

Reparameterising 3D Statistical Shape Models

Haoyang Wang, Stefanos Zafeiriou

Department of Computing, Imperial College London, UK

Abstract—3D statistical shape models are widely used in modelling 3D shapes such as human faces and bodies. The limitation of such model is that, once built, the model can only represent 3D shape instances of a fixed mesh topology. While some applications may require a shape model of a different mesh topology, the model building pipeline has to be repeated with the new template, which could be time and computational resource consuming. In other cases only the statistical model is available and access to the original data is not possible. In this paper, we present a method to reparameterise a given 3D statistical shape model to any topology without using any training data. We also show that the reparameterised model achieves comparable performance as the original model.

I. INTRODUCTION

3D statistical shape models are widely used for modelling human faces [1][2], bodies [3][4][5] as well as objects such as human bones and organs [6]. The idea of building statistical shape model is to perform PCA on a set of registrations aligned to a pre-defined template, the shape space is then parameterised by the principal components. Once a statistical shape model is trained, the topology of the instances in the model space is then fixed. However, in many applications, a model with a different topology might be useful. For example, in [7], in order to reconstruct the 3D structures of human faces in 2D images, an accurate high resolution per-vertex texture model is needed. In [8], the pose dependent shape variations are modelled with a matrix of dimension $3N \times 9n$, where N is the number of vertices in the registration and n is the number of joints in the skeleton structure of the template. For these applications, one may prefer to reuse a model instead of repeating the model training pipeline with another template, since it could be quite demanding in terms of computational time. In other cases it would not even be possible because access to the original training data is prohibited.

Motivated by the aforementioned problems, we propose a simple, yet very efficient, method of reparameterising statistical shape models given a new template of a different topology. Our method is based on the probabilistic nature of statistical shape models. Given a model and a new template, we solve for a covariance matrix for the new model directly without using or generating any training data. We present the formulation and solution of the problem in section III. In section IV, we present both qualitative and quantitative evaluation of our method. Finally in section V, we summarise our contribution and provide ideas for future work.

II. RELATED WORK

The original formulation of 3D Morphable Models (3DMM) was proposed by [1], where they construct 3D face shape models by performing Principal Component Analysis on a set of training face meshes in full correspondence. 3DMMs have since been widely applied to human face modelling. While [2] and [9] showed the linear PCA based 3DMMs can capture the identity dependent variations in human faces, such as gender, ethnicity and age, the work of [7] and [10] also modelled the expression variations. The success of 3DMM is due to the fact that the identity and expression variations in human faces can be well approximated by a hyper-planar manifold in Euclidean space [11], the face shape space can then be parameterised by the axes of the manifold where the captured variance is maximised. The same assumption holds for human body shape variations related to identity [3][4][5][8].

For nonlinear variations such as human body pose changes, the most commonly used approach is blend skinning, where each vertex in the pre-defined template is transformed as a function of the neighbouring bones. In this case the model is parameterised by the parameters of the blend skinning function. We refer the reader to [8] for a detailed review on blend skinning methods. [10] also modelled variations such as jaw movement in faces with blend skinning. [12] used blend skinning method to model the pose changes of hands.

We aim to reparameterise 3D statistical models. For PCA based models, our goal is to reparameterise the principal components. For blend skinning based models, the goal is to reparameterise the per-vertex parameters of the blend skinning function. To the best of our knowledge, this is the first work on reparameterising statistical shape models.

III. METHODOLOGIES

A. Problem Formulation

Assuming a set of D aligned 3D meshes are given, where each mesh \mathbf{m}_i with N vertices is represented by an $3N \times 1$ vector $[x_1, y_1, z_1, \dots, x_N, y_N, z_N]^T$. From the probabilistic view, a PCA-based statistical shape model assumes the shape variations can be modelled with a normal distribution [13]:

$$\mathbf{u} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (1)$$

where the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ are computed from the data as:

$$\boldsymbol{\mu} = \frac{1}{D} \sum_{i=1}^D \mathbf{u}_i \quad (2)$$

$$\Sigma = \frac{1}{D-1} \sum_{i=1}^D (\mathbf{u}_i - \boldsymbol{\mu})(\mathbf{u}_i - \boldsymbol{\mu})^T \quad (3)$$

