



A Semantic Parametric Model for 3D Human Body Reshaping

Dan Song¹, Yao Jin², Tongtong Wang¹, Chengyang Li¹, Ruofeng Tong^{1(✉)},
and Jian Chang³

¹ State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China
trf@tju.edu.cn

² College of Information Science and Technology, Zhejiang Sci-Tech University,
Hangzhou, China

³ NCCA, Bournemouth University, Bournemouth, UK

Abstract. Semantic human body reshaping builds a 3D body according to several anthropometric measurements, playing important roles in virtual fitting and human body design. We propose a novel part-based semantic body model for 3D body reshaping. We adopt 20 types of measurements in regard of length and girth information of body shape. Our approach takes any number (1–20) of measurements as input, and generates a 3D human body. Firstly, all missing measurements are estimated from known measurements using a correlation-based method. Then, based on our proposed semantic model, we learn corresponding semantic body parameters which determine a 3D body from measurements. Our model is trained using a database of 4000 registered body meshes which are fitted with scans of real human bodies. Through experiments, we compare our approach with previous methods and show the advantages of our model.

Keywords: Semantic human model · 3D body reshaping ·
3D body reconstruction · Imputation

1 Introduction

3D human body modeling has been researched for about 20 years in computer graphics and animation, and has various applications in movies, computer games and virtual fitting. Parametric human body models represent 3D body through deforming a template body mesh with a series of parameters, and can be classified into edge-based models and vertex-based models. Edge-based models [1, 2] capture the shape deformation as edge deformations relative to template mesh, usually with a 3×3 matrix. Vertex-based models [3] treat the shape deformation as vertex displacements with a 3-dimensional vector. Our proposed semantic model is a vertex-based model.

There are several ways to create 3D human bodies using parametric models, such as reconstructions from scans, images and anthropometric measurements.

Anthropometric measurements (e.g., height, chest size, waist size, etc.) provide semantic and intuitive controls towards body shapes, so in this paper we propose a semantic parametric model using body measurements to create or edit 3D human bodies.

We review existing related approaches, and conclude the state-of-the-art framework of 3D body reshaping with anthropometric measurements. First of all, a database containing various 3D body shapes with a similar standing pose is prepared and a set of measurements are defined. Then the state-of-the-art framework consists of the following three steps: (1) Using the known measurements, the number of which is not limited, to estimate the missing measurements; (2) Learning a 3D body shape with all measurements; and (3) Optimizing the body shape with the original known measurements as constraints. Zhang et al. [4] use such a framework, while the others [5–7] only put emphasis on part of the framework.

For step 1, most works [1,6,7] take all the defined measurements as input while Zhang et al. [4] and Zeng et al. [5] can take any number of measurements as input, which relaxes the restriction on the input and is user-friendly to body creators. Zhang et al. [4] propose a correlation-based method and Zeng et al. [5] use MICE (Multivariate Imputation by Chained Equations [8]) to estimate missing data from the known one(s). In Sect. 3.1, we compare their methods, KNN (K-Nearest-Neighbor) and a matrix completion method. The method we use for step 1 in our approach is similar to [4] with minor revisions.

For step 2, existing methods map the defined measurements to body parameters which determine a 3D body shape. There are various types of body parameters, such as weights of PCA bases and affine transformations of mesh triangles. Some works [7,9] map the measurements to the weights of PCA bases performed on the whole body shape, which control vertex displacements relative to the corresponding vertices of template mesh. Some researchers [5,6] map the measurements to the triangle deformations relative to the corresponding triangles of template mesh. We propose a vertex-based semantic model consisting of part-based semantic bases which control deformations according to measurements, and whole body non-semantic bases that make the whole body shape coherent. In Sect. 3.2, we compare our method with the two common-used approaches, and our method achieves the best performance for body reshaping while obtaining comparative reconstruction error.

The time-consumption and result quality of step 3 rely on the quality of the body reconstructed from step 2. How to further refine the learned body shape with original known measurements is out of the scope of this paper.

