

# A Pose Invariant Statistical Shape Model for Human Bodies

Matteo COLAIANNI<sup>1</sup>, Michael ZOLLHÖFER<sup>1</sup>, Jochen SÜßMUTH<sup>2</sup>,  
Bettina SEIDER<sup>2</sup>, Günther GREINER<sup>1</sup>

<sup>1</sup> Computer Graphics Group, University Erlangen-Nürnberg, Germany

<sup>2</sup> Adidas Group, Germany

<http://dx.doi.org/10.15221/14.327>

## Abstract

We present a complete pipeline for constructing a statistical shape model that is invariant to deviations in the scan pose while encoding the space of human pose and body shape in an efficient manner. A dense cross-parameterization between a large set of high-quality 3D scans is computed using a fast and robust volume aware non-rigid registration method. Our approach uses a novel encoding that automatically decorrelates shape and pose leading to a statistical model that is oblivious under transformations induced by pose. This allows us to efficiently compensate pose variations in captured input data leading to a compact representation for pose as well as body shape. We present a local as well as a global skeletal encoding and compare both approaches. Finally, we analyze the generalization properties and accuracy of our approach against two state-of-the-art methods. We apply our model to the data clustering problem and use it as a prior for non-rigid shape matching.

**Keywords:** Human Body, Pose Invariance, 3D Statistical Shape Models, PCA, Morphable Model, Shape Aware, Non-Rigid Registration

## 1. Introduction

With the widespread availability of 3D scanning technology in the professional as well as the consumer level sector, the virtualization of real world objects gains more and more importance.

Capturing and analyzing the intra class variability of human bodies by means of a statistical shape model is the basis for a broad field of applications.

### 1.1 Motivation

In the industrial apparel creation process, the statistical analysis of the human body shape is an important tool to understand and learn a target groups' shape properties. This knowledge can be leveraged in the gradation process to find garment sizes that best cover the market. To tackle such design and analysis problems efficiently, human bodies can be mapped to a reduced subspace that still captures most of the variation in the dataset. Another application area is full body reconstruction where generalized body models are used as statistical priors for data completion and denoising of sparse and noisy data captured by consumer level RGB-D sensors. Both of these scenarios require an efficient and compact representation of the space of human pose and body shape.



*Fig. 1. 69 scanned participants: Despite being instructed to take a reference pose, the scans contain a high amount of pose variation (left). Our approach decouples pose and shape (right).*

\* [matteo.colaianni@cs.fau.de](mailto:matteo.colaianni@cs.fau.de); +49- 9131- 85-29923; [www9.cs.fau.de](http://www9.cs.fau.de)

## 1.2 Problem Statement

The basis for building a statistical model of human pose and body shape is a large database of 3D models with a common topology. Such a dataset can be obtained by means of 3D scanning and non-rigid registration to fit a common template model to all acquired input scans. The two biggest challenges in this pipeline are the acquisition of scans of different individuals in exactly the same pose and computing a consistent parameterization of the scans by the use of non-rigid registration. In practice, even if the participants are instructed to take on a specific pose, the variation in pose is enormous (see Figure 1, left). These differences in pose lead to a grave misinterpretation as the underlying representation can then not clearly distinguish between shape and pose variation. Learning a statistical model of the human body shape from such a dataset will introduce unwanted variance as the model will not only cover changes in the body shape but changes in the scanning pose as well.

## 1.3 Related Work

Modern 3D scanning technology allows to acquire a dense set of point samples on a person's body. To obtain a complete reconstruction, multiple views of a subject have to be acquired and transformed into a consistent coordinate system [1, 2, 3, 4, 5]. These registered scans have to be triangulated to obtain a mesh based representation. One possibility is to fit an implicit surface representation [6,7,8,9] followed by a surface extraction step [10].

Scans may contain large holes due to occlusions, these cannot be correctly closed using implicit methods. Therefore, Allen et al. [11] and Amberg et al. [12] solve this problem using a template based hole filling approach. They use a non-rigid registration method to fit a common template mesh to the captured raw input pointclouds. Besides using this prior to faithfully fill in holes, this template based technique also generates a consistent cross-parameterization between all models. Zollhöfer et al. [13] introduced a lattice based deformation paradigm as regularization term. The input geometry is embedded in a regular proxy lattice which serves as a deformation host. This deformation method allows to efficiently apply deformation on the template geometry.

