Shape-based Three-dimensional Body Composition Extrapolation Using Multimodality Registration

Yao Lu[a] and **James K. Hahn**[a]

[a]The George Washington University, Department of Computer Science, 800 22nd St NW, Washington, DC, USA

Abstract

The ubiquity of commodity-level optical scan devices and reconstruction technologies has enabled the general public to monitor their body shape related health status anywhere, anytime, without assistance from professionals. Commercial optical body scan systems extract anthropometries from the virtual body shapes, from which body compositions are estimated. However, in most cases, these estimations are limited to the quantity of fat in the whole body instead of a fine-granularity voxel-level fat distribution estimation. To bridge the gap between the 3D body shape and fine-granularity voxel-level fat distribution, we present an innovative shape-based voxel-level body composition extrapolation method using multimodality registration. First, we optimize shape compliance between a generic body composition template and the 3D body shape. Then, we optimize data compliance between the shape-optimized body composition template and a body composition reference from the DXA pixel-level body composition assessment. We evaluate the performance of our method with different subjects. On average, the Root Mean Square Error (RMSE) of our body composition extrapolation is 1.19%, and the R-squared value between our estimation and the ground truth is 0.985. The experimental result shows that our algorithm can robustly estimate voxel-level body composition for 3D body shapes with a high degree of accuracy.

Keywords

Multimodality; Registration; Medical Imaging; 3D Body Shape; Body Composition

1. INTRODUCTION

The ubiquity of optical body scan systems has enabled large-scale automatic digitalized anthropometry data collection for research institutes as well as daily fitness tracking for the general public. High-end 3D body scanners, such as 3dMDbody[®,1] [*TC*][2®,2] and TELMAT SYMC-AD[®,3] based on high-end optical sensors and sophisticated calibration[4,5] reconstruct sub-millimeter level highly accurate 3D human body shapes. This type of body scanner has been widely used in digitalized anthropometry surveys with large population cohorts.[6–13] Light-weight consumer-level sensor based optical body scanners such as Fit3D[®,14] Styku[®,15] and ShapeScale[®16] scan and reconstruct 3D body shapes from a static pose based

on KinectFusion[17] algorithm. This type of body scanner is typically used by people in the gym or at home to track body shape changes over time. Non-rigid registration based body surface scan and reconstruction systems such as Cui et al.,[18] Li et al.,[19] Dou et al.,[20] and Lu et al.[21] have enhanced the system flexibility by eliminating the need for auxiliary equipment, such as turntable, and arm supporters, which enables users to rotate freely during scan. Full circumference multiple RGB-D sensor systems have also been proposed for 3D body surface reconstruction by Dou et al.[22] and Li et al.[23]

Large-scale body shape datasets have made it possible to parameterize human body shape space using statistical models. Principal Component Analysis (PCA) has been widely explored for shape parameterization.[8,24–26] Body shape variations associated with phenotypes like Body Mass Index (BMI), weight and age have been modeled in the work of Allen et al.[24] and Park et al.[8,27] Digital anthropometries, as summarizations of 3D body shape characteristics, have been explored by Löffler et al. to cluster different body types in a large population cohort.[28] Beyond the traditional anthropometries derived directly from 3D body shape geometries, Lu et al.[29] proposed utilizing high-order shape descriptors, such as surface curvature, for body type classification.

Studies investigating the linkage between 3D body shapes and the underlying health data, such as body fat percentage or fine-granularity fat distribution, are rare since medical-level body composition assessment equipment, such as the BOD POD, Dual-energy X-ray Absorptiometry (DXA), Magnetic Resonance Imaging (MRI) or Computed Tomography (CT), is not readily available for large-scale data collection. Ng et al.[30] studied the relationship between anthropometry features and body fat percentage. Lu et al.[29] explored body shape geometry features in multiple levels for body fat percentage prediction. Lu et al.[21] and Piel[31] further proposed to infer pixel-level DXA-like body composition based on 3D geometry data.

