

Model-based Anthropometry: Predicting Measurements from 3D Human Scans in Multiple Poses

Aggeliki Tsoli, Matthew Loper, and Michael J. Black
Max Planck Institute for Intelligent Systems
Tuebingen, Germany

{aggeliki.tsoli, mloper, black}@tuebingen.mpg.de

Abstract

Extracting anthropometric or tailoring measurements from 3D human body scans is important for applications such as virtual try-on, custom clothing, and online sizing. Existing commercial solutions identify anatomical landmarks on high-resolution 3D scans and then compute distances or circumferences on the scan. Landmark detection is sensitive to acquisition noise (e.g. holes) and these methods require subjects to adopt a specific pose. In contrast, we propose a solution we call model-based anthropometry. We fit a deformable 3D body model to scan data in one or more poses; this model-based fitting is robust to scan noise. This brings the scan into registration with a database of registered body scans. Then, we extract features from the registered model (rather than from the scan); these include, limb lengths, circumferences, and statistical features of global shape. Finally, we learn a mapping from these features to measurements using regularized linear regression. We perform an extensive evaluation using the CAESAR dataset and demonstrate that the accuracy of our method outperforms state-of-the-art methods.

1. Introduction

Measuring the human body from 3D data is gaining increasing importance in applications such as virtual try-on and online shopping. Extracting tailoring measurements directly from 3D scans of people could accelerate the tedious and time consuming process of custom tailoring. That, in turn, offers the potential of personalized sizing in online shopping and a decreased return-rate for web purchases. These applications are currently of interest due to the emergence of low-cost scanning devices such as the Microsoft Kinect [2]. The promise of such methods is that they will be even more accurate than humans at measuring the body.

Here we develop a method for *model-based anthropometry* that accurately predicts measurements; the approach is

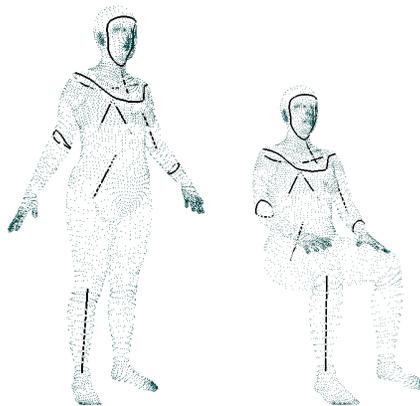


Figure 1: Registered meshes of a person in two poses after fitting a human body model [4, 12]. Measurement prediction is based on local and global shape features. Local shape features comprise circumferences on the mesh surface (shown as curves) and limb lengths (shown as straight lines). Global shape features (not shown here) describe statistics of shape in a database of registered 3D meshes. For visualization clarity, the 3D meshes are displayed as point clouds.

summarized in Fig. 1. Model-based anthropometry has several components: 1) a statistical model of body shape variation across a population of 3D bodies; 2) a deformable 3D body model and a method to fit it reliably to a scan; 3) a method to extract a variety of local and global features from the model; and 4) a method to predict 1D measurements from the features. A significant novelty of our approach is that it allows us to integrate information from multiple scans of a person in different poses. We show that this approach is more accurate than existing methods.

Exactly how accurate are current methods and are they accurate enough for custom tailoring applications? While there have been large studies of the accuracy of human anthropometers [9], there have been no large published studies of automatic methods for deriving measurements from

scans. This is despite the fact that there are relatively large collections of laser body scans with associated hand measurement data [18]. For the first time, we perform an extensive evaluation of existing commercial and research systems using the CAESAR dataset [18].

Previous work on extracting anthropometric measurements from 3D human scans is based on either measuring directly on the raw scan or using a database of registered scans to correlate human shape variation with measurements. In the first case, measurements are extracted by locating anthropometric landmarks on the scan’s surface; this simulates and automates the process of acquiring measurements as performed by a tailor. Measurements typically correspond to straight lines (heights) starting and/or ending on landmarks as well as circumferences along planar slices of the scan based on landmark locations. These approaches are sensitive to acquisition noise and missing data on the scan’s surface that distort the shape of the captured body. They also require high-resolution (e.g. laser scans), making such approaches impractical with today’s low-resolution home scanning systems based on Kinect. Additionally, this need for accurate localization of anthropometric landmarks in 3D typically limits such systems to scans captured in a single canonical pose. We show that different poses are optimal for different measurements and our model-based approach is able to integrate information from multiple poses. Finally, existing commercial solutions are limited in the kinds of measurements they make (linear and circumferential), whereas our model-based approach can regress body shape to any measurement (e.g. weight or even age).

