is used as target and the template is adapted to optimally resemble the target, thereby providing a measurement of joint configurations. In order to reduce the negative influence of wrong correspondences, we weight the correspondences with the dot-product of their normals, respective we skip them, if the angle between the normals is bigger than 90°. In Figure 1 we show the result of fitting SCAPE model 0 to SCAPE model 2. It is clearly visible, that even with this strong difference in target and template mesh, the pose is recovered accurately. After having rigidly pre-aligned the template with the target mesh via PCA axes alignment and standard ICP, the mean distance between corresponding template and target vertices is reduced by kinematic ICP by a factor of 7 in 8 iterations.

Application 2: Pose Tracking from Partial Depth Maps Our method works also perfectly well for tracking with partial meshes, point clouds or depth maps. Here, we present pose tracking results where the template is fitted to depth maps created from multi-view analysis and represented by an unstructured (oriented) point cloud of the upper body, generated with the framework presented in [HE09]. In addition to the above mentioned normal weighting, it turns out to be crucial for this application to weight the correspondences with the reciprocal of their distance in order to reduce the negative influence of non-overlapping regions. The result of one frame in the middle of a sequence is shown in Figure 2. A rough global pre-alignment is not required, because the template configuration of the previous frame is available and suitable. Kinematic ICP reduces the initial mean distance of scan vertices to corresponding template vertices by a factor of 4.

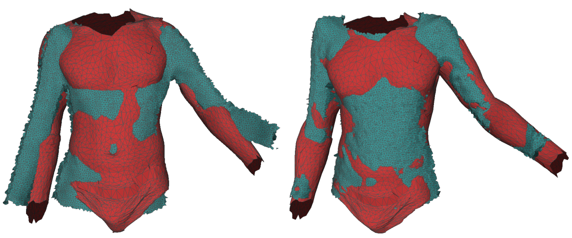

Figure 2: *Result of pose tracking from partial depth map, (left: input, right: output; red: template, green: 3D scan).*

Acknowledgments This research was partially supported by the European Commission under contract FP7-247688 3DLife and FP7-287723 REVERIE.

References

[ASK*05] ANGUELOV D., SRINIVASAN P., KOLLER D., THRUN S., RODGERS J., DAVIS J.: SCAPE: Shape Completion and Animation of People. In *SIGGRAPH* (2005). 1

[FE11] FECHTELER P., EISERT P.: Recovering Articulated Pose of 3D Point Clouds. In *CVMP* (2011). 1

[HE09] HILSMANN A., EISERT P.: Joint Estimation of Deformable Motion and Photometric Parameters in Single View Video. In *ICCV Workshop NORDIA 2009* (2009). 2

[MG03] MOHR A., GLEICHER M.: Building Efficient, Accurate Character Skins from Examples. In *SIGGRAPH* (2003). 1