Clothes Size Prediction from Dressed-human Silhouettes

Dan Song[1], Ruofeng Tong[1*], Jian Chang[2], Tongtong Wang[1], Jiang Du[1], Min Tang[1] and Jian Jun Zhang[2]

[1] State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China
trf@zju.edu.cn
[2] NCCA, Bournemouth University, Bournemouth, UK

Abstract. We propose an effective and efficient way to automatically predict clothes size for users to buy clothes online. We take human height and dressed-human silhouettes in front and side views as input, and estimate 3D body sizes with a data-driven method. We adopt 20 body sizes which are closely related to clothes size, and use such 3D body sizes to get clothes size by searching corresponding size chart. Previous image-based methods need to calibrate camera to estimate 3D information from 2D images, because the same person has different appearances of silhouettes (e.g. size and shape) when the camera configuration (intrinsic and extrinsic parameters) is different. Our method avoids camera calibration, which is much more convenient. We set up our virtual camera and train the relationship between human height and silhouette size under this camera configuration. After estimating silhouette size, we regress the positions of 2D body landmarks. We define 2D body sizes as the distances between corresponding 2D body landmarks. Finally, we learn the relationship between 2D body sizes and 3D body sizes. The training samples for each regression process come from a database of 3D naked and dressed bodies created by previous work. We evaluate the whole procedure and each process of our framework. We also compare the performance with several regression models. The total time-consumption for clothes size prediction is less than 0.1 second and the average estimation error of body sizes is 0.824 cm, which can satisfy the tolerance for customers to shop clothes online.

Keywords: Clothes Size Prediction, Body Size Estimation, Regression Methods

1 Introduction

Suitable clothes size plays a vital role in successful transactions of shopping clothes online. However, it is not easy for customers to achieve fitness when they buy clothes online. On one hand, customers need to possess professional skills to measure themselves with a special tool like a tape in a relatively private space.

* Corresponding author

On the other hand, customers have to check the size charts of items because of various size standards for different countries and different clothes brands. In a word, the techniques for automatic clothes size suggestion are in demand. We adopt 20 body sizes (Figure 2 (b)) which are related to clothes size, including length information and girth information of 3D bodies. Given such 20 body sizes, clothes size can be obtained by searching corresponding size chart. Therefore, we put emphasis on the automatic estimation of 3D body sizes.

Images contain valuable information and are quite convenient to get, so we devote ourselves to automatic image-based methods. Generally speaking, there are two ways to predict body sizes using images. One way is to measure 3D body shapes which are reconstructed from images, while another way is to utilize body landmarks which are extracted from images.

Image-based human body reconstruction has attracted lots of researchers. Some researchers learn the relationship between 2D positions in image and corresponding depth information using a database of 3D bodies. Some investigators train a parametric human body model with a body database and use images as constraints to deform a template mesh. Some methods integrate both ideas illustrated above for body reconstruction. They firstly estimate rough bodies according to the appearances of silhouettes using machine learning methods. Then the bodies are refined geometrically to satisfy the body contours of silhouettes. Body sizes are acquired by measuring 3D bodies reconstructed with these methods.

A few works firstly detect landmarks from body contours and then predict body sizes using body landmarks. Various ways have been tried to detect landmarks positions, which are introduced in section 2.1, but they have their limitations. Manual label requires a professional sense of landmarks locations and costs time and energy. Corner points detection restricts landmarks to be corner points and is very sensitive to the quality of silhouettes. Some researchers extract landmarks using Iterative Closest Point (ICP) for 2D curves which is time-consuming. Song et al. [1] adopt a 3D landmarks regression method which is very efficient and not sensitive to body contour appearances for using global information. However, they need camera calibration for regressing 3D landmarks from 2D silhouettes. Usually, machine learning methods are used to predict 3D body sizes from body landmarks.