In this work, we propose a highly innovative shape-based 3D voxel-level body composition extrapolation model (Fig. 1-d) using multimodality registration based on 3D body shape (Fig. 1-c) derived from a commodity-level optical body scan system,[21] 2D pixel-level body composition reference (Fig. 1-a) derived from DXA body composition assessment, and a generic 3D body composition template with anatomically accurate human skin, muscle and skeleton system (Fig. 1-b). The 3D body shape can be viewed as a boundary constraint for 3D body composition extrapolation, but no information beneath the surface is provided. DXA generates pixel-specific fat distribution and is considered a gold standard for body composition assessment. However, DXA imaging encodes body composition into a 2D orthogonal projection of the subject, and thus, loses information along the depth direction. In other words, we cannot generate a 3D body composition directly based on the fat distribution provided by the DXA and the boundary constraint provided by the 3D body shape due to a lack of shape prior knowledge of the body composition structure in 3D. Therefore, we introduce a generic body composition template as the shape prior.

2. METHOD

Our voxel-level body composition extrapolation framework (Fig. 2) consists of two steps. First, we register the 3D generic body composition template to 3D body shapes for shape compliance optimization. We propose a three-fold method to gradually refine the shape registration condition. In *gender-specific body composition initialization*, we specify the 3D body composition template for different genders. In *conditional anthropometries transfer*, we conditionally pre-adjust the template surface layer according to the 3D body shape anthropometries. This step is critical for the obese since the amount of subcutaneous fat tends to be much higher than the gender-specific body composition template. Directly optimizing the shape compliance can lead to muscle over-stretching. In *3D-3D body composition template registration*, we register the initialized 3D body composition template to the 3D body shape to optimize the shape compliance. Second, we optimize data compliance between the shape-optimized template and the 2D pixel-level body composition reference from DXA imaging. We parameterize the deformation of the muscle layer, and then seek the optimal set of deformation parameters that align local fat distributions in 3D to the local fat distributions assessed by the DXA (*3D-2D data compliance optimization*).

2.1 Shape compliance optimization

2.1.1 Gender-specific body composition initialization—To facilitate shape compliance optimization, we create gender-specific body composition templates for males and females. For each gender, we calculate an average lean body shape. We register the surface layer using non-rigid ICP[21] to each 3D body shape whose body density is larger than a lean density threshold. We set the threshold to 1.05 kg/L for females, which corresponds to the body fat percentage of 21.4%, and 1.07 kg/L for males, which corresponds to the body fat percentage of 12.6% according to the Siri Equation.[32] We consider the body fat percentage of 21.4% for females and 12.6% for males are reasonable thresholds for lean body shapes selection. We calculate an average shape of the registered surface layers as the gender-specific average lean body shape. The next step is to deform body composition template inner layers (fascia, muscle, skeleton) using linear subspace deformation model[33] according to the gender-specific body shape. We generate the tetrahedral mesh[34] from the surface layer to define a deformation space. We register inner layers of the body composition template into this space using barycentric coordinates. We deform the tetrahedral mesh towards the average lean body shape using the linear subspace deformation, treating surface layer correspondences as boundary conditions. The inner body compositions are then deformed accordingly. The female gender-specific body composition template is shown in Fig. 3. We define an abstract skeletal structure Fig. 3-(a) based on the gender-specific body composition template for further shape registrations.

2.1.2 Conditional anthropometries transfer—To achieve a better initial condition for shape compliance optimization, we propose to conditionally deform template skin layers according to anthropometries extracted from 3D body shapes. We have defined the template abstract skeletal structure in Fig. 3-(a). Correspondingly, we define the abstract skeletal structure on the 3D body shape with heuristically estimated skeletal joints. We sample level circumferences along each bone direction of the abstract skeletal structure (Fig. 3-(a) cyan).

We grow the skin layer regionally using skeletal B-spline skinning if the local level circumference on the template is smaller than the one on the 3D body shape. In the B-spline skinning deformation model, we paint skinning weights of the bones on the surface mesh. We set up multiple key points along each bone to parameterize the surface morph. Each key point is associated with an in-plane scale factor to grow or shrink the surface on its planar direction orthogonal to the bone. In between control points, the in-plane scale factor is interpolated using B-spline interpolation. This step is critical to prevent muscle over-stretching during shape compliance optimization, especially for the obese. Fig. 4 shows examples of the conditional anthropometries transfer of the skin layers, compared with the corresponding 3D body shapes.