Given an N_1 -vertexed model $M_1: \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1)$ and a new template \mathbf{m}_2 represented as $\boldsymbol{\mu}_2$ with N_2 vertices, the problem of reparameterising M_1 to be used with \mathbf{m}_2 can be considered as an optimisation problem, where the covariance matrix Σ_2 which minimises the difference between the old model M_1 and the new model M_2 need to be solved. With the probabilistic formulation of PCA-based shape models, we solve for Σ_2 by minimising the Kullback-Leibler divergence between M_1 and M_2 :

$$\Sigma_2 = \underset{\Sigma}{\operatorname{argmin}} D_{KL}(\mathbf{W}\mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1) || \mathcal{N}(\boldsymbol{\mu}_2, \Sigma)) \quad (4)$$

where \mathbf{W} is a $3N_2 \times 3N_1$ matrix which maps M_1 to the same $3N_2$ -dimension as $\boldsymbol{\mu}_2$.

By setting the derivative of Eq 4 w.r.t. Σ to 0, the new covariance matrix can then be obtained:

$$\Sigma_2 = \mathbf{W}\Sigma_1\mathbf{W}^T + (\boldsymbol{\mu}_2 - \mathbf{W}\boldsymbol{\mu}_1)(\boldsymbol{\mu}_2 - \mathbf{W}\boldsymbol{\mu}_1)^T \quad (5)$$

B. Computing the \mathbf{W} matrix

The performance of the new model depends on the choice of the matrix \mathbf{W} . Ideally, the new model M_2 can model the exact distribution of $\mathbf{W}M_1$ if $\mathbf{W}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$, therefore we want $\mathbf{W}\boldsymbol{\mu}_1$ to be as close to $\boldsymbol{\mu}_2$ as possible.

Assuming the new template \mathbf{m}_2 has been non-rigidly aligned to the template \mathbf{m}_1 of M_1 . The vertex to surface correspondences from \mathbf{m}_2 to \mathbf{m}_1 can be computed. We can then define \mathbf{W} as the matrix which maps the N_1 vertices of \mathbf{m}_1 to the N_2 points on its surface corresponding to the vertices of \mathbf{m}_2 , in this way the distance between $\mathbf{W}\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ is minimised while preserving the surface property of \mathbf{m}_1 .

Such matrix \mathbf{W} can be arranged as a sparse matrix of $N_2 \times N_1$ blocks, where each block is a 3×3 submatrix. Suppose the closest point of vertex $\mathbf{v}_i \in \mathbf{m}_2$ on the surface of \mathbf{m}_1 is $(\mathbf{t}_i, \mathbf{x}_i)$, where $\mathbf{t}_i = (a, b, c)$ is the index of the corresponding triangle face defined by three point with index a , b and c , and $\mathbf{x}_i = (u_i, v_i, w_i)$ is the barycentric coordinate of this point within triangle \mathbf{t}_i . We then set block (i, a) , (i, b) and (i, c) of matrix \mathbf{W} as follows:

$$\begin{array}{c} \begin{array}{ccc} a & b & c \\ \dots & \dots & \dots \\ \dots & u_i & 0 & 0 & \dots & v_i & 0 & 0 & \dots & w_i & 0 & 0 & \dots \\ \dots & 0 & u_i & 0 & \dots & 0 & v_i & 0 & \dots & 0 & w_i & 0 & \dots \\ \dots & 0 & 0 & u_i & \dots & 0 & 0 & v_i & \dots & 0 & 0 & w_i & \dots \\ \dots & \dots \end{array} \end{array}$$

Setting the blocks in such way for all N_2 vertices gives us the final \mathbf{W} matrix, where the rest of the entries remain zero.

