ORIGINAL PAPER

Anthropometric clothing measurements from 3D body scans

Song Yan[1] · Johan Wirta[2] · Joni-Kristian Kämäräinen[1]

Abstract
We propose a full processing pipeline to acquire anthropometric measurements from 3D measurements. The first stage of our pipeline is a commercial point cloud scanner. In the second stage, a pre-defined body model is fitted to the captured point cloud. We have generated one male and one female model from the SMPL library. The fitting process is based on non-rigid iterative closest point algorithm that minimizes overall energy of point distance and local stiffness energy terms. In the third stage, we measure multiple circumference paths on the fitted model surface and use a nonlinear regressor to provide the final estimates of anthropometric measurements. We scanned 194 male and 181 female subjects, and the proposed pipeline provides mean absolute errors from 2.5 to 16.0 mm depending on the anthropometric measurement.

1 Introduction

Anthropometric measurements, such as chest and hip circumference or shoulder-to-shoulder distance, provide detailed information about the body shape. The body shape information is essential for industrial design [18], clothing design [8], medical sciences [16] and ergonomics [19]. The measurements have traditionally been made manually from physical subject using a tape measure, but the raise of online shopping and personalized tools set new demand for computerized anthropometric measurements.

A standard pipeline for computerized anthropometric measurements is the following [3,7,10,25–27]: (1) a 2D or 3D body scan producing a 3D point cloud or an initial model, (2) fitting of a pre-defined model and (3) measurements from the fitted model. The main challenge is the step two which should provide an accurate and watertight volumetric model of a subject so that impor-

tant measurements can be made on the model surface. Challenges arise from different sensor modalities, poses and occluded regions. The proposed method in this work shares the main steps of the standard pipeline (Fig. 1), but instead of physiologically valid model fit, we adopt a non-rigid iterative closest point (ICP) registration between the model and captured point clouds. Moreover, we do not make anthropometric measurements directly from the fitted model surface, but extract a set of physiologically meaningful surface features (body circumferences) and use them to train a regressor that provides estimates of the physical anthropometric measurements. Our main contributions are:

– A full processing pipeline from 3D body scans to anthropometric measurements.
– The body model registration step using a non-rigid ICP to fit a pre-defined model to captured body scans.
– Nonlinear regression based anthropometric measurement estimation step from circumference based intermediate features.
– A public benchmark dataset—NOMO3D—with anthropometric measurement ground truth.

Our pipeline is evaluated with the NOMO3D dataset of real male and female subjects (194 plus 181) for which we provide average accuracy and percentage of subjects whose accuracy is below the thresholds in [9].
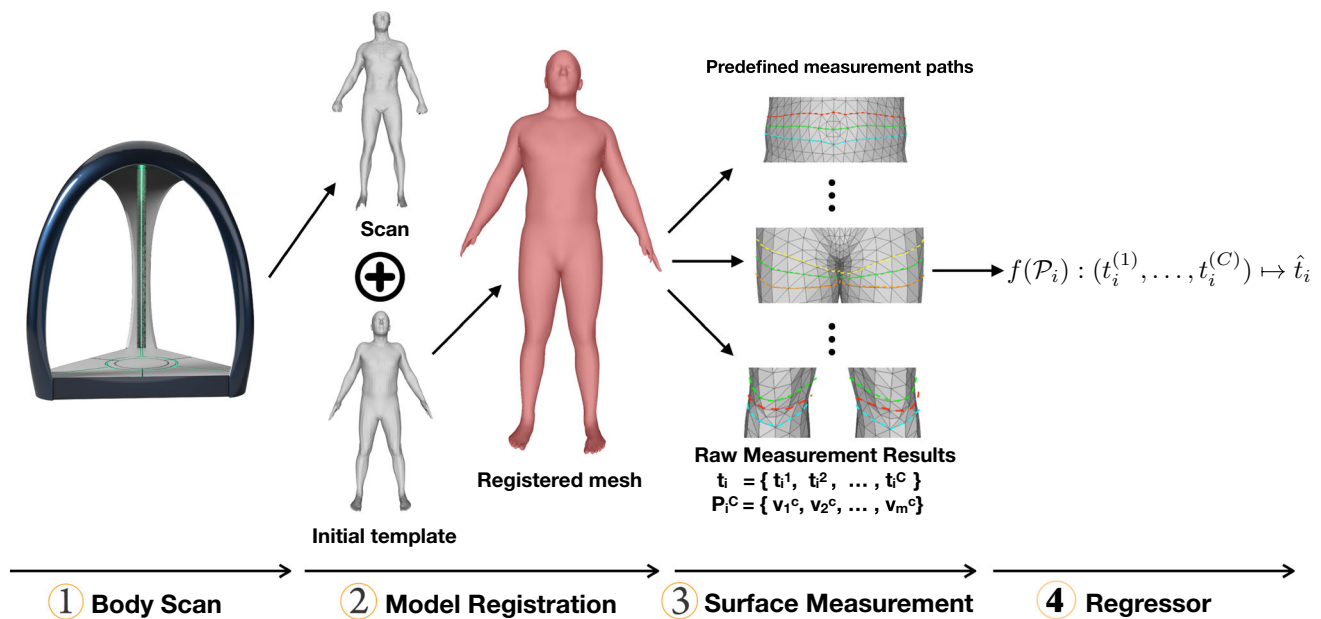
✉ Song Yan
song.yan@tuni.fi

Johan Wirta
johan.wirta@nomo3d.com

Joni-Kristian Kämäräinen
joni.kamarainen@tuni.fi

1 Computing Sciences, Tampere University, Tampere, Finland

2 NOMO Technologies Ltd, Espoo, Finland

Fig. 1   The proposed pipeline for measuring anthropometric clothing measurements from 3D body scans. A 3D point cloud is produced by a set of depth sensors (body scanner). A body template is fitted (registered) to the 3D point cloud (step-2); circumference measurements are computed on the model surface (step-3); supervised regression is adopted to provide estimates of anthropometric measurements (step-4)