Our model-based approach addresses the problems of previous methods. First, given a database of registered 3D scans of humans, together with their measurements, we build a statistical model of shape variation in the population. Features corresponding to 3D shape variation between individuals in the database are correlated with their measurements. Then, given a 3D scan of a new subject, we *register* the model with the scan by deforming the model to match the scan data. Shape features are easily computed from the model and measurements are then predicted from these shape features. A model-based approach could also be used to predict measurements from low-resolution scans [22]. Most previous efforts on correlating measurements with human shape variation assume the human body shape is represented by a single pose ([22] is an exception in that they use several poses). Moreover, previous evaluations are limited either in the set of measurements considered [11] or the number of subjects used [22, 23].

Our contributions are the following: 1) We introduce a set of surface-based shape features that are predictive of standard anthropometric measurements. 2) We optimize over a wide range of features to find the ones most predictive of measurements. 3) We introduce model-based anthropometry for predicting anthropometric measurements from various poses and demonstrate more accurate measurement prediction than the state-of-the-art. 4) We present a comprehensive comparative study between our model-based approach and state-of-the-art commercial and research efforts for measurement prediction. We consider a wide range of standard anthropometric measurements and a large number of subjects using the CAESAR dataset. This evaluation provides a solid foundation for evaluation of commercial and research work in this area.

ometry for predicting anthropometric measurements from various poses and demonstrate more accurate measurement prediction than the state-of-the-art. 4) We present a comprehensive comparative study between our model-based approach and state-of-the-art commercial and research efforts for measurement prediction. We consider a wide range of standard anthropometric measurements and a large number of subjects using the CAESAR dataset. This evaluation provides a solid foundation for evaluation of commercial and research work in this area.

2. Previous Work

There are several studies comparing the performance of commercial 3D scanning systems relative to hand measurements [6, 16, 17]. In particular, [16, 17] provide an evaluation of commercial 3D scanning systems in terms of predicting anthropometric measurements on or around the torso. The measurements are extracted from rigid torso dressforms created to represent variability in the Army population as reported by [9]. Rigid mannequins, however, do not exhibit the things that make real bodies a challenge to measure: pose/posture variation, breathing, soft tissue deformation, body fat hiding anatomical structures, and general ambiguity about where to measure.

Rather than evaluate automated systems, previous work has focused on the accuracy of humans at measuring the body. Human anthropometers remain the gold standard for measurement but there is variability between measurers and by the same measurer over time. The ANSUR study was designed to measure the accuracy of such human measurements [9]. Without ground truth, ANSUR focused on the variability of measurement and used this variability to define a allowed error for each measurement. The ANSUR allowed error is likely unrealistically low and we show that current commercial systems produce errors that significantly exceed the ANSUR standards.

The main study of automated scan measurement analysis is that of Bradtmiller and Gross [6]. Compared to the abovementioned studies, they report a broader set of automatically extracted measurements from real human subjects. They found that these measurements were generally sufficient for garment fitting, but the prediction error was larger than the ANSUR allowable error.

To evaluate measurement prediction, we use the CAESAR dataset [18], which represents the largest publicly available dataset of 3D body scans with associated measurements; here we use the US dataset of approximately 1000 men and 1000 women in both sitting and standing poses. While in wide use, to our knowledge the accuracy of measurement prediction from CAESAR scans has not been evaluated, and nobody has attempted to predict measurements from sitting poses. Robinette and Daanen [19] measured the variance of estimates extracted semi-

