Coregistration: Simultaneous Alignment and Modeling of Articulated 3D Shape

David A. Hirshberg   Matthew Loper   Eric Rachlin   Michael J. Black

MPI for Intelligent Systems, Perceiving Systems Department

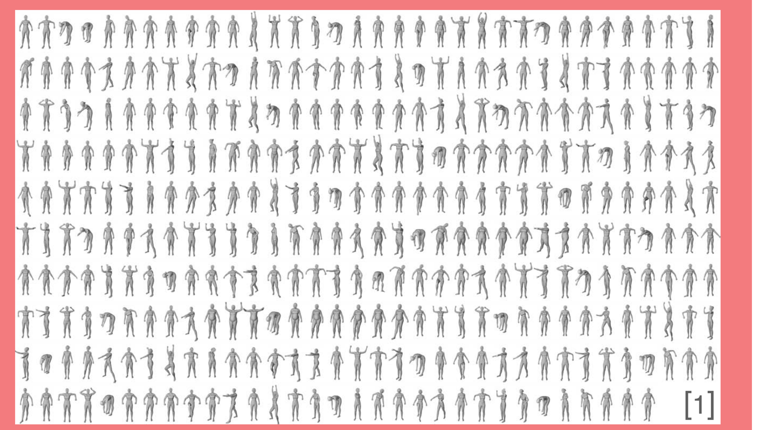
The Corpus Registration Problem

Goal: Align an entire corpus of body scans. Build a highly realistic 3D body model.

Problem: Chicken-and-egg

Model learning relies on scans that are accurately registered to a common 3D template.

Accurate registration is difficult without a good model of how the template is allowed to deform.



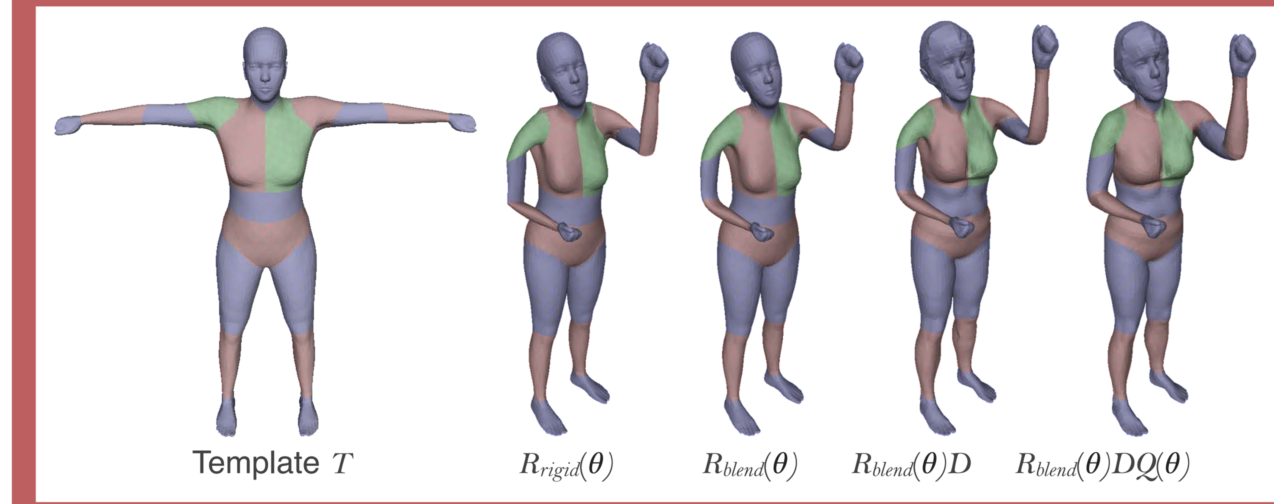
Solution: Coregistration simultaneously registers the corpus and builds the model.