Similar to the 3D landmarks regression method [1], we regress 2D landmarks from silhouettes. However, we do not calibrate camera but require height as input, which is much easier to get. Camera configuration (intrinsic and extrinsic parameters) has an influence on the size and shape of the human silhouette. The shape of the silhouette is affected little when the camera points to the center of human body at a relatively far distance in orthogonal view. The size of the silhouette is estimated using height under a fixed camera configuration. Then we regress the positions of 2D landmarks from resized silhouettes. Different from [1], we do not reconstruct 3D human body but learn body sizes using 2D landmarks. We introduce the framework and main processes of our method in section 3 and evaluate them in section 4.

For the specific task of body sizes estimation, the advantages of our method are summarized as: (a) automatic, (b) efficient (the total time-consumption is within 0.1 second), (c) effective (the average error is 0.824 centimeter) and (d) free of camera calibration. Previous methods can only satisfy part of the advantages stressed above. Additionally, our method is easy to be integrated into clothes shopping websites.

2 Related Work

2.1 Landmarks Detection Methods

Zhu et al. [2] manually label the positions of 2D landmarks. Lin and Wang [3] represent the front and side silhouettes with chain code [3] and detect corner points at 90-degree angles. They define body landmarks as some of these corner points with their criterions. Nguyen [4] define landmarks locations at a template body contour and detect landmarks for other body contours using Iterative Closest Point (ICP) for 2D curves. Cheng et al. [5] detect 2D body landmarks from a depth image of minimally-dressed people with a boosting tree regression method. Similar to [5], Song et al. [1] regress 3D landmarks from the front and side silhouettes of normally-dressed people. They design different feature descriptors for regression. They are inspired by the methods applied to human face such as face alignment [6–8] and face recognition [9].

2.2 Regression Methods

Regression is a statistical process for estimating the relationship among variables. For our tasks, we have tried linear and non-linear regression models and compare their results. Linear regression models are often fitted using the least squares approach, but they may also be fitted in other ways. Ridge regression [10] adds an $L2$-norm regularization to the linear least squares function. They are generally used when multicollinearity is present or when overfitting is a problem. Random forest [11] regression fits a number of classifying decision trees on various subsamples of the dataset and use averaging to improve the predictive accuracy and control over-fitting. Gradient boosting regression [12] produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. Like other boosting methods, gradient boosting combines weak regressors into a strong one in an iterative manner. The successor regressor learns to correct its predecessor by reducing the residual. Support vector machine [13] (SVM) is used for classification and regression. In addition to linear classification, SVM can efficiently perform non-linear classification using kernel trick [14] by implicitly mapping their inputs into high-dimensional feature spaces.

[3] https://en.wikipedia.org/wiki/Chain_code

2.3 Image-based 3D Human Body Reconstruction

Chen and Cipolla [15] reconstruct 3D models from a single view by learning the relationship between 2D positions in an image and their corresponding depth information. They firstly use a template shape to encode the 2D positions in an image and corresponding depth information for 3D bodies in the database. Then they perform Principal Component Analysis (PCA) to reduce the dimension of 2D positions and depth information. Finally, they trained the Gaussian Process Latent Variable Model (GPLVM) from the combinational inputs of both 2D positions and depth information. Balan et al. [17] estimate 3D body shapes from dressed-human silhouettes in 4 views by optimizing a parametric human model (SCAPE [16]). They calibrate cameras for 4 views and set higher weight for exposed-skin parts. They optimize the body model through minimizing the pixels differences between silhouettes and projections of target mesh in 4 views. Boisvert et al. [18] reconstruct 3D human shape from front and side silhouettes by integrating both geometric and statistical priors. Firstly, they train a non-liner function connecting silhouettes appearances and body shapes to make a first approximation. Secondly, with body contours as constraints, body shapes are globally deformed along the principal directions of body database. Finally, they deform body shapes locally to ensure more fitness to input silhouettes.