2.1.3 3D-3D body composition template registration—We have a relatively good initial condition for shape compliance optimization after first generating a gender-specific body composition template and then conditionally transferring anthropometries to the gender-specific template skin layer. We recalculate the tetrahedral mesh under the surface layer after the anthropometries transfer and update the braycentric coordinates for the body composition template inner layers accordingly. To optimize the shape compliance, we first register the template surface layer to the 3D body shape uisng non-rigid ICP.[21] Then, we deform the tetrahedral mesh accordingly using linear subspace deformation,[33] treating the surface layer correspondences as boundary conditions. Body composition template inner layers are then deformed with the new barycentric coordinates calculated inside the new tetrahedral mesh.

2.2 3D-2D data compliance optimization

So far, we have registered the 3D body composition template into a specific 3D body shape by minimizing the shape compliance error. However, fat distribution reflected by current 3D body composition template does not agree with the ground truth fat distribution assessed by the DXA. To align the 3D body composition to the DXA imaging, we propose a data compliance optimization. We parameterize local fat distributions in 3D using B-spline skinning deformation model and quantify the ground truth local fat distributions from the DXA imaging. Then, we minimize the data compliance error.

2.2.1 3D local fat distribution parameterization—To approximate local fat distributions, we uniformly sample level areas on the fascia layer and 3D body shape (Fig. 5) along the bone direction of the abstract skeletal structure (in Fig. 3-(a) cyan). We approximate each level intersection stack as a right cylinder and the height of each stack is identical because of the uniform sampling. Hence, estimating the stack volume is equivalent to estimating the level area. We denote the fat-free stack volume as V^l_{FFM}, which corresponds to the level area of the fascia layer (Fig. 5, area of the red region), and denote the total stack volume as V^l, which corresponds to the level area extracted from the 3D body shape (Fig. 5, yellow + red region), where l is the stack sample index. The 3D local fat distribution $Fat^l_{3D}\%$ can be approximated as a function of V^l_{FFM} and V^l, as shown in Eq. (1), where we approximate the fat tissue density as $\rho FM = 0.9 kg/L$ and the fat-free tissue density as $\rho FFM = 1.1 kg/L$.[32]

$$Fat^l_{3D}\% = \frac{(V^l - V^l_{FFM}) \times \rho_{FM}}{\rho_{FM} \times V^l + (\rho_{FFM} - \rho_{FM}) \times V^l_{FFM}} \quad (1)$$

To control 3D local fat distribution $Fat^l_{3D}\%$, we parameterize the deformation of the fascia layer using B-spline skinning. We define multiple control points along each bone with scaling parameter $\boldsymbol{\alpha}$. Thus, we rewrite the fat-free stack volume as $\widetilde{V}^l_{FFM}(\boldsymbol{\alpha})$.

2.2.2   2D local fat distribution estimation—For each of the 3D local fat distribution parameterized previously, we estimate a corresponding ground truth local fat distribution from the 2D pixel-level body composition assessment of the DXA. First, we register the DXA imaging (Fig. 1-a) to the 2D orthogonal projection of the 3D body shape using free-form deformation[21] to derive the aligned DXA imaging. Second, we estimate local fat percentage $Fat^l_{GT}\%$ on the aligned DXA imaging corresponding to the local fat distribution in 3D. Third, we convert each local ground truth fat percentage $Fat^l_{GT}\%$ to the target fat-free stack volume $\widehat{V}^l_{FFM}$ through Eq. (2).

$$\widehat{V}^l_{FFM} = \frac{\rho_{FM} \times (1 - Fat^l_{GT}\%) \times V^l}{Fat^l_{GT}\% \times (\rho_{FFM} - \rho_{FM}) + \rho_{FM}} \quad (2)$$

2.2.3   Data compliance optimization—We formulate the objective function as Eq. (3) to minimize data compliance error. We search for deformation parameter $\boldsymbol{\alpha}$ such that the fat-free stack volume $\widetilde{V}^l_{FFM}(\boldsymbol{\alpha})$ derived from the morphed fascia layer aligns with the target stack volume $\widehat{V}^l_{FFM}$ estimated by the ground truth local fat distribution derived from the DXA imaging.

