

Automatic Locating of Anthropometric Landmarks on 3D Human Models

Zouhour Ben Azouz Chang Shu
Visual Information Technology Group
Institute for Information Technology
National Research of Canada
Ottawa, K1A 0R6, Canada

{Zouhour.Benazouz, Chang.Shu}@nrc-cnrc.gc.ca

Anja Mantel
Stuttgart University of Applied Sciences
Schellingstr.24
Stuttgart, 70174, Germany
anjamantel@web.de

Abstract

We present an algorithm for automatic locating of anthropometric landmarks on 3D human scans. Our method is based on learning landmark characteristics and the spatial relationships between them from a set of human scans where the landmarks are identified. The learned information is formulated by a pairwise Markov network. Each node of the network is a random variable corresponding to the position of a landmark. The edges of the network represent correlations between the positions of landmark pairs. Probabilistic inference is then performed over the Markov network to locate the landmarks. We evaluated the algorithm on 30 human models with different shapes. The results showed good accuracy for most of the landmarks.

1. Introduction

The remarkable progresses made in body scanning techniques during the last decade has led to increasing interests in 3D anthropometry. Currently, full-length human body can be digitized efficiently, providing tremendous amount of information about the shape of the human body. The 3D data can be used to measure, compare, and conduct statistics and has applications in, for example, ergonomic design of products such as automobiles, furniture, and clothes.

Processing the surface data collected by the sensors, however, proves to be a challenge. The main difficulty is that the digitized models have different number of points and there is no correspondence between different models. This correspondence information is essential to many applications such as variability study and animation [4][1]. Another difficulty is that the data is noisy and often incomplete.

Allen et al. [1] solved the correspondence problem by deforming a template model to fit individual scans. The re-

sulting models all have the same number of triangles and point-to-point correspondences. The deformation relies on a set of anthropometric landmarks provided in the CAESAR (Civilian American and European Surface Anthropometry Resource) database [14]. These landmarks were marked on the measured subjects prior to scanning. They provide the seed correspondences that guide the deformation of the rest of the points. Unfortunately, the marking process considerably increases the time of measurement – from a few seconds to more than 30 minutes. Consequently, this step raises the cost of 3D anthropometric data collection and will most likely be avoided in the future projects.

Ben Azouz et al. [4] used a volumetric representation of the surface models to establish the correspondences between the different models. This landmark-free approach allows the comparison and the extraction of certain modes of variation of the human body. However, it has the limitation that the volumetric representation provides only an approximation of the anthropometric correspondence between different human models.

In this paper we present a technique to locate landmarks without prior marking. Our method is based on statistical learning. We take advantage of the landmarks placed by experts during recent 3D anthropometric surveys such as the CAESAR project. Most anthropometric landmarks represent skeletal features and must be found by palpation. They correlate with surface geometric features in a subtle way. They correspond to a stable set of points on every human body. Through many years of study, anthropometrists have developed strict standards for the locations of these landmarks. With enough instances of these landmarks, we can learn their statistical characteristics and use them to guide our search for their locations on the 3D models.

We model the positions of the landmarks as random variables. The probability of a surface point to be a particular landmark depends on the local surface properties as well

as its relationships with other landmarks. These constraints are naturally modeled by Markov Random Field (MRF) or Markov network. In the training stage, the distributions of surface geometric properties such as SPIN images [10], and the relative positions of landmark pairs are computed. In the subsequent matching stage, landmarks are located on an instance of 3D human model by identifying the surface points that maximizes the joint probability defined by the Markov network. This probability optimization problem is solved efficiently by the *belief propagation* technique developed in statistical learning [18].

The rest of the paper is organized as follows. In Section 2 we review previous work. We describe our algorithm in Section 3. The results of locating landmarks in 3D human scans are presented in Section 4. Finally, we present our conclusions and future work in Section 5.

2. Previous Work

Currently, the dominant way to locate anthropometric landmarks relies on placing markers on the human body prior to scanning [8] [11] [5] [15]. These methods are tedious and time-consuming. Therefore the cost of future projects of 3D anthropometric data collection can be reduced if pre-marking is eliminated or just applied to a few subjects. In the literature there are some attempts to locate anthropometric landmarks automatically without prior marking. Most of this work are limited to locating branching points such as the armpits and the crotch [13] [12] [16] [17]. Dekker [6] extends the identification to a larger subset of landmarks. The approach is based on defining for each landmark a function that combines different attributes. For instance, the acromion (shoulder point) is located as “the first point on the torso, traversing down from the nape on the ridge line of maximum z where the gradient is less than g and then a gradient greater than h ” [6]. The limitation of the existing automatic methods is that they can not be extended to all the anthropometric landmarks. Also, the published results concern the accuracy of traditional measurement that are extracted from the located landmarks but not the accuracy of the landmark positions.