Using the resulting cross-parameterization, statistical analysis can be applied to the dataset. The idea of constructing a statistical shape space for faces goes back to Blanz et al. [14]. They apply a dimensionality reduction technique (PCA) to compress the dataset by spanning a subspace using the directions of highest variance. Allen et al. [11] extended this approach to the space of human body shapes and used this representation for robust markerless tracking. Anguelov et al. [15] encodes pose as well as shape in a model called SCAPE. This idea has been extended by Hasler et al. [16] by scanning different persons also in different poses to sample a dense correlation between shape and pose in a unified model based on local deformation gradients. Instead of analysing the body shape directly they project the geometry in different bases invariant to certain classes of transformations. The problem of using shape spaces based on PCA the coefficients are not bound to semantically meaningful axes (e.g., breast circumference). This is unintuitive for editing purposes. Therefore, a lot of work has been focused on finding semantically meaningful axis (i.e, for changing body weight). Editing approaches based on high-level semantic constraints have been proposed by [17, 18]. In [11, 14, 16] approaches based on statistical regression are proposed to span more intuitive shape spaces. The authors of [14,16] compute semantic directions using linear regression, they are restricted to a single attribute, because the directions are not mutually orthogonal. Allen et al. [11] can directly handle the correlation between multiple attributes by solving for a set of ideal directions in the least squares sense. Non-linear regression has also been applied to the problem of facial aging [19]. Other applications besides semantic editing that build upon the prior information in statistical shape models range from reshaping of bodies in image and video data [20, 21] to the prediction of body shape of dressed humans [22, 23].

## 1.4 Contribution

This work clarifies the issue of differing poses when building a statistical model for human body shapes. A local skeletal encoding is proposed as a preprocessing step when building a statistical model. Three main contributions of the paper are

- A statistical shape model based on a skeletal encoding that is invariant to pose variation. Two alternative encodings for human body shapes are introduced.
- An analysis that clarifies the importance of a pose correcting step when creating generalized shape models. The method is compared against two state of the art approaches with respect to the generalization property.
- A reconstruction method for fitting scanned data using the presented encoding. Decoupling shape from pose yields to a more direct approach of solving the fitting.

## 2. Method

### 2.1 Registration of Scanned Data

To obtain dense correspondences between the input scans, we first fit a common template mesh to our database of body scans. Therefore, we use a non-rigid iterative closest points algorithm which uses the Lattice based As-Rigid-As-Possible (LARAP) deformation paradigm introduced by Zollhöfer et al. [13] as regularization term. The core idea of this algorithm is to embed the input geometry in a regular proxy lattice  $\mathcal{G}$  defined by the nodes  $l_i$  and edges  $e_{i,j}$  and to numerically simulate the deformation on this coarser structure. The vertices  $v_i$  are represented as a convex combination of the surrounding lattice nodes. Due to the regular structure of the proxy lattice, the numerical simulation is easy to implement and highly performant. Figure 3 (left) shows the registration process for one of the test subjects. The deformed lattice  $\mathcal{G}'$  should behave as-rigid-as-possible [24] and thus minimize the following regularization energy:

$$E_{reg}(\mathcal{G}') = \sum_i \sum_{j \in \mathcal{N}_i} \|(l'_i - l'_j) - \mathbf{R}_i(l_i - l_j)\|^2,$$

where  $j \in \mathcal{N}_i$  defines the lattice vertices  $l_j$  connected by an edge  $e_{i,j}$  with  $l_i$ . To drive the deformation, a target position  $t_i$  is specified for some vertices  $v'_i$  within a set of constraints  $\mathcal{C}$ :

$$v'_i = \sum_j \alpha_{i,j} l'_j = t_i.$$

Since this problem is generally overconstrained, Zollhöfer et al. [13] argue to solve for the target positions in a least squares sense as well yielding the following attraction term which minimizes the distance between the (deformed) mesh vertices and their target positions:

$$E_{cons}(\mathcal{G}') = \sum_{i \in \mathcal{C}} \|v'_i - t_i\|^2,$$

Putting both the attraction term and the regularization term together results in the overall registration energy:

$$E_{total}(\mathcal{G}') = \gamma E_{reg}(\mathcal{G}') + E_{cons}(\mathcal{G}').$$

The minimization of  $E_{total}$  is a non-linear least squares problem in the unknown rotations  $\mathbf{R}_i$  and the unknown positions of the lattice nodes  $l'_i$ .

We used this deformation paradigm as the basis to implement a fast and robust non-rigid ICP algorithm for unsupervised non-rigid surface registration. The correspondence search is implemented using a kd-tree to find closest points on the target shape as target constraints. Correspondences are pruned based on distance and normal deviation. Due to the volumetric structure of the proxy geometry, the volume of the input is implicitly preserved. We solve this overconstrained problem in every iteration of the ICP loop in the least squares sense. As proposed by Sorkine and Alexa [24], we minimize the energy by repeatedly solving first for the unknown rotations  $\mathbf{R}_i$  and then using the new rotations to compute better positions for the unknown lattice points  $l'_i$ . In a final step, the deformation of the input geometry can be automatically deduced using the proxy geometry:

$$v'_i = \sum_j \alpha_{i,j} l'_j.$$

The registration pipeline can be seen in Figure 2. Due to its simplicity and robustness, the LARAP deformation model is optimally suited to serve as regularization term in a non-rigid ICP framework. We demonstrate the robustness of our algorithm by fitting a common template mesh to a database of 3D body scans. For the examples in Figure 3 (right), registration takes roughly 5 seconds, where 3 seconds are spend on nearest neighbor search (all computations are performed on the CPU).

