Journal of Visual Languages and Computing 47 (2018) 9–18

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Automatic human body feature extraction and personal size measurement

Tan Xiaohui[a], Peng Xiaoyu[a], Liu Liwen[b], Xia Qing[*,b]

[a] College of Information and Engineering, Capital Normal University, Beijing, China
[b] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

ARTICLE INFO

Keywords:
Size measurement
Feature points extraction
Geodesic distance
Heat kernel

ABSTRACT

It is a pervasive problem to automatically obtain the size of a human body without contacting for applications like virtual try-on. In this paper, we propose a novel approach to calculate human body size, such as width of shoulder, girths of bust, hips and waist. First, a depth camera as the 3D model acquisition device is used to get the 3D human body model. Then an automatic extraction method of focal features on 3D human body via random forest regression analysis of geodesic distances is used to extract the predefined feature points and lines. Finally, the individual human body size is calculated according to these feature points and lines. The scale-invariant heat kernel signature is exploited to serve as feature proximity. So our method is insensitive to postures and different shapes of 3D human body. These main advantages of our method lead to robust and accurate feature extraction and size measurement for 3D human bodies in various postures and shapes. The experiment results show that the average error of feature points extraction is 0.0617cm, the average errors of shoulder width and girth are 1.332 cm and 0.7635 cm, respectively. Overall, our algorithm has a better detection effect for 3D human body size, and it is stable with better robustness than existing methods.

1. Introduction

Automatic human body size measurement is very important to personalized products and customization services, especially for garment application [1]. In principle, feature points extraction is always the first essential step, the definition of features should be application-specific, and the accuracy and effectiveness of feature extraction undoubtedly influence the quality of garment application in the most significant way. In the view of ergonomics, during the design or virtual try-on process of human-oriented garment, the human body should be regarded as a core consideration. The 3D virtual garment is constructed to align with the human body feature points or lines. Extracting human feature points or curves can also drive the garment deformations.

Such design automation of human-centered garment products or virtual try-on applications heavily rely on establishing the volumetric parameter of human bodies and feature points usually serves as anchors to constrain this volumetric parameterization. In prior researches, these feature points are always specified by users or semi-automatically selected by rule-based systems [2,3], but how to automatically and properly map certain kinds of focal features from users' psychological/semantic awareness to the right geometric signals is not yet well known.

The automatic extraction on human body is challenging for two reasons. First, the feature points on human body are not always located at the prominent tip location. However, traditional feature points are usually captured on the shape extremities or protrusions, and inevitably are biased towards high shape saliency in some senses, which tends to ignore non-salient but semantically-meaningful regions. Second, the robustness of local shape matching is more problematic when the postures of human bodies are varying or with different body types.

In this paper, we propose an automatic way of size measurement on 3D human body based on extraction of focal features via random forest regression analysis of geodesic distances. The method proposed in this paper aims at providing an automatic way to extract feature points and measure different sizes on 3D human bodies, which usually serves as a pre-processing step for garment designer or virtual fitting application. Feature points can be identified immediately in a universal way. The primary contribution of this paper is a novel modeling approach based on ensemble learning to extract focal features on 3D human body. The predicted displacements of the randomly sampled points on the new 3D human body surface are used to vote for the desired semantic focal features. So the method is automatic in a data-driven way. The scale-invariant heat kernel signature is exploited to serve as feature proximity, so our method is insensitive to postures and different shapes of 3D human body. These main advantages of our method lead to robust and accurate feature extraction and size measurement for 3D human bodies in various postures and shapes.

This paper is organized as follows: Section 2 is an overview of related work. Section 3 is the semantic focal feature definition and the

* Corresponding author.
E-mail address: xiaqing@buaa.edu.cn (X. Qing).

https://doi.org/10.1016/j.jvlc.2018.05.002
Received 18 December 2017; Received in revised form 10 February 2018; Accepted 2 May 2018
Available online 08 May 2018

detailed methodology of our algorithm. In Section 4, the Random Forest regression for local feature identification is presented. And in Section 5, we present some results along with a brief discussion of specific topics and possibilities for the future work.

2. Related work

In practical personalized garment design and virtual fitting application, feature points usually determine the locations of cutting planes and feature lines of the virtual garment. The pre-process of body size measurement is to identify the feature points of the human body, and then simulate the consumer's morphological shape, which provides references for generation of the virtual garment shape [4,5].

Feature detection on 3D shapes, also known as feature extraction, is a classical problem in geometric modeling and computer graphics [6–9]. Researchers intended to focus on the protrusions on 3D shape. Features are often defined in terms of local curvature. All these local geometric features are described by local Gaussian curvature [9], ridge and valley lines [10], prominent tip points [11], the mean curvature flow [10] and the average geodesic distance distribution on the surface [12]. Benefiting from 2D image feature detection methods, some techniques such as 3D-Harris method [13], mesh-SIFT method [14] were extended in the 3D situation. The aforementioned methods depend on the geometric features and the saliency of the points are fully defined by the 3D shape itself while lacking semantic meaning. Shelling points method [15] can combine the geometric properties with users' intention to locate landmark points on new shapes. However, it requires tedious selecting tasks and the landmark detecting process still relied on Gaussian curvature of local geometric properties.

As we can see from the feature extracting process for the non-rigid 3D shape, the methods can essentially boil down to two key components: the descriptor and assignment algorithm. Different approaches such as contour and edge features [16], local patches [17,18], conformal factor [19], differential operators [10], and local volume properties [20] were used as feature descriptors in shape retrieval literature. Unfortunately, none of them satisfies all of the above desired properties such as intrinsic and some of them is not robust to topology, which are important to personalized garment related applications. Dealing with non-rigid 3D shapes requires compensating for the degrees of freedom resulting from deformations. In [21] and follow-up works [22] , authors processed 3D shapes as metric spaces with intrinsic (e.g. geodesic) distances, which are invariant to inelastic deformations. As to 2D shapes, this frame is used with a metric defined by internal distances in 2D shapes [23]. The Laplacian spectra are used as intrinsic shape descriptors [24]. Feature detection based on intrinsic scale-space analysis such as Sochen et al. [25] would find a few reliable points (usually with high curvature), at which scale estimation can be done. In flat regions, no scale estimation is possible. For this reason, In [26], authors avoided feature detection and used a dense feature descriptor computed at every point of the shape in combination with statistical weighting to reduce the influence of trivial points.

The features always represented by grouping regional information in point as the descriptor. The bag of features paradigm relies heavily on the choice of the local feature descriptor that is used with 3D shapes. In non-rigid shape related applications, an ideal feature descriptor should be first of all intrinsic and thus deformation invariant. Second, it should cope with missing parts, and also be insensitive to topological noise and connectivity changes. Third, it should work across different shape representations and formats and be insensitive to sampling. Finally, the descriptor should be scale-invariant. The last two properties are especially important when dealing with shapes coming from different person in virtual fitting application, where shapes appear in a variety of representations and with arbitrary postures. Heat kernel signatures (HKS) is proposed as a deformation-invariant descriptor based on diffusion of multi-scale heat kernels [27]. HKS is a point based signature satisfying all of the good descriptor properties except for scale invariance. The temperature distribution (TD) of the heat mean signature (HMS) as a shape descriptor is defined for shape matching [28]. Their TD is a global shape descriptor and they used L2 norm which is a very basic matching method to compute the distance between two TD descriptors. With a series of transformations, the HKS scale problem is solved [29]. The same research group has proposed Shape Google approach [30] based on the scaled-invariant HKS. The idea is to use HKS at all points of a shape, or alternatively at some shape feature points, to represent the shape by a Bag of Features (BoF) vector. The aforementioned methods are based on intrinsic metric preservation and simply depend on the local/geometric properties.