3 Method

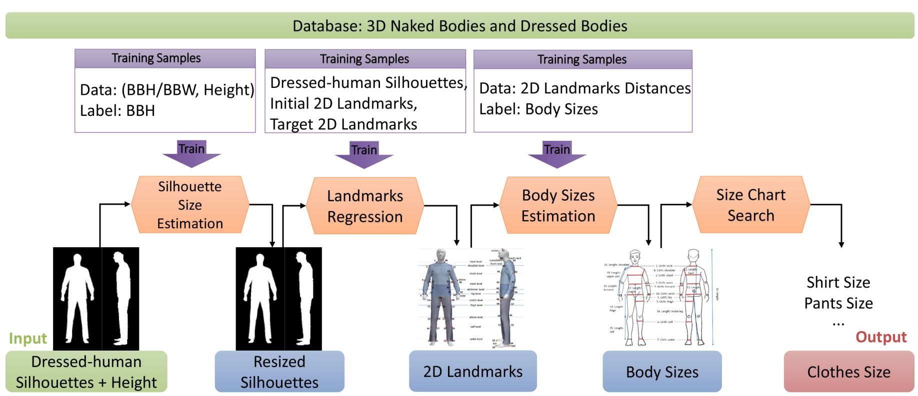


Fig. 1. ***Overview***. We compute the bounding box of human contour in input silhouette and calculate the ratio BBH/BBW (i.e. the height of the bounding box divides the width of the bounding box). According to BBH/BBW and height information, we estimate BBH under our virtual camera configuration and resize silhouette. For both front and side silhouette, we regress the positions of 2D landmarks. 3D body sizes are estimated using corresponding 2D landmarks distances. Finally, clothes sizes are obtained through searching the size chart. We use a database of 3D naked and dressed bodies to construct training samples.

We take a pair of dressed-human silhouettes (front and side) and human height as input and get clothes sizes as output. We propose a data-driven method to efficiently estimate body sizes from dressed-human silhouettes. With body sizes in hand, we obtain clothes sizes by searching the size chart. Figure 1 shows the overview of our method which contains 4 processes. We use a database [1] of 3D naked and dressed bodies to learn: (a) the relationship between height and silhouette size under our virtual camera configuration; (b) the relationship between dressed-human silhouette and 2D body landmarks and (c) the relationship between 2D body sizes and 3D body sizes. The training data is introduced in section 3.1 and the main processes are illustrated in the following sections.

3.1 Training Data Preparation

We use the training database constructed by previous work [1]. The training database has 1081×5 pairs of 3D naked and dressed bodies in total. The 3D naked bodies are synthesized from real bodies with a standard pose in MPI database [19]. The corresponding dressed bodies are acquired by dressing naked bodies with physically based cloth simulation. We get dressed-human silhouettes by projecting 3D dressed bodies in front and side views under our virtual camera configuration. The landmarks of 3D naked bodies are projected to 2D landmarks. 3D body sizes are computed by measuring 3D naked bodies in the database. In summary, our training data contains 3D body sizes, 2D landmarks and dressed-human silhouettes.

3.2 Silhouette Size Estimation

Camera configuration (intrinsic and extrinsic parameters) has an effect on the size and shape of the human silhouette. Therefore, most image-based methods [1][17] need camera calibration to recover 3D information from the silhouette. However, when the camera points approximately to the center of human, the direction of camera view is orthogonal to human plane and the distance between camera and human is relatively far, perspective projection has little effects on the shape of body contour. We set our virtual camera configuration and restrict the input of our method to satisfy these three conditions. We use height information as a clue to estimate the silhouette size under our virtual camera configuration.

We suppose a linear relationship between human height and silhouette size, and use training data to learn the relationship. The bounding box of body contour is computed, and we calculate the height of the bounding box (abbr. BBH) and the width of the bounding box (abbr. BBW). We also compute the height information of bodies in the training database. For silhouettes in the training database, we have their BBH and BBW values and compute the ratio between them (BBH-BBW- Ratio, abbr. HWR). Then we train a linear model (Ridge Regression [10]) with training samples whose data are (height, HWR) and labels are BBH. We also tried several other models to fit the training data and compare their results in section 4.2.