$$\arg \min_{\boldsymbol{\alpha}} \sum_l \| \widehat{V}^l_{FFM} - \widetilde{V}^l_{FFM}(\boldsymbol{\alpha}) \| \quad (3)$$

3.   RESULTS

To evaluate the performance of our method, we test the method on six different data sets, corresponding to six different subjects. The Body Mass Index (BMI) and Body Fat Percentage (BFP) distributions of the subjects are illustrated in Tab. 1. Each data set contains a 3D body shape derived from the commodity-level optical body scan system by Lu et al.,[21] and the corresponding pixel-level body composition reference derived from the DXA. The gender-specific body composition template has been generated by registering the 3D body

composition template into the average lean body shape. We register the reference image to the 2D orthogonal projection of the 3D body shape to derive the aligned reference image. For implementation detail, we refer to the work of Lu et al.[21] We define an abstract skeleton system on the 3D body shape, where we drop the value in z-dimension to unify the skeleton systems of the 2D aligned reference image and the 3D body shape. We extract local fat percentages $Fat_{GT}^l\%$ from the aligned pixel-level body composition reference image. We extract the corresponding level total stack volumes V^l from the 3D body shape. We take 9 key points per bone (Fig. 3-(a) cyan) for local fat estimation and B-spline skinning. We calculate the target fat-free level stack volumes $\hat{V}_{FFM}^l$ through Eq. (2). We search for the deformation parameter $\boldsymbol{\alpha}$ such that level stack volumes of the morphed fascia layer $\tilde{V}_{FFM}^l(\boldsymbol{\alpha})$ minimize the data compliance error in Eq. (3). Finally, we deform the muscle system according to the deformation of the fascia layer. The body composition extrapolation results for the six test subjects are illustrated in Fig. 6.

To evaluate the accuracy of our method, we densely sample the local fat distribution of the extrapolated 3D body composition for each subject with 99 samples on each body part (i.e., the trunk, left leg, right leg). Correspondingly, we calculate the local fat distribution from the aligned pixel-level body composition reference as the ground truth. In Fig. 7, we demonstrate the estimated local fat distribution versus the ground truth fat distribution for each validation sample point. As shown in Tab. 2, the average R-squared value for the different body parts is 0.985, which indicates there is a good agreement between the extrapolated local fat distribution and the ground truth local fat distribution. The average Root Mean Square Error (RMSE) is 1.19% and the average Mean Absolute Error (MAE) is 0.81%, from which we conclude that our 3D body composition extrapolation method has a high degree of accuracy.

4. CONCLUSIONS

We present an innovative method to extrapolate 3D voxel-level body composition with a high degree of accuracy based on the 3D body shape and 2D pixel-level body composition reference. Our work provides a practical solution to assess the voxel-level body composition, which would typically be generated with CT or MRI, using easy-to-access, cost-effective equipment such as the Kinect. The work of Lu et al.[21] has shown the feasibility of predicting the DXA-like 2D pixel-level body composition image based on the 3D body shape. Therefore, our method can be totally DXA independent by adopting the inferred 2D pixel-level body composition. There are several limitations of our work that provide opportunities for further research. First, the 3D body composition template is generated by the artist modeling rather than by the real data derived from the CT or MRI scans. Second, there are small artifacts during registration such as the mesh penetration among different body compositions. Third, we do not model the visceral fat distribution in our 3D body composition extrapolation method. Quantifying and locating the visceral adipose tissue are extremely important since this type of fat has been considered a significant factor for the etiology of various metabolic and cardiovascular diseases.

ACKNOWLEDGMENTS

REFERENCES

[1]. "3dMD," (2016). http://www.3dmd.com/.

[2]. "TC2," (2016). https://www.tc2.com.

[3]. "TELMAT," (2018). http://www.tehnat.fr.

[4]. Wang Q, Wang Z, Yao Z, Forrest J, and Zhou W, "An improved measurement model of binocular vision using geometrical approximation," Measurement. Science and Technology 27(12), 125013 (2016).