To overcome the challenge of automatic feature extraction, some machine learning based methods have been proposed which have shown promising performance. Random Forest (RF) based methods are becoming more and more popular. RF was originally proposed by Breiman [31] for general classification or regression, and the class-specific Hough forest was presented in [32] for object detection. Since then, the methods based on RF have shown very promising results in tasks related to landmark detection or organ localization from medical data [33]. The basic idea is to sample local patches (image patches in 2D and volumes in 3D data) in the image, estimate the displacement from the patches to the landmark by RF regression, and then the landmark position is estimated by a voting scheme considering the individual estimations from all the patches. In particular, Criminisi et al. [34], uses regression forest to predict the offsets between the target anatomy and sampled points in the tomography scans and localizes the anatomy's position based on these offsets, and in [31], they used the same idea to automatically detect landmarks in cephalometric X-Ray images.

Inspired by the former work but having sharp contrasts to them, this paper will simultaneously take into account the local information in the neighborhood of points and the symmetrical feature, with an observation that the randomly-sampled points on the shapes can provide convincing vote sources for locating the desired feature points collectively, and give rise to more precise and robust results. At the same time, the efficiency can be enhanced.

3. Semantic focal feature definition

Given 3D human body model is represented as triangle mesh surface $M_S = (V, T)$ with m vertices $V = \{v_1, v_2, ..., v_m\}$, triangle surfaces $T = \{t_1, t_2, ..., t_n\}$ and a set of predefined feature points G_S. We are going to find the corresponding feature points G_H of an input human model H. Without loss of generality, M_H is also represented by a triangle mesh as M_S. The automatic extraction is challenging for two reasons. First, the feature points on human body are not always located in the shape extremities but defined by the requirement of the virtual try-on application, therefore the local shape matching based methods cannot robustly give satisfactory results. Second, the robustness of local shape matching is more problematic when the postures of human bodies are varied (i.e., the 3D bodies are bent). In this paper, we propose an automatic feature point and size measurement method. Firstly, the randomly-sampled points are spread out over the global shape of M_H. Secondly, the final locations of the unknown feature points are collectively decided by the predicted displacements of these randomly-sampled points.

3.1. Garment related sematic focal feature

The critical constraints of proper 3D body and garment measurement are to align 3D garments with body features to guide the later virtual try-on process. The specific points and lines act as defined focal features, which naturally convey the prior knowledge and can be located anywhere on the 3D body shape. In the proposed method in this paper, anatomical landmarks and body features represented by feature points and lines are accurately recognized on both industrial mannequins and non-ideal real-world bodies of various figure shape types.

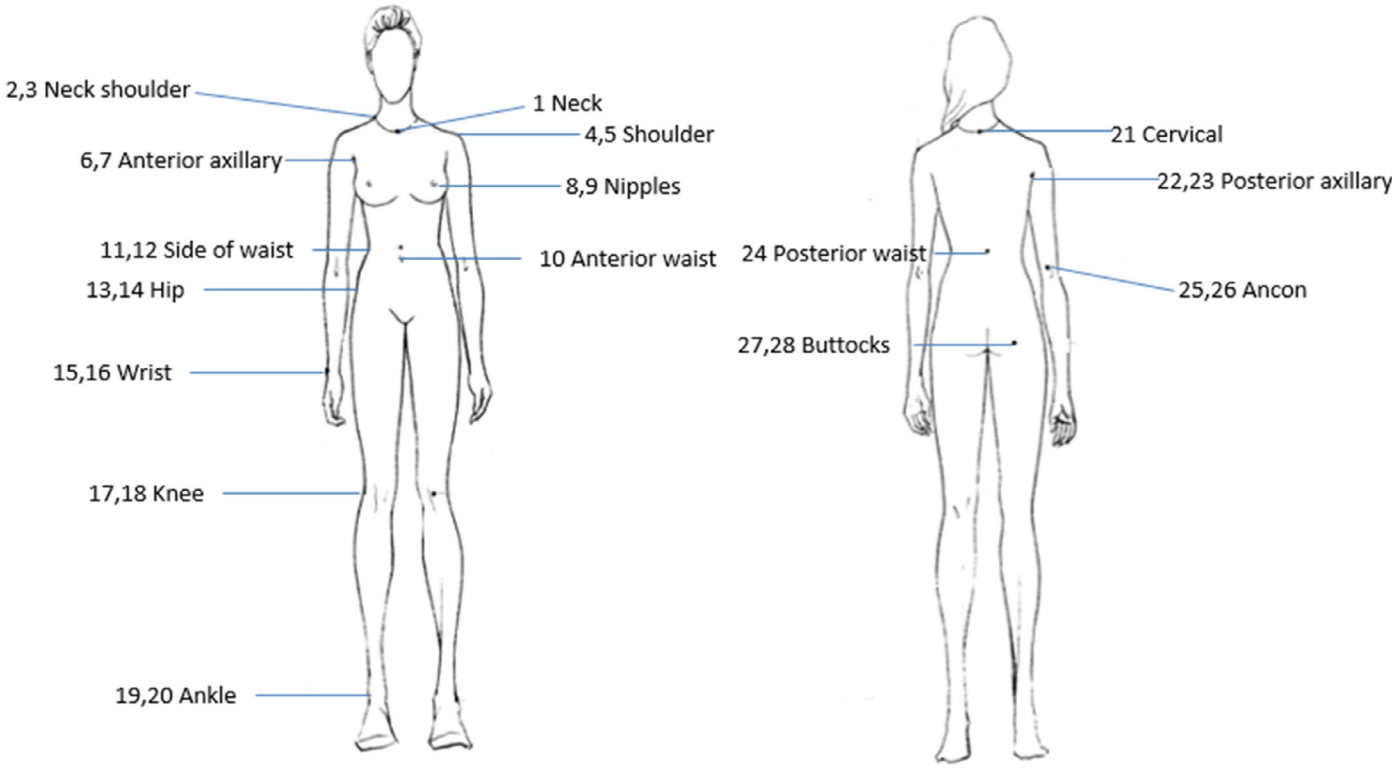

Fig. 1. Defined focal feature points.