Recently, Angelov et al. [2] proposed an unsupervised method to register non-rigid surfaces. This method establishes an embedding of an instance mesh into a template mesh using a Markov network. The algorithm assigns for each vertex of the instance mesh a corresponding vertex on the template mesh. Angelov et al. [3] applied this algorithm to 3D human scans in order to select around 200 corresponding points between the template model and the instance models. These points are then used as landmarks to guide the deformation of the template model to fit their corresponding instance models. The approach can potentially be used to identify anthropometric landmarks of the

instance models, provided that the landmarks on the template model are identified. However, the algorithm requires an initialization by placing 4 to 10 markers manually on each pair of scans. Moreover, the published results did not quantify the accuracy of the correspondences.

The approach we propose in this paper differs from the previous work in two aspects. First, our algorithm is supervised. It is based on a training set of 3D human scans where the position of standard anthropometric landmarks (Figure 1) are identified. Second, our algorithm uses Markov network to represent the landmarks and the relationship between them. In comparison, the Markov network used by Angelov et al. [2] represents the whole mesh of the model, a much larger graph, and is thus harder to solve. Our work can be seen as the extension of the method that uses probabilistic graphical models for object recognition in 2D images [7].

3. Algorithm

We formulate the landmark locating problem as a probabilistic inference problem, where the inference is over a pairwise Markov random field. Here, our goal is to find the most likely assignment to the entire set of the landmarks. Markov random field is an instance of probabilistic graphical model. The central element of this model is a graph in which the nodes correspond to random variables. These models have been used in computer vision for object matching and recognition. For example, Felzenszwalb [7] used undirected graph models to find instances of an object in an image. The object is represented by a pictorial structure model that is a collection of parts with connections between certain pairs of parts.

The landmark locating process consists of two main steps. In the first step, we define and learn the parameters of a pairwise Markov random field. In the second step, we perform probabilistic inference to find the optimal labeling of the landmarks.

3.1 Pairwise Markov Random Field

A Markov random field, or Markov network, is a model of the joint probability distribution of a set of random variables. For our application these random variables are $L = \{l_1, l_2, \dots, l_n\}$ the positions of the n landmarks. A Markov network consists of an undirected graph $G = (V, E)$, where each node v from V represents a random variable in L and each edge $\{l_i, l_j\}$ from E represents a dependency between the random variables l_i and l_j . In a pairwise Markov network a potentials $\phi_i(l_i)$ is associated to each node i . This potential represents the likelihood that a landmark l_i corresponds to a given vertex on the surface. For each edge $\{l_i, l_j\}$ we associate a compatibility potential $\psi_{ij}(l_i, l_j)$ to

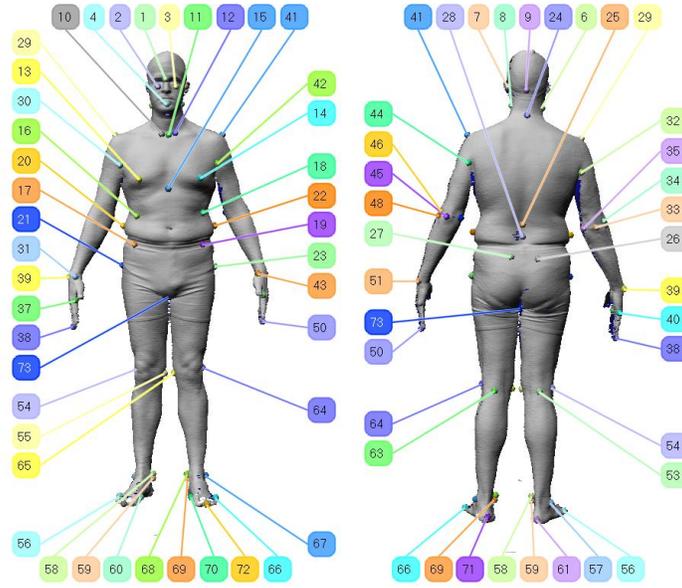


Figure 1. Anthropometric landmarks used in the CAESAR project: the anthropometric nomenclature of these landmarks is given in Table 1.

