Computers & Graphics 81 (2019) 52–60

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on CAD & Graphics 2019

Parametric 3D modeling of a symmetric human body

Yin Chen[a], Zhan Song[b], Weiwei Xu[c], Ralph R. Martin[d], Zhi-Quan Cheng[b,d,*]

[a] College of Defense Engineering, Army Engineering University of PLA, China
[b] Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China
[c] State Key Lab of CAD&CG, Zhejiang University, China
[d] Avatar Science Company, China

ARTICLE INFO

Article history:
Received 7 March 2019
Revised 27 March 2019
Accepted 29 March 2019
Available online 15 April 2019

Keywords:
3D human body
Symmetry
Shape
Pose
Accuracy

ABSTRACT

To realistically represent 3D human body shape in a mathematical way, the parametric model used should incorporate symmetry as displayed by real people. This paper proposes a symmetric parametric model called *symmetricSCAPE*. It successfully incorporates symmetry into a parametric model of the 3D body, formulating body geometric variations of both shape and pose using a triangular mesh representation. The symmetry constraint is imposed on each symmetrically-related triangle pair of the body mesh. Mathematically, symmetry-related constraint matrices are derived, and applied during shape and pose deformation of triangle pairs. By accurately registering a pre-designed symmetrization template mesh to the training dataset, we learn how the *symmetricSCAPE* model causes the body mesh to deform relative to the symmetry. Our experiments demonstrate that the *symmetricSCAPE* model results in a better, more parsimonious, and more accurate parametric model of the 3D human body than traditional non-symmetry-aware representations.

1. Introduction

3D human body modeling is a classical problem in academia and industry. To realistically model the human body, digital full-body avatars need features matching those of real people. The expected parametric model should simultaneously represent a wide range of features to effectively characterize the 3D human body. Many efforts are ongoing, taking into account such features as geometry, texture appearance, articulation, symmetry, etc.

The pioneering SCAPE (*Shape Completion and Animation of PEople*) parametric model [1] represents geometric variations in both shape and pose of 3D human body, in term of articulated deformation. SCAPE has been improved in various ways to provide better articulated models, such as BlendSCAPE [2], RealtimeSCAPE [3] and SMPL [4]. Parametric 3D body models have further been extended to incorporate typical features of the 3D body, such as respiration [5], dynamic soft-tissue motion [6], and texture [7].

Symmetry [8], especially intrinsic reflectional symmetry, is a fundamental property of the human body. Symmetry has been thoroughly studied in 3D and applied in geometry processing to tasks such as symmetrization [9], correspondence computation, segmentation, and geometry repair [10].

Articulation and symmetry are both particular geometric features of the human body, but their processing has mainly developed separately. It is thus a natural idea is to combine both symmetry and articulation to better represent the 3D body. Note that a symmetric template is implicitly used by Hirshberg et al. [2], Loper et al. [4] and Hasler et al. [11], but it still fails to take into account the relationship between the parametric model and symmetry of the human body. In our paper, a combined scheme is carefully used in a unified parametric model. In particular, symmetry of the human body is mathematically incorporated into the unified parametric model, to better express 3D human body shape.

We propose a symmetric parametric 3D body model, called *symmetricSCAPE*, successfully incorporating both articulation and symmetry in a unified model. Our parametric model is built on BlendSCAPE [2], but the extension of BlendSCAPE to incorporate symmetry is not as simple as might be expected, and needs careful consideration. *SymmetricSCAPE* requires reformulating the whole pipeline of BlendSCAPE, from the template and training-data preparation, to formulation of the parametric model in both shape and pose variations.

We focus here on parametric modeling of an ordinary 3D human body with typically symmetric features, and do not consider persons with incomplete bodies, e.g. missing legs or arms. In the implementation, *symmetricSCAPE* model learning and data preparation are both performed during pre-processing. Afterwards,

* Corresponding author at: Avatar Science Company, China.
E-mail address: cheng.zhiquan@avatarscience.com (Z. Cheng).

https://doi.org/10.1016/j.cag.2019.03.013
0097-8493/

symmetricSCAPE can rapidly reconstruct a complete 3D body mesh for a human body instance. Our symmetric template mesh has variable resolution: the face is represented in greater detail, which is important to provide compelling 3D avatars with visually convincing appearances. Parametric model learning is processed by registering the template to high-resolution 3D scans of different subjects in a wide variety of poses. Following Hirshberg et al. [2], during registration of the template to 3D scans, the *symmetricSCAPE* model is simultaneously learned by using principal component analysis (PCA) on the aligned vertices. The key feature of our parametric model is that it is formulated to consistently encode both symmetry and articulation of the 3D body.

Our experiments show that our formulation results in a better, more parsimonious and more accurate parametric 3D model of the human body than traditional non-symmetry-aware representations. The improved performance may be demonstrated in two ways: (i) the *symmetricSCAPE* model results in meshes that are better reconstructed, and exhibit better symmetry across a database of registered body shapes, and (ii) for a fixed number of parameter coefficients, body reconstruction by *symmetricSCAPE* is more accurate than when a non-symmetry-aware model is used.

In the following, we review related work in Section 2, while the definition and training of the *symmetricSCAPE* model are considered in Sections 3 and 4. Body reconstruction is addressed in Section 5. Section 6 presents and analyzes experimental results, and conclusions are given in Section 7.

2. Related work

As a long-standing problem, 3D human body modeling has been studied both theoretically and algorithmically. Researchers aim to build reasonable and realistic 3D representations of the human body, providing concise descriptions of its basic features, such as its articulation, and intrinsic symmetry. In this review of related work, we focus on the most relevant topics: parametric modeling of the 3D human body, and its intrinsic symmetry.

2.1. Parametric modeling of the 3D human body

Parametric 3D modeling of the articulated human body began by studying geometric pose and shape separately, then considered unified models with additional advantages.

The linear blend skinning (LBS) method [12] can represent body pose deformation as a linear function of skinning weights and transformations of articulated bones. Based on this work, Allen et al. [13] used an articulated body model in various poses to fit scanned partial human 3D point clouds.

Considering the opposite problem, Allen et al. [14] proposed a parametric deformation model induced by shape, by fitting a template to 250 different scanned persons in the same pose. Applying PCA decomposition to the displacements of template vertices provided a model of the shape space. At the same time, Seo and Magnenat-Thalmann [15] proposed another shape induced deformation model which decomposes shape deformation into rigid and non-rigid deformation separately. The non-rigid deformation is modeled by PCA of template vertex displacements.