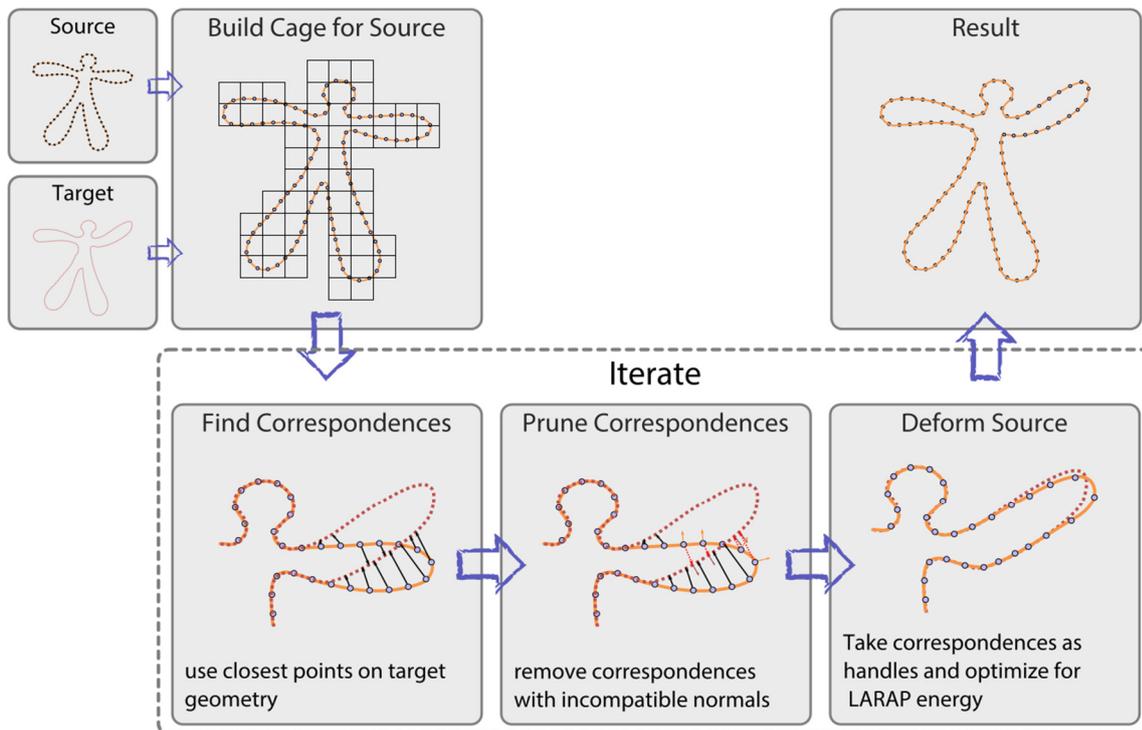


Fig. 2. The template fitting pipeline of the proposed algorithm.

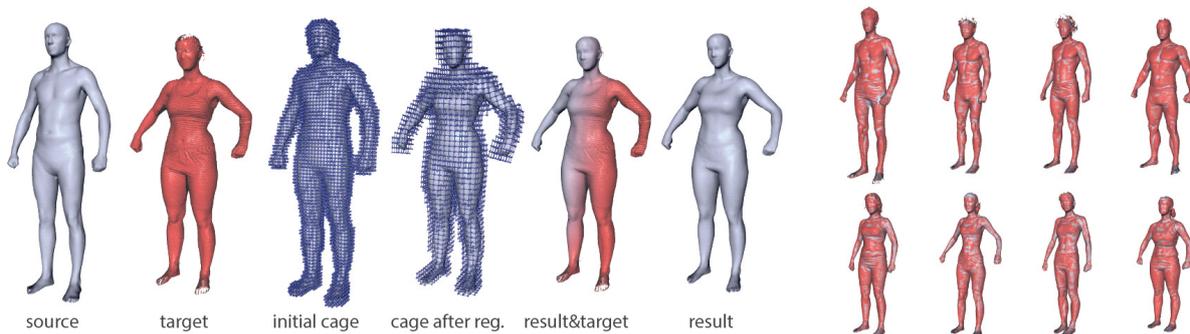


Fig 3. Fitting a template mesh to a target scan using the proposed non-rigid registration algorithm (left). Resulting registration results can be seen on the right side (target pointcloud (red) and deformed template model (gray)).