[5]. Wang Q, Wang Z, and Smith T, "Radial distortion correction in a vision system," Applied optics 55(31), 8876–8883 (2016). [PubMed: 27828289]

[6]. Loeffler M, Engel C, Ahnert P, Alfermann D, Arelin K, Baber R, Beutner F, Binder H, Brähler E, Burkhardt R, et al., "The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany," BMC public health 15(1), 691 (2015). [PubMed: 26197779]

[7]. "Civilian American and European Surface Anthropometry Resource (CAESAR) final report," (2002). https://www.humanics-es.com/CAESARvoll.pdf.

[8]. Park B-K and Reed MP, "Parametric body shape model of standing children aged 3–11 years," Ergonomics 58(10), 1714–1725 (2015). [PubMed: 25933223]

[9]. "SizeUK," (2018). https://www.arts.ae.uk/research/current-research-and-projects/fashion-design/sizeuk-results-from-the-uk-national-sizing-survey.

[10]. "SizeGERMANY," (2018). https://portal.sizegermany.de/SizeGermany/pages/home.seam.

[11]. Hanson L, Sperling L, Card G, Ipsen S, and Vergara CO, "Swedish anthropometries for product and workplace design," Applied ergonomics 40(4), 797–806 (2009). [PubMed: 18977470]

[12]. LEE HP, GARG S, CHHUA N, and TEY F, "Development of an anthropometric database representing the Singapore population,"

[13]. "IBV 3D SURVEYS," (2016). https://antropometria.ibv.org/en/3d-surveys/.

[14]. "FIT3D," (2016). https://www.fit3d.com/.

[15]. "Styku," (2016). https://www.tc2.com.

[16]. "ShapeScale," (2018). http://www.tehnat.fr.

[17]. Newcombe RA, Izadi S, Hilliges O, Molyneaux D, Kim D, Davison AJ, Kohi P, Shotton J, Hodges S, and Fitzgibbon A, "Kinectfusion: Real-time dense surface mapping and tracking," in [Mixed and augmented reality (ISMAR), 2011 10th. IEEE international symposium on], 127–136, IEEE (2011).

[18]. C-ui Y, Chang W, Nöll T, and Stricker D, "KinectAvatar: fully automatic body capture using a single kinect," in [Asian Conference on Computer Vision], 133–147, Springer (2012).

[19]. Li H, Vouga E, Gudym A, Luo L, Barron JT, and Gusev G, "3d self-portraits," ACM Transactions on Graphics (TOG) 32(6), 187 (2013).

[20]. Dou M, Taylor J, Fuchs H, Fitzgibbon A, and Izadi S, "3D scanning deformable objects with a single RGBD sensor," in [Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition], 493–501 (2015).

[21]. Lu Y, Zhao S, Younes N, and Hahn JK, "Accurate nonrigid 3d human body surface reconstruction using commodity depth sensors," Computer Animation and Virtual Worlds, e1807.

[22]. Dou M, Fuchs H, and Frahm J-M, "Scanning and tracking dynamic objects with commodity depth cameras," in [Mixed and Augmented, Reality (ISMAR), 2013 IEEE International Symposium on], 99–106, IEEE (2013).

[23]. Li Wei, X. X and James HK, "3d reconstruction and texture optimization using a sparse set of rgb-d cameras," in [Winter Conference on Applications of Computer Vision (WACV)], IEEE (2019).

[24]. Allen B, Curless B, and Popovi Z, "The space of human body shapes: reconstruction and parameterization from range scans," ACM transactions on graphics (TOG) 22(3), 587–594, ACM (2003).

[25]. Anguelov D, Srinivasan P, Koller D, Thrun S, Rodgers J, and Davis J, "Scape: shape completion and animation of people," in [ACM transactions on graphics (TOG)], 24(3), 408–416, ACM (2005).

[26]. Ballester A, Parrilla E, Uriel J, Pierola A, Alemany S, Nacher B, Gonzalez J, and Gonzalez JC, "3d-based resources fostering the analysis, use, and exploitation of available body anthropometric data," in [5th international conference on 3D body scanning technologies], (2014).