The pioneering work of Anguelov et al. [1] captures correlations of shape deformations between different individual bodies as well as correlations of pose deformations, providing the fundamental unified SCAPE model. Many following works considered how to improve upon this model, aiming to provide highly flexible and realistic body models. The triangle deformation used in [1] is not rotation-invariant, meaning that the same shape is encoded differently depending on its orientation. To alleviate the problem, Hasler et al. [11] encodes each triangle with its three neighbors using a local rotation-invariant transformation, and introduces a statistical model of the 3D body with correlation between shape and pose using scans of several people in different poses. An alternative way of modeling the correlation between shape and pose is formulated in [16] using complex tensor multiplication. Although these two complex encoding approaches both experimentally generate consistent results, articulation of the 3D body is almost overlooked. The template-to-scan alignment scheme of [1] and [11] is a typical two-stage approach, in which one first aligns the template mesh to each training example, then builds a parametric model of the 3D body from the aligned data. Instead, Hirshberg et al. [2] integrates the registration process with the learning of this parametric model, and regularizes alignments using an articulated parametric model, in an approach termed BlendSCAPE. Each triangle's rigid rotation in BlendSCAPE is a linear blend of the rotations of the parts of the skeleton. It outperforms SCAPE in that every triangle within an articulated part has the same rotation. Realtime articulated SCAPE-based models have been proposed by Chen et al. [3], Loper et al. [4] and Ye et al. [17]. They share the same idea that the computation of the pose-dependent deformation can be accelerated by use of a linear skinning function. To realistically model respiration, a new model for breathing was introduced by Tsoli et al. [5]. The shape deformation of SCAPE is separated into the intrinsic personal shape and changes due to breathing; additionally, the pose is separated into a static pose and pose variations due to breathing. A similar shape and pose separation scheme was adopted by Pons-Moll et al. [6] to model dynamic soft-tissue deformations.

The recent SMPL (Skinned Multi-Person Linear) model [4] does not use the triangle-deformation scheme of SCAPE-based approaches, but a unique vertex-based model. SMPL applies 3D displacements directly to vertices to model body changes, avoiding the stitching step needed by SCAPE-based triangle deformation. The generality of linear blend SMPL provides an extensible foundation, e.g. for modeling cloth [18], motion [19], hands [20], integration with sparse IMU sensors [21], physical simulation of soft tissue [22], and single-view realtime dynamic body reconstruction [23]. SMPL is compatible with standard graphics software approaches to skinning of vertices.

2.2. Symmetry of the human body

In addition to (linear) skinning articulation, symmetry is another important feature of the 3D body. It is an interesting question whether symmetry should be incorporated into 3D parametric models of the human body. Symmetry of the human body is of long-standing interest, as attested by the recent survey by Mitra et al. [8] of symmetry research.

Global (approximate) intrinsic symmetry is a salient feature of the human body. In the presence of global symmetry features, stable transformations exist that map the whole body to itself. Furthermore, the human body always undergoes approximately isometric deformations, so use of intrinsic symmetry via measurements in geodesic space is an appropriate approach. Global intrinsic symmetry of the human body can be detected and computed by use of eigenfunctions of Laplace Beltrami operators [24]. Voting for planar reflective symmetry transforms of symmetric triangle pairs provides a symmetry extraction algorithm [10], which can detect partial intrinsic reflectional symmetry of the body. Utilizing unique Möbius transform determination allows stable global intrinsic symmetry detection for the 3D body [25].

In [2,4,11], the hand-created template constructed is (implicitly) symmetric. Indeed, Loper et al. [4] goes further and also uses a symmetry regularization term in the pose-related training process. These papers implicitly use the intuitive idea that enforcing symmetry produces models that are visually more pleasing, although specific experiments were not made to validate this assumption. Indeed, previous 3D parametric body models failed to

explicitly formulate the relationship between the parametric model and symmetry of the 3D body in such a way as to take most advantage of symmetry.

In our paper, symmetry is mathematically incorporated in the 3D parametric modeling of the human body. Furthermore, symmetry is enforced on the data: both the template and training data are prepared using symmetrization [9].

3. Model formulation

SCAPE [1] is a decoupled deformation model which separately accounts for body shape variation between different people, and pose changes. Pose changes are further separated into rigid rotation and non-rigid pose-dependent deformation. In the original SCAPE method, unnatural discontinuities rarely appear at the joints between rigid parts, like the knees, hips and shoulders. To achieve smoother joints, [2] proposed an optimized BlendSCAPE model. Each triangle's rigid rotation in BlendSCAPE is a linear blend of the rotations of those of articulated parts. It outperforms SCAPE, in which every triangle within a rigid part has the same rotation.

Like SCAPE [1] and BlendSCAPE [2], *symmetricSCAPE* parameterizes the 3D body using shape and pose factors. Unlike SCAPE and BlendSCAPE, we take into account symmetry of the 3D body and incorporate symmetry constraints into the parametric model.

We assume that a human body can be represented by a mesh with a set of triangular faces $F = \{f_1, \ldots, f_{|F|}\}$, $|F| = 30,260$ with corresponding vertices $V = \{v_1, \ldots, v_{|V|}\}$, $|V| = 15,132$. Parametric modeling of 3D body shape aims to build a descriptive model $\mathcal{M}(\beta, \theta; \Phi)$ to represent a given human body. A vector of variable parameters specific to the subject indicate both the subject's body pose θ and shape β, while learned model parameters Φ which are constant across all individuals are determined from the training dataset. These variable and learned parameters are given particular values to reconstruct a specific 3D body, by deforming a pre-designed template mesh $\bar{\mathbf{T}}$ which exists in a standard (zero) pose. The template is symmetrized and manually pre-segmented into $|B| = 17$ articulated parts in a kinematic tree structure.

Similar to BlendSCAPE [2], *symmetricSCAPE* is Mathematically formulated by combining a transformation for each triangle, modeling the triangle deformation induced by the body pose, and the shape changes between different individuals. For each triangle face f, the deformation that transforms it from the template mesh $\bar{\mathbf{T}}$ to the target mesh $\mathbf{T}$ is given by:

$$f_{\mathbf{T}} = \left(\sum_{b=1}^{|B|} w_{f,b} R_b(\theta) \right) S_f(\beta) Q_f(\theta) f_{\bar{\mathbf{T}}}, \tag{1}$$

where $w_{f,b}$ is the blend weight for rotation $R_b(\theta)$ of the bth articulated part, following Hirshberg et al. [2], and $R_b(\theta)$ is parameterized in terms of the relative joint angles θ. The non-rigid pose and shape deformations of the face f are encoded as matrices $Q_f(\theta)$ and $S_f(\beta)$ respectively. Following SCAPE [1], $Q_f(\theta)$ (or Q_f for short) is approximated, by principle-component-analysis linearization, as $Q_f = Q_f^0 + \sum_q Q_f^q \theta_f^q$, where $\vec{\theta}_f^q$ is the qth element of the pose vector $\vec{\theta}$ for face f; vectors Q^0 and Q^q contain related coefficients for all faces. $S_f(\beta)$ (S_f for short) can also be approximated by a linear subspace: $S_f = S_f^0 + \sum_s S_f^s \beta_f^s$, where β_f^s is the sth element of the pose vector β for face f; vectors S^0 and S^s contain related coefficients for all faces. All matrices Q_f^0, Q_f^q, S_f^0, and S_f^s have size 3×3.