constrain the positions of these landmarks to be consistent with their spatial relationship. The joint probability associated with this landmark is given by:

$$p(L) = \frac{1}{Z} \prod_i \phi_i(l_i) \prod_{i,j} \psi_{ij}(l_i, l_j) \quad (1)$$

where Z is a normalizing factor. In order to locate the landmarks, the potentials $\phi_i(l_i)$ and $\psi_{ij}(l_i, l_j)$ of the MRF should encode a preference of the local surface properties around the landmarks and maintain the spatial relationship between landmark pairs. In the following two sub-sections we define the surface attributes used to describe the landmarks. We also characterize the spatial relationship between landmark pairs.

3.1.1 Surface Attributes

To characterize the local surface features, we use SPIN Images [10]. A SPIN image is a two-dimensional histogram computed at an oriented point of a surface mesh. An oriented point at a vertex is defined by the 3D position and the surface normal at the vertex. Two cylindrical coordinates (α, β) can be defined with respect to an oriented point as shown in Figure 2.

To compute the SPIN image, a 2D accumulator indexed by α and β is created. The coordinates (α, β) are computed for each vertex of the surface that is within a defined distance of support. The bin indexed by (α, β) in the accumulator is then incremented. The resulting accumulator is

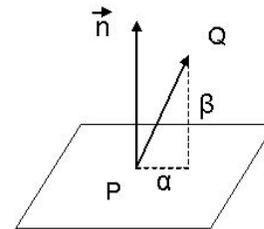


Figure 2. Coordinate system used to compute the spin image on a point surface.

then represented as an image where dark areas correspond to bins that contain many projected points.

Examples of SPIN images computed on 3 different human models for 3 landmarks are shown in Figure 3. We notice the similarity between the SPIN images corresponding to the same landmarks for different models.

SPIN image is a high dimensional descriptor that characterizes the local surface geometry around a vertex. However, it is not directly useful because its dimension is too high. To reduce the dimension to a manageable size, we use the standard Principal Component Analysis (PCA). We first compute an eigenspace from SPIN image of several models. The SPIN image of each model is then characterized by its projection on the most significant eigenvectors.

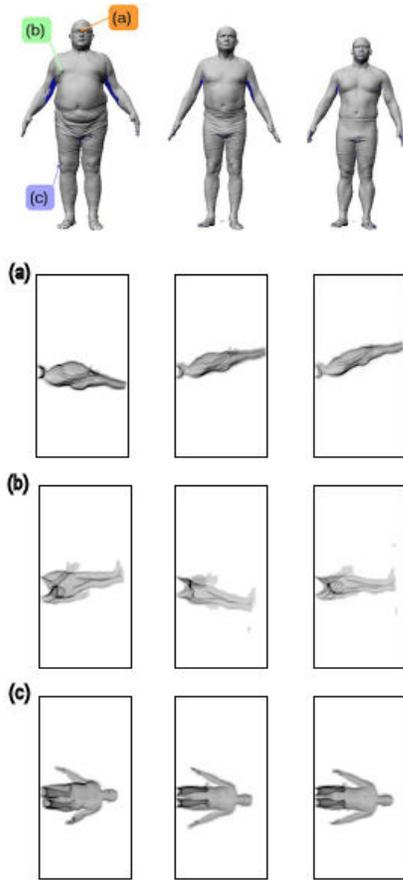


Figure 3. Spin Images computes for three landmarks on three different human models. (a) Sellion. (b) Rt. Thelion. (c) Rt. Femoral Lateral epicon.

3.1.2 Spatial relationship between landmark pairs

The anthropometric landmarks are constrained to a spatial relationship that is the same for all the human models. For instance the positions of the right acromion and the left acromion are approximately symmetric. The right and left iliac crest are below the sellion. In order to encode these constraints in the Markov network, we propose to characterize the edges pairs by the position of a landmark relative to its neighboring landmarks.