## 2.2 Vertex Based Morphable Model

The expressiveness of a statistical model heavily depends on the chosen encoding of the instances. The naive approach proposed by Allen and coworkers [11] represents each instance as a point in a  $3N$  dimensional space by stacking all  $N$  vertices  $(x_i, y_i, z_i)$  into a feature vector

$$\mathbf{x}_i = [x_0, y_0, z_0, \dots, x_{N-1}, y_{N-1}, z_{N-1}]^T$$

representing a single shape. The instance matrix for all  $k$  instances is given by:

$$\mathbf{X} = [\mathbf{x}_0, \dots, \mathbf{x}_{K-1}].$$

The space of plausible human body shapes is then defined as the subspace spanned by all instances. To compute this subspace, dimensionality reduction techniques can be applied. Principal Component Analysis (PCA) [11, 14] computes the directions of highest variance in the dataset by computing an eigenvalue decomposition of the covariance matrix  $COV$ :

$$COV = \mathbf{E}\mathbf{\Sigma}\mathbf{E}^T = (\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T.$$

The resulting set of eigenvectors  $\mathbf{E} = \{\mathbf{e}_0, \dots, \mathbf{e}_{N-1}\}$  spans the space of suitable shapes and can be used to synthesize new instances by a linear combination of  $M$  eigenvectors with scalar coefficients  $\chi_i$ :

$$\mathbf{x}_{syn} = \mu + \sum_{i=0}^{M-1} \chi_i \mathbf{e}_i.$$

### 2.3 Issues for not Normalized Input Data

Note that this pure position based encoding does not only span the space of plausible body shape variations. Instead a space is spanned where shape and pose are tightly correlated. This yields two unwanted effects: First, a significant part of the variation covered by a subset of the eigenvectors is actually used to account for variations in pose. Thus a smaller range of body shapes can be described compared to an encoding that decorrelates pose and shape, since part of the variation goes into the synthesis of the pose. Second, since Allen's approach [11] correlates body shape and pose, the pose of a person has an influence on the recovered shape when fitting the statistical model to a body scan as done in Section 4.2. In practice when projecting the same person in two different poses into such a basis, two different shapes will be generated. To decouple pose and body shape, we propose to use either a global (Section 2.4) or a local (Section 2.5) encoding based on an embedded bone hierarchy. Naturally, skeletal encodings lead to models that are pose invariant. Therefore, the shape representation is more compact.

### 2.4 Global Skeletal Aware Encoding

The first approach to encode skinned geometry is to globally deform each mesh to a common reference pose. Because of the consistent topology for all meshes, we are able to augment the template mesh with a simple common skeleton [25, 26, 27] (shown in Figure 4, left). The joint positions and the bones' local coordinate systems are represented as a linear combination of nearby vertices. As all used instances share the same topology after fitting the scan data, this method allows us to easily embed the skeleton into each given mesh irrespective of the scan pose (shown in Figure 4, right).

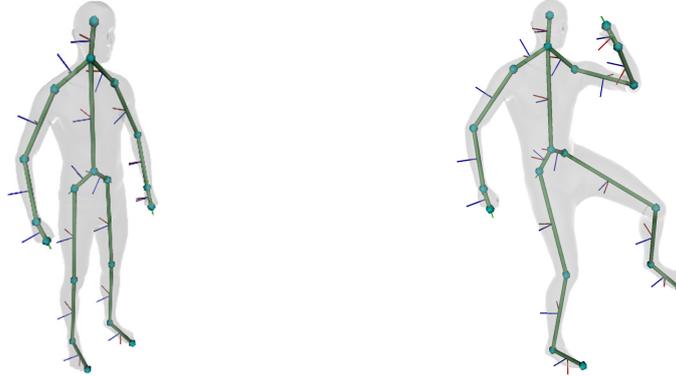


Fig 4. The skeleton is embedded into the mesh using the topology of the instance (left). The bone bases follow the pose deformation (right).

First, we transform each instance such that it matches a canonical target pose while preserving the original bone lengths. Given a skeleton as a set of  $L$  bone's transformations  $\mathcal{S} = \{\mathbf{M}_0, \dots, \mathbf{M}_{L-1}\}$ , a single vertex  $v_i$  can be expressed in the  $j$ -th bone's local space:

$$v_{i,j} = \mathbf{M}_j \cdot v_i.$$

A transformation to the target pose  $\tilde{\mathcal{S}}$  is achieved by:

$$\tilde{v}_i = \sum_{j=0}^{L-1} w_{i,j} \cdot \tilde{\mathbf{M}}_j^{-1} \cdot v_{i,j}.$$

The matrix  $\tilde{\mathbf{M}}_j^{-1} \in \tilde{\mathcal{S}}$  transforms the local vertex  $v_{i,j}$  back to global space.  $w_{i,j}$  is the skinning weight for a single vertex  $i$  with respect to a bone  $j$ . The resulting pose transformed vertices  $\tilde{v}_i$  of mesh  $k$  then are stacked to a feature vector  $\mathbf{x}_{k,norm}$  and all  $k$  instances are combined to an instance matrix  $\mathbf{X}_{norm}$ . Subsequently a model is built as described in Section 2.2. The decoding is straight forward by simply representing the mesh in its original pose while maintaining the bone lengths of the synthesized instance.

