CALVIS: chest, waist and pelvis circumference from 3D human body meshes as ground truth for deep learning

Yansel Gonzalez Tejeda[0000−0003−1002−3815] and Helmut Mayer[1111−2222−3333−4444]

Paris Lodron University of Salzburg, Kapitelgasse 4-65020 Salzburg, Austria
https://www.uni-salzburg.at

Abstract. In this paper we present CALVIS, a method to calculate **C**hest, w**A**ist and pe**LVIS** circumference from 3D human body meshes. Our motivation is to use this data as ground truth for training convolutional neural networks (CNN). Previous work had used the large scale CAESAR dataset or determined these anthropometrical measurements *manually* from a person or human 3D body meshes. Unfortunately, acquiring these data is a cost and time consuming endeavor. In contrast, our method can be used on 3D meshes automatically. We synthesize eight human body meshes and apply CALVIS to calculate chest, waist and pelvis circumference. We evaluate the results qualitatively and observe that the measurements can indeed be used to estimate the shape of a person. We then asses the plausibility of our approach by generating ground truth with CALVIS to train a small CNN. After having trained the network with our data, we achieve competitive validation error. Furthermore, we make the implementation of CALVIS publicly available to advance the field.

Keywords: chest, waist and pelvis circumference · 3D human body mesh · deep learning

1 Introduction

Motivation Predicting 3D human body measurements from images is crucial in several scenarios like virtual try-on, animating, ergonomics, computational forensics and even health and mortality estimation. Researches had named these measurements body intuitive controls [2], biometric measurements [21], body dimensions [7], (European standard EN 13402), semantical parameters [29], traditional anthropometric measurements [28] or only "shape" as in human shape estimation [11,5,16,8,19]. In contrast, we assume a more anthropometric approach motivated by comprehensive compendiums like Panero and Zelnik, 1979 [18]. Throughout this paper the term human body dimensions will be used to refer to the above measurements and we will focus on three of these dimensions, namely chest, waist and pelvis circumference.

The problem of estimating chest, waist and pelvis circumference having only an image of a person is a challenging problem and involve recognizing the corresponding location of these dimensions in the body and estimating their circumference. Additionally, the setting is an under-constrained (or inverse) problem. Information gets lost when a camera is used to capture the human body in 3D space to 'render' a 2D image.

To tackle this problem a supervised learning approach can be used. This approach demands large amount of human body anthropometric measurements and is certainly one of the biggest challenges in the field. Currently there is only one large-scale dataset, the Civilian American and European Surface Anthropometry Resource (CAESAR) [20] with 3D human body scans and their corresponding anthropometric measurements. This survey was extraordinarily large and resource intensive: around nine thousand people from Europe and the USA where scanned, it was conducted over more than five years and costed more than six million dollars.

In the past decade a noticeably amount of researchers have employed this dataset to investigate human shape estimation. Because the measurement acquiring process is resource intensive and requires large amount of human and material resources, this type of studies are rare and the data they offer is expensive. Therefore, it is important to explore alternative methods where human body measurements derived from real data can be obtained for investigation.

3D human body generative models offers such an alternative. In the line of previous work, we start by synthesizing 3D human body meshes using the SMPL [16] generative model. Unlike other works, once we have the 3D meshes we compute chest, waist and pelvis circumference. The next step after obtaining the measurements is to use a camera model to render images. Finally, in possession of this ground truth we can input these images to the learning algorithm and train with the human body dimensions as supervision signal.

Contributions In summary, our contributions are 1) CALVIS: a method capable of automatically outputting chest, waist and pelvis circumference from 3D body meshes. 2) A prototype deep learning framework (CNN) with synthetic images of 3D human body meshes as input and the human body dimensions obtained with CALVIS as supervision signal.

2 Related Work

2.1 Human Dimensions Estimation from Images.

Using metrology techniques [4] and mixture of experts model [21] human body dimensions like height, arm span and even weight had been estimated from images. Chest and waist size had been considered as well [11]. These investigations used human dimensions for validation (ground truth) that were directly recovered from individuals. In contrary, we calculate them consistently from 3D meshes. Another significant research direction studies shape as (non necessarily meaningful) parameters of some model. For example [5] claim describing the

first method to estimate pose and shape from a single image, while [15] estimates shape from 91 keypoints using decision forests. Shape is understood as the parameters of the SMPL model but not human dimensions.

In general, previous work has concentrated on human body shapes not deviating much from the mean. In contrast, we move beyond by exploring 3D meshes reflecting human figures characteristics such as bulky, slim, small and tall subjects.

2.2 Human Body Data Synthesis.

The SCAPE model[3] opened wide possibilities in the field. It provided the scientific community with a method to build realistic 3D human body meshes of different shapes in different poses. In order to synthesize a human body mesh a template body must be deformed. The model exhibits, however, an important limitation, namely the last step on the pipeline is an expensive optimization problem. Other researchers had used variation of the SCAPE model (S-SCAPE [19]) to synthesize human bodies but focused on people detection.