We use MPII database [10] which contains 4301 registered bodies to train and test our approach. The experiment results show that our novel body model can: (1) perform semantic controls towards body shapes, (2) better satisfy the measurements requirement for body reshaping, and (3) keep the whole body shape coherent.

The paper is organized as follows. In Sect. 2, we firstly give an overview of our method, and then introduce measurements estimation and our proposed

semantic model. Experiments are conducted and analyzed in Sect. 3. Section 4 concludes the paper.

2 Method

2.1 Overview

Figure 1 shows the overview of our approach, which contains the online process and the offline process. The online process experiences three stages: (1) estimating all measurements from given limited input, (2) predicting body parameters using all measurements, and (3) reconstructing 3D body shape according to body parameters with our proposed model. The offline process is based on a public database [10] of 4000 3D registered bodies which are fitted with human scans. The following subsections introduce each online stage and corresponding offline preparations.

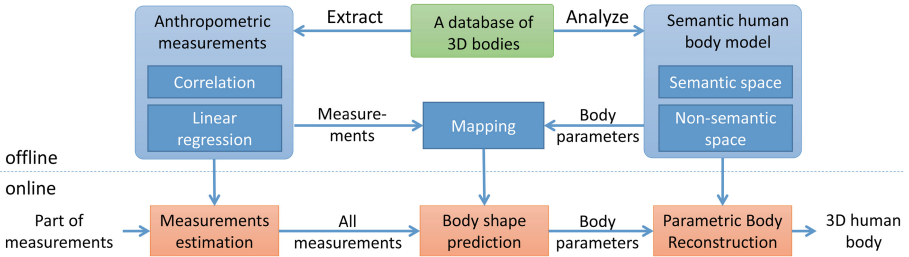


Fig. 1. Overview.

2.2 Measurements Estimation

The anthropometric measurements we use are shown in Fig. 2(a), and we compute the measurements of 4000 3D bodies. How to compute measurements can be found in [11]. Based on the dataset of 4000 sets of measurements, we compute the Pearson’s correlation coefficient and train the linear relationship for any two measurements.

We use a correlation-based method similar to [4] to estimate missing measurements from known ones. Given a subset of 20 anthropometric measurements, which is denoted as S_{in} , we want to get the subset (S_{out}) of unknown measurements. We set a step value s ($s = 0.04$ in our implementation) and iteratively expand S_{in} . Suppose current iteration is $iter$, for measurement i in S_{out} , if there exists any measurement j in S_{in} and the correlation coefficient of i and j is larger than $1 - iter \times s$, we use the trained linear relationship to predict the value of measurement i from measurement j . If there are more than one measurements in S_{in} satisfying the condition, we will compute the weighted average value of predicted values, where the weights are decided by the correlation coefficients.

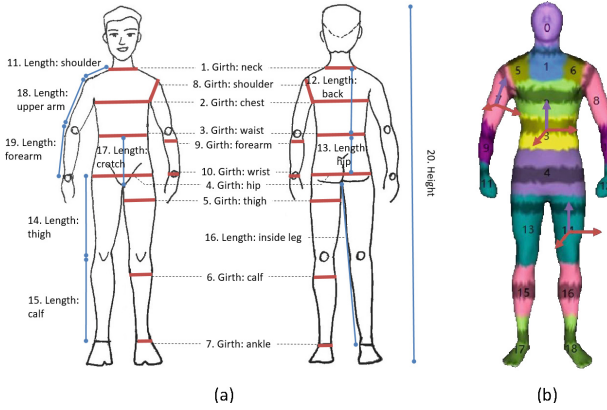


Fig. 2. 20 anthropometric measurements and body partitions. (Color figure online)

2.3 Semantic Human Body Model

We propose a semantic parametric model, consisting of part-based semantic bases which control semantic deformations and whole body shape bases that make the body coherent. Du et al. [12] propose a semantic representation of 3D face model for reshaping with manual semantic bases. Different from their work, we train semantic bases by analyzing the variations of body shapes along semantic directions. We adopt 20 anthropometric measurements including length information and girth information (Fig. 2(a)), and segment human body into 19 partitions in accordance with these measurements (Fig. 2(b)). The vertices of darkened area of each partition in the figure are used for girth calculation.