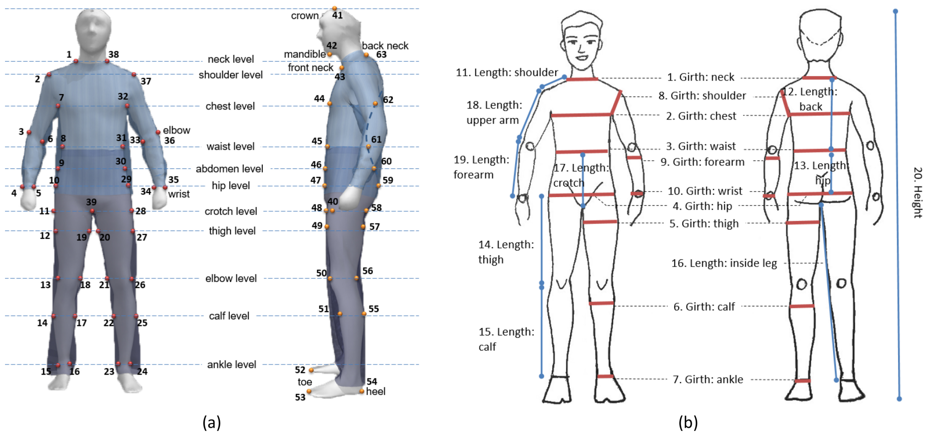


Fig. 2. ***Landmarks and body sizes.*** (a): the locations of 63 landmarks. Landmarks 1-39 are used for front silhouettes while landmarks 40-63 are designed for side silhouettes. (b): 20 body sizes. 10 of them are length information and the others are girth information of body.

3.3 2D Landmarks Regression

The locations of body landmarks we use are shown in figure 2 (a), which are related to the body size measurements. We estimate the positions of 2D landmarks for both front and side silhouettes using a regression method similar to [1].

Song et al. [1] regress 3D landmarks using a pair of silhouettes (front and side) with a boosting tree regression method. They construct training samples as front and side silhouettes, initial and target 3D landmarks, and learn the relationship between the appearance of silhouettes and 3D landmarks movements. They define the feature descriptor as the displacement from projected 2D landmarks and their nearest contour points. For projecting landmarks, they need to calibrate the camera.

We train two separate models for both front silhouette and side silhouette. Our training samples consist of front or side silhouette, initial and target 2D landmarks. We adopt the same boosting tree regression method to estimate 2D landmarks from the silhouettes. For more details of the regression method, please refer to section 6 of [1].

3.4 Body Sizes Estimation

We use the measurements which are closely related to clothes size as our 3D body sizes (figure 2 (b)), and estimate 3D body sizes from 2D landmarks. We first obtain 2D body sizes by computing the distances between related 2D landmarks. The correspondences between 2D body sizes and 2D landmarks are shown in table 1. Then we predict 3D body sizes using 2D body sizes with a linear regression model.

Table 1. The Related Landmarks for 2D Body Sizes

Size No.	Related Landmarks No.	Size No.	Related Landmarks No.
1	(1,38), (43, 63)	11	(1, 2), (37, 38)
2	(7, 32), (44, 62)	12	(1, 8), (31,38)
3	(8, 31), (45, 61)	13	(8, 10), (29, 31)
4	(10, 29), (47, 59)	14	(10, 13), (26, 29)
5	(12, 19), (20, 27), (49, 57)	15	(13, 15), (24, 26)
6	(14, 17), (22, 25), (51, 55)	16	(16, 39), (23, 39)
7	(15, 16), (23, 24)	17	(8, 39), (31, 39)
8	(2, 7), (32, 37)	18	(2, 3), (36, 37)
9	(3, 6), (33, 36)	19	(3, 4), (35, 36)
10	(4, 5), (34, 35)	20	(41, 54)