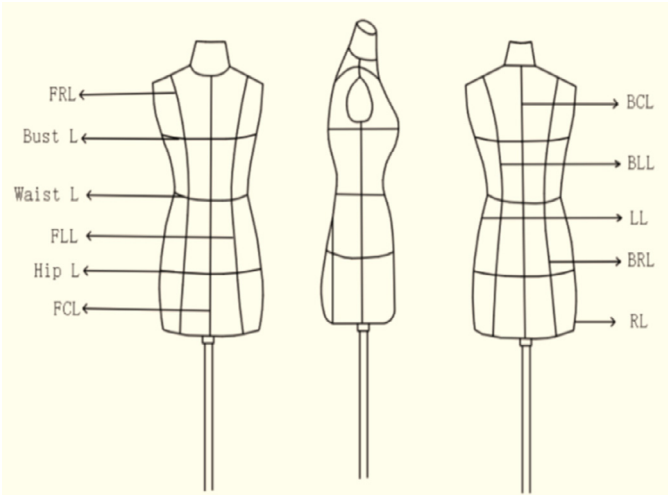

Fig. 2. Defined focal feature lines.

ASTM D5219-09 [35] and ISO 8559:1989 [36] standards are used in the apparel industry for taking anthropometric measurements and locating anatomical landmarks on human bodies. These standards are used to locate body features on the parametric human models in the method. These feature points, as illustrated in Fig. 1, and feature lines, as showed in Fig. 2, are important for which 3D garment should align with.

3.2. SI-HKS for randomly sampled points representation

The features paradigm relies heavily on the choice of the local feature descriptor that is used to represent randomly sampled points. There has been a large amount of works done on designing various local descriptors in the context of 3D shape analysis. Heat kernel signature was introduced as an intrinsic local shape descriptor based on diffusion scale-space analysis. In this paper, scale-invariant version of the heat kernel (SI-HKS) descriptor is adopted to represent the randomly-sampled points.

On compact manifolds, the heat kernel can be presented as [37]

$$K_{X,t}(x, t) = \sum_{i=0}^{\infty} e^{-\lambda_i t}\varphi_i(x)\varphi_i(z), \tag{1}$$

where $\lambda_0, \lambda_1, \ldots \geq 0$ are eigenvalues and $\varphi_0, \varphi_1, \ldots$ are the corresponding eigenfunctions of the Laplace–Beltrami operator, satisfying $\Delta x\varphi_i = \lambda_i\varphi_i$.

For each point v on the shape, the HKS can be compactly represented as a n-dimensional vector:

$$h(v) = (h_1(v), h_2(v), \ldots h_n(v))^T,$$

and

$$h_i(v) = \sum_{j=0}^{\infty} e^{-\lambda_j t}\varphi_j^2(v). \tag{2}$$

The diffusion time t_i is usually logarithmically sampled as suggested in [29]. An ideal feature descriptor should be first of all intrinsic and thus deformation invariant. Second, it should cope with holes, and also be insensitive to topological noise and connectivity changes. SI-HKS is an enhanced version of HKS, for local normalization of the heat kernel signature, which does not suffer from sensitivity to deformations. Fig. 4 shows three kinds of descriptors to extract focal feature points. When the shape is perturbed by Gaussian noise and the shape has many holes, comparison the curve of three kinds of descriptors shows the SI-HKS is still stable while the other two descriptors fluctuate in a wide range, which means the SI-HKS is much more robust than the other two descriptors.

SI-HKS regards HKS as a digital signal and lifts original restrictions in shape scales via Fourier transformation. Moreover, since most of the signal information is contained in the low-frequency components of the Fourier transformation. We can construct a compact descriptor by sampling a small number of low frequencies. We thus have constructed the scale-invariant quantity from the HKS at each point, without

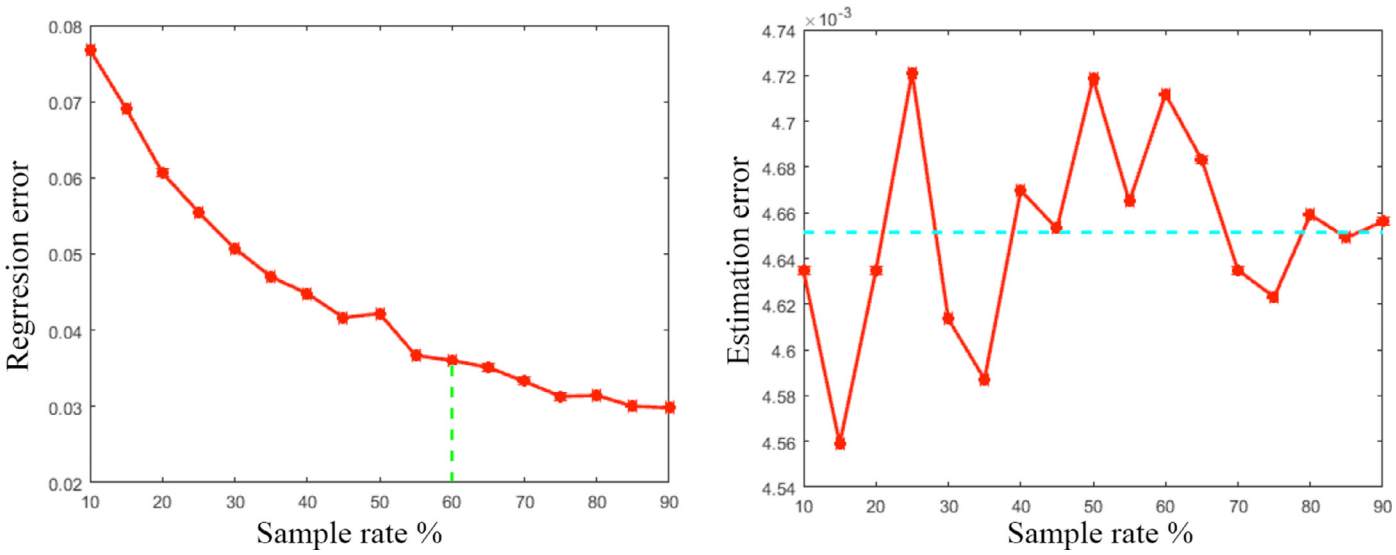

Fig. 3. Sample rate.