## 2.5 Skeletal Subspace Encoding

Motivated by a more local representation, a similar approach is to build the model directly on locally encoded vertices. The vertices are processed once per bone that has a non-zero influence on them. Each affected vertex is transformed into the bone's local space. We scale the corresponding feature dimension by the bone's skinning weight to ensure that the feature dimensions are properly scaled. A so called *bone vertex*  $x_{i,j}$  with global position  $v_i$  and attachment weight  $w_{i,j}$  is calculated by

$$x_{i,j} = w_{i,j} \cdot (\mathbf{M}_j \cdot v_i),$$

the rigid transformation  $\mathbf{M}_j$  maps a vertex into the  $j$ -th bone's local coordinate system. Subsequently, all encoded bone vertices for the  $k$ -th mesh are stacked in the shape vector  $\mathbf{x}_{k,norm}$ . All  $K$  meshes are encoded using this approach and stored in the instance matrix  $\mathbf{X}_{norm}$ . As in the global encoding we extract the subspace of plausible body shape variations using PCA. Given a target pose  $\mathcal{S}$  and the encoded body shape of an instance  $\mathbf{x}_i$  the final mesh can be reconstructed. First the length of the bones are computed. This differential property can be directly inferred from the locally encoded vertices. After re-assembling the skeleton from the pose and the reconstructed bone lengths, each bone vertex can be transformed back to world space:

$$\tilde{v}_i = \sum_j \tilde{\mathbf{M}}_j^{-1} \cdot x_{i,j}.$$

Decoding all instances (Figure 1, left) to a common target pose efficiently cancels out all pose variation (Figure 1, right) while preserving the shape variance.

## 3. Results

### 3.1 Data

The input data used for the evaluation is a set of 69 scanned males partially taken from Hasler's [28] data. The other part is captured by ourselves using a multiple RGB-D camera setup and a subsequent registration step. All persons are scanned in a relaxed default pose. The registration is performed automatically using the method described in Section 2.1. As result we obtain 69 topologically equivalent instances to proceed with. The skeleton including the local bone transformations is fitted to the topology of one instance using manually picked landmarks. Attaching the bones to the topology leads to a skeleton implicitly following deformations (compare Figure 4). The skinning weights are then generated using the method described in [29].

### 3.2 Generalization Property

The quality of the proposed model is tested with respect to its re-projection accuracy. We perform  $k$ -fold cross-validation by randomly splitting the dataset in  $k = 3$  subsets. Our dataset contains 69 scans of human bodies. We build our statistical shape model using  $k - 1$  subsets (46 instances) and project 1 subset (23 instances) in the model space. We measure the average reproduction error in world space for global (Section 2.4) as well as local encoding (Section 2.5) of the pose invariant model. The error simply is the mean per vertex distance between the projected input shapes and the reproduced geometries generated by the basis.

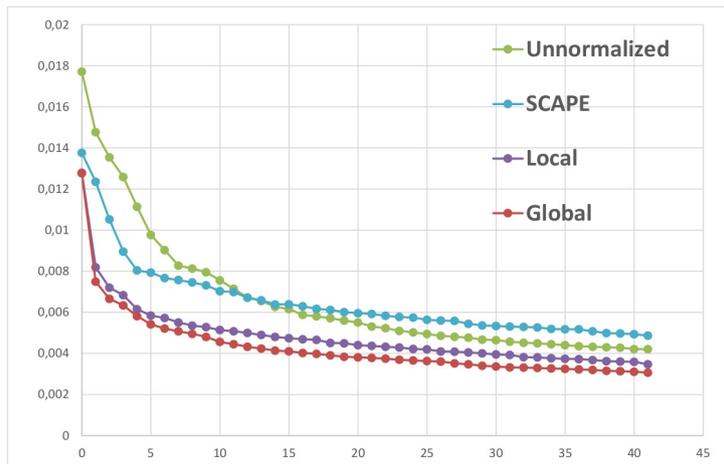


Fig. 5. Re-projection error (in mm per vertex) for an increasing number of used eigenvectors.