C. Blend Skinning Models

For human bodies, PCA based shape models are commonly used to model only the identity dependent shape

changes, while the pose changes are modelled with blend skinning methods. Here we consider the formulation proposed by [8], in which they used linear blend skinning method to model the pose changes of human bodies, as well as pose blend shape to model the pose-dependent shape changes. We use the matrix \mathbf{W} to transform the vertex based parameters in their model such as the blend skinning weights and pose blend shapes. The pose blend shape models the pose dependent deformations of the N vertices in the template while the n joints rotate. The pose blend shape is linear with respect to the joint rotation matrices, therefore it is represented by a $3N \times 9n$ matrix. We compute the pose blend shape of our new model as:

$$\mathbf{P}_2 = \mathbf{W}\mathbf{P}_1 \quad (6)$$

The blend skinning weights define how much the transformation of the n joints affect the transformation of the N vertices. It is represented by an $N \times n$ matrix. We rearrange our \mathbf{W} matrix as a $N_2 \times N_1$ matrix \mathbf{W}' . For each pair of vertex to surface correspondence \mathbf{v}_i and $((a, b, c), (u_i, v_i, w_i))$, we set the entries (i, a) , (i, b) , and (i, c) of \mathbf{W}' as:

$$i \begin{array}{ccc} & a & b & c \\ \left[\begin{array}{ccc} \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & u_i & \dots & v_i & \dots & w_i & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{array} \right]$$

We then compute the new blend skinning weights as:

$$\mathbf{B}_2 = \mathbf{W}'\mathbf{B}_1 \quad (7)$$

IV. EVALUATION

We evaluate the proposed method with statistical models of faces and bodies. For faces, we build a model of 8082 vertices with 2500 randomly selected scans from the MeIn3D [2] dataset following the pipeline proposed in the original paper. We aim to reparameterise the model to have 15176 vertices. 2000 registrations were used for training, and 500 registrations were also aligned to the new template for testing. For bodies, we train a shape model of 6890 vertices with 1657 randomly selected meshes from the MPII Human Shape dataset [14] following the pipeline proposed in [8]. More specifically, we align the meshes to the template used by [8] with non-rigid icp [15], then normalise the registrations to T-pose before performing PCA to build the body model. We then reparameterise the model to have 10412 vertices, and test the models with 500 registrations.

A. PCA-based Shape Models

We first visually inspect the quality of the new model by comparing its first few principal components with the components of the original model. In figure 1 we visualise the first three principal components of the original and new models from -3 to +3 standard deviation. We observe that each principal component of the new models retain similar variations as in the original models.