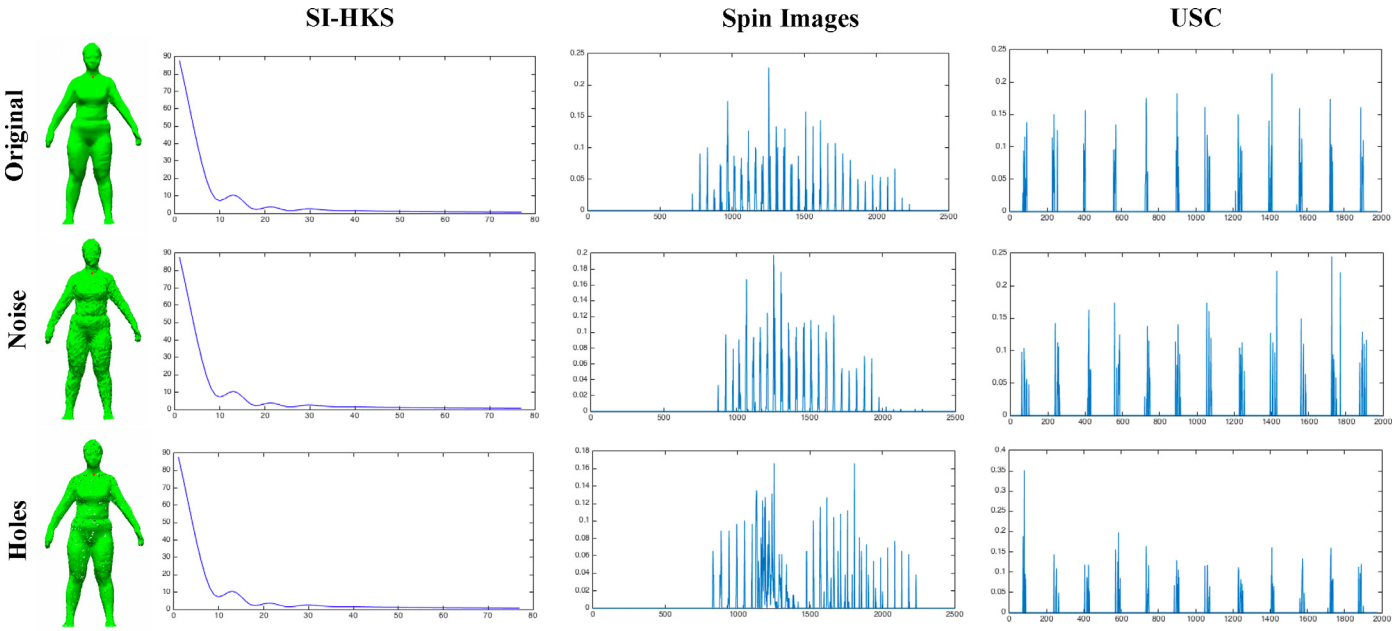

Fig. 4. Stability comparison of SI-HKS [29], SpinImages [39] and unique shape context [40].

performing scale selection. This allows us to compute descriptors at any point of our shape.

Numerical computation of the SI-HKS is done using Formula (2), in which a finite number of terms is taken and the continuous eigenfunctions and eigenvalues of the Laplace–Beltrami operator are replaced by the discrete counterparts. The discretization of the Laplace–Beltrami operator depends on the representation on the shape. For shapes represented as triangular meshes, one of the most common discretization is the cotangent weight.

3.3. Fast relative displacement measurement on 3D body shape

Geodesic distance is the shortest path between two points on a three-dimensional surface, which is analogous to a straight line between two points on a plane. The algorithms like Dijkstra or fast marching can be used to find the shortest path and compute its length. The main drawback of these methods is that they do not reuse information and they highly rely on the triangulation of the shape and is sensitive to noise. In this paper, Geodesics-in-Heat [38] is adopted to be the basic distance metric. The heat method is robust, efficient, and simple to implement since it is based on solving a pair of standard linear elliptic problems. The resulting systems can be prefactor once and subsequently solved in near-linear time. In practice, distance is updated an order of magnitude faster than with state-of-the-art methods, while maintaining a comparable level of accuracy.

The geodesic distance between any pair of points on a Riemannian manifold can be recovered via a simple point-wise transformation of the heat kernel. The heat method consists of the following three basic steps.Step 1. Integrate the heat flow: $\dot{u} = \Delta u$, for fixed time step t. Step 2. Evaluate the vector field $X = -\nabla u/|\nabla u|$. Step 3. Solve the Poisson equation $\Delta\varphi = \nabla \cdot X$.

Here, Δ is the Laplace–Beltrami operator and already been computed when computing the SI-HKS in the sample representation procedure. The function φ represents geodesic distance which is approaching the true distance as t goes to zero. The heat geodesic method asks only that the gradient point is parallel to $\Delta\varphi$. Since the gradient of

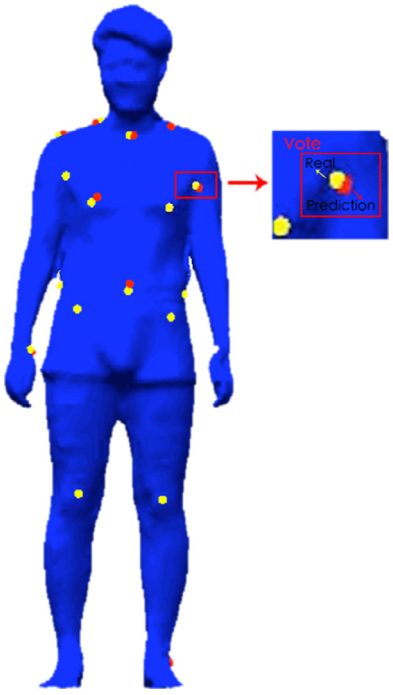

Fig. 5. Defined focal feature points.

the true distance function has unit length, magnitude can safely be ignored. We therefore compute the normalized gradient field $X = -\nabla u/|\nabla u|$ and find the closest scalar potential $\Delta\varphi$.

4. Random forest framework for focal feature identification

Regression forest can map complex input into simpler, discrete or continuous output. Random forests are collections of decision binary trees. The original problem is split into smaller ones by a tree. The smaller problems can be addressed with simple predictors [41]. A binary test containing in the non-leaf node guides sample data towards the left or the right child node. The tests are chosen in a supervised way and training a tree to cluster the training allowing good predictions with simple models. Our purpose is that to estimate the position of the desired feature points on 3D body model by randomly-sampled points on the shape. These points contribute with different confidence to estimating the position of the desired feature points. The displacement from the sampled points to the desired feature points should be similar with those on the training shapes. Randomness is introduced by the subset of training examples provided to each tree.

In the following, we describe briefly the training approach of a random forest. In order to employ such a random forest framework for try-on related feature points recognition and measurement of human body dimensions, we have to acquire annotated training data T, define binary tests, forest training, forest testing and try-on related features localization.

4.1. Training data

Building a forest is a supervised learning problem, before training process, each model in training set should be annotated with labels on the desired output space. In our 3D body measurement setup, a training sample is a 3D model scanning from Kinect, the predefined feature points are annotated by hand. Each point is annotated with real-valued vector represents the location of the feature point, while $\overrightarrow{V_f} = \{v_{fx}, v_{fy}, v_{fz}\}$. These marked points act as defined focal features.

To facilitate the subsequent learning procedure, we can randomly-sample many points on the training shapes and compute their local descriptors. Then we measure the relative displacements from them to the desired feature points. An estimated function maps the descriptors to the displacements. An example of a training point is shown in Fig. 5, The yellow dot represents the ground truth location of the specific feature point and the red dot is the prediction of the point by voting strategy which will be described in detail in 4.3.

4.2. Forest training and testing

Before the training processing, the vertex of 3D human body models in the training dataset are sampled with the fixed rate. The number of the sample points N is and each of the points are calculated by scale-invariant heat kernel signatures(SI-HKS) as their local descriptors. The geodesic distance between the sample point and the user-specified feature point is also computed.