SCAPE and BlendSCAPE

Models or articulated 3D Shape

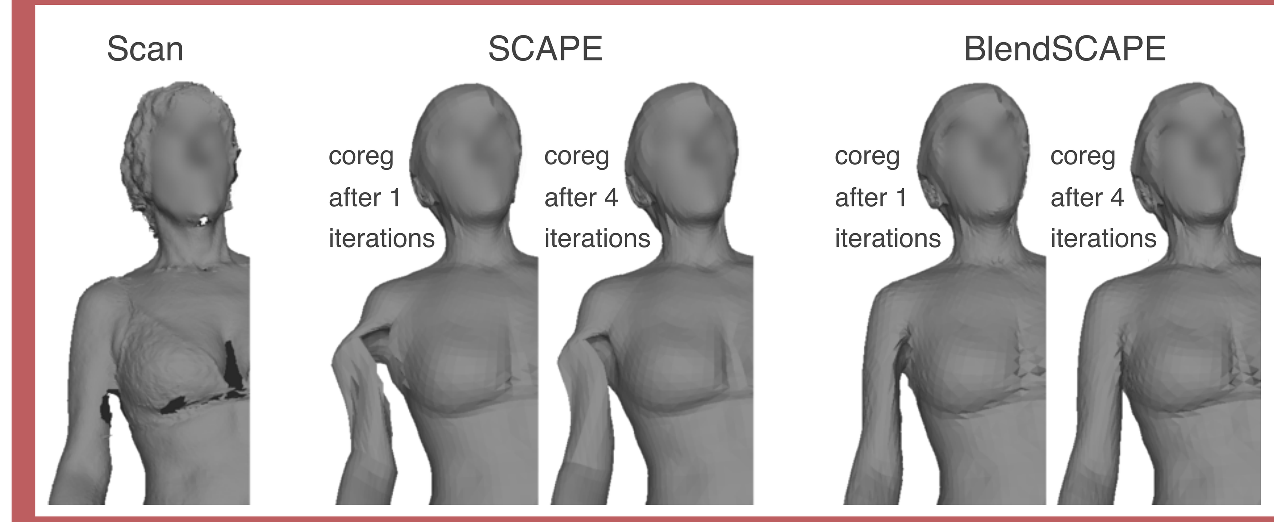
SCAPE [2] and BlendSCAPE realistically deform a segmented template mesh T to have pose θ and shape D.

Before each part of T is rotated by $R(\theta)$, D applies a per triangle shape deformation and $Q(\theta)$ applies a per triangle pose correction.



In SCAPE $R(\theta)$ applies the same rigid rotation to every triangle within each part. In BlendSCAPE, triangles are rotated by weighted averages of the part rotations.

BlendSCAPE smooths away artifacts, making $Q(\theta)$ easier to learn.



Approach

Coregistration approaches model learning and template registration simultaneously. The template is aligned to each scan while an articulated model is estimated to fit all alignments.

Challenge

Register a 3D template mesh to each scan. Each aligned template must closely match the surface of each scan. All alignments must also be in accurate point-to-point correspondence.

