6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

Predicting realistic and precise human body models under clothing based on orthogonal-view photos

Shuaiyin Zhu, P.Y Mok*

Institute of Textiles and Clothing, The HongKong Polytechnic University, Hong Kong

Abstract

Accurate and realistic digital human body models are required by many research applications, for example in the areas of ergonomics, clothing technology, and computer graphics. The already difficult research problem becomes more challenging if the individual subjects to be modelled are dressed in normal or loose-fit clothing. In this study, we present an intelligent two-phase method to customize 3D digital human body models based on two orthogonal-view photos of the customers. It integrates both image-based and example-based modelling techniques to create human body models for individual customers with precise body measurements and realistic appearance. It fills up the research gap of human model customization; without the need of taking body scan, any customers can create their 3D digital body models only based on their orthogonal-view photos in normal or loose-fit clothing. Experimental results have shown that the proposed method can efficiently and accurately customize human models of diverse shapes, meeting the specific needs of the clothing industry.

1. Introduction

An accurate digital human body model is a necessity in clothing related research or many ergonomic applications. Therefore, 3D digital human body modelling has received much research attention in last decades. In general, there are two classes of method for developing digital human body models, including construction methods

---

* Corresponding author. Tel.: +852-2766-4442; fax: +852-2773-1432.
  E-mail address: tracy.mok@polyu.edu.hk

and reconstruction methods. The key difference between the two is the involvement of body scanning. The former class of methods, namely, constructive methods, normally uses some projection devices to detect the customer's body shape and generate shape model. In order to obtain accurate body shape, the subjects being scanned must wear tight-fit clothing. Scanning devices are usually bulky and expensive. To overcome such limitations, researchers proposed reconstructive methods to capture, from images or size measurements, customer's body features, which are used to deform a template model using deformation technologies. Similar to the scanning-based construction methods, reconstruction methods also require customers being nude or dressing in tight-fit clothing for taking pictures or being measured. Recently, some methods were introduced to estimate the human body shape under clothing both from scanner or images. However, the results are not accurate enough for ergonomic or clothing applications.

In this study, we propose an intelligent two-phase method to customize 3D digital human body model based on two orthogonal-view photos of the customer. In our method, the customer needs not be in nude or with specialized tight-fit clothing before taking the photos, but can be dressed in normal or loose-fit clothing. To demonstrate the effectiveness of our method, we recruited a total of 15 female and 6 male subjects for experimental verification. Each subject was asked to have his/her body scanned and also have two photos taken with normal clothing in order to customize a human model using our method. The customized models and scans are compared in several aspects, including size measurements, areas and cross-section shape. We also compare our method with the method of [1], which can customize human models in tight-fit clothing. Experimental results have shown that the proposed method can customize human models of diverse shapes efficiently and accurately, meeting the specific needs of the clothing industry.

2. Related work

Accurate human body models are required in many research areas, such as clothing design, computer vision and ergonomic applications. A large number of research work have been reported in the literature for modelling human subjects in the past two or three decades. Since the 1980s, different types of scanners have been used to obtain accurate models of human body, e.g. head scanners, foot scanners and entire body scanners. Most scanners use either laser or white light to measure the depth information of the body surface for modeling purpose. Although scanning can obtain accurate and detailed 3D models, their applications are restricted by expensive, and often bulky, equipment.

In response to this, many researchers then proposed reconstructive modelling methods, which used information such as partial scans [2], images [4,5] and measurements [2,6] to estimate the 3D shape of the body skin surface by morphing a deformable template model. Most deformable models were developed by the so-called example-based methods [2,3], which learnt statistically the shape models from a large range of scan data. One of most famous methods is SCAPE [2], which combined both shape and pose deformation into one template. Nevertheless, SCAPE [2] can describe general shape features (e.g. slim or fat body type) of human subjects well, but it cannot effectively deform detailed local shape features (e.g. slopy shoulder or the waist level).

Recently, Zhu et al. [1] developed a method to customize body shape of individual subjects from two orthogonal-view photos. They described the human body 3D shape by 17 key feature parts. For each part, they learnt a shape prediction function from a large scale of human body scans. They extracted all 17 local features of individual subject from the orthogonal-view photos, from which they assembled 3D shape according to the relative position of the 17 local features defined in the photos. Their method can customize human models with high measurement accuracy. Unfortunately, their method suffers from a drawback, similar to that of body scanning, that subjects must dress in tight-fit clothing in order to ensure the model accuracy.