[27]. Park B-KD, Reed M, Kaciroti N, Love M, Miller A, Appugliese D, and Lumeng J, "shapecoder: a new method for visual quantification of body mass index in young children," Pediatric obesity 13(2), 88–93 (2018). [PubMed: 27900848]

[28]. Löffler-Wirth H, Willscher E, Ahnert P, Wirkner K, Engel C, Loeffler M, and Binder H, "Novel anthropometry based on 3d-bodyscans applied to a large population based cohort," PloS one 11(7), e0159887 (2016). [PubMed: 27467550]

[29]. Lu Y, McQuade S, and Hahn JK, "3d shape-based body composition prediction model using machine learning," in [2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)], 3999–4002 (7 2018).

[30]. Ng B, Hinton B, Fan B, Kanaya A, and Shepherd J, "Clinical anthropometries and body composition from 3d whole-body surface scans," European journal of clinical nutrition 70(11), 1265 (2016). [PubMed: 27329614]

[31]. Piel M, Predictive Modeling of Whole Body Dual Energy X-Ray Absorptiometry from 3D Optical Scans Using Shape and, Appearance Modeling, PhD thesis, University of California, San Francisco (2017).

[32]. Siri W, "Body composition from fluid spaces and density: analysis of methods. 1961.," Nutrition (Burbank, Los Angeles County, Calif.) 9(5), 480 (1993).

[33]. Wang Y, Jacobson A, Barbi J, and Kavan L, "Linear subspace design for real-time shape deformation," ACM Transactions on Graphics (TOG) 34(4), 57 (2015).

[34]. Si H, "Tetgen, a delaunay-based quality tetrahedral mesh generator," ACM Transactions on Mathematical Software (TOMS) 41(2), 11 (2015).
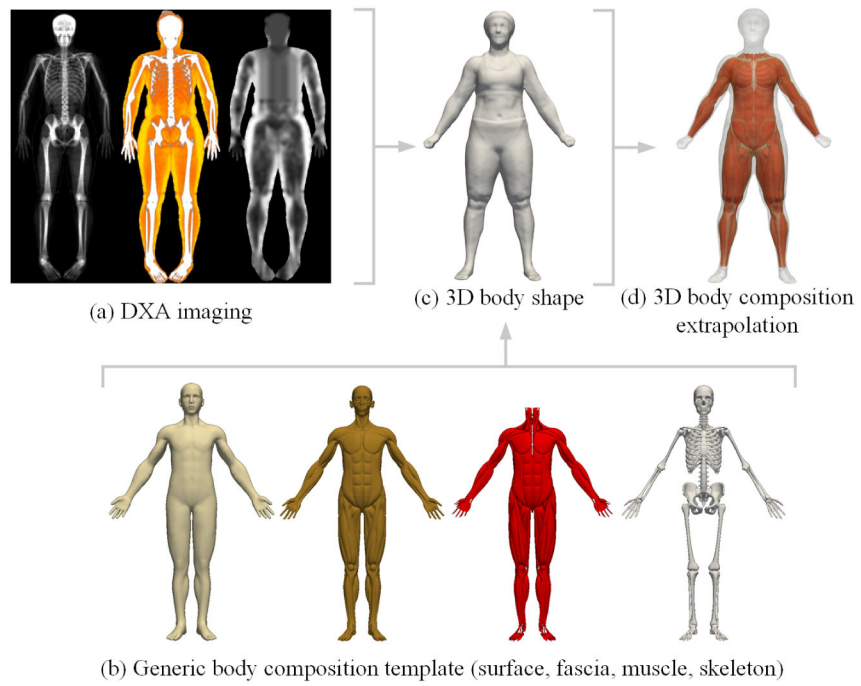
(a) DXA imaging

(c) 3D body shape

(d) 3D body composition extrapolation

(b) Generic body composition template (surface, fascia, muscle, skeleton)

Figure 1.
Schematic of multimodality registration for 3D body composition extrapolation.