Equation 1 illustrates our model, where θ and β are body parameters. \mathbf{V} is a $3N$ -dimensional vector denoting the positions of body vertices, and $\hat{\mathbf{V}}$ represents the corresponding vertex positions of template body. N is the number of vertices, and P is the number of body partitions. \mathbf{B}_i^l represents semantic length bases, \mathbf{B}_i^g represents semantic girth bases, and \mathbf{U} denotes non-semantic bases, the training of which is introduced in the following two paragraphs.

$$\mathbf{V}(\theta, \beta) = \hat{\mathbf{V}} + \sum_{i=1}^P (\mathbf{B}_i^l \theta_i^l + \mathbf{B}_i^g \theta_i^g) + \mathbf{U} \beta \quad (1)$$

We train \mathbf{B}_i^l and \mathbf{B}_i^g separately for each part and \mathbf{U} for the whole body. For each part, we firstly represent vertices using local coordinate whose x-z plane is parallel to girth plane (marked with red arrows in Fig. 2(b)) and y axis corresponds to length direction (marked with purple arrows in Fig. 2(b)). Secondly, we compute the rigid transform from training sample shape to template shape, and let each transformed shape subtract the template shape. Thirdly, we separately perform PCA on the y positions and on the x and z positions to gain semantic bases \mathbf{B}_i^l and \mathbf{B}_i^g respectively. We should mention that \mathbf{B}_i^l is a $3N \times L$ matrix, $3N - N_p$ rows of which are set to zero. \mathbf{B}_i^g is a $3N \times G$ matrix, $3N - 2N_p$

rows of which are set to zero. Here N_p is the number of vertices of the part and L and G represent the number of bases.

For training non-semantic bases of the whole body, we firstly represent each training sample only with semantic bases. Then we make each training sample subtract corresponding semantic represented one, and perform PCA on the vertex residuals of all training samples to obtain non-semantic bases \mathbf{U} (a $3N \times W$ matrix, where W is the number of bases).

2.4 Body Shape Prediction

For each one of 4000 bodies in the training database, we have its body measurements M and body parameters $(\boldsymbol{\theta}, \boldsymbol{\beta})$ as a training example. For each measurement, we learn a linear relationship between the measurement and its corresponding body parameter. The body parameter is L -dimensional for length measurement while G -dimensional for girth measurement. We also train a linear relationship between the semantic parameter $\boldsymbol{\theta}$ and the non-semantic parameter $\boldsymbol{\beta}$.

In online process, given 20 anthropometric measurements, we firstly estimate the semantic parameter according to measurements and then predict the non-semantic parameter. Finally, we reconstruct the 3D body shape with body parameters $(\boldsymbol{\theta}, \boldsymbol{\beta})$ using formula 1.

3 Experiment Results

3.1 Measurements Estimation Error

We prepare 301 testing samples using MPII database [10], which have no overlaps with training samples. The 20 anthropometric measurements are computed for every testing sample. We randomly miss a number of measurements and estimate the missing data from known one(s) using correlation-based method, KNN, SoftImpute [13] with BiScaler [14] and MICE [8]. The correlation-based method is implemented as we describe in Sect. 2.2, and the other three methods are based on the fancyimpute code [15].

Figure 3 shows the mean absolute error of estimated measurements with different numbers of known measurements. Overall, correlation-based method performs best. When we know more than 12 measurements, KNN achieves less estimation error. For the results displayed in Fig. 3, the training and testing samples contain both male and female bodies. If we train separate models for male and female, we will get slightly less error, but the trends and comparisons of these methods are the same.