In recent years, a few research work were reported that modelled human bodies under clothing. Hasler et al. [7] introduced a method for estimating the naked body shape from a dressed people's scan. It deformed an example-based deformable model to fit the dressed scan iteratively until meeting some defined constraints. In 2010, Guan et al. [8] described a generative model that combines the contour of 2D human body and the deformation of clothing. This model can be used to estimate 2D body shape and underlying poses from images. Zhou et al. [9] presented a method that reshapes human bodies in images. It first detects body profile in the input image, and matches the profile with a morphable 3D model. Next, it morphs the 3D model so as to drive the rendering of images for

reshaping the body of the subject in the image. The above two research work focused on estimating or deforming body shape under clothing in images, thus they mainly deal with 2D images. Extending the work to 3D space, Hasler et al. [10] developed a method based on a multi-linear model for 3D human pose and body shape deformation. It can estimate deformation parameters that drive the deformation of the 3D multi-linear deformable model, based on the silhouette captured from images. However, it is important to note all these statistical learnt deformable models can only capture the average shape deformation, but cannot reach the accuracy of body model customization. Therefore, none of these research applications is aim at obtaining an accurate 3D body shape model for a dressed subject in the images.

3. Methodology

In this paper, we proposed a two-phase method to customize human body models from photos in which subjects are dressed in normal or loose-fit clothing. In the first phase, we predict a 2D feature of the subject's body shape under clothing. Based on the predicted 2D feature, we construct a 3D body shape feature in second phase, and such 3D shape feature can be used to customize a detailed 3D model of the subject.

3.1. 2D feature prediction

Our method aims to create accurate body models under clothing based on customer's photos. Since photos only contain 2D information, similar to other related studies, we define customer's 2D feature as body profiles – front-view and side-view profiles. Since most part of body profile is being covered by clothing in the photos, the most challenging task is to predict a complete profile based on some cues not being covered by clothes.

To do so, we first establish a database of normalized body profiles. The profiles are extracted from more than 5000 scans of real subjects with different body shapes. The database covers a wide range of body shapes; some example profiles are showed in Fig.1.

Second, we extract body feature cues from photos and use which to predict the complete profiles using the profile database. Due to the complex situations of input images, e.g. different clothing and noisy background, we realize the cue extraction by manually selecting features points on the photos. These feature points are categorized into two types: reference points and boundary points. The reference points are used to define locations of key body features on the images, such as neck and ankle locations. The boundary points are used to define the potential body shape of the subjects. Generally speaking, boundary points should be defined at locations where the body contours are not being covered by clothing. To allow more flexibility, users can manually define, based on their experience, the boundary point positions anywhere on the photos, even at positions where body contours are being covered by clothing (e.g. the red point in Fig.2(a)). It is another reason why manual feature extraction is preferred in our method. We normally define a total of 7-9 boundary points on the front-view and side-view photos, which are used to predict the complete 2D front- and side-view profiles. In our method, the extracted boundary points are normalized, using the defined reference points, to match the format of profiles in the database. The normalized boundary points are used to search profiles with similar features in the profile database by three steps: 1) calculating the difference