2 Related work

Anthropometric measurement datasets   There have been several campaigns to collect 3D body scans and anthropometry ground truth for them. For example, the UMTRI dataset was collected to find the safest sitting posture of young children in cars [12]. ANSUR 88 (1988) and ANSUR 2012 datasets contain 3D body scans and tape measured anthropometric measurements of US Army Force soldiers. ANSUR 2012 contains 4082 male and 1986 female subjects of varying age and 93 ground truth anthropometric measurements for each of them. Unfortunately, UMTRI and ANSUR datasets are not publicly available. CAESAR dataset [21] is a commercial counterpart of ANSUR and contains 3D scans of 2400 US & Canadian and 2000 European civilians with tape measured ground truth. CAESAR (http://store.sae.org/caesar/) has been used in various scientific works but has not been widely adopted in benchmarking due to its price. The main usage of UMTRI, ANSUR and CAESAR datasets is to make "virtual tape measurements" on the point cloud surface. In the follow-up work of the CAESAR, Robinette and Daanen [20] compared virtual tape measurements over two different scanners and scanning teams and showed that measurements are highly reproducible within the US Army defined error limits (cf. ANSUR experiments). Reproducibility error in their experiments was less than ±5 mm for the most measurements. However, these were relative accuracies over repeated tests. Simmons and Istook [23] noted that there

is substantial variation in available softwares how to measure anthropometric measurements from 3D data. Paquette et al. [17] demonstrated much larger errors for 3D measurements as compared to manual tape measurements. They reported systematic errors of up to 30–40 mm despite the fact that standard measurement procedures were implemented to the softwares (ISO-8559 and US Army).

3D Human body models   The early works following the data campaigns above were based on "virtual tape measurements" where the anthropometric measurements were made manually with the help of 3D measurement software. If this step needs to be automated, then 3D scan data need to be aligned with a model for which the measurement paths can be predefined using 3D model vertex ids. However, first a good 3D human body model needs to be devised. The model should contain intuitive parameterization for shape and pose and provide realistic body shapes. There are several options for scientific work. The most popular parametric body model is MakeHuman which is an open source project (http://www.makehuman.org/) based on an artistic body model and aiming at high quality rendering for games and movies.

However, better models are based on statistics of real human data. These require a single artist made initial point model which is iteratively matched to scanned point clouds in a normalized pose. Principal component analysis (PCA) over the matched model points provides natural parameterization for the shape. The pose can be intuitively defined by

a skeleton joint model, but the final quality depends on how well the model can represent pose-specific shape deformations. One of the first attempts to create a 3D human body from PCA shape and skeleton pose is the SCAPE body model by Anguelov et al. [2]. Hirshberg et al. [11] proposed a better parametric body model for SCAPE and introduced the Blend-SCAPE model. Other attempts are by Baek and Lee [3] and more recently the SMPL model by Loper et al. [15]. SMPL provides high-quality models where the shape is divided to pose invariant and pose-dependent deformations and the model parameters are optimized using a combination of their own dataset of 1786 scans and 3800 scans from CAESAR. For this work, we adopt the SMPL model due to its good overall quality.

Computerized anthropometric measurements   There have been several attempts to infer 3D body models from 2D RGB images. For example, Guan et al. [10] proposed a method and compared their measurements to the ground truth. However, for many industrial and commercial applications, the accuracy of 2D measurements is insufficient. For better accuracy, 3D scans are needed.

Weiss et al. [26] propose a Kinect-based 3D body scan method that uses the SCAPE body model. The method requires manual pose initialization and then optimizes the model mesh using a standard ICP. Tsoli et al. [25] propose a pipeline that is similar to ours. They use the BlendSCAPE model to register a 3D scan, and then, they compute various local and global features which are used in regression. A different approach was proposed by Zuffi et al. [32] in their "stitched puppet" model where the body model is divided to local templates where "local PCA" matching is performed and then the local parts are globally aligned in the next optimization step. Wuhrer et al. [27] introduce an inverse problem of ours where a 3D body model is estimated from the given 1D anthropometric measurements.

The above works particularly address the problem of unknown pose. However, we believe that a fixed pose can be assumed for many applications since customers can be assumed co-operative. Therefore, the process can be drastically simplified and provide accurate results.

3 3D body scanning

Recently, novel single depth sensor-based body scanning approaches have been proposed, for example, BodyFusion [28] and DoubleFusion [29], but since 3D scanning is out of the scope of this work, a commercial 3D body scanner was used. Our dataset was collected using a commercial TC2 body scanner (https://www.tc2.com) that uses off-the-shelf depth sensors (Intel RealSense R200). Inside the scanner, subjects were instructed to step on the rotating
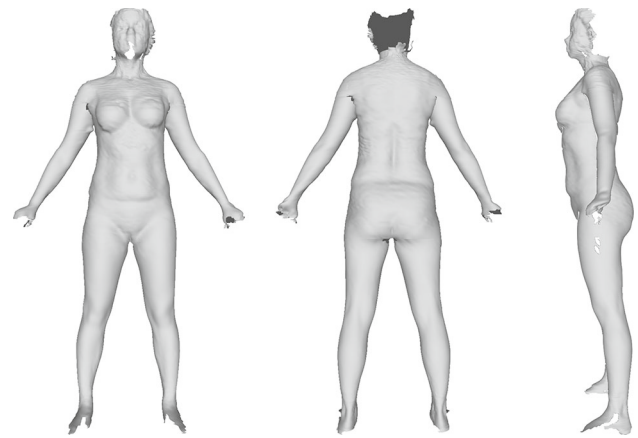

Fig. 2 A 3D body scan (point cloud) captured by TC2 body scanner. The scanner covers most of the body surface and missing parts occur only in the head and feet regions

platform and take a standing pose with the feet at around their shoulder width apart and the arms slightly raised to create a gap between the arms and torso. The platform then rotates around once, during which three depth sensors produce a raw 3D scan of the customer, and the process takes a few seconds (Fig. 2). The test subjects wore tight fitting underwear-like sport costumes. The scanner outputs a triangulated mesh structure in the regular OBJ file format. Each triangulated mesh contains on average 57,000 vertices and around 113,000 faces. For our experimental studies, we scanned 194 men and 181 women. Scanned persons were instructed to wear tight underwear.