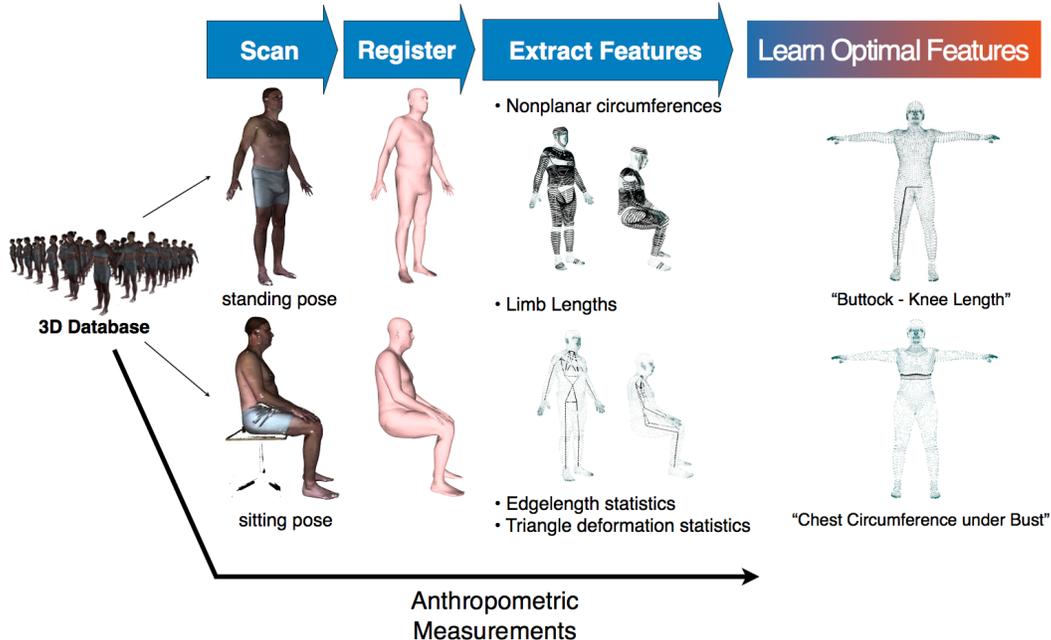


Figure 2: The training stage of our method. We start with a database of 3D scans in multiple poses (standing, sitting) with corresponding anthropometric measurements [18]. Initially, we register the scans using prior knowledge about human body shape. Then, we extract shape features. We consider local features, such as body circumferences and limb lengths, as well as global features, such as statistics on edge lengths and triangle deformations of the registered meshes. Finally, we learn optimal features for predicting each measurement.

automatically from CAESAR but did not evaluate the error with respect to manual measurements.

The model-based anthropometry approach we introduce relies on human shape features extracted from a database of registered 3D human bodies. There are a wide range of 3D shape descriptors for assessing similarity or matching 3D shapes in applications such as content-based 3D shape retrieval. Examples include spin images, spherical harmonics, mesh HoG, heat kernel signatures; these are summarized in the following overview papers [7, 13, 21]. These generic 3D shape descriptors are useful for classification but are less useful for detailed analysis of shape within a class. Given objects of a specific category, a common way to characterize variation in shape is using dimensionality reduction techniques. Employing Principal Component Analysis (PCA) [14] to describe 3D human body shape has shown encouraging results on generating and describing 3D human bodies [4]. Weiss et al. [22] use PCA coefficients to predict anthropometric measurements of 3D human models from Kinect data by combining information from multiple poses. More recent work takes standard PCA basis vectors and rotates these to better predict body measurements [11]. Alternatively Guan et al. [10] constrain body shape variations related to a particular measurement and exclude these from the orthogonal subspace. The work in [5] correlates body shape variation with measurements by perform-

ing PCA on the joint space of bodies and measurements. We derive global features of shape variation using PCA and augment them with features such as circumferences around limbs and limb lengths to represent local shape details.

Although we are interested in extracting measurements from 3D scans of humans, there has also been work on measuring synthetically generated 3D human bodies. Wuhler et al. [23] present a hybrid of the above mentioned paradigms by measuring, in a consistent way, paths along the surface of 3D human meshes with the same topology. However, the goal of the authors is mainly to synthesize 3D human bodies that conform to a set of input measurements. In addition, [8] reports measurement prediction on 3D bodies generated from photos of real humans and measurements on these photos; they evaluate, however, a very limited set of four measurements.

3. Measurements from 3D scans

Our approach for measurement prediction consists of a training and a testing stage. In the training stage, we register a set of high-resolution 3D human scans to a common 3D template mesh, learn a statistical model of shape deformations in the training set, extract shape features for each registered scan, and learn the optimal shape features for measurement prediction. Figure 2 provides an illustra-

tion of the training stage. In the testing stage, given a new high-resolution 3D scan, we register it with a common template mesh using the learned statistical shape model, derive shape features and use them for predicting standard anthropometric measurements.

3.1. Registration

Registration refers to the fitting of a template body mesh to a scan. This brings the scan into alignment with a database of pre-aligned meshes and our statistical body shape model. We only briefly summarize the mesh registration process as it is not the main contribution and has been described elsewhere [12].

The registration procedure serves two goals: (a) it provides a hole-free mesh that accurately captures the shape in the scan; and (b) it provides correspondences between 3D meshes, which facilitates statistical analysis.

Figure 3 shows an example of holes in the armpit area of a 3D scan. In this scenario, a generic surface reconstruction approach, such as [15], creates an unrealistic human shape by lowering the height of the reconstructed armpit. Our approach, which takes into account prior knowledge about the human shape and articulation gets much closer to the true shape of the scanned human subject.