2D distances (unit: pixel) between related 2D landmarks are computed as 2D body sizes. Our 3D body sizes are classified into length information and girth information. Length information is acquired by computing 3D distances between related 3D landmarks. The correspondences between length values and related 3D landmarks are similar to 2D situation (table 1). Each girth information of body is related to a group of vertices which are marked with pink color in figure 3 (a). These 3D vertices are projected to a plane and we compute the convex-

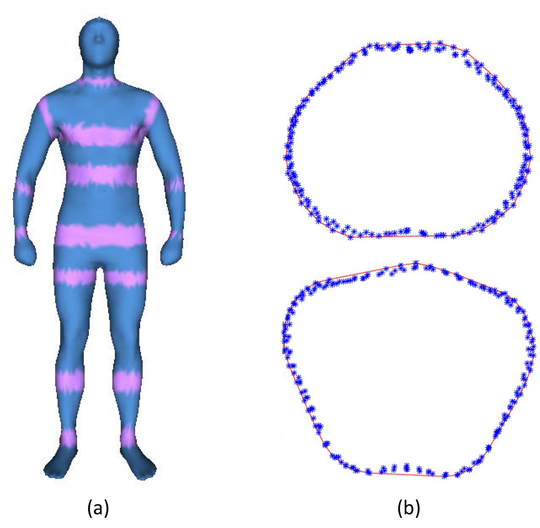


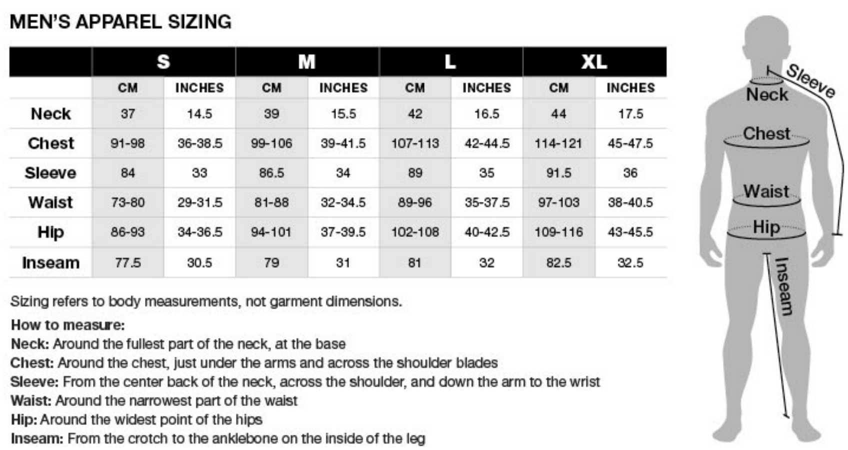
Fig. 3. ***Girth calculation***. (a): related vertices for girth information. The related vertices are marked with pink color. (b): Two examples of girth calculation. Firstly, the related 3D vertices are projected to a plane. Then, we compute the convex-hull of 2D projected points. Girth information is the circumference of the convex-hull.

hull of 2D projected points. We define the girth value as the circumference of the convex-hull. Figure 3 (b) shows two examples of girth calculation.

We suppose a linear relationship between 2D sizes and 3D sizes, and use training data to train the regression model (ridge regression). 2D body sizes are prepared as the data of training samples, while 3D body sizes are used as the label of training samples. We train a separate model for each body size. We also try several other regression models to fit training data, and compare their results in section 4.4.

3.5 Size Chart Searching

Figure 4 shows an example of men's apparel sizing for online shopping.[4] According to the measurements of neck, chest, sleeve, waist, hip and inseam, customers determine their clothes sizes. Our proposed body sizes contain these information and are consistent with the common size chart.