After some attempts on improving the human models quality, for example, to make more easily capturing the correlation between body shape and pose [12], or to better understand body dimensions (SPRING model [29]), Loper et al., 2015 [16] developed the SMPL generative human body model. In this work we synthesize 3D human meshes using this model. We will briefly describe the model in subsection 3.1.

Our approach to synthesize images from 3D meshes has been strongly influence by the recent publication of the large scale dataset SURREAL [26]. This work uses the SMPL model to generate images from humans with random backgrounds. However, no human body dimensions are computed of estimated.

More recently, a human model with added hands and face had been presented [13]. We do not use this more complex model because the human dimensions we estimate do not require that level of detail.

2.3 Human Body Dimensions from 3D Models.

Although extensive research has been carried out on human shape estimation, methods to consistently define shape based on 3D mesh anthropometric measurements have been little explored. Only a handful of researchers have reported calculating 1D body dimensions from 3D triangular meshes to use them as ground truth for training and validation in a variety of inference processes.

Early research performed feature analysis on what they call body intuitive controls, e.g. height, weight, age and gender [2] but they do not calculate them. Instead they use the CAESAR demographic data. Recording human dimensions beyond height like body fat and the more abstract “muscles” are described by [12].

Pertinent to our investigation is also the inverse problem: generating 3D human shapes from traditional anthropometric measurements [27]. Like this work

we use a set of anthropometric measurements that we call dimensions, unlike them we calculate 1D measurements from 3D human bodies to use them as ground truth for later inference.

Strongly related to our work are methods that calculate waist and chest circumference by slicing the mesh on a fixed plane and compute the convex hull of the contour [11] or path length from identified (marked) vertices on the 3D mesh [6,8,9]. However, is not clear how they define the human body dimensions.

In contrast, we do not calculate the dimensions from fix vertices on the template mesh. Instead, we take a more anthropometric approach and use domain knowledge to calculate these measurements.

2.4 Human Shape Estimation with Artificial Neural Networks

A huge amount of research has been conducted in recent years to address the problem of 3D/2D human shape estimation using ANNs. A state-of-the-art method is [14] where they estimate human shape and pose. While these ANN's output are human body models parameters (i.e., βs in [8] and [25]) our network is capable to output human dimensions directly.

3 Chest, Waist and Pelvis Measurements

While human body meshes are traditionally acquired by specialized 3D scanners, we use a generative model to synthesize them, which we briefly review in subsection 3.1. To calculate the actual human body dimensions, we first need to formalize the notion of chest, waist and pelvis of a 3D human body mesh. Our strategy consist of segmentating the mesh in five regions, three of them of interest, which we discuss in subsection 3.2. Finally, within these regions we identify the intended dimensions employing a human body mesh signature (HBMS), defined in subsection 3.3.

3.1 SMPL Generative Model of the Human Body

In this work we synthesize 3D human meshes using the SMPL model [16]. The generative nature of our approach establishes this model as starting point (and not the 3D mesh). Nevertheless, our method is flexible enough to begin the pipeline with a 3D mesh. In that case, an SMPL model can be fitted to the volumetric data, using the method described by Varol et al., 2018 [25] for example.

The SMPL model is at its top level a **skinned articulated model**, i.e., consists of a surface mesh $\mathcal{M}$ that mimics the skin and a corresponding skeleton $\mathbf{J}$. The mesh $\mathcal{M}$, which is a boundary representation stores both the body geometry (vertex position) and topology (how the vertices are connected). The skeleton $\mathbf{J}$ is defined by its joints location in 3D space $j_i \in \mathbb{R}^3$ and its kinematic tree. Two connected joints define a 'bone'. Moreover, a child bone is rotated relative to its connected parent. The pose $\vec{\theta}$ is the specification of every bone rotation plus an orientation for the root bone.

The SMPL model is also a **deformable model** [23]. In order to synthesize a new human mesh one has to deform the provided template mesh by setting shape parameters $\vec{\beta}$. Pose parameters $\vec{\theta}$ are used for animation.

More specifically, the model is defined by a template mesh (mean template shape) $\bar{\mathbf{T}}$ represented by a vector of $N = 6890$ concatenated vertices in the zero pose, $\vec{\theta}^*$. The skeleton has $K = 23$ joints.

As an **statistical model**, SMPL provides learned parameters

$$\Phi = \{\bar{\mathbf{T}}, \mathcal{W}, \mathcal{S}, \mathcal{J}, \mathcal{P}\}. \tag{1}$$

As mentioned above $\bar{\mathbf{T}} \in \mathbb{R}^{3N}$ is the mean template shape. The set of blend weights $\mathcal{W} \in \mathbb{R}^{N \times K}$ represents how much the rotation of skeleton bones affects the vertices. In addition, the matrices $\mathcal{S}$ and $\mathcal{P}$ define respectively linear functions $B_s(\vec{\beta}; \mathcal{S}) : \mathbb{R}^{|\vec{\beta}|} \rightarrow \mathbb{R}^{3N}$ and $B_p(\vec{\theta}; \mathcal{P}) : \mathbb{R}^{|\vec{\theta}|} \rightarrow \mathbb{R}^{3N}$ that are used to deform $\bar{\mathbf{T}}$; and the function $\mathcal{J} : \mathbb{R}^{|\vec{\beta}|} \rightarrow \mathbb{R}^{3K}$ predicts skeleton rest joint locations from vertices in the rest pose. A new mesh $\mathcal{M}_{new}$ can be then generated using the SMPL model M

$$\mathcal{M}_{new} = M(\vec{\beta}, \vec{\theta}, \Phi). \tag{2}$$

Since we held fix parameters Φ during the synthesis and either we change $\vec{\theta}$ because our approach focuses on the zero pose $\vec{\theta}^*$, we manufacture a new mesh

$$\mathcal{M}_{new} = M(\vec{\beta}). \tag{3}$$

$$\mathcal{M}_{new} = \bar{\mathbf{T}} + B_s(\vec{\beta}; \mathcal{S}). \tag{4}$$