Goal

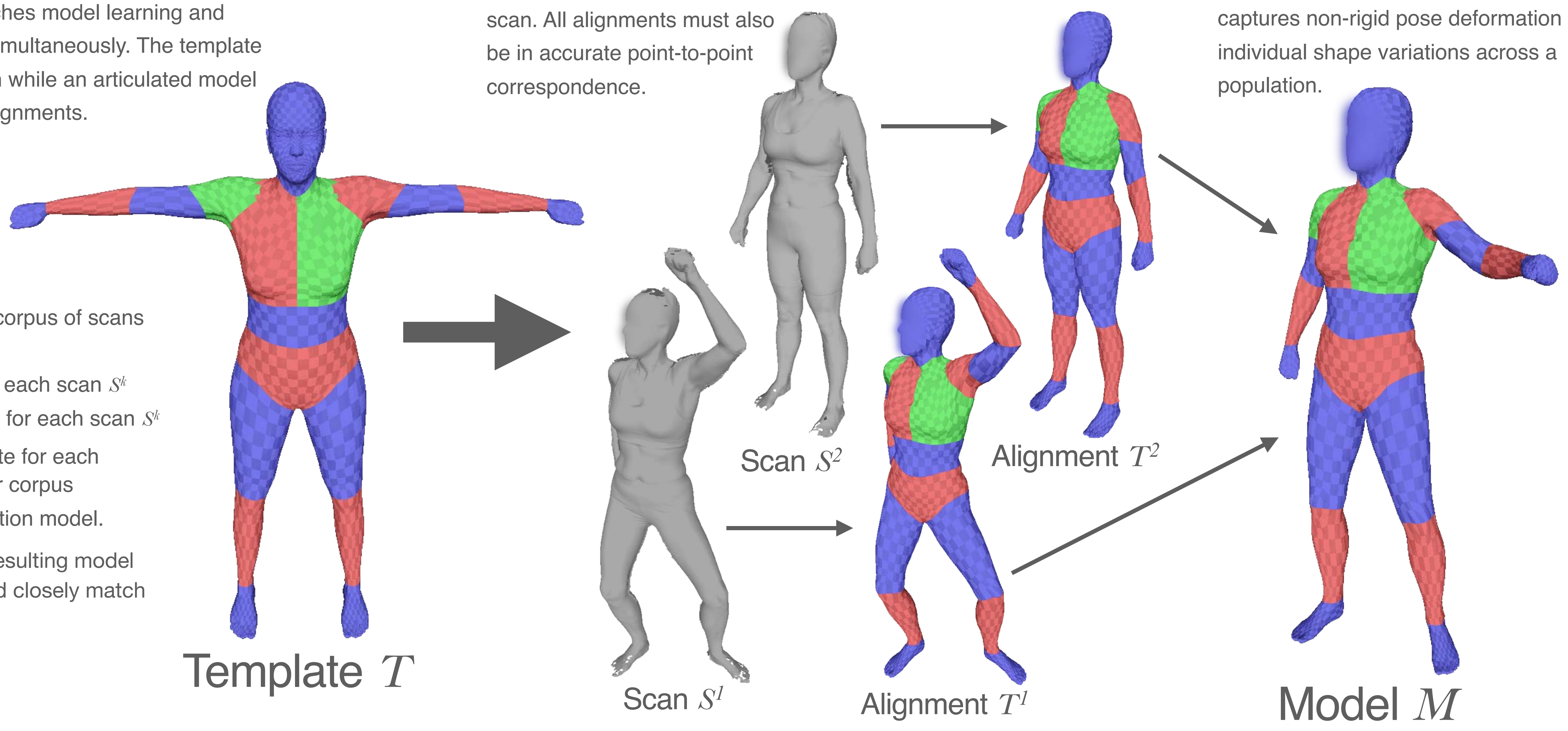
An accurate 3D model of the body that captures non-rigid pose deformation and individual shape variations across a population.

Outputs

When coregistering a corpus of scans we obtain:

- T^k : An alignment to each scan S^k
- θ^k : A pose estimate for each scan S^k
- D^{p_k} : A shape estimate for each person p in our corpus
- Q : A pose deformation model.

For each scan S^k the resulting model $M(\theta^k, D^{p_k}, Q(\theta^k))$ should closely match each alignment T^k.



The Optimization

By minimizing a single objective function, we reliably obtain high quality alignments to noisy scans while simultaneously learning a realistic articulated body model.

$$E = \sum_{S_k} E_k(T^k, D^{p_k}, Q; S^k)$$

$$E_k = \int_{x_s \in S^k} \rho\left(\min_{x_t \in T^k} ||x_s - x_t||\right) + \quad \text{Alignment-to-Scan Distance}$$