MEN'S APPAREL SIZING

	S		M		L		XL	
	CM	INCHES	CM	INCHES	CM	INCHES	CM	INCHES
Neck	37	14.5	39	15.5	42	16.5	44	17.5
Chest	91-98	36-38.5	99-106	39-41.5	107-113	42-44.5	114-121	45-47.5
Sleeve	84	33	86.5	34	89	35	91.5	36
Waist	73-80	29-31.5	81-88	32-34.5	89-96	35-37.5	97-103	38-40.5
Hip	86-93	34-36.5	94-101	37-39.5	102-108	40-42.5	109-116	43-45.5
Inseam	77.5	30.5	79	31	81	32	82.5	32.5

Sizing refers to body measurements, not garment dimensions.

How to measure:
Neck: Around the fullest part of the neck, at the base
Chest: Around the chest, just under the arms and across the shoulder blades
Sleeve: From the center back of the neck, across the shoulder, and down the arm to the wrist
Waist: Around the narrowest part of the waist
Hip: Around the widest point of the hips
Inseam: From the crotch to the anklebone on the inside of the leg

Fig. 4. ***An example of size chart.*** An online shopping website uses this figure to illustrate the size chart for men's apparel.

4 Results

The hardware environment for our experiments is a 64-bit desktop with 32GB RAM. The processor is Intel(R) Core(TM) i7-4790K CPU at 4.00GHz. The whole procedure completes within 0.1 second. The prediction time-consumption of size regression is within $1e^{-4}$ second. 2D landmarks regression costs about

[4] https://blackdiamondequipment.com/en/size-chart-apparel-mens-f13.html

$1.60e^{-2}$ second. We illustrate our testing data in section 4.1. The whole procedure and each process of our method are evaluated in the following sections. We should mention that currently we use male bodies in 1 clothes type to validate our framework.

4.1 Testing Data Illustration

We use the testing database of previous work [1] containing 637 pairs of 3D naked and dressed bodies, which are separate from training database. The resolution of testing silhouettes is 800×600. The average BBH (i.e., height of human bounding box) value of front silhouettes is 362.98 pixels and the average BBH value of side silhouettes is 372.26 pixels. We use testing data to evaluate each process of our pipeline and the whole procedure in the following sections.

4.2 Estimation Error of Silhouette Size

We compute BBH (i.e., height of human bounding box), HWR (i.e., BBH/BBW) and height values for testing data. With HWR and height values of testing data as input, we estimate BBH value with several regression models and compare the results with the ground truth. Table 2 and 3 compare the estimation error for BBH value. The results show that ridge regression performs best for both front and side conditions.

Table 2. Estimation Error for Front Silhouette Size (unit: pixel)

Error \ Model	Ridge	SVR	Random Forest	Gradient Boosting
Average	1.58	3.36	2.22	1.81
Stdev	1.12	1.78	1.66	1.25

Table 3. Estimation Error for Side Silhouette Size (unit: pixel)

Error \ Model	Ridge	SVR	Random Forest	Gradient Boosting
Average	1.63	1.93	1.89	1.81
Stdev	1.18	1.70	1.63	1.35

4.3 Regression Error of 2D Landmarks

We project 3D dressed bodies and 3D landmarks of naked bodies in the testing database to dressed-human silhouettes and 2D landmarks with the same virtual camera. Given dressed-human silhouette, we regress 2D landmarks positions and compare them with the ground truth. The average landmarks regression error is 0.64 pixel for font silhouette and 0.57 pixel for side silhouette.

4.4 Estimation Error of Body Sizes

For testing data, we compute 2D body sizes and 3D body sizes. Then we use 2D body sizes to estimate 3D body sizes with several regression models, and compare the results with the ground truth. Table 4 illustrates the error of different regression models.