3.1.3 Learning

The advantage of representing landmarks by a graph is to model the correlation between the positions of landmarks. In our current implementation, we manually construct a graph to link pairs of landmarks (Figure 4). This graph does not necessarily express optimal correlations between land-

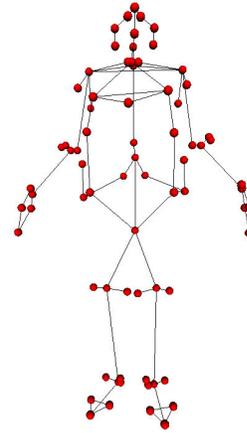


Figure 4. Structure of the landmark graph.

marks. This aspect will be addressed in our future work.

In the learning step we identify the parameters of the potentials attributed to the nodes and the edges of the Markov network. We model the potential associated to each node $\phi_i(l_i)$ by a Gaussian distribution of the SPIN image computed on the corresponding landmark. The parameters of the potential, (μ_i, Σ_i) , are the mean vector and the covariance matrix. The potential that a landmark i is located on the vertex v_k is:

$$\phi_i(l_i = v_k) = \mathcal{N}(S(v_k), \mu_i, \Sigma_i) \quad (2)$$

where \mathcal{N} is a multivariate Gaussian with mean vector μ_i and covariance matrix Σ_i , and $S(v_k)$ is the SPIN image computed on vertex v_k .

The potential $\psi_{ij}(l_i, l_j)$ is the Gaussian distribution of the relative position of landmark l_j with respect to landmark l_i .

We normalize the height of 3D human scans. This normalization has two main advantages. First, it eliminates the variability in the landmark positions that is due to a scale variation and not to the body shape variation. Second, after normalization we can learn a distribution for the position of each landmark. This distribution can be used to limit the search space in the landmark locating stage.

3.2 Probabilistic Inference

The goal of probabilistic inference over a Markov network is to find assignments for the random variables that maximize the joint probability defined in equation 1.

When the number of nodes is large, exact inference is computationally infeasible. Instead, we apply the approximate inference method known as *loopy belief propagation* [18]. This method is based on iteratively propagating messages between adjacent nodes.

In each iteration, a node i updates the messages sending to each of its neighbors j about what position they should be in:

$$m_{ij}(l_j) = \sum_{l_i} \phi_i(l_i) \psi_{ij}(l_i, l_j) \prod_{k \in N(i) \setminus \{j\}} m_{ki}(l_i) \quad (3)$$

After convergence, for each node we compute the belief that it be at a given position as follows:

$$b_i(l_i) = K \phi_i(l_i) \prod_{j \in N(i)} m_{ji}(l_i) \quad (4)$$

where K is a normalization constant.

We approximately obtain the most likely configuration of the nodes by assigning for each node the position corresponding to the highest belief.

4. Experimental Results

We applied our algorithm to 230 human scans from the CAESAR database, in which the positions of 73 anthropometric landmarks, as shown in Figure 1, are available for all the models. We used 200 scans for learning and the rest 30 scans for validation. The Euclidean distances between the predicted landmark positions and their corresponding positions placed by the anthropometry experts are calculated. When computing the locations of the landmarks, the heights of all the models are normalized to 1 metre, but the errors are calculated using the models in their original scale. Figure 5 illustrates the results of the landmark locating for a wide variety of human body shapes. Table 1 shows the average and standard deviation of the location errors for each landmark.

We notice that for most of the landmarks the error is less than $2cm$. It is important to mention that the landmarks we use as reference are subject to uncertainties, since different measurer place the same landmark at slightly different positions. Gordon et al. [9] reported the accuracy of the traditional 1D measurements. This study used the mean absolute differences to indicate the variances between repeated measurements of the same subjects. A maximum allowable error is derived from these values. The variation of results produced by different measurer is due in part to the difference in interpreting the anthropometric landmarks when they palpate the body to take the measurements. For most of the traditional measurements the accepted error is around $1cm$.

5. Conclusions

We have presented an algorithm for automatically locating anthropometric landmarks in human body scans. Our

method makes use of the landmark data placed by anthropometric experts in the recent anthropometric surveys. Through the use of these data, we train a set of parameters that relates the landmark location with local surface properties. The assigning of landmark labels to the vertices of each instance model is achieved by solving a graphical probabilistic optimization problem. Our experiments show that accurate landmarks can be found by using this method.

The use of anthropometric landmark data for training is the key to our success. Since anthropometric landmarks correspond to skeletal positions, they are a set of stable positions on the human body. Exploiting the knowledge embedded in these data allows us to locate the landmarks accurately.