Formally, Eq. (1) is similar to the approach used in [2], but the formulation is substantially different. In the *symmetricSCAPE* model, (intrinsic) symmetric deformation constraints are imposed on the part weights $w_{f,b}$, and linear coefficients Q_f and S_f, so that the deformed mesh is also intrinsically symmetric. Specifically, given a pair of symmetrically-related triangle faces f_l and f_r in the body mesh, f_l being located in (left) part b and f_r in right part d, we have:

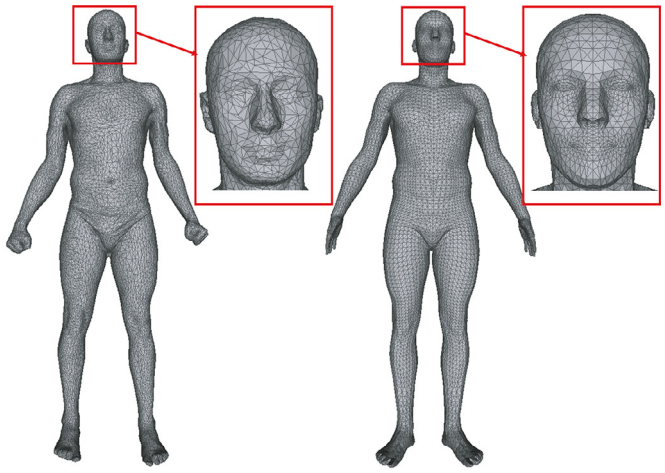

Fig. 1. Template mesh symmetrization. Left: the template used by SCAPE [1]. Right: our symmetrization version.

$$\begin{aligned} w_{f_l,b} &= w_{f_r,d} \\ Q_{f_l} &= \mathfrak{Q} Q_{f_r} \\ S_{f_l} &= \mathfrak{S} S_{f_r} \end{aligned} \tag{2}$$

where parts b and part d are symmetrically-related, $\mathfrak{Q}$ and $\mathfrak{S}$ are the same 3×3 matrices given by:

$$\mathfrak{Q} = \mathfrak{S} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}. \tag{3}$$

Please see the Appendix for the derivation of how the two matrices $\mathfrak{Q}$ and $\mathfrak{S}$ encode symmetry constraints on pose and shape deformation.

Eqs. (1) and (2) together define the *symmetricSCAPE* model. We need to learn the model in Eq. (1) with respect to the constraints in Eq. (2).

Specifically, *symmetricSCAPE* is a parametric model representing a given symmetric body, using variable parameters specific to the subject and constant learned parameters. The variable parameters represent pose $\theta = [\theta^1, \ldots, \theta^{|\theta|}]^T \in \mathbb{R}^{3 \times |B|}$ and shape $\beta = [\beta^1, \ldots, \beta^{|\beta|}]^T \in \mathbb{R}^{3 \times 3 \times |V|}$. The learned parameters include pose-related weights $W = [w_{1,1}, \ldots, w_{f,b}, \ldots, w_{|F|,|B|}] \in \mathbb{R}^{|F| \times |B|}$ coefficients $Q = [Q_0^0, \ldots, Q_f^q, \ldots]$, and shape-related coefficients $S = [S_0^0, \ldots, S_f^s, \ldots]$.

4. Model training

Training the *symmetricSCAPE* model requires computation of the learned parameters Φ from the prepared data. In Section 4.1 we explain how both template and training datasets are prepared, then in Sections 4.2 and 4.3 respectively how the pose-related and shape-related parameters are learned.

4.1. Training data preparation

Firstly, the template (mesh) is prepared using a symmetrization method [9]. As shown in Fig. 1, the template is symmetrized with respect to the reflectional plane $x = 0$. For any vertex v, symmetrization guarantees that a corresponding symmetric vertex v' exists on the mesh, related by reflection in the plane $x = 0$. In addition, the symmetric template mesh has variable resolution in which the face is represented in greater detail, which is important

to ensure a compelling 3D avatar with visually intuitive appearance.

We train the *symmetricSCAPE* model parameters by minimizing reconstruction error on two datasets: a multi-pose dataset, and a multi-shape dataset. The multi-pose dataset is built using the SMPL dataset [4], consisting of 1786 registrations of 40 individuals (891 for 20 females, and 895 for 20 males); The multi-shape dataset is built using the CAESAR dataset [26] with 1500 males and 1500 females. Each mesh in these training datasets has the same topology as our template, since it was formed by aligning the template to a high-quality 3D scan in the SMPL or CAESAR dataset. Note that each set of training data in the CAESAR dataset was also symmetrized by the same method as the pre-designed template.

4.2. Pose-related training

During pose-related training, we learn the weights W and coefficients $Q = [Q_0^0, \ldots, Q_{3\times|B|}^{|M_Q|}] \in \mathbb{R}^{3\times3\times(1+3\times|B|)\times|M_Q|}$, where $|M_Q|$ is the number of registered meshes in the multi-pose dataset.

Given a specific individual, pose-related training is performed on each face f with vertices (v_1, v_2, v_3) with known shape-related parameters (S_f). Following Anguelov et al. [1], the deformation of all poses works on vertex-difference vectors. Specially, we formulate the pose-related training as minimizing a non-linear least-squares error:

$$\min_{W,Q} \sum_{j=1}^{|M_Q|} \sum_{k=2,3} \|(\sum_{b=1}^{|B|} w_{f,b} R_b(\theta)) S_f Q_f(\theta) \bar{e}_{f,k} - e_{f,k}^j\|^2 \tag{4}$$

where j indicates the pose index of the specific individual, $\bar{e}_{f,k} = \bar{v}_k - \bar{v}_1$ and $e_{f,k}^j = (v_k^j - v_1^j)$ are known vertex-difference vectors, and vertex $\bar{v}$ in the template corresponds to vertex v in the registered mesh of a given pose instance. The pose deformation values θ and $R_b(\theta)$ are already known, so the unknown parameters are the pose-related weights and coefficients.