Table 4. Estimation Error for Body Sizes (unit: centimeter)

Error \ Model	Ridge	SVR	Random Forest	Gradient Boosting
Average	0.42	1.03	0.46	0.42
Stdev	0.38	0.90	0.47	0.40

4.5 Overall Estimation Error of Body Sizes

In this section, we test the overall estimation error of 3D body sizes. For each testing data, we take dressed-human silhouettes and height information as input, and use 3D body sizes as the ground truth. Firstly, we estimate silhouette

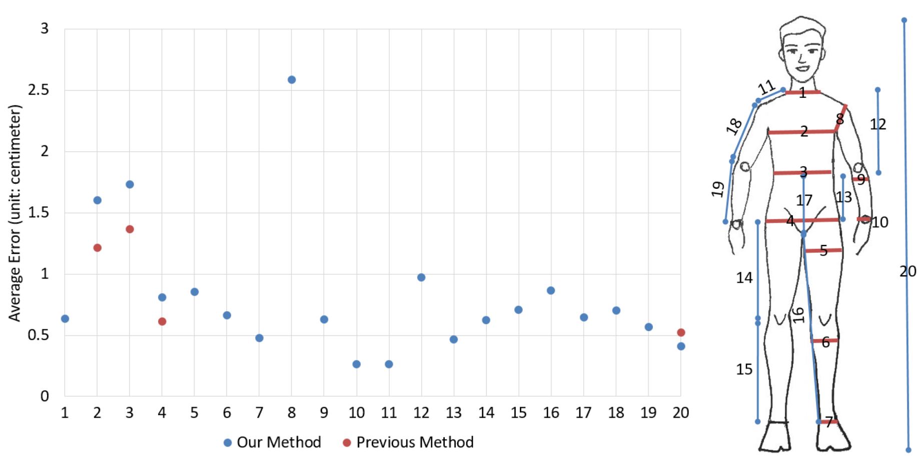


Fig. 5. ***Average error for 20 body sizes.*** (a): average error for 20 body sizes. The blue points show the error of our method while the red points illustrate the error of the previous method [1]. (b): the body size that each number stands for.

size with ridge regression method and resize silhouette. Secondly, we regress 2D landmarks from resized silhouette. Thirdly, based on estimated 2D landmarks, we compute 2D body sizes. Finally, 3D body sizes are acquired using 2D body sizes with ridge regression method. Figure 5 shows the average error for each one of 20 body sizes. With the same testing database, Song et al. [1] estimate chest/waist/hip girth and height. The results show that we can acquire comparative accuracy. We should mention that the time-consumption of [1] is within 4 seconds while ours is within 0.1 second. Our method is more efficient for clothes size suggestions for online shopping.

4.6 Implementation Details

The configurations of regression methods for size estimation are shown in table 5. For 2D landmarks regression, the configuration is the same as [1] whose regression target is 3D landmarks.

Table 5. The Configuration Details for Regression Models

Regression Method	Configuration
Ridge	(alpha = 1.0, fit_intercept = True, normalize = False, copy_X = True, max_iter = None, tol = 0.001, solver = 'auto', random_state = None)
SVR	(kernel = 'rbf', degree = 3, gamma = 'auto', coef0 = 0.0, tol = 0.001, C = 1.0, epsilon = 0.1, shrinking = True, cache_size = 200, verbose = False, max_iter = -1)
Random Forest	(n_estimators = 10, criterion = 'mse', max_depth = None, min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_features = 'auto', max_leaf_nodes = None, min_impurity_split = 1e-07, bootstrap = True, oob_score = False, n_jobs = 1, random_state = None, verbose = 0, warm_start = False)
Gradient Boosting	(loss = 'ls', learning_rate = 0.1, n_estimators = 100, subsample = 1.0, criterion = 'friedman_mse', min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_depth = 3, min_impurity_split = 1e-07, init = None, random_state = None, max_features = None, alpha = 0.9, verbose = 0, max_leaf_nodes = None, warm_start = False, presort = 'auto')

5 Conclusion, Limitations and Future Work

We adopt 20 body sizes which are closely related to clothes size, including length information and girth information of 3D body. We explore an automatic framework to efficiently estimate 3D body sizes from dressed-human silhouettes. We get rid of camera calibration through estimating the size of silhouette under known virtual camera configuration with human height information. We learn the relationship between the positions of 2D landmarks and the appearance of silhouette. Several regression models are tried to estimate 3D body sizes from 2D body sizes. We have compared several regression models and ridge regression is most suitable for our tasks. The whole procedure and each process of our framework are evaluated. Our method completes in less than 0.1 second and the average estimation error of body sizes is 0.824 cm. This satisfies the real-time and accuracy requirements for customers to shop clothes online.