As a result we obtain a mesh that realistically represents a human body. Next, we focus on establishing reasonable chest, waist and pelvis regions on this mesh.

3.2 Chest, Waist and Pelvic Regions Segmentation

Let us consider a human body mesh $\mathcal{M}$. Our method requires that $\mathcal{M}$ is standing with arms raised parallel to the ground at shoulder height at a 90° angle (SMPL zero pose $\vec{\theta}^*$). Additionally, we assume that the mesh has LSA orientation, e.g., x, y and z axis are positively directed from right-to-left, inferior-to-superior and posterior-to-anterior, respectively. If the mesh has another orientation we can always apply rigid transformations to LSA-align it.

We observe that the chest circumference is usually measure below the armpits, also known as axilla, but above the natural waist. Similarly, the pelvis circumference is measured often around the hips. This observation suggests that we should consider segmenting the mesh in meaningful regions. Moreover, we can use the skeleton when determining the region boundaries. For example, there is consensus regarding the axilla definition as the area of the human body underneath the shoulder joint ([24], [22], [10]). Therefore, we can use the shoulder

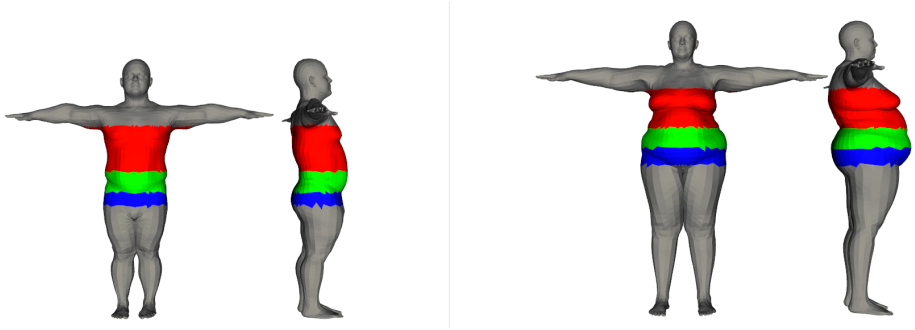

Fig. 1. Mesh segmentation in chest (red), waist (green) and pelvic (blue) regions. We show male (left) and female (right) subjects in frontal and lateral view. Note how the regions are correctly segmented for different gender and body shapes.

joint as a hint to locate the axilla and establish a upper bound for the chest region. More generally, we can use the skeleton joints defined in $\mathbf{J}$ to identify the regions boundaries.

Let the axilla location be the point $p_a = (p_a^x, p_a^y, p_a^z)$ and j_{nw}, j_{pl}, $j_{rh} \in \mathbf{J}$ be the natural waist, pelvis and right hip joints (joints with identifier `Spine1`, `Pelvis`, `R_Hip` in $\mathbf{J}$), respectively.

Recall that mesh $\mathcal{M}$ has N vertices. Let the set of vertices be $\mathcal{V}$ and v_i^y the y-coordinate of vertex v_i, we can segment the mesh in chest, waist and pelvic regions $\mathcal{CR}, \mathcal{WR}, \mathcal{PR}$, respectively

$$\mathcal{CR} = \left(v_i^y \,|\, p_a^y > v_i^y \geq j_{nw}^y\right) \tag{5}$$

$$\mathcal{WR} = \left(v_i^y \,|\, j_{nw}^y > v_i^y \geq j_{pl}^y\right) \tag{6}$$

$$\mathcal{PR} = \left(v_i^y \,|\, j_{pl}^y > v_i^y \geq j_{rh}^y\right) \tag{7}$$

Figure 1 shows the result of applying these equations on two meshes. Note how the regions are correctly segmented for different gender and body shapes. For example, the male subject on the left is smaller than the female subject on the right. Additionally, the subjects have very different muscular and fat build.