3.2 Evaluation of Semantic Body Model

In this section, we compare our approach, which predicts 3D body shape as introduced in Sect. 2.4, with two common-used approaches. One maps measurements to the weights of whole body PCA bases (abbr. PCA weight mapping),

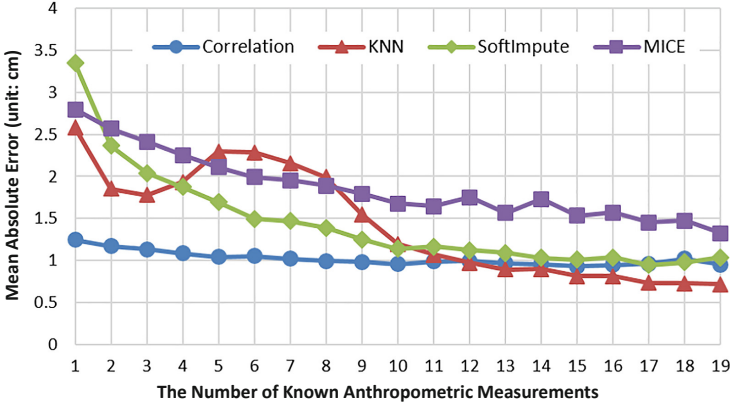


Fig. 3. Measurements estimation error with different numbers of measurements as input.

and the other maps measurements to the triangle deformations (abbr. triangle deformation mapping). PCA weight mapping method is adopted by many researchers such as [7] and [9], which learns the linear relationship between measurements and the weights of PCA bases. Triangle deformation mapping method [5,6] learns the linear relationship between the affine deformation and the corresponding measurement for each triangle. After obtaining the affine transformation for each triangle, we adopt the vertex formulation proposed by [16], which satisfies the shared vertex constrains, to solve the positions of vertices.

We take the 20 anthropometric measurements of 301 testing samples as input to predict 3D bodies using these three methods, and Table 1 compares the mean absolute vertex-to-vertex error in x/y/z direction. Our approach achieves comparative reconstruction accuracy with PCA weight mapping method, while triangle deformation mapping method falls behind.

Table 1. Reconstruction error of different methods (unit: mm)

Method	X direction	Y direction	Z direction
Our method	2.82	3.56	2.40
PCA weight mapping	2.80	3.54	2.37
Triangle deformation mapping	5.12	5.44	6.40

We further compare the performance of these three methods for body reshaping by changing sizes. Figure 4 shows examples of the shape changes when we adjust chest size to 80 mm more or waist size to 50 mm less. We measure the increment of chest size or the decrement of waist size for these three methods. We also compute the absolute vertex-to-vertex distance between the shape before

adjustment and that after adjustment, and show the distance with colors. The blue color denotes smaller distance while the red color illustrates far distance.

The shape deformation performed by our method is the closest to the requirement, while triangle deformation mapping method barely changes the shape. Triangle deformation mapping method maps measurements to the triangle deformations relative to the corresponding triangles of template body mesh. The vertex positions of triangles affected by the adjusted size rely on the positions of neighboring triangles controlled by the unchanged sizes, so this method cannot get required shape change when we only adjust partial sizes. Our method and PCA weight mapping method use vertex-based human body models, and learn vertex displacements relative to the corresponding vertices of template mesh. Compared with PCA weight mapping method, our method achieves better size changes, and we suppose that it owes to the part-based semantic bases which improve the expressive ability of model for local deformations.

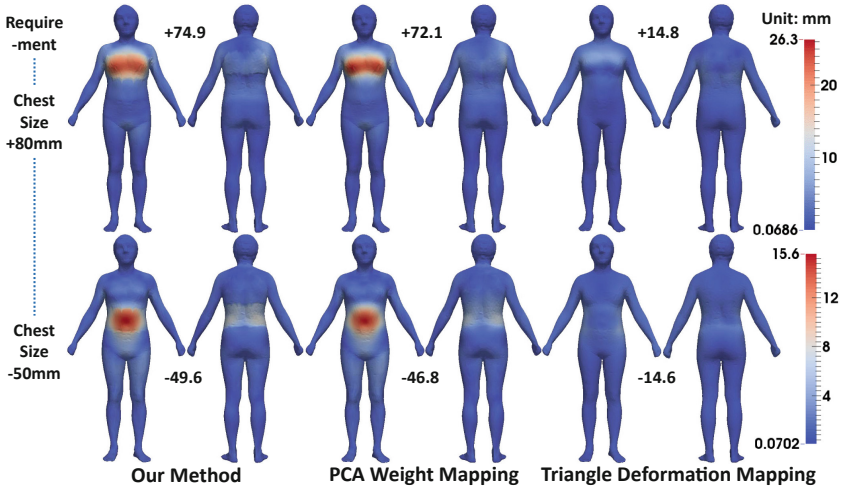


Fig. 4. Shape changes when we adjust chest size to 80 mm more or waist size to 50 mm less. (Color figure online)