4 Model registration

4.1 SMPL body model

The popular 3D human body models MakeHuman, SCAPE [2], BlendSCAPE [11] and SMPL [15] (see Sect. 2 for details) share similar model parameterization $\{\mathcal{T}, \mathcal{S}, \boldsymbol{\theta}\}$ where $\mathcal{T}$ is the initial model in a "canonical shape" and "canonical pose," $\mathcal{S}$ defines the shape deformation and $\boldsymbol{\theta}$ defines the pose. Pose parameterization is intuitive and typically based on a skeleton rig of K skeleton joints. A pose is encoded to the 3D rotation angles of K joints in $\boldsymbol{\theta}$. Each vertex location in $\mathcal{T}$ is relative to a specific skeleton part or parts, and therefore, the whole point cloud deforms. Parameterization of the shape is more difficult to model since parameters need to capture shape statistics of the human population. The standard approach is to use principal component analysis (PCA) where principal components represent the most important axes of variation in the population. In the PCA space, any shape can be reconstructed by linearly adding $|\boldsymbol{\beta}|$ principal directions to a mean shape $\mathcal{T}$ (the zero shape):
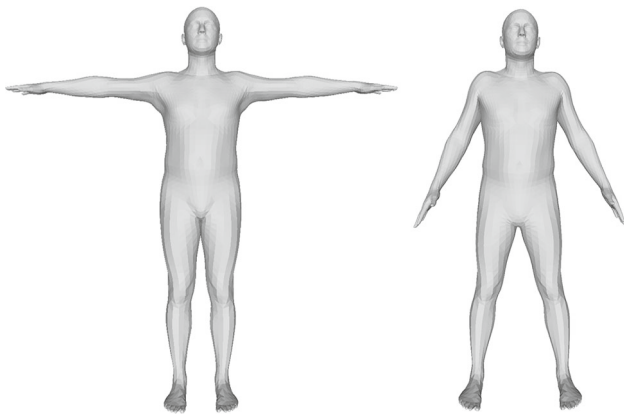
Fig. 3 We adopted the skinned multi-person linear model (SMPL) [15] for our framework since it provides intuitive model parameterization and high-quality models. A SMPL model in its canonical (zero-pose) position (left) and the model in the initial position that corresponds to the instructed pose in our body scans (right)

$$\mathcal{T} + B(\boldsymbol{\beta}) = \mathcal{T} + \sum_{n=1}^{|\boldsymbol{\beta}|} \beta_n \mathcal{S}_n \ . \tag{1}$$

Often as few as $|\boldsymbol{\beta}| = 10$, principal component vectors provide sufficient accuracy for applications where subtle details are not important. For our work, we selected the *skinned multi-person linear model* (SMPL) by Loper et al. [15] since it provides very competitive accuracy and the original implementation is publicly available.

SMPL mesh model contains $N = 6890$ vertices (13,766 faces) and $K = 23$ skeleton joints. The mesh has the same topology for men and women, spatially varying resolution, a clean quad structure, segmentation into parts, initial blend weights and a skeletal rig. A particular detail that makes SMPL registration more accurate than its competitors is that it divides the shape deformation to pose-independent deformation $B_S(\boldsymbol{\beta})$ and pose-specific deformation $B_P(\boldsymbol{\theta})$ which are summed to define the final shape. Notably, the shape deformation parameters are also used to predict the rotations of the $K = 23$ skeleton joints $J(\boldsymbol{\beta}) : \mathbb{R}^{|\boldsymbol{\beta}|} \rightarrow \mathbb{R}^{3K}$. We re-defined the SMPL zero-pose to correspond to the pose subjects were instructed to take (Fig. 3).

4.2 Non-rigid ICP registration

The goal of the body model registration to the scanned point cloud is to provide "skin-level registration" where the two surfaces, the model and the scan, overlay almost perfectly (Fig. 4). This is a challenging task since (a) points contain measurement noise, (b) large point regions may be missing and (c) the model points do not exactly match the scan point locations. To make the final anthropometric measurements accurate in the next processing stage, we need a registration
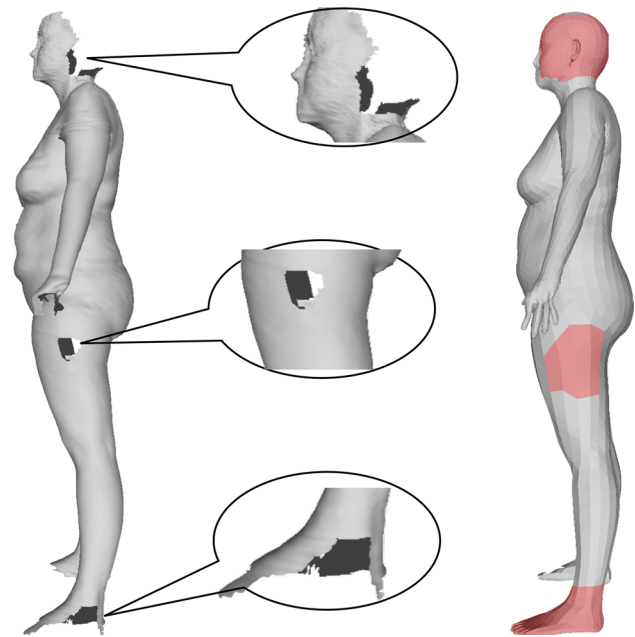

Fig. 4 A scanned point cloud contains holes and measurement noise, but registration of the 3D body model (red) is robust to these distortions and achieves an accurate—"skin level"—registration which is essential for accurate anthropometric measurements in the next stage (color figure online)

method that is accurate and robust to the aforementioned non-idealities.

A core component in constructing the SCAPE, Blend-SCAPE and SMPL datasets is an artistic-generated point model and an algorithm to register the model to real human scans. However, these algorithms perform complex optimization and must be manually initialized. Therefore, the artistic models and special algorithms have not been used outside body model generation. However, the final body models, SCAPE, BlendSCAPE and SMPL, provide intuitive parameterization as discussed in Sect. 4.1 and registration can be defined as an optimization problem where a few pose and shape parameters $\{\mathcal{S}, \boldsymbol{\theta}\}$ are optimized to minimize a registration error. Skin-level registration requires a large number of PCA components for the shape, and therefore, we take an alternative approach from the generic point cloud matching literature.

Several comparison of generic registration methods exist. For example, Bogo et al. [4] introduced the FAUST dataset for comparing non-rigid registration methods. In their experiments, several popular methods, e.g., generalized multi-dimensional scaling (GMDS) [5], Möbius voting [14] and blended intrinsic maps (BIM) [13], did not perform well since these methods assume that both inputs are watertight and have the same topology. However, the baseline point cloud matching method, iterative closest point (ICP), does not require such assumptions.

There are two extensions of the baseline ICP that are suitable for human body point clouds: Amberg et al. [1] and Schneider et al. [22]. Since the 3D scans often contain holes (Fig. 2), we adopted the Amberg et al. approach that explicitly handles missing points. The challenge is twofold—we want to retain the global convergence properties of ICP while still allow local deformations to the skin level. Local deformations make this ICP non-rigid.