The landmarks found in the human body scans are essential to subsequent processing of the 3D human models. They can be used for deforming a template model to each instance scan so that all models are in correspondence. They can also be used to make traditional anthropometric measurements.

The current results of our algorithm can be improved in several ways. There are different parameters that influence the results. One of these parameters is the support distance for computing the SPIN images. In our current experiments we choose a small support distance. This means that few vertices around an oriented point participate in the computation of the SPIN image. The accuracy of landmarks can also be improved by integrating other surface descriptors.

Currently we characterize the relationship between landmarks uniquely by the relative position of a given landmark with respect to its neighbor. Landmark locating may be improved by encoding other constraints such as the relative orientation of surface normals and principal curvature directions between the landmarks.

The use of Gaussian distribution to define the potentials attributed to the nodes and edges of the Markov network is only suggestive. Other distributions such as a mixture of Gaussians can improve the accuracy of landmark location.

For the current experiment we manually construct a graph to link pairs of landmarks. Identifying automatically the most correlated landmarks will potentially enhance the algorithm.

References

- [1] B. Allen, B. Curless, and Z. Popovic;. The Space of Human Body Shapes: Reconstruction and Parametrisation from Range Scans. *ACM Transactions on Graphics (ACM SIGGRAPH'2003)*, 22(3):587–594, 2003.
- [2] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, H. Pang, and J. Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. *In Advances in Neural Information Processing Systems*, 17:33–40, 2005.

- [3] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. Scape: shape completion and animation of people. *ACM Trans. Graph.*, 24(3):408–416, 2005.
- [4] Z. B. Azouz, C. Shu, R. Lepage, and M. Rioux. Extracting Main Modes of Human Shape Variation from 3-D Anthropometric Data. In *Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM'05)*, Ottawa, Canada, June 2005.
- [5] D. Burnsides, M. Boehmer, and K. Robinette. 3-D Landmark Detection and Identification in the Caesar Project. In *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling (3DIM'2001)*, pages 393–398, Quebec City, Canada, May 2001.
- [6] L. Dekker, I. Douros, B. Buxton, and P. Treleaven. Building Symbolic Information for 3D Human Body Modeling from Range Data. In *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling (3DIM'99)*, pages 388–397, Ottawa, Canada, October 1999.
- [7] P. Felzenszwalb and D. Huttenlocher. Pictorial structures for object recognition. *International Journal of Computer Vision*, 61(1), 2005.
- [8] G. Geisen, C. Mason, V. Houston, J. Whitestone, B. McQuiston, and A. Beattie. Automatic Detection, Identification and Registration of Anatomical Landmarks. In *Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society*, volume 2, pages 750–753, San Diego, 1995.
- [9] C. Gordon, B. Bradtmiller, C. Clausen, T. Churchill, J. McConville, I. Tebbets, and R. Walker. Anthropometric survey of US Army personnel. Method & summary statistics. Technical Report TR-89-044, US Army Natick Research Development and Engineering Center, Natick, MA, 1989.
- [10] A. Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.
- [11] F. Lewark and J. Nurre. Automated Fiducial Labeling on Human Body Data. In *Proceedings of the Three-Dimensional Image Capture and Applications, SPIE*, volume 3313, pages 82–89, San Jose, CA, January 1998.
- [12] J. Nurre. Locating Landmarks on Human Body Scan Data. In *Proceedings of the International Conference on Recent Advances in 3-D Digital Imaging and Modeling (3DIM'97)*, pages 289–295, Ottawa, Canada, May 1997.
- [13] R. Pargas, N.J.Staples, and J.S.Davis. Automatic Measurement Extraction for Apparel from a Three-Dimensional Body Scan. *Optics and Lasers in Engineering*, 28(2):157–172, 1997.
- [14] K. Robinette, H. Daanen, and E. Paquet. The Caesar Project: A 3-D Surface Anthropometry Survey. In *Second International Conference on 3-D Digital Imaging and Modeling (3DIM'99)*, pages 380–386, Ottawa, Canada, October 1999.
- [15] Y. Tai, S. Chao, and S. Yang. Automatic Detection and Labeling of Anatomical Landmarks on Virtual Human Model from 3D Laser Digitizer. In *Proceedings of the International Workshop on Advanced Image Technology (IWAIT)*, 2002.
- [16] C. Wang, T. Chang, and M. Yuen. From Laser-Scanned Data to Feature Human Model: A System Based on Fuzzy Logic Concept. *Computer Aided Design*, 35:241–253, 2003.
- [17] Y. Xiao, J. Siebert, and N.Werghi. Topological Segmentation of Discrete Human Body Shapes in Various Postures Based on Geodesic Distance. *The 17th International Conference on Pattern Recognition (ICPR'04), IEEE Computer Society Press*, pages 131–135, 2004.
- [18] J. Yedidia, W. Freeman, and Y. Weiss. *Understanding Belief Propagation and Its Generalizations*. Science & Technology Books, 2003.