Our registration energy and procedures are similar to those in [12]; as in that work, a BlendSCAPE body model is used (whose form was heavily influenced by SCAPE [4]). As in [12], the data term encourages the template to match the scan and the prior term encourages deformations that are consistent with a learned statistical body shape model. The shape space of our model was trained from approximately 800 aligned CAESAR scans. While the registration process used 73 landmarks (part of CAESAR) for initialization, we observed that a Gaussian prior on body pose parameters worked equally well.

While not necessary, here we assume the subject is scanned in both standing and sitting poses, which means that the registration procedure produces two registered meshes per subject. Registered meshes across subjects are in correspondence by construction.

3.2. Feature Extraction

We extract global and local features of shape variation from a set of registered 3D scans. Let $M_{ia} = (V_{ia}, E_{ia})$, $i = 1, \dots, N$, $a = 1, 2$ denote the registered scan of the i -th human subject in pose a where N is the total number of human subjects in our dataset. $a = 1$ corresponds to the standing pose and $a = 2$ to the sitting pose. Each registered 3D scan is represented as a mesh with vertices V_{ia} and edges E_{ia} .

Triangle deformations from a template mesh provide a common representation of 3D shape [4, 20]. Previous authors have shown the triangle deformations carry informa-

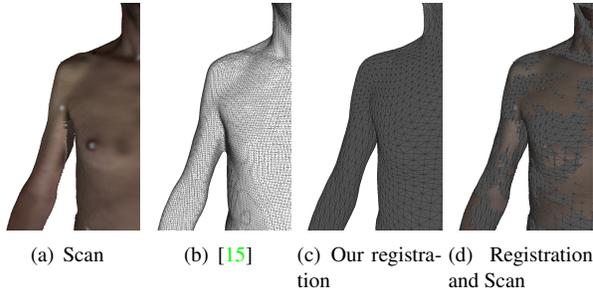


Figure 3: Model-based alignment effectively deals with holes. (a) A 3D scan showing a hole in the armpit area. Generic surface reconstruction approaches, such as (b) [15], create “webbing” effects and unrealistic human shapes. Our model-based approach (c) is robust to holes and captures the body shape. Subfigure (d) shows an overlay of our registration outcome on the initial 3D scan. Scan and registered mesh interleave with high frequency which indicates that the two surfaces are very close to each other.

tion about measurements [11, 22]. In particular, the coefficients of a low-D PCA representation can be used to predict linear measurements. Here we go further and consider deformations from three scenarios: 1) from a standing pose only; 2) from a seated pose only; and 3) using deformations from both poses.

Triangle deformations, however, are non-linearly related to geodesic distances on the body, and geodesics are similar to many standard tailoring measurements. Consequently we also consider the length of mesh edges, which are directly related to 3D length and circumference measurements, as the foundation for additional shape features.

Given triangle deformations and edge lengths of template meshes registered with the training dataset, we compute low dimensional representations for each using PCA. Given a new registered test mesh, it is projected onto these low-D spaces, yielding a set of coefficients that characterize the shape; in both cases we use 300 principal components. Let \mathbf{t}_{ia} , $a = 1, \dots, 3$ and \mathbf{d}_{ia} , $a = 1, 2$ denote respectively the coefficients of the triangle deformations and the edge lengths for the meshes of the i -th test mesh across different poses. Here we have extended our notation of a to account for standing and sitting poses together ($a = 3$). Considering both poses together is possible because the triangle deformations are taken with respect to the intrinsic shape of the human subject. According to [4], the intrinsic shape is abstracted from effects due to pose, such as muscle bulging, etc.

Low-dimensional representations of the body necessarily remove fine shape details. Additionally these global linear shape models capture correlations in the population and an individual may differ from this. We address this by adding extra local features that we observe on the sur-

face of the registered bodies. We hypothesize that these features may be more directly related to tailoring measurements. More specifically, we consider circumferences around limbs and the trunk as well as limb lengths. A circumference feature is calculated as a piecewise linear path over the edges of the mesh. Limb lengths are defined as Euclidean distances between neighboring joint locations, where joint locations are defined as a linear combination of pre-defined mesh vertices. Let c_{ia} , l_{ia} respectively be the features corresponding to circumferences and limb lengths associated with the i -th mesh in pose a , $a = 1, 2$.