Firstly, we set the number of decision binary trees as $T = 30$, then the number of sampled points N is randomly-sampled. The set of signatures denoted as P stored in the root of every single tree. Starting from the root, each tree is built recursively by assigning a binary test to each non-leaf node J. The binary test denoted as $h(p, j)$, where $\theta_j = (\varphi_j, \psi_j, \tau_j)$ as

$$h(p, j) = [f(:\varphi_j, \psi_j) > \tau_j], \tag{3}$$

where $\varphi = \varphi(p)$ is a map equivalent to select some elements from vector $\overrightarrow{p}$. ψ customizes the geometric primitives to separate the data, we

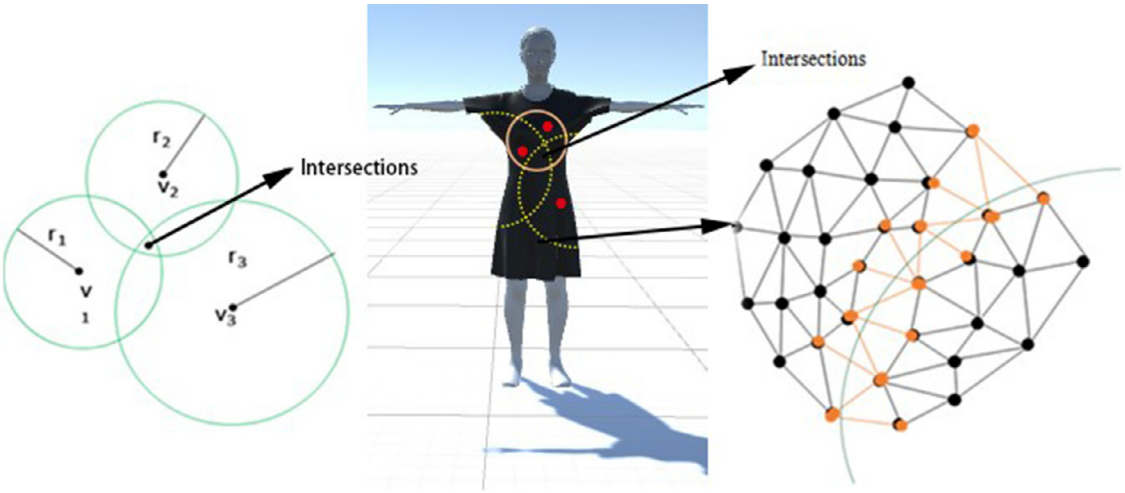

Fig. 6. Locating method.

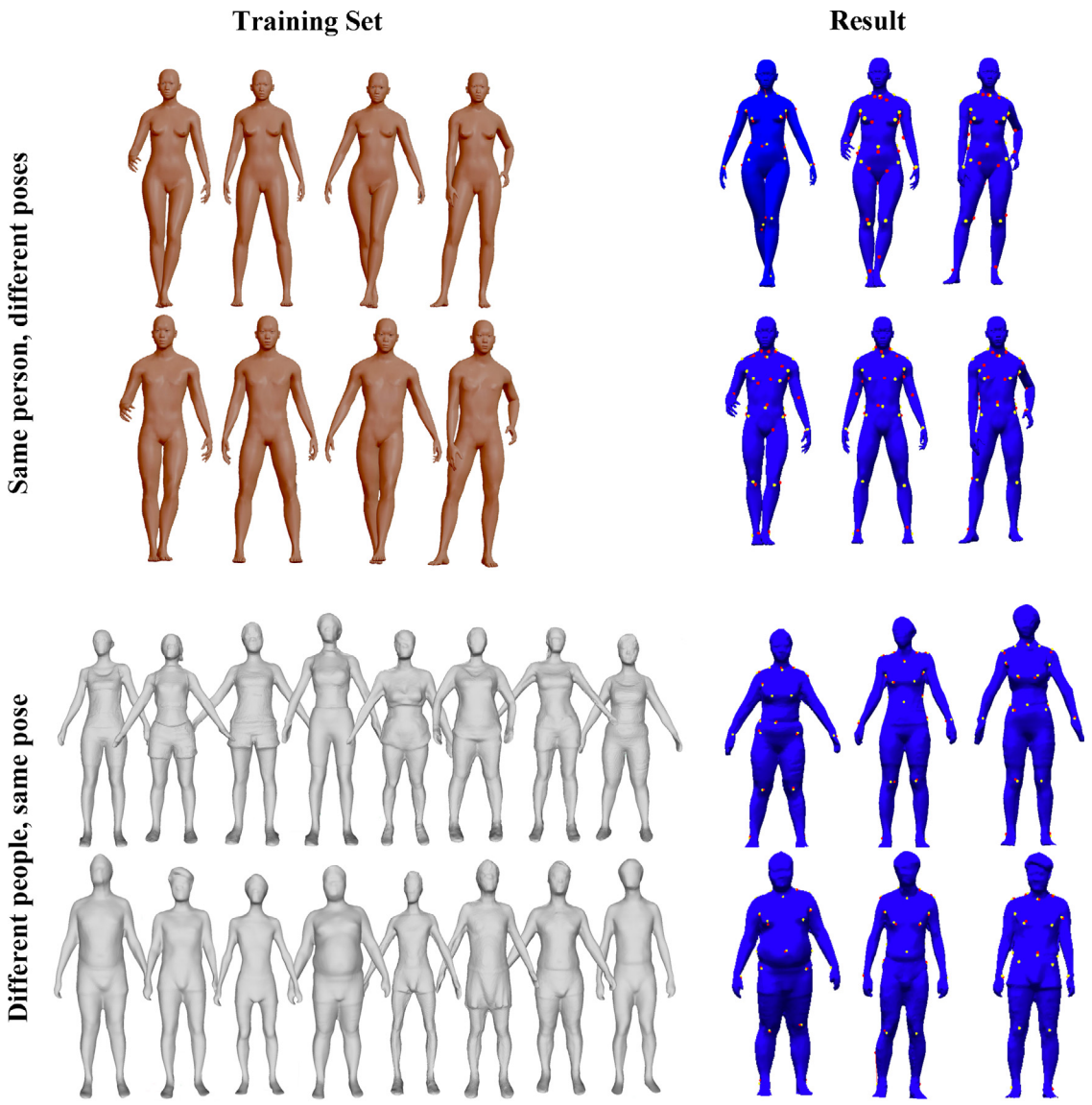

Fig. 7. Defined focal feature points.

express it by a linear expression as 4

$$f = \varphi(p)\psi \tag{4}$$

The expression (3) represents a logical judgment that returns true when the inequality satisfied. Such test sends each point either to the left or right child of the tree according to the binary test on local descriptor. The process continues with the left and the right child using the corresponding training sets until a leaf is created when n_s the number of $\overrightarrow{p}$ on the current node is less than 5, then the process of the training should be quit. We could obtain a decision model to estimate the geodesic distances.

During the testing processing, the vertexes of models in the testing set are sampled with the fixed rate . The signature of the sample points is q .We could get the normal distribution estimation of the sub-node, then we get an even distribution of all binary trees as (5)

$$p_\varepsilon(D) = \frac{1}{T}\sum_{t=1}^{T} p_t(D), \tag{5}$$

and this normal distribution is used to the final estimates, the expression of the parameter correspond to q and geodesic distance D_t as 6

$$E(D_t) = \int_D D p_\varepsilon(D) dD, \tag{6}$$

where $E(D_t)$ is the mathematical expectation of normal distribution, the value of it is used to predict geodesic distance D_t.