Shape compliance optimization

Data compliance optimization

Gender-specific body composition initialization → Conditional anthropometries transfer → 3D-3D body composition template registration → 3D-2D data compliance optimization
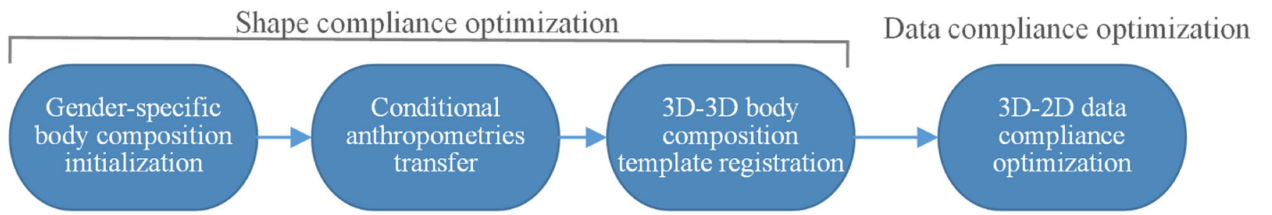
Figure 2.
3D body composition extrapolation pipeline.
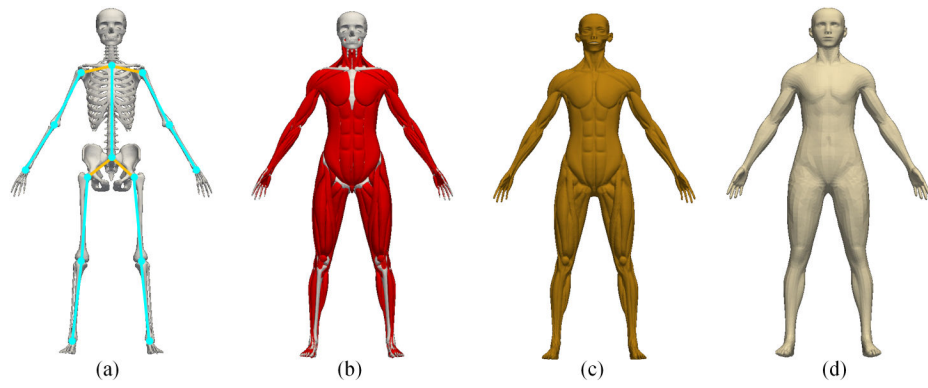
Figure 3.
The gender-specific body composition template of females. (a) The skeleton system and abstract skeletal structure. (b) The muscle system. (c) The fascia layer. (d) The skin layer.
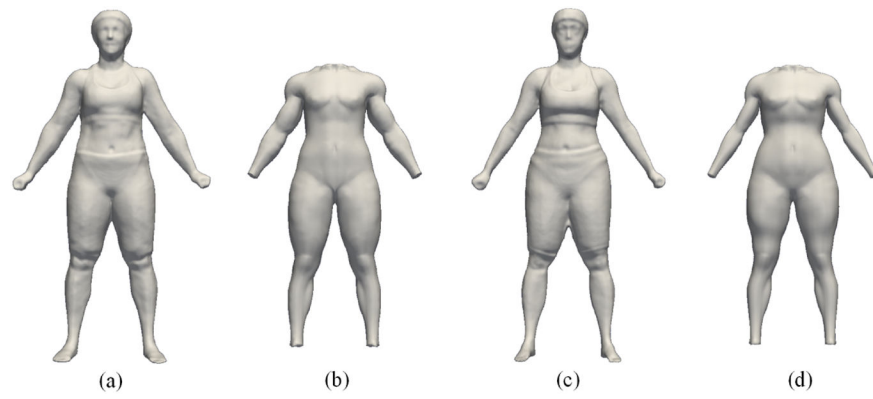
(a)　　　　　(b)　　　　　(c)　　　　　(d)

Figure 4.
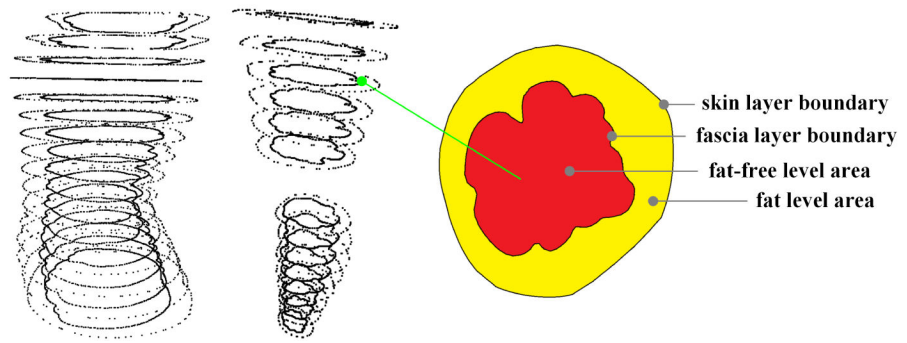The 3D body shapes (a, c) and corresponding measurements-transferred skin layers (b, d).