The minimization problem in Eq. (4) is solved by alternating optimization of the weight $w_{f,b}$ and $Q_f(\theta)$ for each face f with symmetric constraints:

- We fix the transformation $Q_f(\theta)$ and optimize the weight $w_{f,i}$. We minimize the weights for symmetrically-related triangles f_l and f_r at the same time, since these 2 weights should be equal. This can be formulated as:

$$\begin{aligned} \min_{w_{f_l,b}} & \sum_{j=1}^{|M_Q|} \sum_{k=2,3} (\|(\sum_{b=1}^{|B|} w_{f_l,b} R_b(\theta)) S_{f_l} Q_{f_l}(\theta) \bar{e}_{f_l,k} - e_{f_r,k}^j\|^2 \\ & + \|(\sum_{b=1}^{|B|} w_{f_l,b} R_i(\theta)) S_{f_r} Q_{f_r}(\theta) \bar{e}_{f_r,k} - e_{f_r,k}^j\|^2) \\ \text{such that} \quad & \sum_{b=1}^{|B|} w_{f_l,b} = 1, w_{f_l,b} > 0. \end{aligned} \tag{5}$$

  The weights for both triangles are represented by the same optimization variable $w_{f_l,i}$ to enforce weight equality.
- We fix the weights and optimize for $Q_f(\theta)$ under the symmetry constraint to obtain the linear coefficients Q. Again, symmetric triangles f_l and f_r are considered simultaneously in this step. The subproblem is now:

$$\begin{aligned} \min_{Q_{f_l}, Q_{f_r}} & \sum_{j=1}^{|M_Q|} \sum_{k=2,3} (\|(\sum_{b=1}^{|B|} w_{f_l,b} R_i(\theta)) S_{f_l} Q_{f_l}(\theta) \bar{e}_{f_l,k} - e_{f_l,k}^j\|^2 \\ & + \|(\sum_{d=1}^{|B|} w_{f_r,d} R_d(\theta)) S_{f_r} Q_{f_r}(\theta) \bar{e}_{f_r,k} - e_{f_r,k}^j\|^2) \\ \text{such that} \quad & Q_{f_l} = \mathfrak{Q} Q_{f_r} \end{aligned} \tag{6}$$

The training algorithm iterates these 2 steps until convergence.

Following Bogo et al. [7], if we sort pose directions according to their eigenvalues and plot the cumulative explained variance, we observe that the explained variance saturates rather rapidly. Instead of using all elements of Q (there are thousands of them), we reduce the number of pose-related elements used to about 300 coefficients, resulting in a mean absolute vertex distance error (M_{abs}) less than 3 mm between the deformed template and registered mesh.

4.3. Shape-related training

Shape-related training aims to compute the learned parameters $S = [S_0^0, \ldots, S_{|F|\times|M_S|}^{|M_S|}] \in \mathbb{R}^{3\times3\times|F|\times|M_S|}$, where $|M_S|$ is the number of registered meshes in the multi-shape dataset, modeling the nonlinear deformation at each triangle across different individuals. With known pose deformation information (weight $w_{f,b}$, rigid rotation $R_b(\theta)$, and coefficient $Q_f(\theta)$), the computation of S on face f may be formulated as:

$$\arg\min_S E(S) = E_1 + \lambda_1 E_2 + \lambda_2 E_3, \tag{7}$$

where λ_1 and λ_2 are weights experimentally set to 0.1 and 1000, and the terms E_1, E_2 and E_3 are defined as follows:

- E_1 measures matching between the template and the registered mesh j in the multi-shape dataset:

$$E_1 = \sum_{j=1}^{|M_S|} \sum_{k=2,3} \|(\sum_{b=1}^{|B|} w_{f,b} R_b(\theta)) S_f(\beta) Q_f(\theta) \bar{e}_{f,k} - e_{f,k}^j\|^2. \tag{8}$$

- E_2 is a regularization constraint to enforce smoothness between adjacent faces $adj(f_l, f_r)$:

$$E_2 = \sum_{adj(f_l,f_r)} \|S_{f_l} - S_{f_r}\|^2. \tag{9}$$

- E_3 represents the symmetry constraint for symmetrically-related faces $sym(f_l, f_r)$:

$$E_3 = \sum_{sym(f_l,f_r)} \|S_{f_l} - \mathfrak{S} S_{f_r}\|^2. \tag{10}$$

As for the pose-related coefficients, the number of shape-related elements used is also reduced to about 300 coefficients, so that M_{abs} is less than 3 mm.

5. 3D body reconstruction

Like SCAPE [1] and BlendSCAPE [2], *symmetricSCAPE* can reconstruct both a complete 3D body mesh $\mathbf{T}$ representing both shape and pose, from 3D scanned data $Z = \{z_1, \ldots, z_{|Z|}\}$, even given incomplete single-view data.

The full 3D body mesh $\mathbf{T}$ is deformed from the template $\bar{\mathbf{T}}$ to best fit the input scanned data, while also being consistent with the *symmetricSCAPE* model. The learned parameters Φ, including W, Q, and S, have already been determined during training. The remaining unknown variables are the positions of vertices $V_{\mathbf{T}}$ in $\mathbf{T}$, β,

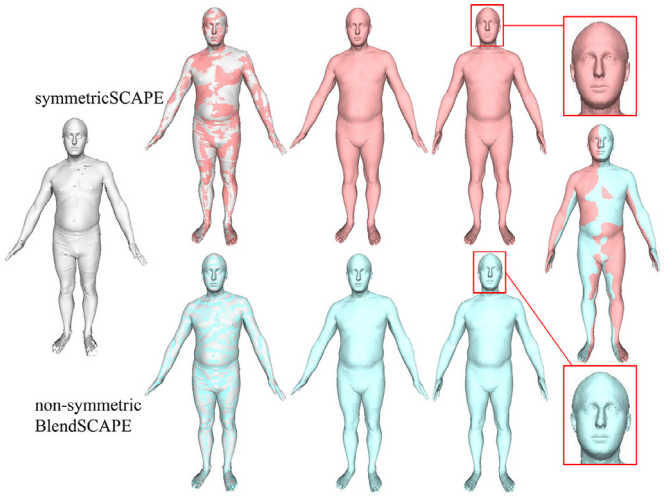

Fig. 2. Model fitting results. Using the instance data (gray, left), we fit both *symmetricSCAPE* (above) and non-symmetry-aware BlendSCAPE (below) models to reconstructed meshes $\mathbf{T}(\beta, \theta)$ (center) by optimizing pose and shape. The fits between the reconstructed meshes and input data are shown in the left-center column. The right-center column shows the estimated pure-shape results $\mathbf{T}(\beta)$ without pose information, parts of the face are magnified and there is a clear difference especially in the nose region. The alignment between them is shown in the right column. It is easy to see that symmetry of the body is preserved better by the *symmetricSCAPE* model.