We provide an effective and efficient solution for clothes size prediction when customers buy clothes online, but we still face some difficulties to overcome. It will be more convenient if we take natural photos instead of silhouettes as input without the limitation of clothes types. We validate our framework through male bodies and should extend our work to female situations.

Acknowledgements:

The research is supported in part by NSFC (61572424) and the People Programme (Marie Curie Actions) of the European Unions Seventh Framework Programme FP7 (2007-2013) under REA grant agreement No.612627-"AniNex". Min Tang is supported in part by NSFC (61572423) and Zhejiang Provincial NSFC (LZ16F020003).

References

1. Song, D., Tong, R., Chang, J., Yang, X., Tang, M., Zhang, J. J.: 3D Body Shapes Estimation from Dressed Human Silhouettes. Computer Graphics Forum 35(7), 147-156 (2016)
2. Zhu, S., Mok, P. Y.: Predicting Realistic and Precise Human Body Models Under Clothing Based on Orthogonal-view Photos. Procedia Manufacturing 3, 3812–3819 (2015)
3. Lin, Y. L., Wang, M. J. J.: Automatic feature extraction from front and side images. Industrial Engineering and Engineering Management, IEEM 2008. IEEE International Conference on (pp. 1949-1953). IEEE. (2008)
4. Nguyen, H. T.: Automatic Anthropometric System Development Using Machine Learning. BRAIN. Broad Research in Artificial Intelligence and Neuroscience 7(3), 5-15 (2016)
5. Cheng, K. L., Tong, R. F., Tang, M., Qian, J. Y., Sarkis, M.: Parametric human body reconstruction based on sparse key points. IEEE transactions on visualization and computer graphics 22(11), 2467-2479 (2016)
6. Dollar, P., Welinder, P., Perona, P.: Cascaded Pose Regression. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1078-1085 (2010)
7. Cao, X. and Wei, Y., Wen, F., Sun, J.: Face Alignment by Explicit Shape Regression. International Journal of Computer Vision, 107(2, SI), 177-190 (2014)
8. Cao, C., Weng, Y., Lin, S., Zhou, K.: 3D Shape Regression for Real-time Facial Animation. ACM Transactions on Graphics, 32(4) (2013)

9. Shao, H., Chen, S., Zhao, J., Cui, W., Yu, T.: Face recognition based on subset selection via metric learning on manifold. Frontiers of Information Technology & Electronic Engineering 16(12), 1046-1058 (2015)
10. Hoerl, A. E., Kennard, R. W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1), 55-67 (1970)
11. Liaw, A., Wiener, M.: Classification and regression by randomForest. R news 2(3), 18-22 (2002)
12. Friedman, J. H.: Greedy function approximation: a gradient boosting machine. Annals of statistics (pp. 1189-1232) (2001)
13. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning 20(3), 273-297 (1995)
14. Boser, B. E., Guyon, I. M., Vapnik, V. N.: A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152) (1992)
15. Chen, Y., Cipolla, R.: Learning shape priors for single view reconstruction. Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on (pp. 1425-1432). IEEE. (2009)
16. Anguelov, D., Srinivasan, P., Koller, D, Thrun, S, Rodgers, J., Davis, J.: SCAPE: Shape Completion and Animation of People. ACM Transactions on Graphics 24(3), 408-416 (2005)
17. Balan, A. O., Black, M. J.: The Naked Truth: Estimating Body Shape Under Clothing. Computer Vision - ECCV 2008, PT II, Proceedings, Lecture Notes in Computer Science, 5303, 15-29 (2008)
18. 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)
19. Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., Seidel, H. -P: A Statistical Model of Human Pose and Body Shape. Computer Graphics Forum 28(2), 337-346 (2009)