The starting point of our algorithm is a pre-aligned model defined by $\{\mathcal{T}, \beta_i, \theta_k\}_{i=1,...,|\beta|, k=1,...,K}$ that brings the SMPL template to approximate correspondence with the obtained scan point cloud $\mathcal{T}_{\text{scan}}$. A simple procedure for pre-alignment is described in Sect. 4.3. If we define the pre-aligned model as $\mathcal{V}$, then the problem is to find optimal values for the alignment parameters $\mathcal{X}$ so that $\mathcal{V}(\mathcal{X})$ registers the template points to the surface points $\mathcal{T}_{\text{scan}}$.

To solve the optimal parameters $\mathcal{X}$, an energy function of three terms is defined [1]:

$$E(\mathcal{X}) = E_d(\mathcal{X}) + \alpha E_s(\mathcal{X}) + \beta E_l(\mathcal{X}) . \tag{2}$$

E_d is the standard ICP *distance term* between the model and scan points

$$E_d(\mathcal{X}) = \sum_{v_i \in \mathcal{V}} w_i \text{dist}^2 (\mathcal{T}_{\text{scan}}, \mathcal{X}_i v_i) \tag{3}$$

where $\mathcal{X}_i$ is a linear mapping of a single model vertex v_i to correspondence in $\mathcal{T}_{\text{scan}}$. w_i defines whether a model point has a correspondence in scan ($w_i = 1$) or not ($w_i = 0$). E_s is a local *stiffness term*

$$E_s(\mathcal{X}) = \sum_{i \in \mathcal{N}_j} \|(\mathcal{X}_i - \mathcal{X}_j)\text{diag}(1, 1, 1, \gamma)\|_F^2 \tag{4}$$

where $\| \cdot \|_F^2$ is the matrix Frobenius norm. The stiffness term enforces similar transformations between neighbor vertices $\mathcal{N}_j$ of the model vertex v_j. γ is used to weight differences in the rotational and skew part of the deformation against the translations part of the deformation ($\gamma = 1$ in the experiments). The third energy term is a *landmark term*

$$E_l(\mathcal{X}) = \sum_{v_i, l \in \mathcal{L}} \|\mathcal{X}_i v_i - l\|^2 . \tag{5}$$

The landmarks $\mathcal{L}$ are pre-defined and important positions in the model and this term enforces them to be registered accurately. The landmark term improves registration significantly, but requires manual labeling of selected keypoints and is therefore omitted in our experiments.

The algorithm in [1] uses locally affine regularization which assigns an affine transformation to each vertex and minimizes the difference in the transformation of neighboring vertices. The deformation parameters $\mathcal{X}$, which would be applied on source vertices to generate the target surface deformation, are obtained by minimizing the cost function in Eq. 6 directly and exactly.

$$\bar{E}(\mathcal{X}) = \left\| \begin{bmatrix} \alpha M \otimes G \\ WD \\ \beta D_L \end{bmatrix} \mathcal{X} - \begin{bmatrix} 0 \\ WU \\ U_L \end{bmatrix} \right\|_F^2 \tag{6}$$
$$= \|A\mathcal{X} - B\|_F^2$$

The cost function $\bar{E}(\mathcal{X})$ takes its minimum at $\mathcal{X} = (A^T A^{-1})A^T B$. In the above equation, M is the node-arc incidence matrix of the template mesh topology, and $G := \text{diag}(1, 1, 1, \gamma)$ is a weighting matrix, $W := \text{diag}(w_1, \ldots, w_n)$ represents the weighting matrix in which $w_i = 0$ if template vertices v_i corresponds to missing data in the target mesh and n represents the number of template vertices, D is the sparse matrix of template vertices mapping the $4n \times 3$ deformation parameters $\mathcal{X}$, U is the matrix of the correspondence points on the target mesh, D_L and U_L are the pre-defined landmarks on the template mesh and their correspondence points on the target mesh, respectively, the Kronecker product is denoted by $\otimes$. α and β are the penalty terms that balance the two corresponding energy functions with respect to the standard ICP term E_d.

The whole registration process consists of two loops. In the outer loop, a series of deformations of the template are performed for each stiffness $\alpha^i \in \{\alpha^1, \ldots, \alpha^n\}$, where $\alpha^i > \alpha^{i+1}$. These α values guarantee the registration process from a global deformation to more localized ones. In our experiments, α values are set to from 100 to 1 by step size 1. In the inner loop, a deformation $\mathcal{X}$ for a fixed stiffness term α^i and preliminary correspondences is found. Preliminary correspondences are found by a nearest point search. The optimal deformation $\mathcal{X}$ is determined until $\|\mathcal{X}^j - \mathcal{X}^{j-1}\| < \epsilon$, where ϵ is the threshold.

4.3 Pre-alignment and initialization procedures

A simple pre-alignment procedure is performed before non-rigid ICP registration. Generally, the mis-alignment of registration is partly raised by wrong scales, face orientations, the different center points of subjects. To depreciate it, first, we scale all scans into the same unit of measurement (meter) as the SMPL model meshes; we then rotate all scans to make sure that they face the same direction. Compared to the previous works which adopt the mean coordinate of vertices as the center points and align all meshes into the same center point, we additionally align all samples into the same lowest point (Z-axis). The center points change dramatically since the missing parts on scans and bring negative effects on reg-

istration. A standard point $(x, y, 0)$ is employed as the lowest point for all meshes. After the pre-alignment procedure, all scans and the SMPL models are standing on the X–Y plane and facing to Y-axis direction with the same scale.

The height of the SMPL model is controlled by the first shape parameter β_1. To obtain a suitable initial value for β_1, we utilize a simple linear function over the heights of the training set scans to estimate the parameter $\hat{\beta}_1 \approx \beta_1$. To initialize the pose parameters, we start from the pose $\boldsymbol{\theta}$ (on the right in Fig. 3) and iteratively test a number of arm angle shifts to match with the target scan. These initialization procedures aid convergence and improve accuracy, but their effect is not significant.

5 Anthropometric measurements

The proposed pipeline outputs estimates of the target physical anthropometric measurements from a fitted model (Sect. 4) by first calculating *circumference paths* through the model points (Sect. 5.1) and then estimating the physical measurements from the path distances by nonlinear regression (Sect. 5.2).

5.1 Surface measurements

The registration process brings two main benefits: (a) it produces a hole-free mesh without missing body parts and reduces the point cloud noise; and (b) registered meshes of all subjects are in the same topology that facilitates finding the corresponding vertices of the pre-defined circumference paths.