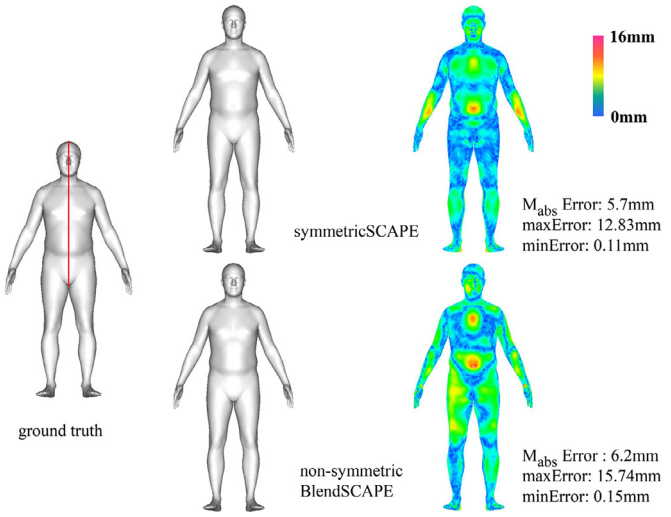

Fig. 3. Reconstructions produced using 10 shape-related components from *symmetricSCAPE* and BlendSCAPE models for the symmetrized data on the left (line of symmetry shown in red). Absolute vertex-to-vertex distance for each vertex is rendered in pseudo-color. It is clear that symmetry is not perfectly preserved in the non-symmetry-aware BlendSCAPE result, especially on the cheeks and legs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and θ, which are determined by minimizing the objective function:

$$\underset{V_{\mathbf{T}},\beta,\theta}{\arg\min} \quad (E_{cost} + \lambda_Z \sum_{i=1}^{|Z|} \|v_{closest}(z_i) - z_i\|^2), \tag{11}$$

where E_{cost} is defined by:

$$\sum_{f=1}^{|F|} \sum_{k=2,3} \|(\sum_{b=1}^{|B|} w_{f,b} R_b(\theta)) S_f Q_f(\theta) \bar{e}_{f,k} - e_{f,k}\|^2. \tag{12}$$

Here, λ_Z are marker weights, and $v_{closest}(z_i)$ is the closest vertex to z_i in the reconstructed mesh $\mathbf{T}$.

The unknown variables are found using an iterative process, which optimizes each of the three sets of parameters ($V_{\mathbf{T}}$, β, and θ) separately in turn, keeping the others fixed. This optimization process converges to a local optimum of Eq. (11).

6. Experimental results and discussion

So far, we have focused on the *symmetricSCAPE* model formulation and the training scheme. In this section, we firstly evaluate the *symmetricSCAPE* model by showing results qualitative and quantitative results, then give a further discussion.

6.1. Results

We evaluate qualitative and quantitative results obtained by fitting the *symmetricSCAPE* model to various meshes representing new people and poses. Fitting involves both shape-based and pose-based deformations, optimizing variable parameters of shape β and pose θ to find the best fitting mesh $\mathbf{T}$ as given by Eq. (11). For clarity, the fitted mesh is denoted by $\mathbf{T}(\beta, \theta)$, while the reconstructed mesh purely related to shape is expressed as $\mathbf{T}(\beta)$. For comparison, a comparable non-symmetry-aware model (e.g. SCAPE, BlendSCAPE) was also trained on the symmetrized template and training data, as described in Section 4.1.

Fig. 2 shows meshes reconstructed by *symmetricSCAPE*, and non-symmetric BlendSCAPE models for a typical instance. The template is unique as shown in Fig. 1. One instance, processed by *symmetricSCAPE* and non-symmetric BlendSCAPE models respectively, is shown in Fig. 2, including one single input scanned data (left) and final reconstructed results (center). The left side shows the real data, which was not used during training of the parametric models. The left-center column indicates the fit between the reconstructed mesh and data; it demonstrates that both models do a good job of fitting the data. The center gives the reconstructed results $\mathbf{T}(\beta, \theta)$, showing desirable appearance. The right-center shows the estimated pure-shape results $\mathbf{T}(\beta)$ without pose information. It is easy to notice salient visual differences between the magnified face regions for different models, especially around the nose. The reconstructed nose result from the BlendSCAPE model lacks symmetry in its visual appearance. A comparison of pure-shape results is shown after alignment, which makes it easier to see that the symmetry results from *symmetricSCAPE* are much more acceptable than those from the non-symmetry-aware BlendSCAPE.

For the specific set of symmetrized data illustrated in Fig. 3 (with line of symmetry shown in red), *symmetricSCAPE* and non-symmetric BlendSCAPE models were used to reconstruct intermediate examples using the first ten principal components (PCs) of shape-related coefficients. For this specific instance, the initial template and final registration ground-truth result are both symmetrized data, so effects of non-symmetric deformation (especially pose-related) are greatly alleviated. The reconstructed results are shown in Fig. 3, to demonstrate the effects of symmetry-preserving functionality on our symmetricSCAPE model. To evaluate the ability of the model to fit a mesh, the error was measured as absolute vertex-to-vertex distances between the intermediate examples ($\mathbf{T}(\beta, \theta)$) and the registration ground-truth. This distance is rendered in pseudo-color for each vertex. For *symmetricSCAPE*, the mean absolute vertex-to-vertex distance (M_{abs}) error is 5.7 mm, with a maximal error (*maxError*) of 12.83 mm, and a minimal error (*minError*) of 0.11 mm. In comparison, for non-symmetry-aware BlendSCAPE, M_{abs}, *maxError*, and *minError* are respectively 6.2 mm, 15.74 mm, and 0.15 mm. The pseudo-color results show clearly that BlendSCAPE does not perfectly preserve symmetry (especially on the cheeks and legs), while *symmetricSCAPE* does.

Fig. 4 demonstrates how M_{abs} error varies for the instance data in Fig. 3 as a function of the number principal components of