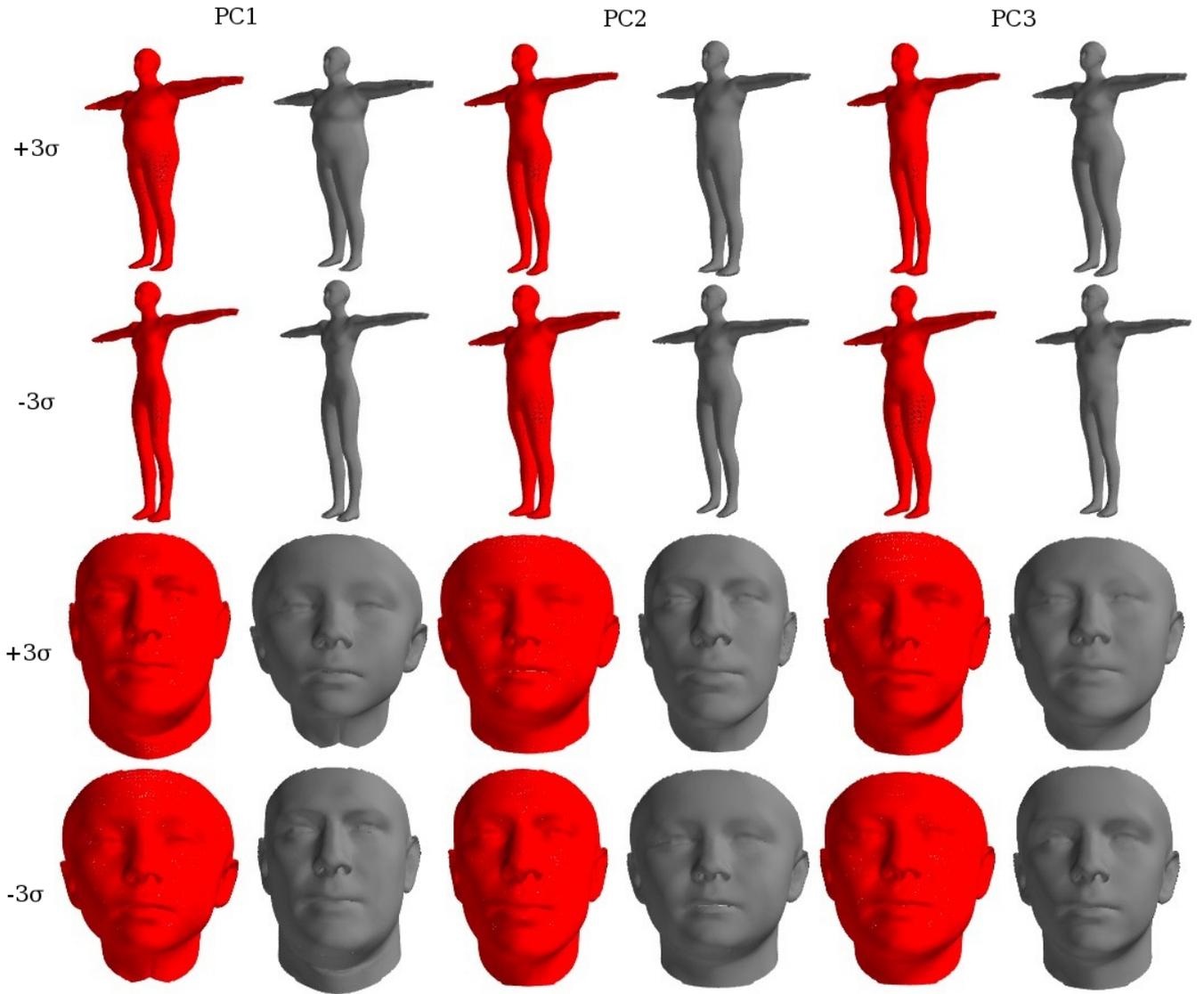


Fig. 1. Visualisation of the first three principal components at -3 and $+3$ standard deviations. The red meshes indicate the original model, while the gray meshes indicate the new model.

We then compare the original model and the new model in terms of their compactness, generalisation, and specificity as proposed by [16]. Compactness is the percentage of variance in the training data that is explained by the model. Figure 2 shows the variance retained by the original and new model while a certain number of principal components are kept. For both the body and face models, the principal components of the new model are able to explain nearly the same percentage of variance as the original model. Therefore we consider the new model to be as compact as the original model.

Generalisation measures the model’s ability of generalising to new instances that are unseen during the training. We project each of the models to the test set, and compute the mean of the average per-vertex euclidean distance between the model reconstruction and the testing instances to obtain the generalisation error for each model. Note that during training, we scale our templates to fit inside box of diagonal

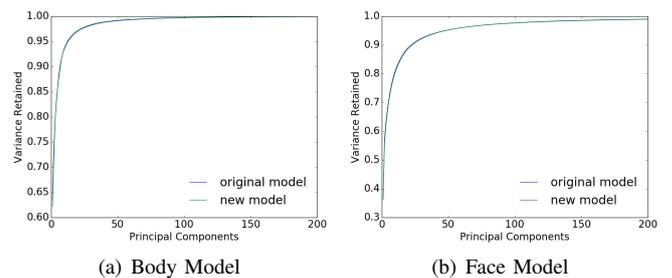


Fig. 2. Compactness of the Original and Reparameterised Models.

1. Thus the generalisation and specificity error are measured at this scale. For both the body and face models, the difference between the generalisation error of the new and original model is smaller than 0.001. Therefore we believe that the new models computed with our method can achieve comparable generalisation ability as the original models.