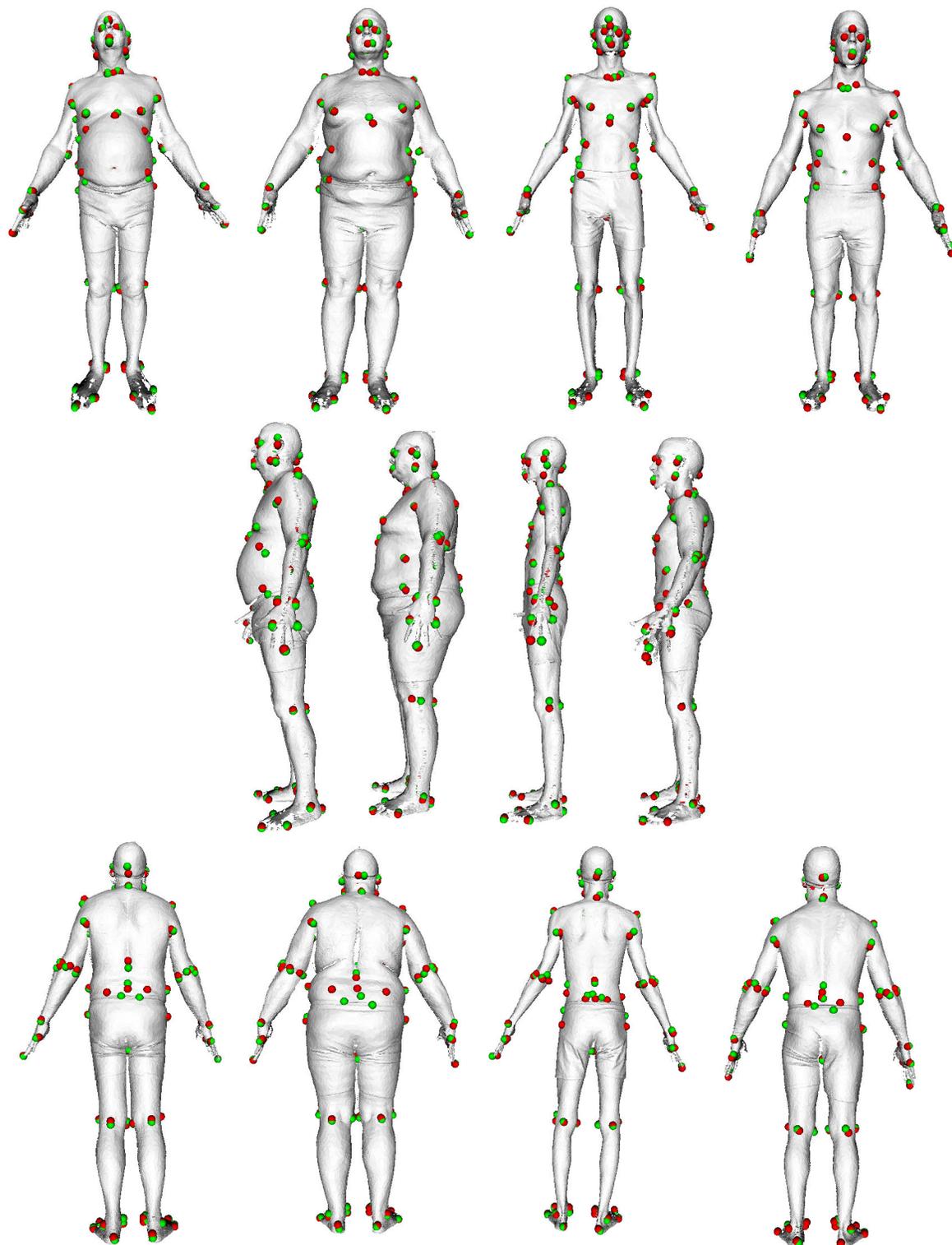


Figure 5. Results of landmark locating on 4 human models. The red spots correspond to landmarks located in the CAESAR survey by prior marking and the green ones correspond to the landmarks located using our approach. Each row corresponds to a different view of the 4 human models.