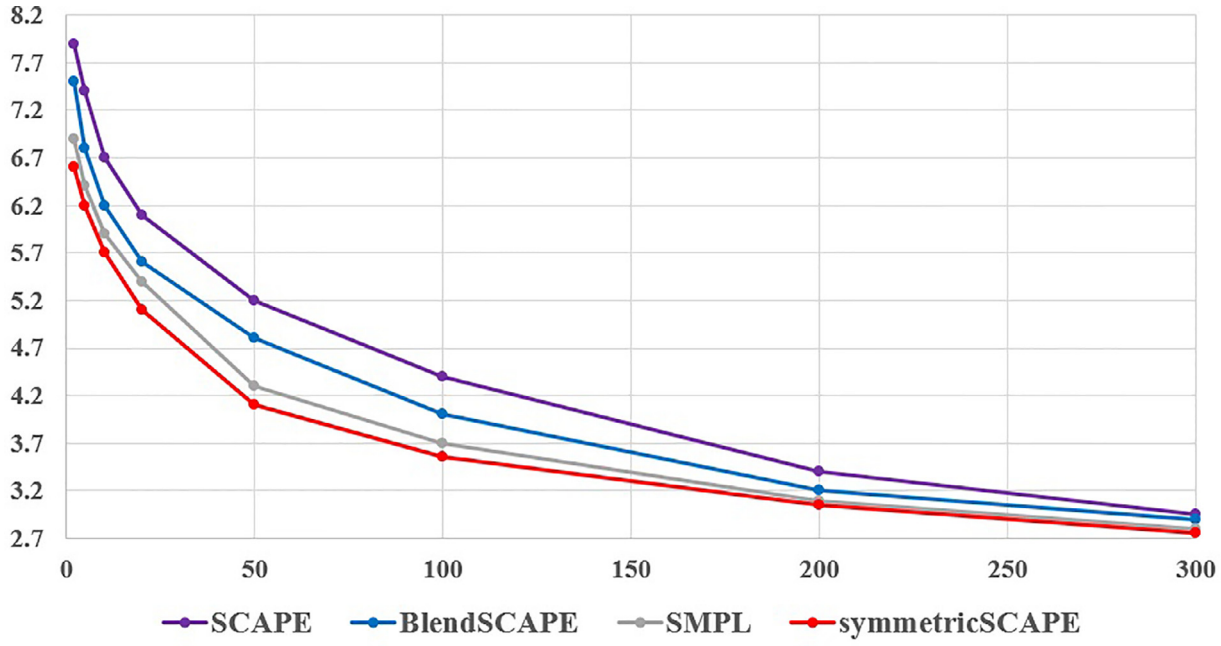

Fig. 4. Variation in mean absolute vertex (M_{abs}) error for five typical models with varying numbers of principal components representing shape-related coefficients, for the instance data in Fig. 3. Errors were measured with 2, 5, 10, 20, 50, 100, 200, and 300 components.

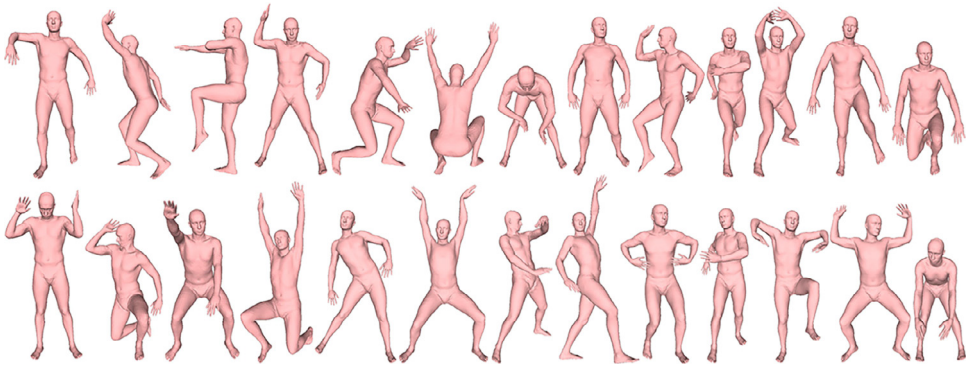

Fig. 5. Reconstructed results produced by the *symmetricSCAPE* model using the multi-pose SCAPE dataset.

shape-related coefficients used for reconstruction. Generally, the error between the examples and final $\mathbf{T}(\beta, \theta)$ gradually decreases with the number of principal components. *symmetricSCAPE* performs better than SCAPE, BlendSCAPE, and SMPL for this instance data.

Using 71 poses from the SCAPE dataset, we tested pose-related reconstruction, to evaluate the ability to generalize a shape of an individual to new poses of the same person. We took the first pose to estimate body shape, then used this parametric model to fit the same subject in different poses without readjusting shape parameters. Fig. 5 demonstrates some reconstructed results produced by the *symmetricSCAPE* model.

Fig. 6 shows variation in pose-related M_{abs} error, primarily to test how well the reconstructed pose-related meshes fit the scanned mesh in the given pose. Results show that *symmetricSCAPE* performs better than BlendSCAPE on this test.

We now consider the run time of the body reconstruction algorithms for both *symmetricSCAPE* and BlendSCAPE models, using a desktop PC with a 3.3 GHz Intel Core i9 processor, without GPU acceleration. To produce the result in Fig. 2 from the given 3D point cloud, the *symmetricSCAPE* model takes about 12.5 s while BlendSCAPE takes 12.3 s to finish the fitting process. Both approaches use the same number (200) of iterations to converge to the final optimum of Eq. (11). For the multi-pose SCAPE dataset, *symmetricSCAPE* and BlendSCAPE models take more or less the same time on average, 31 ms, to compute each individual result. For the pose testing application, shape-related processing is done once and the shape is then kept fixed for pose deformation, so the cost of animating the mesh is just that for the pose-related computation: realtime performance can be achieved by use of the linear skinning scheme.

6.2. Discussion

Note that we only consider a general 3D symmetric body with complete parts. In practice, some people are very clearly not at all symmetric, e.g. having lost an arm or leg in an accident. A direct application of *symmetricSCAPE* in such cases is to recover the missing parts and restore the body's symmetry, as in [10].

Less drastically, even ordinary people typically display slight asymmetry of their body parts. In particular, enforcing too much symmetry on the face region can result in quite visible modifications to a person's appearance. Ideally a modeling approach should remove local, point- or triangle-wise, *random* differences in symmetrically related parts of the body, but at the same time preserve small *coherent* differences which occur across regions (left

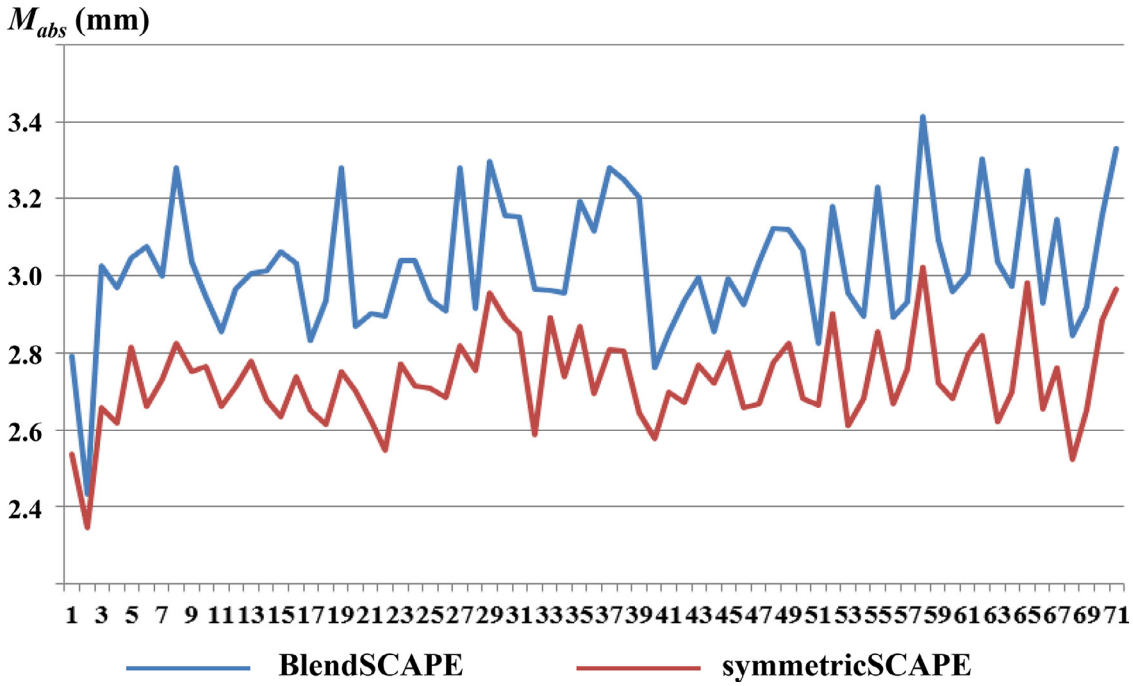

Fig. 6. Variation in mean absolute vertex (M_{abs}) error for the reconstructed results from BlendSCAPE (blue) and symmetricSCAPE (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and right cheeks for example), to preserve such asymmetry. A suitable model to capture this idea needs further consideration, and is a topic for future research.