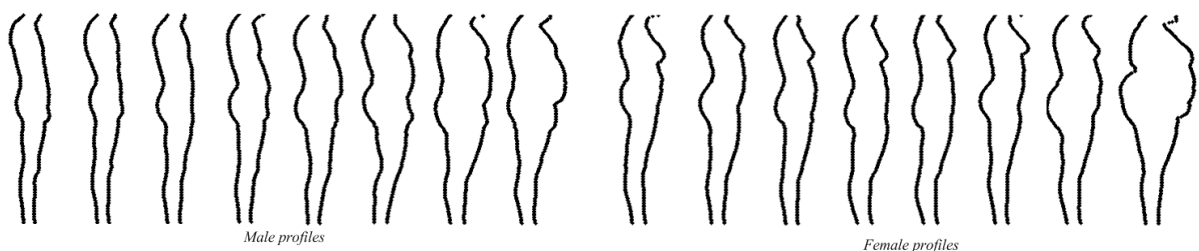*Male profiles*   *Female profiles*

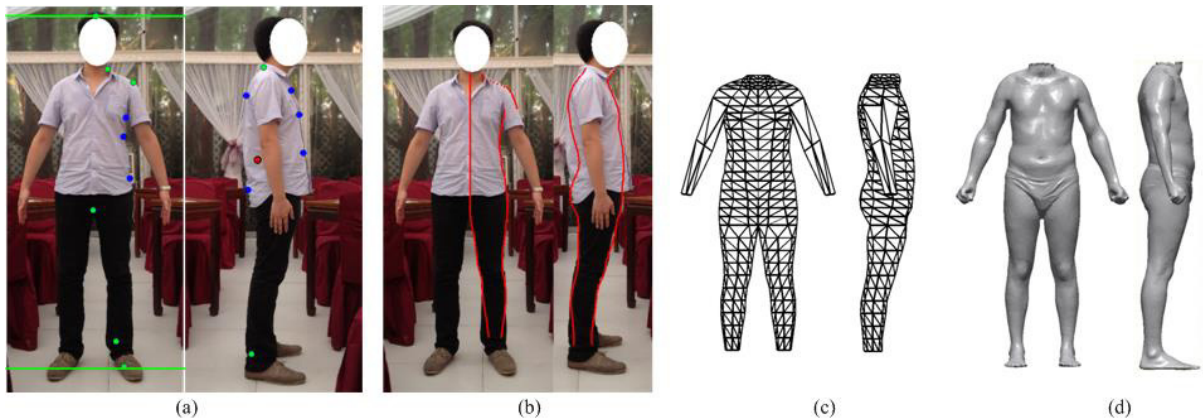Fig. 1. Example profiles in the profile database.

Fig. 2. Overall methodology: (a) cues extraction; (b) predicted profiles; (c) reconstructed 3D body shape feature; (d) customised model.

between boundary points and relevant feature points at the same level of the profile; 2) searching a number of *N* profiles with least total differences of all defined boundary points; and 3) synthesizing one single profile by combining the *N* selected profiles. Fig.2 (b) shows the example of estimated profile based on the points extracted in Fig.2(a).

3.2. 3D feature reconstruction

Similar to the work of [1], we define a framework to represent 3D body shape feature. In [1], the framework was constructed by defining the locations of 17 key cross-sections of body shape model from the photos. However, this method is not suitable for modelling clothed subjects because of two following reasons: (1) most of the body contour is covered by clothing, which makes it very difficult to define all the 17 cross sections; (2) the location identification of a number of cross-sections is very tedious. In this paper, we define a framework involving 30 cross-sections (as shown in Fig.2 (c)). The framework is built by first automatically recognising a small number of key cross-sections from 2D profiles, and then by interpolating extra cross-sections between recognized key cross-sections.

The second phase of the proposed method on clothed subject body model customization is to reconstruct the subject's 3D body feature from the estimated 2D profiles. To do so, we first extract local and global body features from predicted profiles, obtained in the phase one. The local features refer to widths and depths of the 30 cross-sections defining the 3D shape feature, and such cross-sectional widths and depths can be obtained by locating the relevant levels of the predicted front-view and side-view profiles. The global features refer to the relative positions of the key cross-sections. The cross-sectional widths and depths are then used to predict 3D shape of the particular cross-section. The prediction is based on relationship models learnt between local features and cross-sectional shape from a large scale of real human scanned models. With the predicted cross-sectional 3D shapes, we assemble or reconstruct 3D body shape feature using the global features extracted from the profiles. With the reconstructed 3D shape feature, a template model is then deformed to the shape of the particular subject using combined triangular Free Form Deformation (ct-FFD) algorithm [1]. Fig.2 (d) shows the deformed model of the example subject.

4. Experiment result

4.1. Loose-fit results

An experiment was carried out to evaluate the effectiveness of our model customization method. A total of 21 subjects, including 6 males and 15 females, were recruited. The subjects were found with diverse body builds that can be classified as underweight, normal and overweight. All subjects have their front-view and side-view photos

dressing in loose-fit clothes taken for model customization. All subjects also had their body scanned by [TC]² NX16 scanning system for comparison purpose. Fig. 3 compares some customized results with corresponding scanned models. Table 1 shows the ranges of discrepancy between the extracted girth measurement of the customized models and that of the scanned models. It can be found that all mean discrepancies are lower than 2.0cm, which is within the size tolerance of the clothing industry. Apart from girth measurements, the key cross-section at chest/bust, waist and hips are compared in Fig. 4.