We compare the generalization properties with the vertex based statistical model [14] without pose correction and the SCAPE model [15]. The error is shown for an increasing number of used eigenvectors (Figure 5). The results demonstrate the importance of a pose invariant encoding. Both of the presented encodings are able to reconstruct a shape with ten eigenvectors while the vertex based model needs 25 and SCAPE needs almost 40 eigenvectors to reach the same reconstruction quality. In applications that know the exact skeleton beforehand a local encoding can be used to reach best generalization results. Even a pose-aware model like SCAPE is not able to generalize as good as our pose invariant method. The average projection error for the pose aware encoding model is visualized in Figure 6.

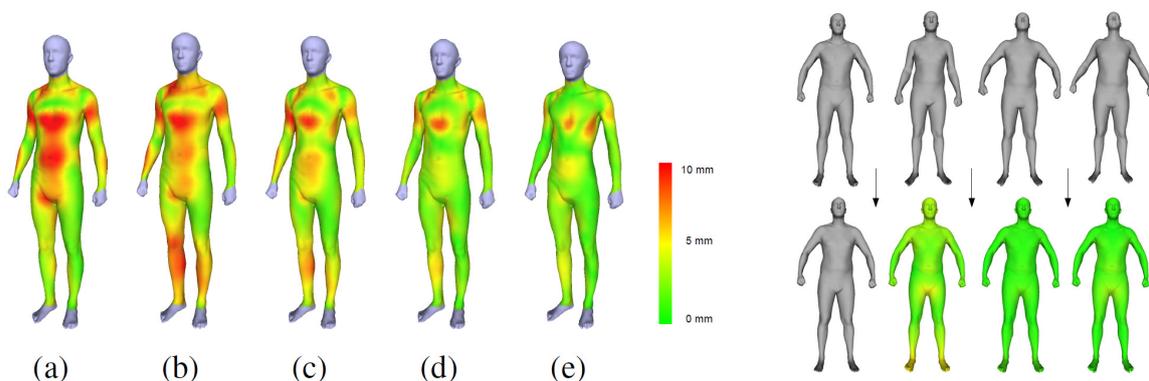


Fig. 6. Model reproduced with 1 (a), 5 (b), 10 (c), 20 (d) and 42 (e) eigenvectors. Average distance to test mesh is colored. Right: The same person in four different poses (first row) is reconstructed to one target pose (second row).

### 3.3 Variance Comparison and Deformation Error

The contained variance in the spanned statistical space shows the influence of the pose error introduced during data acquisition. The original data exhibit a comparably high pose variation (Figure 1, left). Extracting a subspace out of these data as proposed by Allen et al. [11] contains this variance as well. A pose correction step (Figure 1, right) leads to a much more compact subspace. We project both data sets (the normalized and the original one) into a common basis to compare them. The total variance of the not normalized shapes reach 15.56 while the pose normalized instances have a significantly reduced total variance of 7.27. Thus, roughly half of the variance in the input data does not cover changes in shape but changes in pose.

An important advantage of a model locally encoding shape (as proposed in Chapter 2.5) is the possibility to decode the shape to different similar poses. Because the encoding stores the vertices in local bone coordinates, we can efficiently decouple pose and shape by implicitly normalizing the

instances. We show the local encoding to be oblivious under pose transformations by scanning one person in different poses and then decoding the instances back to one common pose for a comparison in world space. In theory this produces identical meshes. In our experiments the measured average per vertex error is less than 1 mm (Figure 6, right). Thus, our model effectively eliminates little pose deviations without noticeable loss of information.

### 3.4 Limitation

Since the proposed algorithm uses linear blend skinning to account for pose variations, it suffers from the typical artifacts of this technique: collapsing elbows and the candy wrapper effects. Especially for large deformations, the results significantly depend on the bone weights, which usually have to be tweaked by hand. When we encode a body that is in a pose which significantly differs from our standard scanning pose, we implicitly propagate the aforementioned artifacts into the encoding. Projecting the encoded body shape into the space of plausible human body shapes eliminates the effects. However, when decoding the body in its initial pose, the artifacts become apparent again. To overcome these artifacts, one could utilize more elaborate skinning techniques such as dual quaternion skinning [30] or spherical blend skinning [31].

## 4. Applications

### 4.1 Clustering Instances for Measurement Extraction

A key step in apparel creation process is the gradation to find sizes that represent a population well. Statistical shape models can be used to accelerate this tedious and time consuming step by synthesizing sizes according to the statistical distribution. The task has high demands on both accuracy and precision. This scenario clarifies the need for a shape model minimizing the pose error. In case the assumption of equally posed instances is not fulfilled, errors caused by pose deviation are propagated to the synthesized shapes. Our model meets these requirements well since it considers pose variation. We clustered the available instances into  $k$  sets by sampling the first eigenvector equidistantly. Subsequently, the corresponding cluster representatives are synthesized. The generated meshes for the cluster centers are shown in Figure 7.