Axilla recognition As mentioned above, we would like to measure the chest circumference below the armpit, also known as axilla. This raises the question how we can recognize the axilla. What we want is a point p_a in the mesh at the proper location under the arms. One point suffices because we require only a

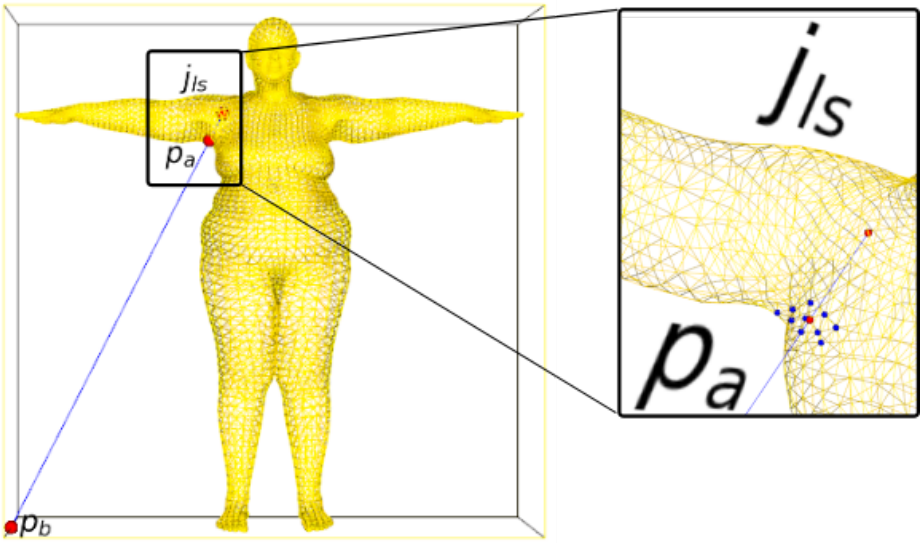

Fig. 2. Axilla recognition based on the skeleton joint. We cast a ray from the left shoulder joint j_{ls} in direction to the middle left inferior edge of the bounding box (left). The ray intersects the mesh at point p_c (red) that we consider the axilla center. We plot the 10 nearest vertices to p_c (right). Point p_a is the nearest vertex with smallest y coordinate. As it can be observed, the axilla is properly recognized.

reference p_a^y along the y-axis to define the chest region according to equation 5. Therefore, we focus on the right shoulder joint $j_{rs} \in \mathbf{J}$. Since the axilla is related to the shoulder joint, we can cast a ray $\boldsymbol{\iota}$ from it in direction to the middle left inferior edge of the bounding box outside the mesh at point

$$p_c = \left(v_{min}^x, v_{min}^y, v_{min}^z + \frac{|v_{max}^z - v_{min}^z|}{2} \right). \tag{8}$$

Here $v_{min}^x, v_{min}^y, v_{min}^z$ are vertex minimum x,y and z coordinates, respectively; analog is v_{max}^z the maximum vertex z coordinate. The ray ι is then determined by points j_{rs} (inside the mesh) and p_b (outside the mesh). Therefore, the ray ι intersects the mesh at one point p_c (Figure 2), which we consider the axilla center. Furthermore, we can consider an strategy to increase the method robustness against the large variability of the human body. Since this approach relies on one point to establish the chest upper bound, we identify the 10 nearest neighbors in the set of vertices $\mathcal{V}$ to p_c and define p_a to be the vertex with the smallest y coordinate of the neighbors.

Once we have a segmented mesh, we search the dimensions within these regions. The body dimensions we want to calculate are geodesic on the mesh. Next subsection discuss how.

3.3 Human Body Mesh Signature

Intuitively, we would like to measure the chest circumference below the arms at the *widest* part of the torso and the waist circumference at the *narrowest* part beneath the chest but above the hips. This intuition is compliant with prior body dimensions standardized definition, for example, the European standard EN 13402-1 [1]. Similarly, the pelvis circumference is measured often around the *widest* part of hips and buttocks. The general idea is to cast the chest, waist and pelvis circumference estimation problem as a constrained optimization problem. If we are able to establish a function that measures meaningful geodesics on the mesh, we can search for its maximum, e.g., chest in the chest region.

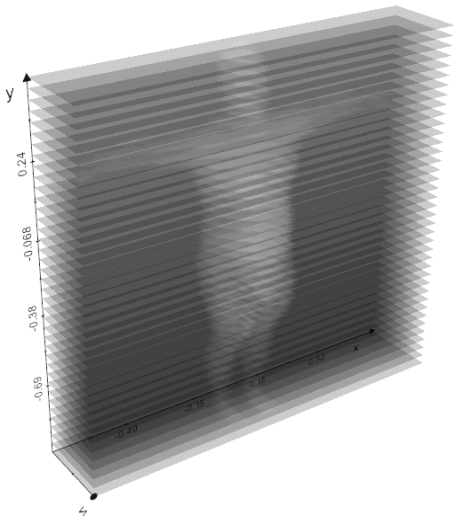

Fig. 3. 3D mesh iterative slicing. We can use a plane $\boldsymbol{\pi_j}$ parallel (with normal $\vec{n}_{||}$) to the floor to intersect the mesh at point $q_j = (0, q_j^y, 0)$, $q_j^y \in \mathbb{R}$ along the y-axis. By varying point q_j (we show here 40), we can obtain slices.