Table 1. Range of discrepancy of six cross-section girth measurements between deformed models and scan models.

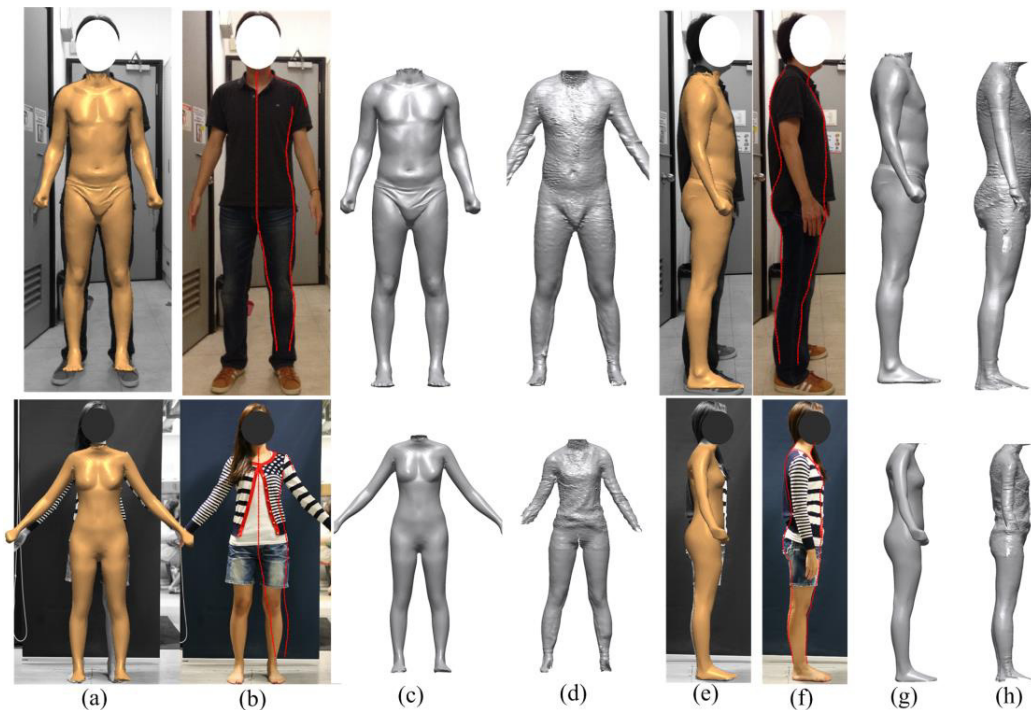| Cross-section	Range of size discrepancy (cm)	Mean absolute size discrepancy/Standard deviation(cm)	Mean absolute area variations (%)
Bust/Chest	(-1.88, 1.54)	1.063/0.54	2.18%
Waist	(-0.82, 1.19)	0.834/0.36	2.71%
Hip	(-1.41, 1.28)	0.954/0.63	2.64%
Shoulder	(-1.75, 1.77)	1.039/0.49	3.54%
Max. Thigh	(-1.13, 1.06)	1.182/0.59	3.91%
Calf	(-0.89,0.809)	0.742/0.55	3.29%

Fig. 3.(a)(e) Customized model mapping on photos; (b)(f) predicted profiles; (c)(g) customized models; (d)(h) scanned models.
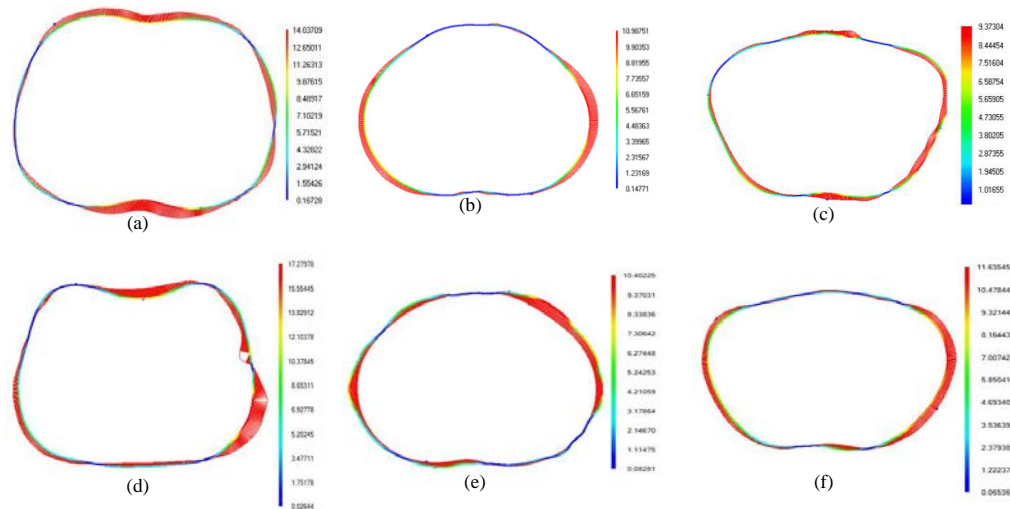
Fig. 4. Cross-sectional comparison between deformed and scanned models at (a)(d) chest/bust; (b)(e) waist; and (c)(f) hips level of a male subject and a female subject.