4.3. Focal features localization

With the forest constructed above, we can randomly sample points on the new 3D body model and compute their descriptors. Then the trained random forest regressor are invoked to predict the relative displacements from these points to the unknown predefined feature ones. Here a locating method like GPS (Global Positioning System) is used in this paper. As shown in the left side of Fig. 6, this is an example of 2D illustration. Three satellites are located at v_1, v_2, v_3, as to each satellite there are cycles with radius r_1, r_2, r_3 respectively, the receiver lies in the intersection region of these three cycles. A lot of predicted displacements are not exactly equal to the ground-truth ones, but the values are oscillating around the ground truth ones, that means the predicted positions may fall into a small neighborhood of the target feature point. A voting strategy to determine the location of the feature points are used here. We compute a sphere with the sample point as

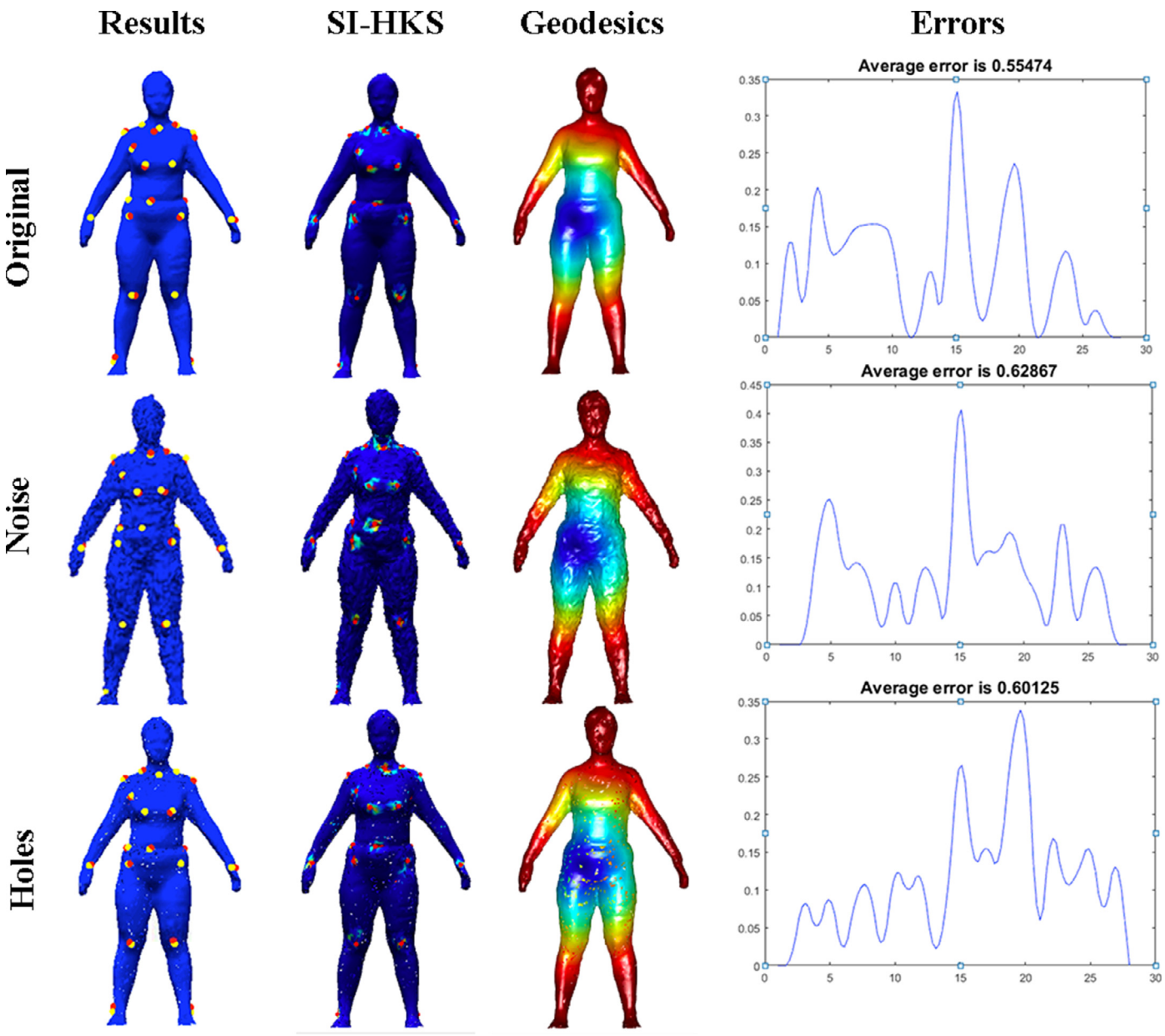

Fig. 8. Comparisons of feature extraction with holes and noise.

Table 1
Accuracy statistics.

Model	#v	#s	Training				Testing			Error
			sr	t_s	t_d	t_t	sr	t_e	t_v	
Man(fat)	13.6K	5	0.6	18.7908	20.8172	16.3090	0.3	0.0432	1.4666	0.6913
Man(slim)	13.6K	5	0.6	16.6587	16.9882	16.3090	0.3	0.0383	1.4376	0.9354
Man(normal)	13.6K	5	0.6	17.5806	18.7353	16.3090	0.3	0.3648	1.4773	1.3859
Woman(fat)	13.6K	5	0.6	16.4709	19.0484	13.8240	0.3	0.0353	1.4247	1.7888
Woman(slim)	13.6K	5	0.6	12.6102	7.8153	13.8240	0.3	0.0265	0.9338	0.9696
Woman(normal)	13.6K	5	0.6	16.7238	16.6073	13.8240	0.3	0.0401	1.4132	1.2655

From left to right, #v: number of shape vertices, #s: number of body shapes, sr: randomly sampling rate on each shape, t_s: average time for descriptors' computation on each shape, t_d: average time for geodesics' computation on each shape, t_t: time for forest training, sr: randomly sampling rate on testing shape, t_e: time for forest testing, t_v: time for voting, Error: average error of all feature points to the predefined feature points of the testing shape.

center and displacement as radius. And then, the sampled point will give one vote to the two points of each intersected edge. We weight the vote of this sampled point according to its corresponding displacement as $w = e^{-\mu * d * d}$, here μ indicates the influence range. Finally, we count the votes on each point, and the average votes in a small neighborhood can be described as the probability of being the feature point, the feature point is the point with highest average votes.