We can formalize this intuition by considering the cross-sectional length of the 2D intersection curves along the y-axis. Moreover, we can use a plane $\boldsymbol{\pi_j}$ parallel (with normal $\vec{n}_{||}$) to the floor to intersect the mesh at point $q_j = (0, q_j^y, 0)$, $q_j^y \in \mathbb{R}$ along the y-axis. Since $\mathcal{M}$ is triangulated, the boundary of this intersection is a polygonal curve consisting in segments with length s_i. Therefore, we can determine the intersection boundary length as

$$\mathcal{BL}(\mathcal{M}, \vec{n}_{||}, q) = \sum_{i=1}^{i=N} s_i \tag{9}$$

Note that for cross sections where $\boldsymbol{\pi_j}$ intersects the legs, two disconnected curves will appear. This is not a problem, because a set of disconnected curves is, in our case, a collection of segments with non-zero length.

Starting from the top q_t of the bounding box we slice iteratively mesh $\mathcal{M}$ with plane $\boldsymbol{\pi_j}$ every m-meters along the y-axis until we reach the bounding box bottom q_b. Figure 3 shows 40 of these slices. Next, we assemble the mesh slicing vector $\vec{\mathcal{L}}$. The elements of this vector are slice points

$$\vec{\mathcal{L}}(\mathcal{M}, \vec{n}_{||}, m) = [q_t, q_t - m, q_t - 2m, \cdots, q_b] \tag{10}$$

Finally, we apply equation 9 to every slice defined in $\mathcal{L}$ in equation 10. In other words, for every slice at point q_j, we compute intersection boundary length $\mathcal{BL}(q_j)$. Here we drop from the notation mesh $\mathcal{M}$ and plane normal $\vec{n}_{||}$ because they remain constant. We then can define the **human body mesh signature** $\mathcal{MS} : \vec{\mathcal{L}} \to \mathbb{R}$ that maps every slice at point q_j to the corresponding boundary length $\mathcal{BL}(q_j)$.

$$\mathcal{MS}(\vec{\mathcal{L}}) = [\mathcal{BL}(\mathcal{L}_1), \mathcal{BL}(\mathcal{L}_2), \cdots, \mathcal{BL}(\mathcal{L}_{|\vec{\mathcal{L}}|})] \tag{11}$$

$$\mathcal{MS}(\mathcal{M}, \vec{n}_{||}, m) = [\mathcal{BL}(q_t), \mathcal{BL}(q_t - m), \cdots, \mathcal{BL}(q_b)] \tag{12}$$

Figure 4 shows the $\mathcal{MS}$ of two meshes (male and female) for $m = 0,001$ and plane parallel to the floor (with normal $\vec{n}_{||} = (0, 1, 0)$) using the library trimesh [17]. Note that the function as defined by equation 12 is bounded and not continuous. It resembles a rotated silhouette of the human body and exhibits several *extrema*.

In general, we expect these *extrema* to be adequate features to calculate the human dimensions. More specifically, we assume that:

1. The chest circumference cc is the local maximum within the chest region.
2. The waist circumference wc (not to be confused with the natural waist line based on the joint, see subsection 3.2) is the minimum within the waist region.
3. The pelvis circumference pc is the maximum within the pelvis region.

$$cc = \underset{\mathcal{BL}(q_j)}{\arg\max}\ \mathcal{MS} \quad \text{subject to} \quad p_a^y > q_j \geq j_{nw}^y \tag{13}$$

$$wc = \underset{\mathcal{BL}(q_j)}{\arg\min}\ \mathcal{MS} \quad \text{subject to} \quad j_{nw}^y > v_i^y \geq j_{pl}^y \tag{14}$$

$$pc = \underset{\mathcal{BL}(q_j)}{\arg\max}\ \mathcal{MS} \quad \text{subject to} \quad j_{pl}^y > v_i^y \geq j_{rh}^y \tag{15}$$

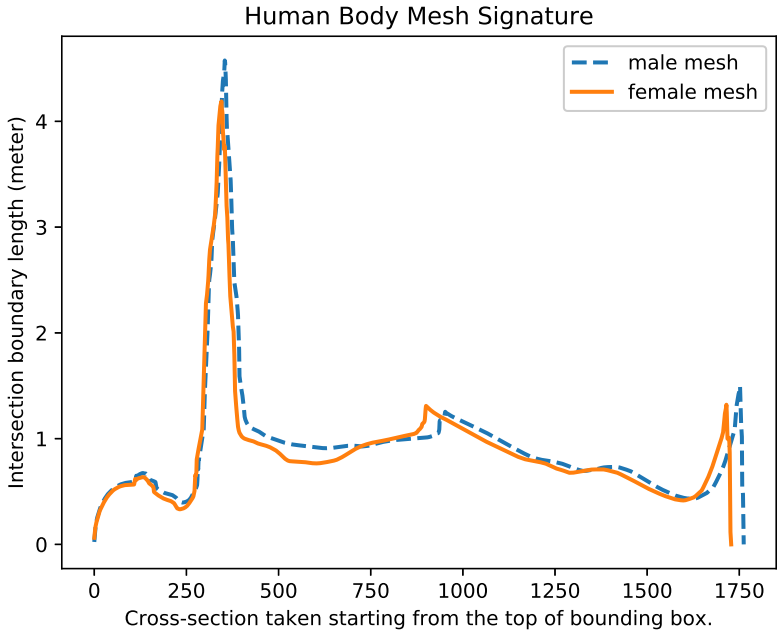

Fig. 4. Human Body Mesh Signature for male and female meshes. The function resembles a rotated silhouette of the human body and exhibits several *extrema*.