For each anthropometric measurement t_i, we define a set of surface circumference paths. The path lengths $t_i^{(1)}, \dots, t_i^{(C)}$ are used as features for regression. The paths are defined as sets of vertices in the model $\mathcal{P}_i^c = \{\boldsymbol{v}_1^c, \boldsymbol{v}_2^c, \dots, \boldsymbol{v}_m^c\}$. The length of a circumference path is the sum of edge lengths through the defined path (Fig. 5). The selected circumference paths were not optimal, but manually set near the true anthropometric measurement locations. It was assumed that multiple paths provide extra robustness to shape deformations (see the ablation study in the experimental part of our work).

5.2 Nonlinear regression

The purpose of a suitable regressor is to find a mapping $f(\cdot)$ such that

$$f(\mathcal{P}_i) : \left(t_i^{(1)}, \dots, t_i^{(C)} \right) \mapsto \hat{t}_i \tag{7}$$
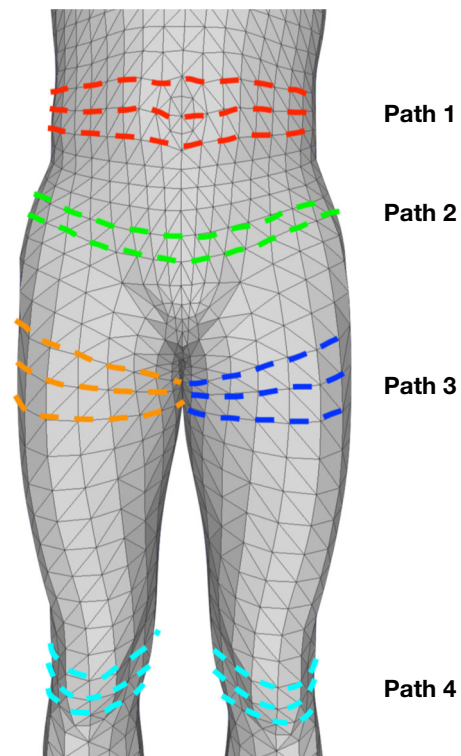
Fig. 5 Distances of circumference paths through mesh vertices of a registered SMPL body model are used as features for regression. Multiple paths (dotted red, green and blue lines) are used to estimate a single anthropometric measurement. Example circumference paths: Path 1: *NaturalWAIST*; Path 2: *Hip*; Path 3: *Thigh*; Path 4: *Knee* (color figure online)

where $\hat{t}_i$ is the estimate of the true anthropometric clothing measurement t_i. The most straightforward solution is the ordinary least squares (linear regression) which finds a solution $\boldsymbol{\omega} = (\omega_0, \omega_1, \dots, \omega_C)^{\mathrm{T}}$ that minimizes the squared loss over training subjects i

$$\sum_i \left(t_i - \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{t}_i \right)^2 \tag{8}$$

where t_i is a training set the ground truth value and $\boldsymbol{t}_i = (t_i^{(1)}, \dots, t_i^{(C)})^{\mathrm{T}}$ are the computed circumference path distances for this specific anthropometric measurement. Linear regression with regularization (ridge regression) minimizes the squared loss with a weight penalty term λ

$$\sum_i \left(t_i - \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{t}_i \right)^2 + \lambda ||\boldsymbol{\omega}|| \ . \tag{9}$$

There are also more advanced extensions of linear regression, such as elastic net regression [31], and other learning-based regressors such as support vector regression (SVR) [24]. We compare several popular regression methods in our ablations studies.

6 Experiments

6.1 Dataset and settings

We collected a set of 3D scans using the commercial scanner (Sect. 3). The dataset—NOMO3D—consists of 194 male and 181 female scans. For each subject, a clothing expert (tailor) made the actual anthropometric measurements (15 male and 19 female). All results are average performance over fivefold cross-validation.

Method evaluation We employ the mean absolute error (MAE) as the error metric between the ground truth and estimated anthropometric measurements. For each measurement i, mean absolute error ϵ_i, over all subjects j was obtained as

$$\epsilon_i = \frac{1}{|j|} \sum_{j=1}^{|j|} \left| t_i^{(j)} - \hat{t}_i^{(j)} \right|. \tag{10}$$

In addition to the measurement-specific MAEs, we also computed the average MAEs over all measures. All numbers were measured in millimeters (mm). Moreover, for each measurement, we also report the proportion of the test samples for which the accuracy was below the defined error limits in [9] as *Success rate*.

Computational complexity The most time-consuming part is the non-rigid ICP registration. Matlab code was adapted from [1] and it runs approximately 2 min on each scan. The pre-alignment and initialization procedures are very fast, less than a second, as well as the regression which is also computationally fast.

6.2 Results

The average fivefold errors for each anthropometric measurement and their accuracy thresholds and success rates are shown in Table 1. In all cases, the number of surface measurements were optimized for each anthropometric measurement and the best performing regressor (nonlinear SVR) was used. For the both male and female subjects the best performing measurement was neck circumference with 93% test cases below the threshold (6 mm) for men and 81% for women. The worst performing measure was ankle circumference for which only 28% of male 24% of female success rates were achieved. The error distributions for the male and female neck and ankle circumferences and male chest and female natural waist circumferences are shown in Fig. 6. The distributions reveal that there exists a small amount of test samples with a large error. It turned out that the main source of large esti-

mation errors yields from the body scanner that often misses certain body parts. For example, feet regions often lack point cloud points which makes the registration fail in these regions (Fig. 7).

6.3 Ablation study

Number of circumference paths In the first ablation study, we investigated the effect of adding multiple surface measurements (circumference paths) to the anthropometric regression. The results for three well and three poorly performing measurements for the both male and female are shown in Fig. 8. Results are for *nonlinear SVR* regressor with fivefold cross-validation. The most important findings are that additional paths always improve the accuracy and depending on the measurement the results saturate at 3–9 surface circumference paths. In particular, paths close to the physical anthropometric measurement location strongly contribute to the estimation accuracy. The best single paths ($C = 1$) were also selected using cross-validation and the results with and without SVR regression are shown in Table 1. These results indicate that (i) the multi-path regression is superior to single path regression and (ii) SVR significantly improves the estimation performance.

Non-rigid ICP To validate the importance of non-rigid ICP, we conducted an experiment where the SMPL model was directly fitted to the point clouds. SMPL parameter optimization was done using the popular L-BFGS-B optimizer [30]. Similar to the non-rigid ICP, the *distance term E_d* with the normal direction constraints was used as the target function. The stop criterion was set to 10^{-6} to keep the computation times reasonable, and the same pre-alignment procedure was adopted. The results are shown in Table 1 and are clearly inferior to the proposed non-rigid ICP registration.