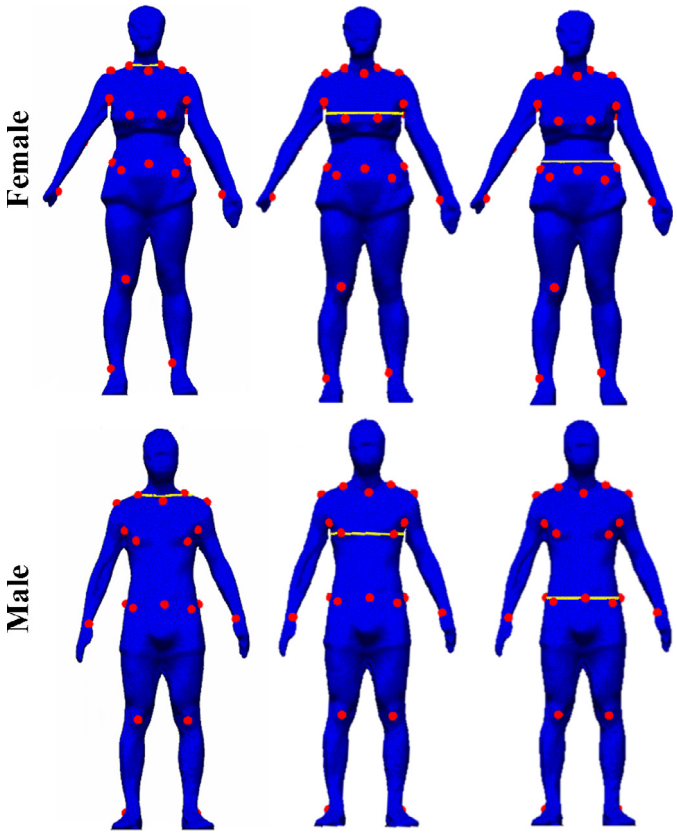

Fig. 9. Feature lines for computing garment size.

Table 2
Measured human body parameters.

Model	Parameters	Cs	Gs
$Model_1$	shoulder width	34.900	34.1
	Bust girth	97.041	95.7
	Hips girth	104.052	98.5
	Waist girth	87.488	80.5
$Model_2$	shoulder width	46.154	49.8
	Bust girth	93.412	92.1
	Hips girth	91.775	97.1
	Waist girth	89.063	82.1
$Model_3$	shoulder width	47.700	46.7
	Bust girth	103.132	98.9
	Hips girth	102.664	101.2
	Waist girth	90.804	89.3
$Model_4$	shoulder width	40.327	39.8
	Bust girth	92.095	85.2
	Hips girth	95.995	96.4
	Waist girth	78.991	77.8
$Model_5$	shoulder width	47.931	45.6
	Bust girth	121.984	114.5
	Hips girth	119.679	122.5
	Waist girth	121.068	116.5

From left to right,Cs: computation size of the human body model after normalization, Gs:ground truth size of the human body model.

5. Results and discussion

We have implemented the proposed feature extraction and automatic body size computation framework in a prototype program by C++/Matlab. The experimental tests are carried out on a PC with Intel Core i5 430 CPU (2.27GHz) plus 4GB main memory running 64bit MS Windows 7. **Data.**

In our experiments we use both synthetic and real data. The synthetic dataset is made up of shapes by MakeHuman. In order to make the dataset more realistic, each model is made according to the revised garment standard GB/T1335-81 of China in 2008. The second dataset is composed of real scans acquired with a Kinect sensor. The 3D body models covered four types of shapes Y, A,B and C. As for the real data in the training dataset, we use the KinectFusion algorithm to reconstruction human body with a single Kinect depth camera. There are 10 men and 10 women in our experiments. The people sampled in our experiments are 20 to 45 years old, their weights range from 40 kg to 88 kg and their heights range from 1.55 m to 1.85 m. The data was scanned using ICP tracking and scene fusion process and Fig. 7 shows examples of some models of human database.

Parameters Setting. Sample rate is the crucial parameter which play important role on the accuracy in testing the training phase. Fig. 3(a) indicates that when the sample rate in training dataset approach 60% the rate of regression error turned slow. Similarly, the plots in Fig. 3(b) show the variation of estimation error rate in test phase.

In training and testing phase, we adopt the optimization value of sample rate is 60% and 30% in our training and testing experiments. We implement the random forest based on the Sherwood Library[45], 30 decision trees are used for each regression, the tree construction stops when the leaf node contains fewer than 10 descriptors or the information gain is no longer increased.

5.1. Robust feature extraction with different body shapes

Fig. 7 shows the results of our approach on male and female examples of human bodies with different postures made by MakeHuman software. All the results are generated automatically. As shown in Fig. 7, our proposed method can accurately extraction the predefined body features of humans with different postures. In other words, our proposed method can be easily applied in public situations, such as shopping malls and can be widely used in home situations. As mentioned above, SI-HKS and geodesics are both invariant under isometric deformations, so our feature extraction also inherits this powerful property. Both for the man and woman examples, we have 10 shapes, for each training shape, we randomly sampled 60% points as the training points, and only 30% points are sampled to estimate the location of the feature points on the testing shape. From the results, we can see our method can precisely predict the focal feature points on a new 3D body shape. Fig. 8 shows the results of feature points extraction under perturbations like noise and many holes. As we know, heat diffusion on a 3D model is insensitive to noise or holes, thus, the descriptions and geodesics derived from heat inherit this property obviously and make our method robust under these perturbations as show in Fig. 8.

5.2. Accuracy

Table 1 documents the accuracy statistics of the experiments we have conducted above. We can see that, as for a given new 3D body shape, it can predict the location of predefined garment related feature points. We can see that, it only takes a few seconds to predict the location of our focal feature points, which makes our method very suitable for practical use. And the training time is also acceptable, because our idea is to establish connections between the randomly-sampled local points of the global shape with the location of the desired focal features, which takes advantage of information from many other sampled points and makes our method require much less samples to be trained. The error distance in this table is measured by the Euclidean distance between the predicted feature point and its desired location of the testing shape. The predicted feature points always lie in the 1-ring or 2-ring neighborhood of the ground truth points. It may be noted that, some errors are naturally caused by the inaccuracy of users' selection action due to the lack of accurate pointing device.

Fig. 10. 3D Body size measurement on application of virtual-fitting.

5.3. Automatic garment related human size computation

For measuring the shoulder width and bust, hips and waist girth, we specify the indices of the start and end points according to the pre-defined feature points in advance. The algorithm can automatically measure these parameters from an obtained annotated model. For measuring the bust, hips and waist girth, we specify the indices of a circle of points around a corresponding location according to the standard definition in advance. Then the related geodesics distance is summed up to obtain the size of the girth. As shown in Fig. 9, the girth is drawn on the 3D human body model. Table 2 shows the human body parameters that our system measures. Absolute size of the human body model is the size of parameters on the scanned 3D model computed by our algorithm. After normalization of the parameters of the body model, we can get the computation size of the model based on the absolute value. From the result of Table 2, we can see the computation size of the 3D human body is really close to the ground truth value. The average relative errors of shoulder width and girth are 1.332 cm and 0.7635 cm.

5.4. Why in 3D

Measurement of the human body size is precedent to accurate cloth fitting and therefore fundamental to garment productions. There are many problems when simply measuring body size in 2D images. The first one is to compute the girth. In 2D case, the girth of waist can only be computed by doubling the width of waist and this is actually not accurate, because people with more fat at their abdomen obviously have bigger girth of waist even though they share the same width of waist. However, in 3D case, the girth can be simply computed by the sum of geodesic distances between the left most and right most points at the waist. In addition, it is easier to build correspondence between human and cloth using 3D model, and the deformations of cloth driven by anchors defined on 3D human shape are more plausible and realistic, especially when one moves forward and back (with depth varying), and this is hard in 2D case.