Summarizing, we consider nine types of features, global and local, per body with $\mathbf{o}_i = \{\mathbf{t}_{i1}, \mathbf{t}_{i2}, \mathbf{t}_{i3}, \mathbf{d}_{i1}, \mathbf{d}_{i2}, \mathbf{c}_{i1}, \mathbf{c}_{i2}, \mathbf{l}_{i1}, \mathbf{l}_{i2}\}$ being the features of the i -th body in the dataset.

3.3. Feature Type Selection per Measurement

To find the most predictive feature types for each measurement, we learn the relationship between shape features and each measurement using Elastic Net linear regression [24]. For computational efficiency, we examine only unary and pairwise combinations of feature types.

Let S denote the set of unary and pairwise combinations of the feature types described above. Let also $\mathbf{z} = \{z_{qi}\}$, $q = 1, \dots, Q$, $i = 1, \dots, N$ be the set of anthropometric measurements for all human bodies in the dataset where Q is the total number of measurements. We select the optimal combination of feature types \mathbf{s}^q for the q -th measurement as

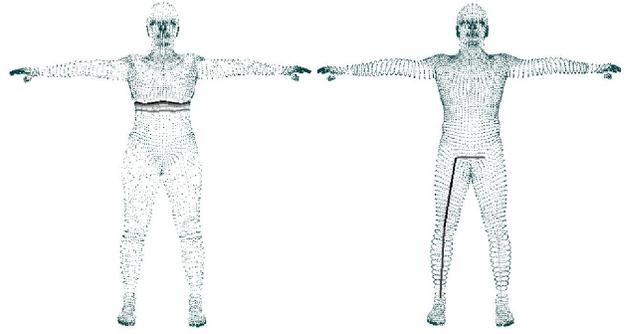
$$\mathbf{s}^q = \arg \min_{\mathbf{s} \in S} \sum_{i=1}^N |f_{\mathbf{s}}(\mathbf{o}_i^{\mathbf{s}}) - z_{qi}| \quad (1)$$

where $f_{\mathbf{s}}$ is an Elastic Net regression function for predicting measurements trained on the subset of features \mathbf{s} . $\mathbf{o}_i^{\mathbf{s}}$ denotes the feature values of the i -th human subject from the subset of features \mathbf{s} .

3.4. Measurement Prediction

Measurement prediction is performed independently for each measurement. Let $\mathbf{o}^{\mathbf{s}^q}$ be the values of the optimal features for predicting the q -th measurement (Sec. 3.3) extracted after registering the input scan. The predicted measurements for the human subject is the set $\{f_{\mathbf{s}^q}(\mathbf{o}^{\mathbf{s}^q})\}$, $q = 1, \dots, Q$.

An alternative approach for predicting measurements would be to predefine manually (i.e. by an anthropometer) curves or lines on the reference template mesh for each measurement. Measuring their length directly after the registration procedure would yield an estimate of the corresponding measurement. However, given a reference mesh with a pre-defined topology, it is not certain that there is a sequence of edges or vertices that match exactly the measurement we are interested in. We believe that using a learning framework that correlates shape features with measurements will



(a) Chest Circumf. under Bust (b) Buttock Knee Length

Figure 4: Examples of feature selection in females and males for the following measurements: (a) Chest Circumference under Bust, (b) Buttock Knee Length. The higher the importance of each feature, the darker its color. Features are displayed on the average female or male shape from our database posed in the T-pose. Optimal local features usually correspond to circumferences near the desired circumference measurement or to a set of limb lengths related with the desired height measurement.

introduce, to some degree, invariance to the topology of the registered meshes. As a by-product, we are able to predict attributes of human shape, such as weight, that cannot be measured directly from a 3D scan, but could potentially be useful in virtual try-on applications.

4. Results

4.1. Method Evaluation

We compare our method with the commercial software Anthroscan [1] as well as with [11]. Anthroscan predicts measurements directly from a 3D scan in the standing pose (only). Hasler et al. [11] performs registration-based measurement prediction. The data we use for training and testing are registered high-resolution 3D scans and the 40 anthropometric measurements from the CAESAR dataset. We use approximately 800 bodies per gender for training and 200 for testing (see [3]). The error metrics that we use are the Mean Absolute Difference (MAD) between extracted measurements and hand measurements in CAESAR, e_{MAD}^q , for each measurement q as well as the Average Mean Absolute Difference (AMAD), e_{AMAD} , over all measurements:

$$e_{\text{MAD}}^q = \frac{1}{N_s} \sum_{i=1}^{N_s} |f_{\mathbf{s}^q}(\mathbf{o}_i^{\mathbf{s}}) - z_{qi}| \quad (2)$$

$$e_{\text{AMAD}} = \frac{1}{Q} \sum_{i=1}^Q e_{\text{MAD}}^q \quad (3)$$

where N_s is the number of bodies in the test set.