Regression methods We compared a number of publicly available regression methods for the regression step. The standard linear regressors were *linear regression*, *stepwise linear regression* and *ridge regression*, and more recent regression methods are *elastic net linear regression*, *Gaussian process regression (GPR)*, *binary regression decision tree (BRDT)*, *linear support vector regression (SVR)* and *Nonlinear SVR*. The mean accuracy and success rates for these methods are shown in Table 2. The results show that even the basic linear regressors (linear regression, ridge regression and stepwise linear regression) perform well indicating that the proposed registration step performs well. Nonlinear SVR and Gaussian process regression also perform well. They are all safe choices for regressing anthropometric measurements from surface measurements, but we

Table 1 Average fivefold (80% for training and 20% for testing) performance (mean absolute error) and success rate (a proportion of the test samples within the error limits in [9]) of anthropometric measurements

Measure	Best single (C = 1)		Best single + SVR		L-BFGS-B + SVR			Multiple paths + SVR			Limit [9]
	MAE (mm)	%	MAE (mm)	%	C	MAE (mm)	%	C	MAE (mm)	%	(mm)
Male											
Ankle circ.	36.4	0	8.4	**37.6**	6	12.3	27.5	6	**7.7**	28.6	4
Bicep circ.	7.8	45.8	6.5	**57.8**	8	19.4	16.2	8	**6.1**	57.6	6
Calf circ.	6.9	41.2	4.0	70.3	6	17.7	16.8	6	**3.0**	**82.2**	5
Chest circ.	15.6	60.0	15.5	61.6	5	43.3	22.9	5	**14.3**	**63.7**	15
Elbow circ.	4.0	56.5	3.7	62.8	8	11.4	16.2	8	**2.6**	**77.9**	4
Hip circ.	9.2	71.7	9.3	**75.4**	4	32.6	22.9	4	**8.8**	73.3	12
Knee circ.	8.5	28.3	6.0	44.0	6	15.5	12.9	6	**5.1**	**46.6**	4
NaturalWaist circ.	15.8	49.0	13.2	57.3	4	50.0	15.6	4	**12.8**	**57.6**	12
NeckBase circ.	35.1	4.2	10.2	61.0	3	15.7	43.6	3	**8.0**	**72.6**	11
Neck circ.	3.0	92.1	3.0	91.1	4	16.3	22.4	4	**2.5**	**93.7**	6
Thigh circ.	10.6	31.4	10.5	32.5	8	27.9	16.8	8	**7.9**	**48.7**	6
TrouserWaist circ.	25.5	–	12.0	–	3	36.4	–	3	**9.1**	–	–
Wrist circ.	7.2	43.2	5.2	57.8	6	6.6	49.2	6	**4.5**	**67.2**	5
Shoulder_to_shoulder	13.7	–	13.8	–	4	18.0	–	4	**12.0**	–	–
Shoulder_to_wrist	40.3	–	14.7	–	6	27.3	–	6	**12.7**	–	–
Avg.	16.0	43.6	9.1	59.1		23.4	23.6		**7.8**	**64.1**	
Female											
Ankle circ.	18.8	14.4	14.3	23.0	6	17.7	14.7	6	**13.4**	**24.7**	4
Bicep circ.	19.7	8.5	7.9	48.3	8	15.9	25.4	8	**4.9**	**73.9**	6
Calf circ.	7.3	37.4	3.8	70.7	6	18.0	19.8	6	**3.0**	**82.8**	5
Bust circ.	17.3	44.0	15.2	60.6	3	42.1	19.2	3	**12.0**	**71.4**	15
Elbow circ.	4.5	57.4	4.5	59.7	6	11.7	22.0	6	**3.4**	**70.5**	4
Hip circ.	18.7	26.3	8.9	70.9	4	37.0	21.5	4	**8.9**	**71.4**	12
Knee circ.	9.9	21.1	6.9	39.4	6	17.3	22.0	6	**5.9**	**41.1**	4
NaturalWaist circ.	13.7	55.7	12.8	56.3	5	41.0	16.4	5	**12.0**	**59.7**	12
NeckBase circ.	58.8	0.6	10.6	**63.6**	3	13.0	54.2	3	**10.2**	62.5	11
Neck circ.	6.3	67.1	5.5	74.0	5	13.4	32.2	5	**4.8**	**81.5**	6
Thigh circ.	10.1	35.8	9.7	39.2	8	29.9	13.6	8	**7.9**	**46.3**	6
TrouserWaist circ.	15.6	–	15.4	–	3	38.0	–	3	**14.8**	–	–
Wrist circ.	6.0	49.1	5.0	59.4	8	6.9	40.7	8	**4.4**	**65.7**	5
UnderBust circ.	14.2	69.5	14.3	69.5	2	34.3	27.1	2	**13.4**	**71.8**	16
Shoulder_to_shoulder	26.5	–	13.8	–	4	17.9	–	4	**12.7**	–	–
Shoulder_to_wrist	22.4	–	16.8	–	4	25.5	–	4	**13.7**	–	–
Bust_to_bust	12.2	46.0	11.6	54.6	9	15.7	39.6	9	**10.4**	**57.5**	10
NeckSide_to_wrist	26.4	–	16.8	–	4	25.8	–	4	**16.0**	–	–
NeckSide_to_bust	13.9	30.9	13.4	35.4	6	17.7	24.9	6	**13.0**	**36.6**	8
Avg.	17.0	37.6	10.9	55.0		23.1	26.2		**9.7**	**61.2**	