4 Experiments and Results

We conduct two experiments. In the first experiment we synthesize eight (four female and four male) human body meshes using shape parameters provided by SURREAL [26]. The meshes reflect human figures characteristics such as bulky, slim, small and tall subjects. Then we apply CALVIS to calculate chest, waist and pelvis circumference. Since we do not have ground truth, we evaluate the results qualitatively. The second experiment serves to asses the plausibility of our approach to use the synthetic data for deep learning. We input the calculated human dimensions to an artificial neural network. After having trained the network with our data, we show that the network is able to learn the task.

4.1 Qualitative evaluation

Here we evaluate our method qualitatively.

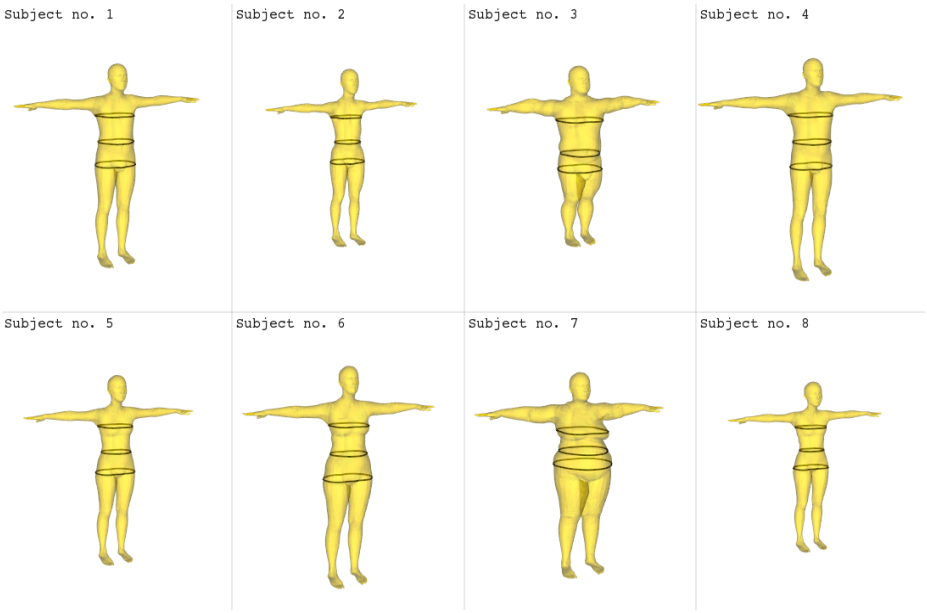

Fig. 5. Qualitative evaluation for male and female meshes.

Figure 5 shows the calculated dimensions on eight subjects, male subjects no. 1-4 and female subjects no. 5-8. The black paths represent chest, waist and pelvis circumference (from top to bottom) as defined by equations 13-15, respectively. We slightly rotate the meshes and add transparency for visualization purposes. Note that our method is able to estimate automatically the human body dimensions for all subjects. One particularity can be observed regarding the chest circumference: while for all male and female subjects except for subject no. 7 the chest circumference is estimated above the breast, for subject no.7 the chest circumference is estimated around the breast.

4.2 Learning with Calvis-Net

This experiment demonstrates the plausibility of employing CALVIS to obtain data that can be used as a ground truth for machine learning (in general) and specifically as supervising signal for CNN.

Experimental setup We annotate the CMU dataset with CALVIS. The automatic annotation process took approximately $1\,h\,47\,min$ in an enhanced modern personal computer.

Calvis-CMU Dataset We synthesize x images of every person using the package Blender. To increase the realism of these images we adopt a yansel lighting

model. This makes sense because it resembles the way how a person is standing in a room when she/he goes to the tailor for her/his body dimensions to be taken.

Synthetic images as input We synthesize x images of every person using the package Blender. To increase the realism of these images we adopt a yansel lighting model. This makes sense because it resembles the way how a person is standing in a room when she/he goes to the tailor for her/his body dimensions to be taken.

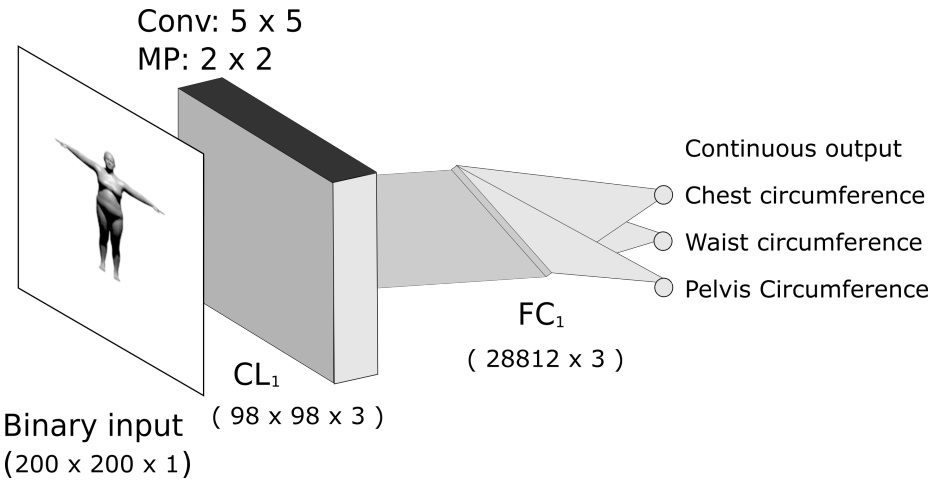

Fig. 6. Calvis-Net.