Optimal groups of features per measurement are derived using 20-fold cross validation in the training set of 800 bodies. For the 11 common anthropometric measurements between Anthroscan and CAESAR, we compute a linear correction for each Anthroscan measurement using leave-one-out cross validation. This effectively adapts the Anthroscan measurements to the CAESAR measurements. We compare with [11] by using their proposed features and Elastic Net regression. Elastic Net regression compares favorably to the linear prediction framework presented in [11]. We additionally compare our performance with the ANSUR inter-observer error [9]. Recall that ANSUR reports the median absolute deviation between measurements made by experts rather than actual measurement errors.

Figure 5(a) shows aggregate statistics of the performance of our method relative to [1] and [11]. More specifically, we report the e_{AMAD} both in terms of absolute values (millimeters) as well as expressed relative to the ANSUR allowable error (AE). Our e_{AMAD} is around 1 cm, which translates to 1.2 to 1.3 times the AE. The error using the features from [11] is around 10% higher than our error. In 50% to 65% of the cases our improved prediction accuracy is also statistically significant. Statistical significance was assessed per measurement using a paired t-test. Comparing our method with the Anthroscan software for the measurements that are common between CAESAR and Anthroscan, we see that the e_{AMAD} error for Anthroscan is 10-15% higher than the error in our case. For the measurements that are common between CAESAR and Anthroscan we come very close to the allowable error. In 65% to 80% of the cases our improved prediction accuracy is also statistically significant.

Figure 5(b) shows the performance of the above mentioned approaches for the subset of CAESAR measurements that is common between CAESAR and Anthroscan for the female test set. A detailed overview of performance for females and males is found in the supplemental material [3]. We observe that for most of the measurements, our approach performs better than previous work and the errors are close to the allowable error for each measurement. Most of the differences between our measurement predictions and predictions from previous work are statistically significant. Measurement predictions with no statistical significance between our method and at least one of the other approaches are denoted with small font size in Figure 5(b).

4.2. Prediction from Multiple Poses

Most of previous work has focused on predicting measurements from a single standing pose. However, it is unclear which scanning pose is the optimal one or whether different poses would give different prediction results. Muscles bulge, soft tissue deforms, and joints vary in different ways depending on the pose, which is why anthropometric

	Standing	Sitting	Stand+Sit
Male	10.21 mm (1.3 x AE)	10.8 mm (1.38 x AE)	10.09 mm (1.28 x AE)
Female	10.24 mm (1.27 x AE)	11.44 mm (1.37 x AE)	10.02 mm (1.23 x AE)

Table 1: Effect of **pose** on measurement prediction. We report the Average Mean Absolute Difference (AMAD) error over the 40 anthropometric measurements in CAESAR [18]. AE denotes the allowable error based on ANSUR [9]. The error using the standing pose is lower than the error using the sitting pose. Best results are obtained combining both poses.

or tailoring measurements are typically acquired using multiple poses. Table 1 shows a comparison of measurement predictions from two poses that were available in CAESAR. The AMAD error over the 40 anthropometric measurements that come with CAESAR is around 1 cm for each single pose which translates to around 1.3 or 1.4 times the variance of the measurements that expert anthropometers would report based on ANSUR. The measurement prediction error using the standing pose is lower than the error using the sitting pose; this is expected given that most of the CAESAR measurements were taken in the standing pose. Combining the two poses by selecting automatically the pose that gives the best prediction, using cross-validation on the MAD error, results in lower measurement prediction error. This confirms our hypothesis that combining information from multiple poses is beneficial.

4.3. Features for Measurement Prediction

In Sec. 3.2 we proposed a set of global and local features for measurement prediction. On one hand, we proposed generative global features, coefficients of PCA components that can be used to generate new bodies. On the other hand, we proposed local features that resemble measurements. Do we really need both and which set of features is more influential? To answer these questions we predict measurements only from global and only from local features. The results are summarized in Table 2. The AMAD error using only local features is almost 10% higher than using global features. This is interesting because it shows that global shape features are good for predicting linear measurements. But, as with multiple poses, we find that the combination of global and local features yields the best results.