As noted in Section 2, parametric 3D body modeling generally relies on triangle-based SCAPE-like schemes or vertex-based SMPL-like schemes. Our formulation in Eqs. (1)–(3) captures geometric variations in both shape and pose using a triangular mesh representation. Our approach to symmetry could also be incorporated in other methods also using triangle-based SCAPE-like schemes, e.g. SCAPE [1], BlendSCAPE [2], RealtimeSCAPE [3]). However, doing so in SMPL [4] is still a challenging problem, since it is a vertex-based scheme. A symmetry regularization term is used in its pose-related training process, but is not clearly expressed in the basic SMPL model. Despite our attempts, we have been unable to appropriately capture and enforce symmetry mathematically in the SMPL model in a unified way.

Recently, Zuffi et al. [27] extended SMPL to model the 3D shapes and poses of animals. This extension could also be applied to triangle-based SCAPE-like methods, implyings that our *symmetricSCAPE* could be likewise extended to other symmetric living things such as animals, and beyond.

7. Conclusions

Previous models of 3D human body shape have had the ability to represent a range of identity-dependent body shapes, and to deform them naturally into various poses, even exhibiting soft-tissue motions. However, they have failed to integrate that symmetry which is typically inherent in a human body.

This paper has successfully formulates a symmetric 3D parametric model of the human body, which we call *symmetricSCAPE*. Symmetry-related constraint matrices are applied during shape and pose deformation of triangle pairs to ensure symmetry is preserved when modeling the human body.

In future, we expect research efforts to be devoted towards development of more plausible parametric models, which realistically model a 3D body, including the notion that symmetry is only approximate. More generally, a unified deformation model was recently proposed by Joo et al. [28] for the markerless capture of multiple scales of human movement, including facial expressions, body motion, and hand gestures. Various deep learning methods have also been tried for 3D body modeling, producing reasonable results. More importantly, novel applications are needed to make good use of these models. An ideal modeling result would be realistically clothed, detailed and textured, making it appropriate for many more applications.

Acknowledgment

This work was supported by the Shenzhen Science and Research Plan (JCYJ20170818160448602 and KQJSCX20170731165108047), Shenzhen Scientific and Technological Innovation Project (2018–320), Natural Science Foundation of China (61602507 and 61732016), China Postdoctoral Science Foundation (2016M602555).

Appendix: the symmetry matrices

A.1. The pose-constraint matrix

In the symmetrized template, the pose deformation of two symmetrically-related faces f_l (the left, red, triangular face in Fig. 7(a)) and f_r (right, blue, faces in Fig. 7(a)) satisfies $f_r = A f_l$, where A is a transformation matrix representing reflection in the plane $x = 0$:

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{A.1}$$

When the template is deformed by the pose-related variable parameters θ as shown by the left model in Fig. 7(b), pose deformation acts on face f_l with vertices (v_{l1}, v_{l2}, v_{l3}). Consider for example the vertex-difference vector $e_l = v_{l2} - v_{l1}$, $\hat{e}_l$ with a non-rigid deformation $\hat{e}_l = Q_{f_l}(\theta)e_l$. The face f_r has symmetric pose θ_{sym} for θ, with $\hat{e}_r = Q_{f_r}(\theta_{sym})e_r$, resulting in the right-hand model

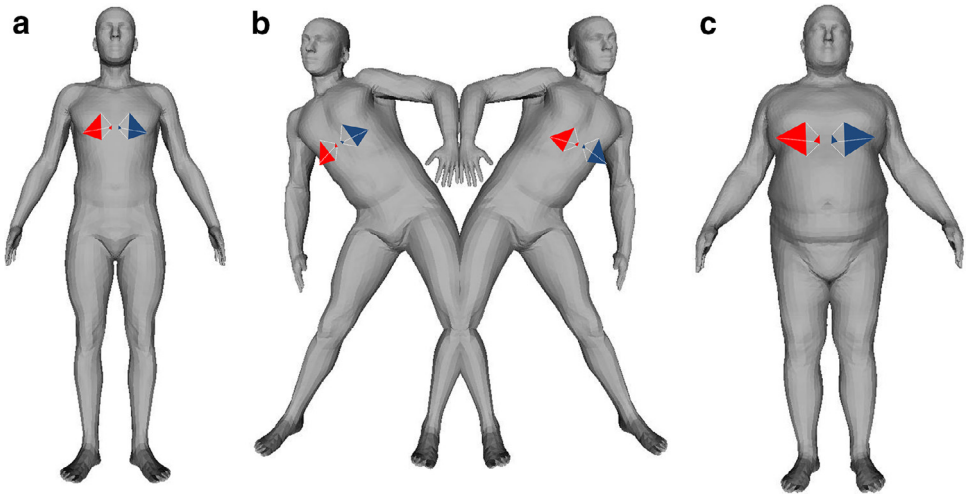

Fig. 7. Deformation with symmetry constraint. (a) Left (red) and right (blue) symmetrically-related faces in the template, (b) changes under pose deformation, (c) changes under shape deformation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in Fig. 7(b). The vertex-difference vector $\hat{e}_l$ for pose θ and the vector $\hat{e}_r$ for pose θ_{sym} are also symmetric about the plane $x = 0$. This means that:

$$\hat{e}_r = A\,\hat{e}_l. \tag{A.2}$$

Substituting $\hat{e}_l$ and $\hat{e}_r$ into the former Eq. (14) gives:

$$Q_{f_r}(\theta_{sym})e_r = A\,Q_{f_l}(\theta)e_l. \tag{A.3}$$

Then, substituting $e_r = Ae_l$, we obtain:

$$Q_{f_r}(\theta_{sym})Ae_l = A\,Q_{f_l}(\theta)e_l. \tag{A.4}$$

In order to satisfy above equation, we must have:

$$Q_{f_r}(\theta_{sym}) = A\,Q_{f_l}(\theta)\,A^{-1}, \tag{A.5}$$

so: $Q_{f_r}(\theta_{sym}) = \mathfrak{Q}\,Q_{f_l}(\theta)$, where $\mathfrak{Q}$ is

$$\mathfrak{Q} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}. \tag{A.6}$$