4.2. Tight-fit results

Since the proposed method can customize models using two-view photos, it can customize body models for subjects being dressed in loose-fit clothing in the photos, and it also can customize body models for subjects dressed in tight-fit clothing. We therefore customized all 30 models reported in Zhu et al. [1]. Fig. 5 compares some results of our method, that of [1] and the scans. The size discrepancy between deformed models and scans at six girth measurements are shown in Table 2. For comparison, Table 3 lists the size discrepancy between the customization results of [1] and the scanned models. As shown, we can find that our method has smaller mean size discrepancy and standard deviation. It is probably because 30 cross-sections better describe the 3D shape of human models. Moreover, our method reduces the amount of manual operation than [1], which reduces the human errors.

Table 2. Measurements comparison between our models and scan models.

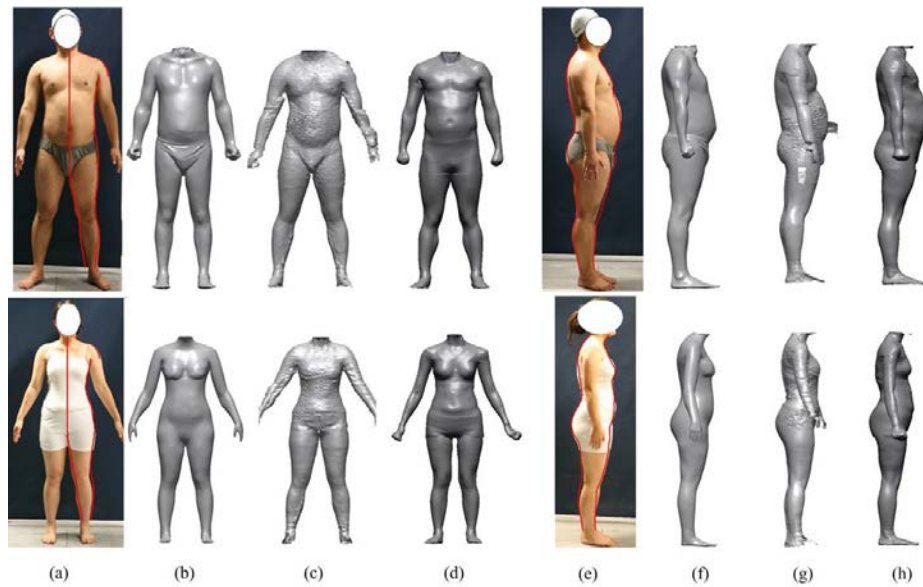| Cross-section	Range of size discrepancy - males (cm)	Range of size discrepancy - females (cm)	Mean absolute size discrepancy/ Standard deviation - male (cm)	Mean absolute size discrepancy/ Standard deviation - female (cm)	Mean absolute size discrepancy/ Standard deviation - all (cm)
Bust/Chest	(-1.8, 0.9)	(-1.6,1.6)	0.715/0.54	0.958/0.50	0.832/0.52
Waist	(-1.8, 2.0)	(-2.1,1.4)	0.863/0.62	0.855/0.58	0.858/0.60
Hip	(-2.3, 1.5)	(-1.3,1.7)	0.904/0.67	0.896/0.42	0.891/0.55
Shoulder	(-2.3, 1.3)	(-1.8,2.2)	0.764/0.61	0.835/0.61	0.798/0.59
Max. Thigh	(-1.8, 2.2)	(-2.1,1.4)	0.925/0.57	1.150/0.59	1.035/0.58
Calf	(-1.1, 1.0)	(-1.7,1.3)	0.624/0.27	0.843/0.49	0.728/0.41

Fig. 5. (a)(e) tight-fit photos and predicted profile;(b)(f) our customized model;(c)(g) scanned model;(d)(h)customized model of Zhu et al.[1].