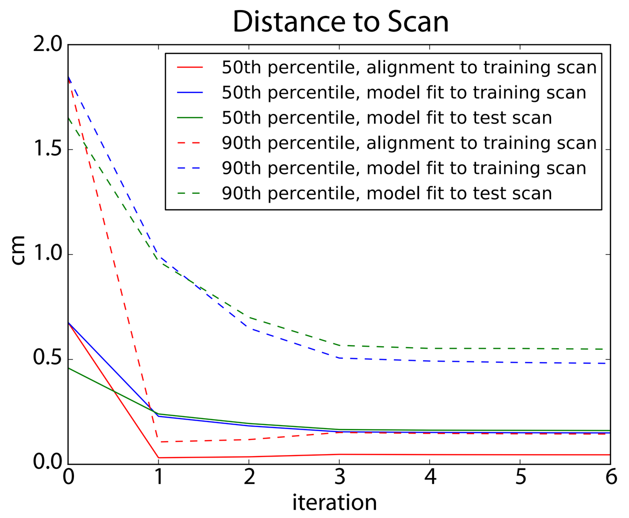
$$\sum_{f \in faces} \min_{\theta^k} ||T_f^k - M_f(\theta^k, D^{p_k}, Q(\theta^k))||_F^2 + \quad \text{Alignment-to-Model Distance}$$

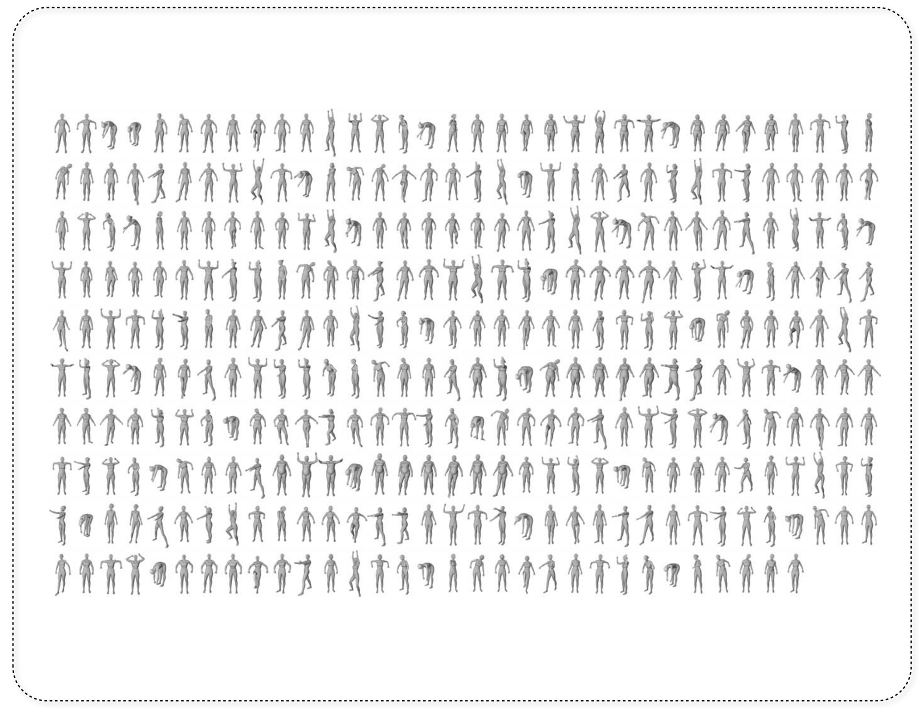
$$smoothness(D^{p_k}) + ||Q - I||_F^2 \quad \text{Model Regularization}$$

Coregistration alternates between updating each alignment T^k and pose θ^k, each person's body shape D^{p_k}, and the global pose deformation model $Q(\theta)$.

Results

We coregister a corpus of several hundred 3D scans featuring multiple individuals in a wide range of poses. Coregistration improves the model's ability to fit test and training data and yields visually accurate alignments.





References

[1] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, H.-P. Seidel, *A statistical model of human pose and body shape.* Computer Graphics Forum. 28:2 (2009) 337–346

[2] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, J. Davis. *SCAPE: Shape completion and animation of people.* ACM ToG. 24 (2005) 408–416

Scan	Alignment	Model
S^k	T^k	$M(\theta^k, D^{p_k}, Q(\theta^k))$