Landmark	Average (mm)	Standard Deviation (mm)	Landmark	Average (mm)	Standard Deviation (mm)
1 Sellion	10	0.73	38 Rt. Dactylion	12	1.4
2 Rt. Infraorbitale	7.1	0.057	39 Rt. Ulnar Styloid	11	0.11
3 Lt. Infraorbitale	11	1.3	40 Rt. Metacarpal-Phal. V	7.6	0.41
4 Supramenton	12	0.00097	41 Lt. Acromion	12	2.1
5 Rt. Tragion	13	0.26	42 Lt. Axilla, Ant	13	0.27
6 Rt. Gonion	16	0.08	43 Lt. Radial Styloid	11	0.66
7 Lt. Tragion	16	0.53	44 Lt. Axilla, Post.	15	1.7
8 Lt. Gonion	15	0.76	45 Lt. Olecranon	13	0.32
9 Nuchale	20	0.79	46 Lt. Humeral Lateral Epicn	13	0.32
10 Rt. Clavicale	10	0.68	47 Lt. Humeral Medial Epicn	18	0.57
11 Suprasternale	12	0.12	48 Lt. Radiale	16	0.47
12 Lt. Clavicale	13	0.56	49 Lt. Metacarpal-Phal. II	7.6	0.16
13 Rt. Thelion/Bustpoint	4.8	0.64	50 Lt. Dactylion	9.5	0.64
14 Lt. Thelion/Bustpoint	7.2	0.84	51 Lt. Ulnar Styloid	27	2
15 Substernale	16	0.54	52 Lt. Metacarpal-Phal. V	8.9	0.47
16 Rt. 10th Rib	27	2.4	53 Rt. Knee Crease	11	0.076
17 Rt. ASIS	33	0.42	54 Rt. Femoral Lateral Epicn	14	0.79
18 Lt. 10th Rib	21	0.75	55 Rt. Femoral Medial Epicn	15	1.7
19 Lt. ASIS	27	0.32	56 Rt. Metatarsal-Phal. V	8	0.94
20 Rt. Iliocristale	17	0.87	57 Rt. Lateral Malleolus	12	1.1
21 Rt. Trochanterion	16	1.8	58 Rt. Medial Malleolus	12	0.53
22 Lt. Iliocristale	16	1.4	59 Rt. Sphyrion	9.3	0.33
23 Lt. Trochanterion	16	0.83	60 Rt. Metatarsal-Phal. I	8.3	0.44
24 Cervicale	12	0.38	61 Rt. Calcaneous, Post.	11	0.44
25 10th Rib Midspine	24	1.5	62 Rt. Digit II	8.2	0.96
26 Rt. PSIS	62	2	63 Lt. Knee Crease	13	1.5
27 Lt. PSIS	55	3.2	64 Lt. Femoral Lateral Epicn	15	0.67
28 Waist, Preferred, Post.	19	0.71	65 Lt. Femoral Medial Epicn	19	2.1
29 Rt. Acromion	9.5	2.8	66 Lt. Metatarsal-Phal. V	6.8	0.47
30 Rt. Axilla, Ant	14	0.14	67 Lt. Lateral Malleolus	11	0.63
31 Rt. Radial Styloid	9.4	0.48	68 Lt. Medial Malleolus	6.6	0.39
32 Rt. Axilla, Post.	16	0.46	69 Lt. Sphyrion	4.6	0.53
33 Rt. Olecranon	13	0.41	70 Lt. Metatarsal-Phal. I	8.6	0.045
34 Rt. Humeral Lateral Epicn	14	0.26	71 Lt. Calcaneous, Post.	11	0.084
35 Rt. Humeral Medial Epicn	16	0.085	72 Lt. Digit II	9.2	1.1
36 Rt. Radiale	15	0.38	73 Crotch	17	1.3
37 Rt. Metacarpal Phal. II	8	0.42			

Table 1. Error of localisation computed over 30 test human scans.