A.2. The shape-constraint matrix

As for the pose case, the shape-constraint matrix is derived using symmetrically-related faces f_l and f_r in the template, which satisfy $f_l = A\,f_r$ with reflectional symmetry about plane $x = 0$. Using example vertex-differences e_l and e_r again, shape deformation transforms them as $\hat{e}_l = S_{f_l}(\beta)e_l$ and $\hat{e}_r = S_{f_r}(\beta_{sym})e_r$ respectively, as shown in Fig. 7(c). $\hat{e}_l$ and $\hat{e}_r$ are still related by reflectional symmetry, and have the property $\hat{e}_r = A\,\hat{e}_l$. By substituting $\hat{e}_l$, $\hat{e}_r$, and $e_r = Ae_l$ as before, we can derive the following equation:

$$S_{f_r}(\theta_{sym})Ae_l = A\,S_{f_l}(\theta)e_l. \tag{A.7}$$

Thus:

$$S_{f_r}(\theta_{sym}) = A\,S_{f_l}(\theta)\,A^{-1}, \tag{A.8}$$

so: $S_{f_r}(\theta_{sym}) = \mathfrak{S}\,S_{f_l}(\theta)$, where $\mathfrak{S}$ is

$$\mathfrak{S} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}. \tag{A.9}$$

References

[1] Anguelov D, Srinivasan P, Koller D, Thrun S, Rodgers J, Davis J. SCAPE: shape completion and animation of people. ACM Trans Graph 2005;24(3):408–16. (special issue of SIGGRAPH).

[2] Hirshberg DA, Loper M, Rachlin E, Black MJ. Coregistration: simultaneous alignment and modeling of articulated 3d shape. In: Proceedings of the twelfth European conference on computer vision; 2012. p. 242–55.

[3] Chen Y, Cheng Z, Lai C, Martin RR, Dang G. Realtime reconstruction of an animating human body from a single depth camera. IEEE Trans Vis Comput Graph 2016;22(8):2000–11.

[4] Loper M, Mahmood N, Romero J, Pons-Moll G, Black MJ. SMPL: a skinned multi-person linear model. ACM Trans Graph 2015;34(6):248:1–16.

[5] Tsoli A, Mahmood N, Black MJ. Breathing life into shape: capturing, modeling and animating 3D human breathing. ACM Trans Graph 2014;33(4):52:1–11.

[6] Pons-Moll G, Romero J, Mahmood N, Black MJ. Dyna: a model of dynamic human shape in motion. ACM Trans Graph 2015;34(4):120:1–14.

[7] Bogo F, Black MJ, Loper M, Romero J. Detailed full-body reconstructions of moving people from monocular RGB-D sequences. In: Proceedings of the international conference on computer vision; 2015. p. 2300–8.

[8] Mitra NJ, Pauly M, Wand M, Ceylan D. Symmetry in 3D geometry: extraction and applications. Comput Graph Forum 2013;32(6):1–23.

[9] Mitra NJ, Guibas LJ, Pauly M. Symmetrization. ACM Trans Graph 2007;26(3):63:1–8.

[10] Xu K, Zhang H, Tagliasacchi A, Liu L, Li G, Meng M, et al. Partial intrinsic reflectional symmetry of 3D shapes. ACM Trans Graph 2009;28(5):138:1–10.

[11] Hasler N, Stoll C, Sunkel M, Rosenhahn B, Seidel HP. A statistical model of human pose and body shape. Comput Graph Forum 2009;2(28):337–46. (Proc Eurographics 2009).

[12] Lewis JP, Cordner M, Fong N. Pose space deformation: a unified approach to shape interpolation and skeleton-driven deformation. In: Proceedings of the SIGGRAPH; 2000. p. 165–72.

[13] Allen B, Curless B, Popović Z. Articulated body deformation from range scan data. ACM Trans Graph 2002;21(3):612–19. (special issue of SIGGRAPH).

[14] Allen B, Curless B, Popović Z. The space of human body shapes: reconstruction and parameterization from range scans. ACM Trans Graph 2003;22(3):587–94.

[15] Seo H, Magnenat-Thalmann N. An automatic modeling of human bodies from sizing parameters. In: Proceedings of the ACM symposium on interactive 3D graphics; 2003. p. 19–26.

[16] Chen Y, Liu Z, Zhang Z. Tensor-based human body modeling. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2013. p. 105–12.

[17] Ye M, Wang H, Deng N, Yang X, Yang R. Real-time human pose and shape estimation for virtual try-on using a single commodity depth camera. IEEE Trans Vis Comput Graph 2014;20(4):550–9.

[18] Pons-Moll G, Pujades S, Hu S, Black MJ. ClothCap: seamless 4D clothing capture and retargeting. ACM Trans Graph 2017;36(4):73:1–15.

[19] Romero J, Tzionas D, Black MJ. Embodied hands: modeling and capturing hands and bodies together. ACM Trans Graph 2017;36(6):245:1–17.

[20] Bogo F, Romero J, Pons-Moll G, Black MJ. Dynamic FAUST: registering human bodies in motion. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 5573–82.

[21] von Marcard T, Rosenhahn B, Black MJ, Pons-Moll G. Sparse inertial poser: automatic 3D human pose estimation from sparse IMUs. Comput Graph Forum 2017;36(2):349–60.

[22] Kim M, Pons-Moll G, Pujades S, Bang S, Kim J, Black MJ, et al. Data–driven physics for human soft tissue animation. ACM Trans Graph 2017;36(4):54:1–12.

[23] Yu T, Zheng Z, Guo K, Zhao J, Dai Q, Li H, et al. DoubleFusion: real-time capture of human performances with inner body shapes from a single depth sensor. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 7287–96.

[24] Ovsjanikov M, Sun J, Guibas L. Global intrinsic symmetries of shapes. Comput Graph Forum 2008;27(5):1341–8.
[25] Kim VG, Lipman Y, Chen X, Funkhouser TA. Möbius transformations for global intrinsic symmetry analysis. Comput Graph Forum 2010;29(5):1689–700.
[26] Robinette KM, Daanen H. The caesar project: a 3-D surface anthropometry survey. In: Proceedings of the international conference on 3-D digital imaging and modeling; 1999. p. 380–6.
[27] Zuffi S, Kanazawa A, Jacobs DW, Black MJ. 3D menagerie: modeling the 3D shape and pose of animals. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 5524–32.
[28] Joo H, Simon T, Sheikh Y. Total capture: a 3D deformation model for tracking faces, hands, and bodies. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 8320–9.