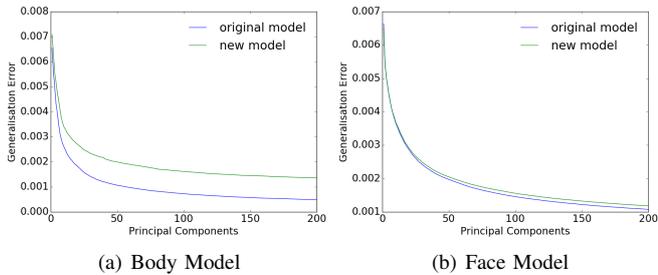


Fig. 3. Generalisation Error of the Original and Reparameterised Models.

Specificity evaluates the validity of instances generated by the model. For each model, we randomly generate 1000 samples from the model, and compute the averaged per-vertex euclidean distance between each instance and its nearest neighbour in the test set. We then average this error over all 1000 samples as the specificity error of this model. The specificity errors for the face and body models are plotted in figure 4. We observe that for the face model, the new model computed with our method have smaller specificity errors. And for the body model, the new model have specificity error larger than the original model, but the difference between the specificity error is smaller than 0.001. Therefore we believe that the new models can generate similar random instances as the original models.

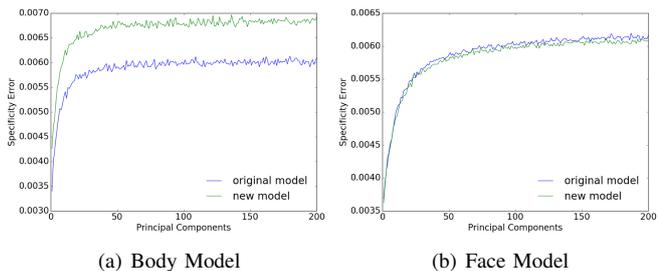


Fig. 4. Specificity Error of the Original and Reparameterised Models.

B. Blend Skinning Models

We animate the body models with poses from the UP-3D dataset [17] to inspect the quality of the new blend skinning parameters. Some examples of the animated meshes are presented in figure 5. We observe that the animated meshes have smooth surface, which suggests the reparameterised blend skinning weights are able to represent the relations between each vertex in the new template and the joints correctly. Also the pose blend shapes are able to preserve the muscle bulging details given by the original model and the bending artefact around the joints resulted from the linear blend skinning method is corrected accordingly.

V. CONCLUSION

In this paper we have presented a method to reparameterise 3D statistical shape models given a new template. We have computed a transformation matrix to transform the model template that we wish to reparameterise to be at the same

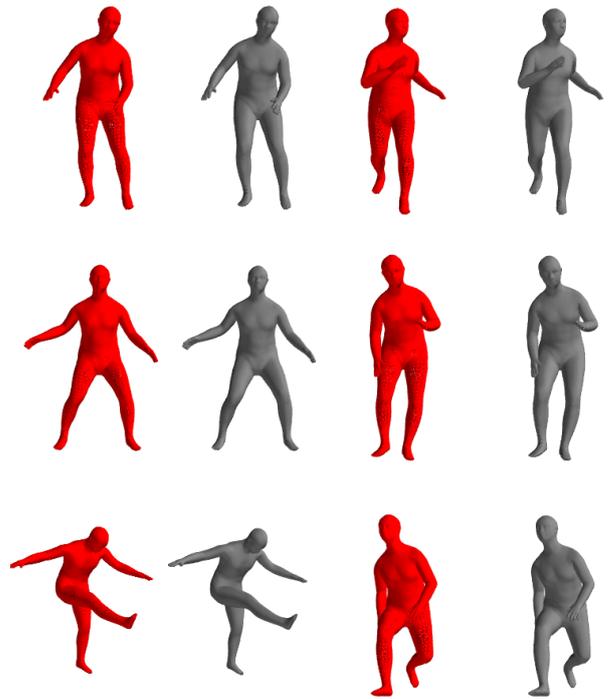


Fig. 5. Visualisation of the animated blend skinning models. Red meshes: animation result of the original model with 6890 vertices. Gray meshes: animation result of the reparameterised model with 10412 vertices.

dimension as the new template. For PCA based models, we then computed a new covariance matrix by minimising the KL-divergence between the original model and the reparameterised model. For blend skinning based models, we have transformed the per-vertex parameters with our transformation matrix. We have showed that the reparameterised models are as compact as the original model, and generalise equally well while being more specific than the original model. We have also demonstrated that the reparameterised blend skinning models are able to preserve the surface details of the original model.