Figure 4 shows the most influential local features for a representative sample of measurements. It is interesting to see that the most influential local features for each measurement are in areas of the body close to where a tailor would choose to take the specific measurement. Importantly, these features are automatically discovered. In Fig. 5(b), we see

	[11]		Our Method			[1]		Our Method	
Male	11.11 mm (1.41 x AE)	0%	10.09 mm (1.28 x AE)	51%	Male	12.65 mm (1.51 x AE)	9%	10.78 mm (1.15 x AE)	64%
Female	11.25 mm (1.42 x AE)	0%	10.02 mm (1.23 x AE)	65%	Female	11.11 mm (1.24 x AE)	9%	10.28 mm (1.06 x AE)	82%

(a) Average Mean Absolute Difference (AMAD) Error

Measurement	[1]	[11]		Our method		AE [9]
Ankle Circumference	7.55	6.59	6.19	deformations, stand	edges, stand	4
Arm Length Shoulder - Elbow	11.26	8.42	6.65	limbs, stand	edges, stand	6
Arm Length Shoulder - Wrist	11.67	10.42	10.05	limbs, stand	edges, stand	
Arm Length Spine - Wrist	13.19	13.40	11.87	girths, stand	limbs, stand	
Chest Circumference	12.43	13.02	12.73	girths, stand	edges, stand	15
Crotch Height	7.45	7.53	5.50	limbs, stand	deformations, stand	10
Head Circumference	7.44	7.45	5.87	girths, sit	limbs, stand	5
Hip Circ Max Height	17.05	18.96	18.59	girths, stand	limbs, stand	
Hip Circumference, Maximum	7.47	16.15	12.35	girths, stand	edges, stand	12
Neck Base Circumference	21.13	16.96	15.79	limbs, sit		11
Stature	5.60	10.21	7.51	girths, stand	limbs, stand	10

(b) Mean Absolute Difference (MAD) Error per Measurement in Females

Figure 5: Comparative evaluation. (a-left) Comparison with [11] on 40 CAESAR measurements for 200 test bodies of each gender. (a-right) Comparison with Anthroscan on the measurements that Anthroscan and CAESAR have in common. In both cases we report Average Mean Absolute Difference (AMAD) error between extracted and hand measurements. AE denotes the allowable error based on the ANSUR study [9]. Percentages represent the fraction of measurements where the predictions of the best performing method are statistically significant. (b) Mean Absolute Difference (MAD) between extracted and hand measurements on females for the common measurements between CAESAR and Anthroscan (in mm). Bold is best. Results that are not statistically significant are denoted with smaller font size. For our method, we additionally show the optimal types of features (deformations, edges, girths, limbs) and pose (stand, sit) selected for each measurement.

	Global	Local	Global+Local
Male	10.29 mm (1.3 x AE)	11.47 mm (1.46 x AE)	10.09 mm (1.28 x AE)
Female	10.34 mm (1.27 x AE)	10.98 mm (1.40 x AE)	10.02 mm (1.23 x AE)

Table 2: Effect of our **global and local features** on measurement prediction. We report the Average Mean Absolute Difference (AMAD) error over the 40 anthropometric measurements in CAESAR [18]. The error using only local features is almost 10% higher than using global features. AE denotes the allowable error based on ANSUR [9].

the features selected by our approach for a subset of measurements. More detailed results are presented in the supplemental material [3]. For most of the measurements, optimal measurement prediction is achieved through a combination of global and local features.

5. Conclusions

In this paper we present an alignment-based approach for extracting anthropometric measurements from high-

resolution 3D human body scans. Representing the 3D scan using a reference mesh deformed based on a human body model allows us to capture effectively the shape of the scanned subject. Additionally, it provides correspondences to a database of 3D humans scans (already in correspondence) varying in intrinsic shape and pose. Shape features derived consistently across the registered scan and all the bodies in the database are fused using a regularized linear learning framework that leads to more accurate measurement prediction than state-of-art approaches. Our approach generalizes easily to scans in arbitrary poses (as long as registration with a reference mesh can be performed). We perform an extensive evaluation and find that our method significantly outperforms the state of the art.

The accuracy numbers reported here are critical for industrial applications in which clothing manufacturers must know how accurate automated methods can be. Our accuracy is significantly better than existing methods, including commercial solutions. The accuracy of current commercial methods has never before been demonstrated. While our errors are above the inter-observer errors in ANSUR, this does not mean they are insufficient for real applications. Inter-observer variance ignores observer bias and does not

directly quantify measurement accuracy.

Future work includes experimenting with 3D human scans of lower resolution or partial 3D scans of humans. Because our method can provide correspondences between 3D scans, it facilitates the use of a great variety of 3D shape descriptors and learning methods. With that in mind, we are interested in extending our approach to predict more intrinsic attributes of the human body, such as age, muscularity, *etc.*

Acknowledgements. The authors thank P. Gehler, E. Rachlin, A. Weiss, D. Hirshberg for insightful discussions. This work was supported in part by the Office of Naval Research, Contract W911QY-10-C-0172.