4 Conclusion

We propose a novel semantic parametric model for 3D human body reshaping. Our model contains part-based semantic bases which control deformations according to measurements, and whole body non-semantic bases that make the whole body shape coherent. The experiment results show that we obtain comparative reconstruction accuracy, and can perform desired shape deformations with sizes changing.

Acknowledgments. The research is supported in part by NSFC (61572424, 61832016) and the Science and Technology Department of Zhejiang Province (2018C01080). Yao Jin is supported in part by Zhejiang Provincial Natural Science Foundation of China(LY17F020031).

References

1. Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., Seidel, H.P.: A statistical model of human pose and body shape. *Comput. Graph. Forum* **28**(2), 337–346 (2009)
2. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: SCAPE: shape completion and animation of people. *ACM Trans. Graph. (TOG)* **24**(3), 408–416 (2005)
3. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. *ACM Trans. Graph. (TOG)* **34**(6), 248 (2015)
4. Zhang, Y., Zheng, J., Magnenat-Thalmann, N.: Example-guided anthropometric human body modeling. *J. Vis. Comput.* **31**(12), 1615–1631 (2015)
5. Zeng, Y., Fu, J., Chao, H.: 3D human body reshaping with anthropometric modeling. In: Huet, B., Nie, L., Hong, R. (eds.) *ICIMCS 2017*. CCIS, vol. 819, pp. 96–107. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-8530-7_10
6. Yang, Y., Yu, Y., Zhou, Y., Du, S., Davis, J., Yang, R.: Semantic parametric reshaping of human body models. In: *2014 2nd International Conference on 3D Vision (3DV)*, vol. 2, pp. 41–48. IEEE, Tokyo (2014)
7. Wuhler, S., Shu, C.: Estimating 3D human shapes from measurements. *J. Mach. Vis. Appl.* **24**(6), 1133–1147 (2013)
8. Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J.: Multiple imputation by chained equations: what is it and how does it work? *Int. J. Methods Psychiatr. Res.* **20**(1), 40–49 (2011)
9. Chen, Y., Cheng, Z.Q., Martin, R.R.: Parametric editing of clothed 3D avatars. *Vis. Comput.* **32**(11), 1405–1414 (2016)
10. Pishchulin, L., Wuhler, S., Helten, T., Theobalt, C., Schiele, B.: Building statistical shape spaces for 3D human modeling. *Pattern Recogn.* **67**, 276–286 (2017)
11. Song, D., et al.: Clothes size prediction from dressed-human silhouettes. In: Chang, J., Zhang, J.J., Magnenat Thalmann, N., Hu, S.-M., Tong, R., Wang, W. (eds.) *AniNex 2017*. LNCS, vol. 10582, pp. 86–98. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-69487-0_7
12. Du, J., Song, D., Tang, Y., Tong, R., Tang, M.: “Edutainment 2017” a visual and semantic representation of 3D face model for reshaping face in images. *J. Vis.* **21**(4), 649–660 (2018)
13. Mazumder, R., Hastie, T., Tibshirani, R.: Spectral regularization algorithms for learning large incomplete matrices. *J. Mach. Learn. Res.* **11**(Aug), 2287–2322 (2010)
14. Hastie, T., Mazumder, R., Lee, J.D., Zadeh, R.: Matrix completion and low-rank SVD via fast alternating least squares. *J. Mach. Learn. Res.* **16**, 3367–3402 (2015)
15. <https://github.com/iskandr/fancyimpute>
16. Sumner, R.W., Popović, J.: Deformation transfer for triangle meshes. *ACM Trans. Graph. (TOG)* **23**(3), 399–405 (2004)