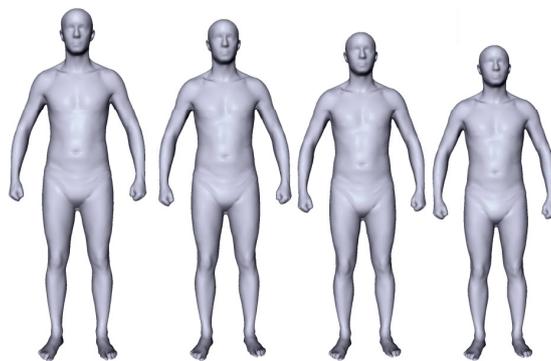


Fig. 7. The reconstructed meshes for the cluster centers.

### 4.2 Reconstructing Scanned Data

Another application is to use the proposed statistical model as a regularizer for non-rigid registration. The advantage of this method is the ability to solve for pose and shape independently, but within the same system. The optimization solves a non-linear least squares problem for all joint positions  $\alpha$  and all model coefficients  $\beta$ . The parameter vector  $\gamma = [\{\alpha\}, \{\beta\}]$  stacks the unknowns. Solving

$$\hat{\gamma} = \operatorname{argmin}_{\mathbf{x}} \sum_{i=0}^{N-1} \|r_i(\gamma) - t_i\|^2 + \sum_i^M \|\beta_i\|^2.$$

where  $r_i$  are the reconstructed vertices for the model parameters  $\gamma$  and the target correspondences  $t_i$ , yields the parameters  $\hat{\gamma}$  for the best fitting reconstruction in shape and pose. The correspondences between the template mesh and the target points are updated in an ICP like manner. The process of fitting the model to a high resolution point cloud is shown in Figures 8 a-d. We also fit our model to low resolution data captured by a consumer level RGB-D sensor setup (six cameras, shown in Figures 8 e-h).

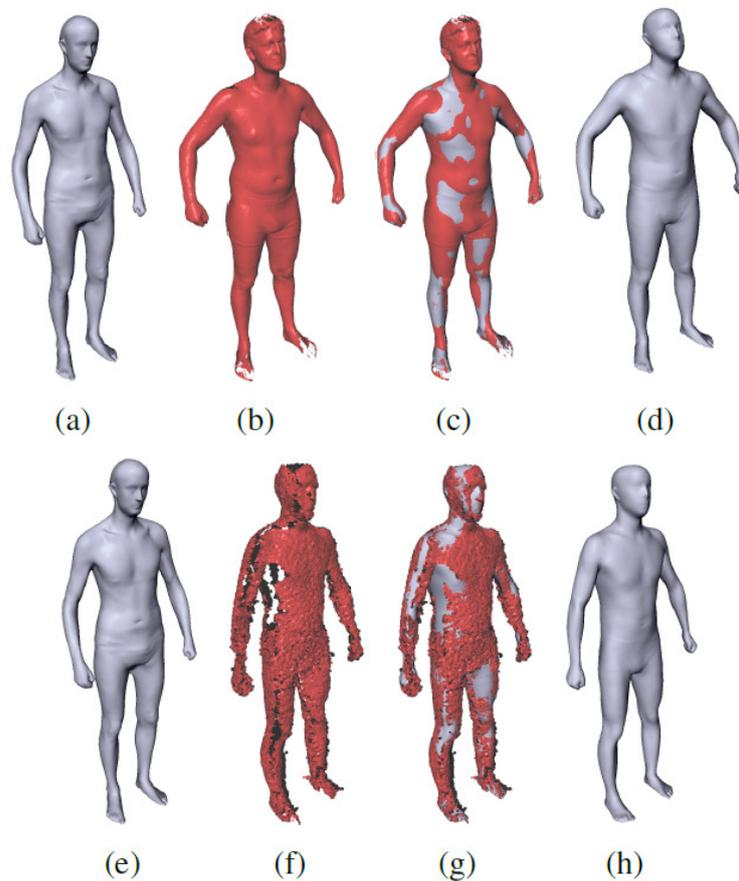


Fig. 8. A template (a) is registered the scan data (b,c). The result is shown in (d). The second row shows the registration for a low resolution target (e-h).

## 5. Conclusion

We introduced a new conceptually simple statistical model for human bodies and evaluates two alternative ways to encode the human body base. Due to its property of being invariant to pose errors, our method performs significantly better than previous methods in terms of generalization. This allows us to reconstruct a huge variety of shapes by using only a few degrees of freedom. This makes the model perfectly suited for reconstructing plausible body shapes from sparse and noisy scan data.