Bold denote the best result among several methods

"*C*" denotes the number of circumference paths used in estimation. "best single" is the best single path performance. "+SVR" uses SVR regression for the estimates. "L-BFGS-B + SVR" uses the SMPL model fitted by the L-BFGS-B optimizer
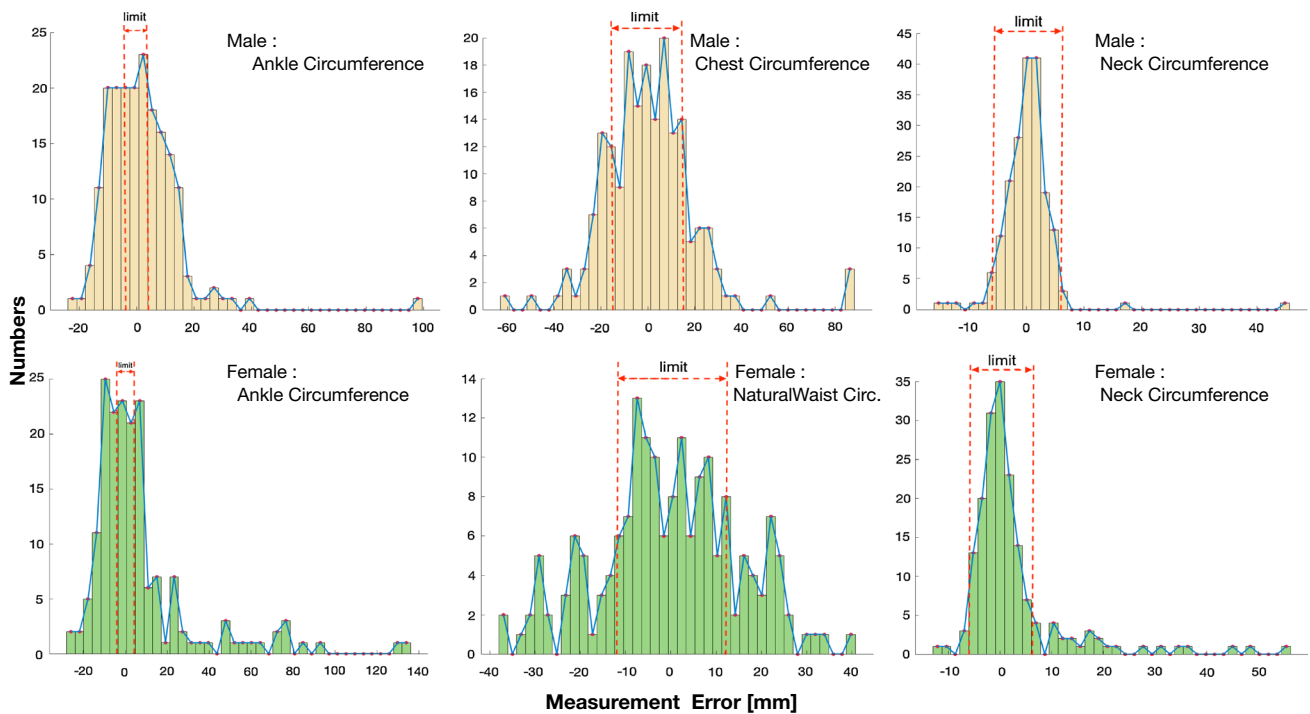
Fig. 6 Error distributions illustrating low, moderate and well performing estimates. Top (male): *ankle circumference* (low), *chest circumference* (moderate), *neck circumference* (high); bottom (female): *ankle circumference* (low), *natural waist circumference* (moderate) and *neck circumference* (high). The red vertical lines denote the acceptance thresholds in [9] (color figure online)
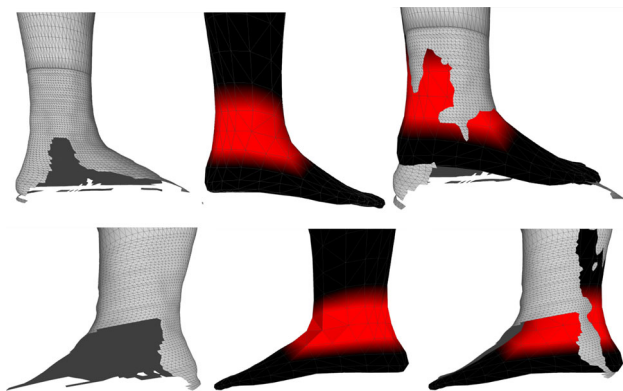

Fig. 7 Two examples of registration failures due to missing points in the scanned point clouds: scanned point cloud (left), model emphasizing the ankle circumference location (middle) and output of the registration process (right) that illustrates the failure cases (color figure online)

Fig. 8 Test set errors (MAEs) as functions of the number of surface measurements (circumference paths) for three well performing (success > 50%) (top) and three poorly performing (< 50%) anthropometric measurements (bottom)

selected the nonlinear SVR due to its best overall performance.

7 Conclusions

This work introduced a full processing pipeline for estimating physical anthropometric measurements from 3D body scans. The pipeline consisted of a commercial 3D scanner, a
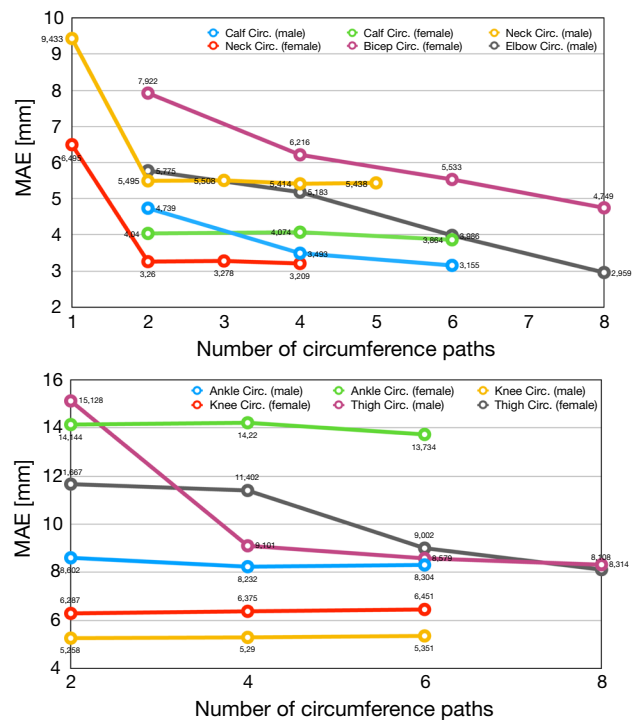
Table 2 Average MAEs and success rates of several regression methods

Non-lin SVR		Ridge reg.	Lin reg.	Stepw. reg.	GPR	ElasticNet	BRDT	Lin SVR
Male								
7.81	**64%**	8.28 62%	8.32 62%	8.35 62%	8.33 62%	9.16 56%	10.94 52%	31.80 19%
Female								
9.73	**61%**	10.55 57%	10.62 56%	10.54 58%	10.60 59%	12.05 50%	14.05 46%	28.72 23%

Bold denote the best result among several methods

deformable SMPL body model, non-rigid ICP-based model registration, computation of circumference path features and nonlinear regression for anthropometric measurement estimation. Depending on the measurement, our pipeline provided success rates from 28% to 93% for male and from 24 to 82% for female subjects. The proposed pipeline works in practice and shows that an affordable scanning system can be built for clothing industry.

In the future work, we will further investigate and refine each step of the pipeline. For example, selection of better surface features in addition to the circumference paths, fast-to-compute alternatives for the slow ICP algorithm (e.g. Chen et al. [6]) and better scanners and scanning procedures.