In future work, we will explore further the performance of our method by applying the reparameterised models to mesh registration, 3D reconstruction of 2D images, and dense shape regression problems.

VI. ACKNOWLEDGEMENTS

Dr. Zafeiriou acknowledges support from a Google Faculty award, as well as from the EPSRC fellowship Deform (EP/S010203/1).

REFERENCES

- [1] Volker Blanz and Thomas Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Transactions on pattern analysis and machine intelligence*, 25(9):1063–1074, 2003.
- [2] James Booth, Anastasios Roussos, Allan Ponniah, David Dunaway, and Stefanos Zafeiriou. Large scale 3d morphable models. *International Journal of Computer Vision*, 126(2-4):233–254, 2018.
- [3] Brett Allen, Brian Curless, and Zoran Popović. The space of human body shapes: reconstruction and parameterization from range scans. In *ACM transactions on graphics (TOG)*, volume 22, pages 587–594. ACM, 2003.

- [4] Hyewon Seo, Young In Yeo, and Kwangyun Wohn. 3d body reconstruction from photos based on range scan. In *International Conference on Technologies for E-Learning and Digital Entertainment*, pages 849–860. Springer, 2006.
- [5] Jonathan Boisvert, Chang Shu, Stefanie Wuhrer, and Pengcheng Xi. Three-dimensional human shape inference from silhouettes: reconstruction and validation. *Machine vision and applications*, 24(1):145–157, 2013.
- [6] Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3d medical image segmentation: a review. *Medical image analysis*, 13(4):543–563, 2009.
- [7] James Booth, Epameinondas Antonakos, Stylianos Ploumpis, George Trigeorgis, Yannis Panagakis, Stefanos Zafeiriou, et al. 3d face morphable models in-the-wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [8] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. *ACM Transactions on Graphics (TOG)*, 34(6):248, 2015.
- [9] Pascal Paysan, Reinhard Knothe, Brian Amberg, Sami Romdhani, and Thomas Vetter. A 3d face model for pose and illumination invariant face recognition. In *Advanced video and signal based surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*, pages 296–301. Ieee, 2009.
- [10] Tianye Li, Timo Bolkart, Michael J Black, Hao Li, and Javier Romero. Learning a model of facial shape and expression from 4d scans. *ACM Transactions on Graphics (TOG)*, 36(6):194, 2017.
- [11] Chao Zhang, Behrend Heeren, Martin Rumpf, and William AP Smith. Shell pca: Statistical shape modelling in shell space. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1671–1679, 2015.
- [12] Javier Romero, Dimitrios Tzionas, and Michael J Black. Embodied hands: Modeling and capturing hands and bodies together. *ACM Transactions on Graphics (TOG)*, 36(6):245, 2017.
- [13] Marcel Lüthi, Thomas Gerig, Christoph Jud, and Thomas Vetter. Gaussian process morphable models. *IEEE transactions on pattern analysis and machine intelligence*, 2017.
- [14] Leonid Pishchulin, Stefanie Wuhrer, Thomas Helten, Christian Theobalt, and Bernt Schiele. Building statistical shape spaces for 3d human modeling. *Pattern Recognition*, 67:276–286, 2017.
- [15] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [16] Martin A Styner, Kumar T Rajamani, Lutz-Peter Nolte, Gabriel Zsemlye, Gábor Székely, Christopher J Taylor, and Rhodri H Davies. Evaluation of 3d correspondence methods for model building. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 63–75. Springer, 2003.
- [17] Christoph Lassner, Javier Romero, Martin Kiefel, Federica Bogo, Michael J Black, and Peter V Gehler. Unite the people: Closing the loop between 3d and 2d human representations. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, volume 2, page 3, 2017.