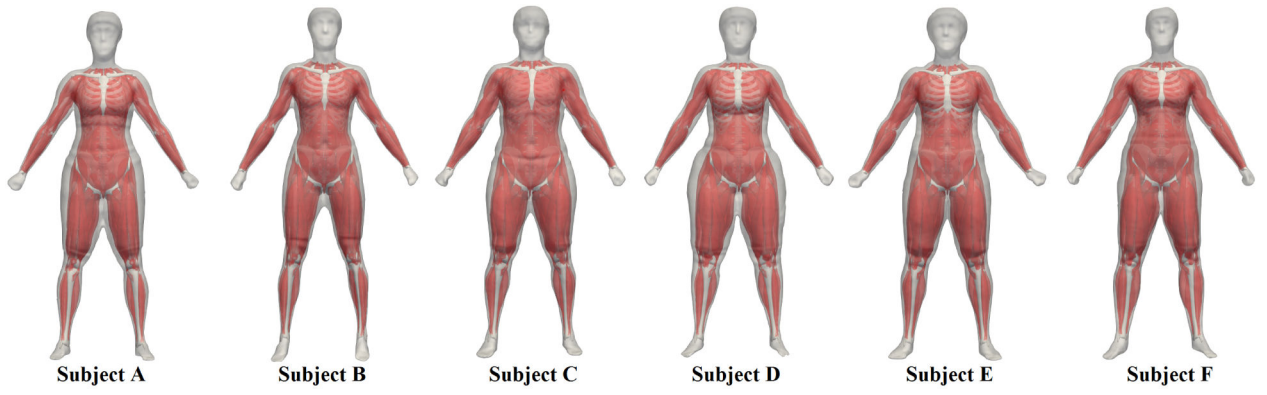**Figure 5.**
The mechanism for fat-free volume estimation.

Subject A **Subject B** **Subject C** **Subject D** **Subject E** **Subject F**

Figure 6.
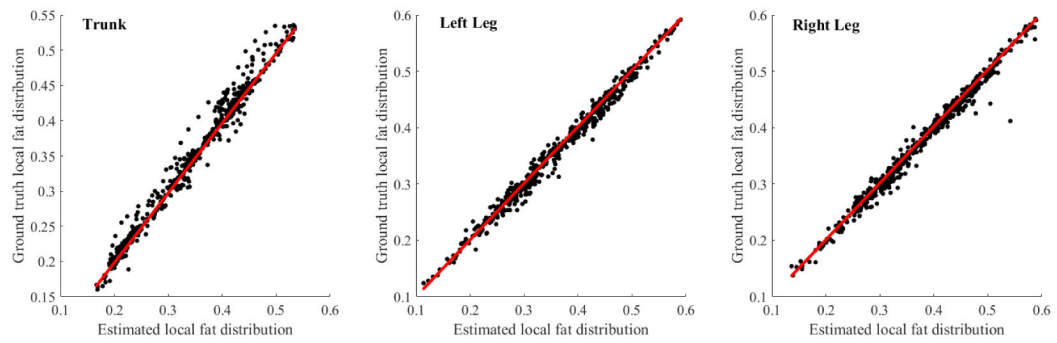Visualization of 3D body composition extrapolation results.

Figure 7.

Estimated local fat distribution versus the ground truth fat distribution for the trunk, left leg and right leg.

Table 1.

The Body Mass Index (BMI) and Body Fat Percentage (BFP) distributions of the subjects.

	Subject A	Subject B	Subject C	Subject D	Subject E	Subject F
BMI	21.5	28.2	21.9	31.0	29.0	29.9
BFP	24.8%	41.2%	15.1%	42.7%	33.0%	38.1%

Table 2.

3D body composition extrapolation accuracy evaluation.

	Trunk	Left Leg	Right Leg	Mean
R-squared	0.981	0.988	0.986	0.985
Root Mean Square Error	1.307%	1.086%	1.189%	1.19%
Mean Absolute Error	0.853%	0.799%	0.792%	0.81%