References

1. Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid ICP algorithms for surface registration. In: CVPR (2007)
2. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: SCAPE: shape completion and animation of people. In: SIGGRAPH (2005)
3. Baek, S.Y., Lee, K.: Parametric human body shape modeling framework for human-centered product design. Comput. Aided Des. **44**(1), 56–67 (2012)
4. Bogo, F., Romero, J., Loper, M., Black, M.J.: FAUST: dataset and evaluation for 3D mesh registration. In: CVPR (2014)
5. Bronstein, A., Bronstein, M., Kimmel, R.: Numerical Geometry of Non-rigid Shapes. Springer, Berlin (2008)
6. Chen, Q., Koltun, V.: Robust nonrigid registration by convex optimization. In: ICCV (2015)
7. Chen, Y., Robertson, D.P., Cipolla, R.: A practical system for modelling body shapes from single view measurements. In: BMVC (2011)
8. Daanen, H., Hong, S.A.: Made-to-measure pattern development based on 3D whole body scans. Int. J. Cloth. Sci. Technol. **20**(1), 15–25 (2008)
9. Gordon, C.C., Churchill, T., Clauser, C.E., Bradtmiller, B., McConville, J.T., Tebbetts, I., Walker, R.A.: Anthropometric survey of us army personnel: summary statistics, interim report for 1988. Tech. rep, Anthropology Research Project Inc, Yellow Springs, OH (1989)
10. Guan, P., Weiss, A., Balan, A.O., Black, M.J.: Estimating human shape and pose from a single image. In: ICCV (2009)
11. Hirshberg, D.A., Loper, M., Rachlin, E., Black, M.J.: Coregistration: simultaneous alignment and modeling of articulated 3D shape. In: ECCV (2012)
12. Kim, K., Jones, M., Ebert, S., Malik, L., Manary, M., Reed, M., Klinich, K.: Development of virtual toddler fit models for child safety restraint design. Tech. Rep. UMTRI-2015-38, University of Michigan Transportation Research Institute (2015)
13. Kim, V., Lipman, Y., Funkhouser, T.: Blended intrinsic maps. In: SIGGRAPH (2011)
14. Lipman, Y., Funkhouser, T.: Möbius voting for surface correspondence. In: SIGGRAPH (2009)
15. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. ACM Trans. Graph. (Proc. SIGGRAPH Asia) **34**(6), 248:1–248:16 (2015)
16. Ogden, C., Fryar, C., Carroll, M., Flegal, K.: Mean body weight, height, and body mass index, United States 1960–2002. Examination Surveys 347, Division of Health and Nutrition (2004)
17. Paquette, S., Brantley, J.D., Corner, B.D., Li, P., Oliver, T.: Automated extraction of anthropometric data from 3D images. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 44, pp. 727–730. SAGE Publications, Los Angeles (2000)
18. Park, B.K.D., Ebert, S., Reed, M.: A parametric model of child body shape in seated postures. Traffic Inj. Prev. **18**(5), 533–536 (2017)
19. Pheasant, S., Haslegrave, C.: Bodyspace: Anthropometry, Ergonomics and the Design of Work, 3rd edn. Taylor & Francis, New York (2005)
20. Robinette, K., Daanen, H.: Precision of the CAESAR scan-extracted measurements. Appl. Ergon. **37**, 259–265 (2006)
21. Robinette, K.M., Blackwell, S., Daanen, H., Boehmer, M., Fleming, S.: Civilian American and European surface anthropometry resource (CAESAR) final report. Tech. Rep. AFRL-HE-WP-TR-2002-0169, US Air Force Research Laboratory (2002)
22. Schneider, D.C., Eisert, P.: Fast nonrigid mesh registration with a data-driven deformation prior. In: ICCV Workshop (2009)
23. Simmons, K., Istook, C.: Body measurement techniques: comparing 3D body-scanning and anthropometric methods for apparel applications. J. Fash. Mark. Manag. Int. J. **7**, 306–332 (2003)
24. Vladimir, V., Golowich, S.E., Smola, A.J.: Support vector method for function approximation, regression estimation and signal processing. In: Proceedings of the 9th International Conference on Neural Information Processing, pp. 281–287 (1997)

25. Tsoli, A., Loper, M., Black, M.: Model-based anthropometry: predicting measurements from 3D human scans in multiple poses. In: WACV (2014)
26. Weiss, A., Hirshberg, D., Black, M.J.: Home 3D body scans from noisy image and range data. In: ICCV (2011)
27. Wuhrer, S., Shu, C.: Estimating 3D human shapes from measurements. Mach. Vis. Appl. **24**(6), 1133–1147 (2013)
28. Yu, T., Guo, K., Xu, F., Dong, Y., Su, Z., Zhao, J., Li, J., Dai, Q., Liu, Y.: Bodyfusion: real-time capture of human motionand surface geometry using a single depth camera. In: ICCV (2017)
29. Yu, T., Zheng, Z., Guo, K., Zhao, J., Dai, Q., Li, H., Pons-Moll, G., Liu, Y.: Doublefusion: real-time capture of human performances with inner bodyshapes from a single depth sensor. In: CVPR (2018)
30. Zhu, C., Byrd, R.H., Lu, P., Nocedal, J.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans. Math. Softw. (TOMS) **23**(4), 550–560 (1997)
31. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. R. Stat. Soc. **67**, 301–320 (2005)
32. Zuffi, S., Black, M.: The stitched puppet: a graphical model of 3D human shape and pose. In: CVPR (2015)

Song Yan is a second-year PhD student at Department of Computing, Tampere University, under the supervision of Associate Professor Joni-Kristian Kämäräinen at the Computer Vision Group. He received BE degree from Department of Software Engineering, Sichuan University, in June 2015 and masters degree from Department of Information Technology, Communication Systems and Networks, Tampere University of Technology, in August 2017. As an enthusiastic researcher, his goal is to combine state-of-the-art deep learning algorithms with real-life industrial problems. He is currently solving problems of human body reconstruction, pose and shape estimation.

Johan Wirta is the CEO and Founder of NOMO Technologies Ltd who has 12 years of experience in various positions in the retail business and 10 years of international experience. He participated in several information system development projects.

Joni-Kristian Kämäräinen is an associate professor at Computing Sciences Department at Tampere University where he is part of the Vision Group. He did his post-doc in Center of Vision, Speech and Signal Processing, University of Surrey, with Josef Kittler. After the post-doc, he spent five years in the Faculty of the LUT School of Engineering Science, LUT University, after which he was selected to the current tenure-track position in 2012.