## References

- [1] Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(2):239–256, February 1992.
- [2] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the ICP algorithm. In *Proceedings of 3DIM'01*, pages 145–152, 2001.
- [3] Szymon Rusinkiewicz, Olaf Hall-Holt, and Marc Levoy. Real-time 3d model acquisition. *ACM Trans. Graph.*, 21(3):438–446, July 2002.
- [4] Kari Pulli. Multiview registration for large data sets. In *Proceedings 3DIM'99*, pages 160–168, 1999.
- [5] Yang Chen and Gérard Medioni. Object modelling by registration of multiple range images. *Image Vision Comput.*, 10(3):145–155, 1992.
- [6] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of SIGGRAPH '96*, pages 303–312. ACM, 1996.
- [7] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. Mc-Callum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of SIGGRAPH' 01*, pages 67–76, 2001.

- [8] Greg Turk and James F. O'Brien. Variational implicit surfaces. Technical report, 1999.
- [9] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. *SIGGRAPH Comput. Graph.*, 26(2):71–78, July 1992.
- [10] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Proceedings of SIGGRAPH '87*, pages 163–169, 1987.
- [11] Brett Allen, Brian Curless, and Zoran Popović. The space of human body shapes: Reconstruction and parameterization from range scans. *ACM Trans. Graph.*, 22(3):587–594, July 2003.
- [12] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. In *Proceedings of CVPR'07*, pages 1–8, 2007.
- [13] Michael Zollhöfer, Ezgi Sert, Günther Greiner, and Jochen Süßmuth. Gpu based arap deformation using volumetric lattices. In *Eurographics (Short Papers) 2012*, pages 85–88, 2012.
- [14] Volker Blanz and Thomas Vetter. A morphable model for the synthesis of 3d faces. In *Proceedings of SIGGRAPH '99*, pages 187–194, 1999.
- [15] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: Shape completion and animation of people. In *Proceedings of SIGGRAPH'05*, pages 408–416, 2005.
- [16] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H.-P. Seidel. A statistical model of human pose and body shape. In *Computer Graphics Forum (Proc. Eurographics 2008)*, volume 2, 2009.
- [17] An example-based approach to human body manipulation. *Graphical Models*, 66(1):1 – 23, 2004.
- [18] Mustafa Kasap and Nadia Magnenat-Thalmann. Parameterized human body model for real-time applications. In *Proceedings of the CW'07, CW '07*, pages 160–167, 2007.
- [19] Kristina Scherbaum, Martin Sunkel, Hans-Peter Seidel, and Volker Blanz. Prediction of individual non-linear aging trajectories of faces. *Comput. Graph. Forum*, 26(3):285–294, 2007.
- [20] Arjun Jain, Thorsten Thormählen, Hans-Peter Seidel, and Christian Theobalt. Moviereshape: Tracking and reshaping of humans in videos. *ACM Trans. Graph.*, 29(6):148:1–148:10, December 2010.
- [21] Shizhe Zhou, Hongbo Fu, Ligang Liu, Daniel Cohen-Or, and Xiaoguang Han. Parametric reshaping of human bodies in images. In *ACM SIGGRAPH 2010 Papers*, pages 126:1–126:10, 2010.
- [22] Alexandru O. Bălan and Michael J. Black. The naked truth: Estimating body shape under clothing. In *Proceedings of ECCV'08*, pages 15–29, 2008.
- [23] Niels Hasler, Carsten Stoll, Bodo Rosenhahn, Thorsten Thormählen, and Hans-Peter Seidel. Estimating body shape of dressed humans. *Computers & Graphics*, 33(3):211–216, 2009.
- [24] Olga Sorkine and Marc Alexa. As-rigid-as-possible surface modeling. In *Proceedings of SGP'07*, pages 109–116, 2007.
- [25] Xiaohuan Corina Wang and Cary Phillips. Multi-weight enveloping: Least-squares approximation techniques for skin animation. In *Proceedings of the SCA'02*, pages 129–138, 2002.
- [26] J. P. Lewis, Matt Cordner, and Nickson Fong. Pose space deformation: A unified approach to shape interpolation and skeleton-driven deformation. In *Proceedings of SIGGRAPH'00*, pages 165–172, 2000.
- [27] Nadia Magnenat-Thalmann and Daniel Thalmann. Making them move. Chapter *Human Body Deformations Using Jointdependent Local Operators and Finiteelement Theory*, pages 243–262. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1991.
- [28] Niels Hasler. <http://www.mpiinf.mpg.de/resources/scandb>, 2014.
- [29] Ilya Baran and Jovan Popović. Automatic rigging and animation of 3d characters. *ACM Trans. Graph.*, 26(3), July 2007.
- [30] Ladislav Kavan, Steven Collins, Jiří Žára, and Carol O'Sullivan. Geometric skinning with approximate dual quaternion blending. *ACM Trans. Graph.*, 27(4):105:1–105:23, November 2008.
- [31] Ladislav Kavan and Jiří Žára. Spherical blend skinning: A real-time deformation of articulated models. In *Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games*, pages 9–16, 2005.