Conflict of interest disclosure. ML is a consultant and MJB is a shareholder and member of the board at Body Labs Inc., a company commercializing 3D body shape technology.

References

- [1] Human Solutions GmbH. Anthroscan. <http://www.human-solutions.com/> 5, 6, 7
- [2] Microsoft Corp. <http://www.xbox.com/kinect> 1
- [3] Supplemental Material. http://ps.is.tue.mpg.de/project/Model-based_Anthropometry 5, 6, 7
- [4] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. SCAPE: Shape Completion and Animation of PEople. *SIGGRAPH*, 24(3):408–416, 2005. 1, 3, 4
- [5] S. Baek and K. Lee. Parametric human body shape modeling framework for human-centered product design. *Computer-Aided Design*, 44(1):56–67, 2012. 3
- [6] B. Bradtmiller, M. Gross, B. Bradtmiller-Anthrotech, and M. Gross-Anthrotech. 3D whole body scans: Measurement extraction software validation. *Digital Human Modeling for Design and Eng. Conf.*, 1999. 2
- [7] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape Google: Geometric words and expressions for invariant shape retrieval. *ACM ToG*, 30(1):1–20, 2011. 3
- [8] Y. Chen, D. Robertson, and R. Cipolla. A practical system for modelling body shapes from single view measurements. *BMVC*, pp. 82–91, 2011. 3
- [9] C. Gordon, T. Churchill, C. Clauser, B. Bradtmiller, and J. McConville. Anthropometric survey of US Army personnel: Methods and summary statistics 1988. Tech. Rep., DTIC, 1989. 1, 2, 6, 7
- [10] P. Guan, A. Weiss, A. Balan, and M. J. Black. Estimating human shape and pose from a single image. *ICCV*, pp. 1381–1388, 2009. 3
- [11] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H. Seidel. A statistical model of human pose and body shape. *Computer Graphics Forum*, 28(2):337–346, 2009. 2, 3, 4, 5, 6, 7
- [12] D. Hirshberg, M. Loper, E. Rachlin, M. J. Black. Coregistration: Simultaneous Alignment and Modeling of Articulated 3D Shape. *ECCV*, pp. 242–255, 2012. 1, 4
- [13] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani. Three-dimensional shape searching: State-of-the-art review and future trends. *Computer-Aided Design*, 37(5):509–530, 2005. 3
- [14] I. Jolliffe. *Principal Component Analysis*. Wiley Online Library, 2005. 3
- [15] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. *Symposium on Geometry Processing*, pp. 61–70, 2006. 4
- [16] L. Mckinnon and C. Istook. Comparative analysis of the Image Twin system and the 3T6 body scanner. *J. Textile and Apparel, Technology and Management*, 1(2):1–7, 2001. 2
- [17] S. Paquette, J. Brantley, B. Corner, P. Li, and T. Oliver. Automated extraction of anthropometric data from 3D images. *Human Factors and Ergonomics*, v. 44, pp. 727–730, 2000. 2
- [18] K. Robinette, S. Blackwell, H. Daanen, M. Boehmer, S. Fleming, T. Brill, D. Hoeflerlin, and D. Burnsides. Civilian American and European Surface Anthropometry Resource (CAESAR) final report. AFRL-HE-WP-TR-2002-0169, 2002. 2, 3, 6, 7
- [19] K. Robinette and H. Daanen. Precision of the CAESAR scan-extracted measurements. *Appl Ergon*. 37(3):259–265, 2006. 2
- [20] R. Sumner and J. Popović. Deformation transfer for triangle meshes. *ToG*, 23(3):399–405, 2004. 4
- [21] J. Tangelder and R. Veltkamp. A survey of content based 3D shape retrieval methods. *Shape Modeling Applications*, pp. 145–156, 2004. 3
- [22] A. Weiss, D. Hirshberg, and M. Black. Home 3D body scans from noisy image and range data. *ICCV*, pp. 1951–1958, 2011. 2, 3, 4
- [23] S. Wuhrer and C. Shu. Estimating 3D human shapes from measurements. *Machine Vision and Applications*, 24(5):1133–1147, 2013. 2, 3
- [24] H. Zou and T. Hastie. Regularization and variable selection via the Elastic Net. *J. R. Stat. Soc. B*, 67:301–320, 2005. 5