Table 3. Measurements comparison between their models [1] and scan models.

Cross-section	Range of size discrepancy - males (cm)	Range of size discrepancy - females (cm)	Mean absolute size discrepancy/ Standard deviation - male (cm)	Mean absolute size discrepancy/ Standard deviation - female (cm)	Mean absolute size discrepancy/ Standard deviation - all (cm)
Bust/Chest	(-1, 1.5)	(-3.7, 1.8)	0.627 /0.39	1.427/0.96	1.027/0.83
Waist	(-1.6, 0.6)	(-4.4, 2.3)	0.659/0.40	1.480/1.36	1.070/1.07
Hip	(-2.8, 0.7)	(-2.8, 1.8)	1.113/0.80	1.113/0.82	1.113/0.80
Shoulder	(-2.8, 2.7)	(-2.3, 1.8)	1.052/0.93	0.748/0.65	0.897/0.81
Max. Thigh	(-1.9, 2.4)	(-2.7, 2.2)	1.108/0.643	1.268/0.86	1.187/0.75
Calf	(-0.8, 1.7)	(-1.7, 1.9)	0.563/0.41	0.894/0.59	0.730/0.53

5. Conclusion

We have proposed in this paper a rapid method for reconstructing precise 3D body models from customer's photos. Compared with the work of Zhu et al [1], it reduces tedious feature extraction operation on customer's photos. Moreover, it can reconstruct customer's detailed geometric characteristics even when subjects are dressed in loose-fit clothing in the photos. Experimental results have proved that (1) the method can customize customers' body models based on the photos of the customers who wore loose-fit clothing in the photos; (2) the resulting models have realistic appearance and accurate size measurements; (3) the customization process is efficient with minimal interactive operations; and (4) the process meets the requirement for real-time applications. To conclude, the method contributes in accurate human body model customization from photos.

Acknowledgements

References

[1] Zhu, S., Mok, P. Y., Kwok, Y. L. (2013). An efficient human model customization method based on orthogonal-view monocular photos. CAD Computer Aided Design, 45(11), 1314–1332.

[2] Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., Davis, J. (2005). Scape. ACM Transactions on Graphics, 24(3), 408.

[3] Seo H. and N. Magnenat-Thalmann (2004). An example based approach to human body manipulation. Graphical Models, 66(1): 1-23.

[4] Wang, C. C. L., Wang, Y., Chang, T. K. K., Yuen, M. M. F. (2003). Virtual human modeling from photographs for garment industry. CAD Computer Aided Design, 35(6), 577–589.

[5] Hilton A., Beresford D., Gentils T., Smith R., Sun W., Illingworth J. (2000). Whole-body modelling of people from multiview images to populate virtual worlds. Visual Computer, 16(7): 411-36.

[6] Wang, C. C. L. (2005). Parameterization and parametric design of mannequins. CAD Computer Aided Design, 37(1), 83–98.

[7] Hasler, N., Stoll, C., Rosenhahn, B., Thormahlen, T., Seidel, H. P. (2009). Estimating body shape of dressed humans. Computers and Graphics (Pergamon), 33(3), 211–216.

[8] Guan, P., Freifeld, O., Black, M. J. (2010). A 2D human body model dressed in eigen clothing. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6311 LNCS(PART 1), 285–298.

[9] Zhou, S., Fu, H., Liu, L., Cohen-Or, D., Han, X. (2010). Parametric reshaping of human bodies in images. ACM Transactions on Graphics, 29(4), 1.

[10] Hasler, N., Ackermann, H., Rosenhahn, B., Thormählen, T., Seidel, H. P. (2010). Multilinear pose and body shape estimation of dressed subjects from image sets. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1823–1830.

[11] Guan, P. G. P., Weiss, A., Balan, A. O., Black, M. J. (2009). Estimating human shape and pose from a single image. Computer Vision, 2009 IEEE 12th International Conference on.