Figure 6 shows our prototypical CNN Calvis-Net where the input is a synthetic image from a 3D mesh and the output are three body dimensions.

Figure 7 shows that after epoch number 5 the estimation error decreases and the learning converges.

We achieve mean absolute validation error of 0.061698 m plus +-.

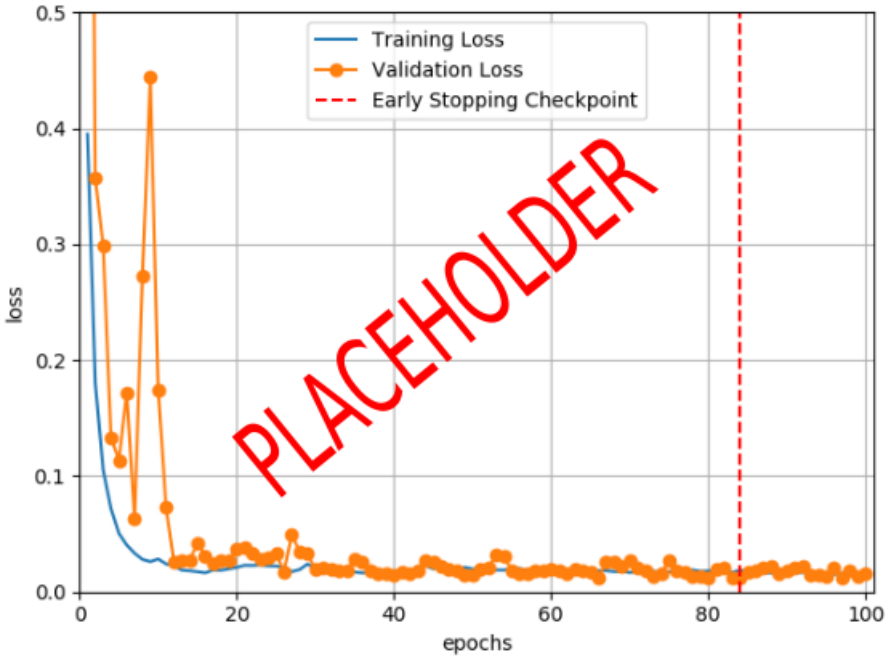

Fig. 7. Learning convergence evaluation.

5 Conclusion

In this paper we presented CALVIS, a method to calculate chest, waist and pelvis circumference from 3D human meshes. We demonstrate that our approach can be used to train a CNN by imputing synthetic images of humans and using the measurements calculated with CALVIS as supervision signal. Furthermore, we contribute with a prototype CNN CALVIS-NET. Our experiments show that the CNN is able to learn how to perform this task, achieving mean absolute validation error of 0.061698 m plus +-. The code and data use in this paper will be made publicly available for researchers.

References

1. EN 13402-1 : 2001. Size designation of clothes - PART 1: TERMS, DEFINITIONS AND BODY MEASUREMENT PROCEDURE. European Committee for Standardization (2001) 8
2. Allen, B., Curless, B., Popović, Z.: The space of human body shapes. In: Rockwood, A.P. (ed.) ACM SIGGRAPH 2003 Papers on - SIGGRAPH '03. p. 587. ACM Press, New York, New York, USA (2003). https://doi.org/10.1145/1201775.882311 1, 3

3. Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J.: Scape: Shape completion and animation of people. ACM Transactions on Graphics (2005) 3
4. BenAbdelkader, C., Yacoob, Y.: Statistical body height estimation from a single image. In: 8th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2008), Amsterdam, The Netherlands, 17-19 September 2008. pp. 1–7 (2008). https://doi.org/10.1109/AFGR.2008.4813453, https://doi.org/10.1109/AFGR.2008.4813453 2
5. Bogo, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., Black, M.J.: Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In: Computer Vision – ECCV 2016. Lecture Notes in Computer Science, Springer International Publishing (Oct 2016) 1, 2
6. Boisvert, J., Shu, C., Wuhrer, S., Xi, P.: Three-dimensional human shape inference from silhouettes: Reconstruction and validation. Machine Vision and Applications **24**(1), 145–157 (2013). https://doi.org/10.1007/s00138-011-0353-9 4
7. Chen, Y., Robertson, D.P., Cipolla, R.: A practical system for modelling body shapes from single view measurements. In: British Machine Vision Conference, BMVC 2011, Dundee, UK, August 29 - September 2, 2011. Proceedings. pp. 1–11 (2011). https://doi.org/10.5244/C.25.82, https://doi.org/10.5244/C.25.82 1
8. Dibra, E., Jain, H., Öztireli, C., Ziegler, R., Gross, M.: Hs-nets: Estimating human body shape from silhouettes with convolutional neural networks. In: 2016 Fourth International Conference on 3D Vision (3DV). pp. 108–117. IEEE (2016). https://doi.org/10.1109/3DV.2016.19 1, 4
9. Dibra, E., Öztireli, A.C., Ziegler, R., Gross, M.H.: Shape from selfies: Human body shape estimation using cca regression forests. In: ECCV (2016) 4
10. Federative Committee on Anatomical Terminology (FCAT), International Federation of Associations of Anatomists (IFAA): Terminologia anatomica (ta) (2019) 5
11. Guan, P.: Virtual Human Bodies with Clothing and Hair: From Images to Animation. Ph.D. thesis, Department of Computer Science at Brown University (2013), http://cs.brown.edu/pguan/publications/thesis.pdf 1, 2, 4
12. Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., Seidel, H.: A statistical model of human pose and body shape. Comput. Graph. Forum **28**(2), 337–346 (2009). https://doi.org/10.1111/j.1467-8659.2009.01373.x, https://doi.org/10.1111/j.1467-8659.2009.01373.x 3
13. Joo, H., Simon, T., Sheikh, Y.: Total capture: A 3d deformation model for tracking faces, hands, and bodies. CoRR **abs/1801.01615** (2018), http://arxiv.org/abs/1801.01615 3
14. Kanazawa, A., Black, M.J., Jacobs, D.W., Malik, J.: End-to-end recovery of human shape and pose. In: Computer Vision and Pattern Regognition (CVPR) (2018) 4
15. Lassner, C., Romero, J., Kiefel, M., Bogo, F., Black, M.J., Gehler, P.V.: Unite the people: Closing the loop between 3d and 2d human representations. CoRR **abs/1701.02468** (2017), http://arxiv.org/abs/1701.02468 3
16. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: Smpl: A skinned multi-person linear model. ACM Trans. Graph. **34**, 248:1–248:16 (2015) 1, 2, 3, 4
17. Michael Dawson-Haggerty: trimesh, https://trimsh.org/ 9
18. Panero, J.: Human dimension & interior space : a source book of design reference standards. Whitney Library of Design, New York (1979) 1
19. Pishchulin, L., Wuhrer, S., Helten, T., Theobalt, C., Schiele, B.: Building statistical shape spaces for 3d human modeling. Pattern Recognition **67**, 276–286 (2017). https://doi.org/10.1016/j.patcog.2017.02.018 1, 3

20. Robinette, K.M., Daanen, H., Paquet, E.: The caesar project: a 3-d surface anthropometry survey. In: Second International Conference on 3-D Digital Imaging and Modeling (Cat. No. PR00062). pp. 380–386. IEEE (1999) 2
21. Sigal, L., Balan, A., Black, M.J.: Combined discriminative and generative articulated pose and non-rigid shape estimation. In: Platt, J.C., Koller, D., Singer, Y., Roweis, S.T. (eds.) Advances in Neural Information Processing Systems 20, pp. 1337–1344. Curran Associates, Inc (2008), http://papers.nips.cc/paper/3271-combined-discriminative-and-generative-articulated-pose-and-non-rigid-shape-estimation.pdf 1, 2
22. Structural Informatics Group, University of Washington: Foundational model of anatomy ontology (fma) (2019) 5
23. Terzopoulos, D., Platt, J., Barr, A., Fleischer, K.: Elastically deformable models. SIGGRAPH Comput. Graph. **21**(4), 205–214 (Aug 1987). https://doi.org/10.1145/37402.37427, http://doi.acm.org/10.1145/37402.37427 5
24. U.S. National Library of Medicine: Medical subject headings (mesh) (2019), https://meshb.nlm.nih.gov/record/ui?ui=D001365 5
25. Varol, G., Ceylan, D., Russell, B., Yang, J., Yumer, E., Laptev, I., Schmid, C.: BodyNet: Volumetric inference of 3D human body shapes. In: ECCV (2018) 4
26. Varol, G., Romero, J., Martin, X., Mahmood, N., Black, M.J., Laptev, I., Schmid, C.: Learning from synthetic humans. In: CVPR (2017) 3, 10
27. Wuhrer, S., Shu, C.: Estimating 3d human shapes from measurements. Mach. Vis. Appl. **24**(6), 1133–1147 (2013). https://doi.org/10.1007/s00138-012-0472-y, https://doi.org/10.1007/s00138-012-0472-y 3
28. Wuhrer, S., Shu, C., Xi, P.: Landmark-free posture invariant human shape correspondence. The Visual Computer **27**(9), 843–852 (Sep 2011). https://doi.org/10.1007/s00371-011-0557-z, https://doi.org/10.1007/s00371-011-0557-z 1
29. Yang, Y., Yu, Y., Zhou, Y., Du, S., Davis, J., Yang, R.: Semantic parametric reshaping of human body models. In: 2nd International Conference on 3D Vision (3DV), 2014. pp. 41–48. IEEE, Piscataway, NJ (2014). https://doi.org/10.1109/3DV.2014.47 1, 3