With advancement in technology, 3D depth camera is designed to capture the shape of a human body in seconds and further produce its true-to-scale 3D body model. The use of data generated is extensive, including enabling a pragmatic approach to the offer of mass customization and facilitating virtual model fit trials that enhance online clothing shopping experiences. Consumers have become savvier than ever and as the demand for well-fitted garments is increasing, 3D body size measurement is being viewed as a significant bridge between craftsmanship and computer-aided design technologies. The structural lines of the 3D block garments are properly aligned with the features of the body. Garment fit is guaranteed by controlled ease distribution to the blocks in three-dimensional space. It is expected to facilitate consumer satisfaction and reduce commercial waste due to 'ill-fit' returns. With the body-size measurement, we can visually get the virtual-fitting result, as show in Fig. 10.

6. Conclusion

The proposed method in this paper presents an automatic human body feature extraction and size measurement by random forest regression analysis of geodesics distance. Our automatic feature extraction and size measurement methodology involve scale-invariant random sample representation, robust relative displacement measurement on 3D human body, well-designed random forest regressor to map

the random sample's descriptors to the corresponding displacements, and a voting strategy for locating the focal feature points. The presented approach is designed for engineering garment applications that require feature point identification on the surface of 3D human bodies. The experimental tests have verified the effectiveness and correctness of our approach. The automatic method can further shorten the time of garment product design cycle.

Nevertheless, our method still suffers from some limitations, such as computation efficiency. We aim to tackle this problem by taking advantages of the GPU techniques with parallel computing and making our method applicable to interactive or online applications with near-realtime performances

Acknowledgment

This work is supported by National Natural Science Foundation of 61602324, Scientific Research Project of Beijing Educational Committeen KM201710028018, the open funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (Grant No. BUAA-VR-17KF-12)

References

[1] H. Lim, C.L. Istook, N.L. Cassill, Advanced mass customization in apparel, J. Text. Apparel Technol.Manage. 6 (1) (2009).
[2] C.C. Wang, K.-C. Hui, K.-M. Tong, Volume parameterization for design automation of customized free-form products, IEEE Trans. Autom. Sci. Eng. 4 (1) (2007) 11–21.
[3] C.C. Wang, T.K. Chang, M.M. Yuen, From laser-scanned data to feature human model: a system based on fuzzy logic concept, Comput.-Aided Des. 35 (3) (2003) 241–253.
[4] S. Petrak, D. Rogale, Systematic representation and application of a 3d computer-aided garment construction method: part i: 3d garment basic cut construction on a virtual body model, Int. J. Clothing Sci.Technol. 18 (3) (2006) 179–187.
[5] P. Guan, L. Reiss, D.A. Hirshberg, A. Weiss, M.J. Black, Drape: dressing any person. ACM Trans. Graph. 31 (4) (2012) 1–35.
[6] T. Hou, H. Qin, Efficient computation of scale-space features for deformable shape correspondences, Computer Vision–ECCV 2010, (2010), pp. 384–397.
[7] L.M. Lui, S. Thiruvenkadam, Y. Wang, P.M. Thompson, T.F. Chan, Optimized conformal surface registration with shape-based landmark matching, SIAM J. Imaging Sci. 3 (1) (2010) 52–78.
[8] T. Hou, H. Qin, Robust dense registration of partial nonrigid shapes, IEEE Trans. Visual. Comput. Graphics 18 (8) (2012) 1268–1280.
[9] Y. Liu, X.-L. Wang, H.-Y. Wang, H. Zha, H. Qin, Learning robust similarity measures for 3d partial shape retrieval, Int. J. Comput. Vision 89 (2) (2010) 408–431.
[10] A. Zaharescu, E. Boyer, K. Varanasi, R. Horaud, Surface feature detection and description with applications to mesh matching, Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009, pp. 373–380.
[11] E. Zhang, K. Mischaikow, G. Turk, Feature-based surface parameterization and texture mapping, ACM Trans. Graphics (TOG) 24 (1) (2005) 1–27.
[12] H. Zhang, A. Sheffer, D. Cohen-Or, Q. Zhou, O. Van Kaick, A. Tagliasacchi, Deformation-driven shape correspondence, Computer Graphics Forum, 27 Wiley Online Library, 2008, pp. 1431–1439.
[13] I. Pratikakis, M. Spagnuolo, T. Theoharis, R. Veltkamp, A robust 3d interest points detector based on harris operator, Eurographics workshop on 3D object retrieval, vol 5, (2010).
[14] C. Maes, T. Fabry, J. Keustermans, D. Smeets, P. Suetens, D. Vandermeulen, Feature detection on 3d face surfaces for pose normalisation and recognition, Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on, IEEE, 2010, pp. 1–6.
[15] X. Chen, A. Saparov, B. Pang, T. Funkhouser, Schelling points on 3d surface meshes, ACM Trans. Graphics (TOG) 31 (4) (2012) 29.
[16] M. Kolomenkin, I. Shimshoni, A. Tal, On edge detection on surfaces, Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009, pp. 2767–2774.
[17] N.J. Mitra, L.J. Guibas, J. Giesen, M. Pauly, Probabilistic fingerprints for shapes, in: Symposium on Geometry Processing, EPFL-CONF-149326, 2006, pp. 121–130.
[18] I. Pratikakis, M. Spagnuolo, T. Theoharis, R. Veltkamp, et al., Visual vocabulary signature for 3d object retrieval and partial matching.
[19] M. Ben-Chen, C. Gotsman, Characterizing shape using conformal factors. 3DOR, (2008), pp. 1–8.
[20] N. Gelfand, N.J. Mitra, L.J. Guibas, H. Pottmann, Robust global registration, Symposium on Geometry Processing, vol 2, (2005), p. 5.
[21] A. Elad, R. Kimmel, Bending invariant representations for surfaces, Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol 1, IEEE, 2001, p. 1.
[22] A.M. Bronstein, M.M. Bronstein, R. Kimmel, Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching, Proc. Natl. Acad. Sci. 103 (5) (2006) 1168–1172.
[23] A.M. Bronstein, M.M. Bronstein, A.M. Bruckstein, R. Kimmel, Analysis of two-dimensional non-rigid shapes, Int. J. Comput. Vision 78 (1) (2008) 67–88.
[24] M. Reuter, S. Biasotti, D. Giorgi, G. Patanè, M. Spagnuolo, Discrete Laplace–Beltrami operators for shape analysis and segmentation, Comput. Graphics 33 (3) (2009) 381–390.
[25] N. Sochen, R. Kimmel, A.M. Bruckstein, Diffusions and confusions in signal and image processing, J. Math. Imaging Vision 14 (3) (2001) 195–209.
[26] M. Ovsjanikov, A.M. Bronstein, M.M. Bronstein, L.J. Guibas, Shape google: a computer vision approach to isometry invariant shape retrieval, Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, IEEE, 2009, pp. 320–327.
[27] J. Sun, M. Ovsjanikov, L. Guibas, A concise and provably informative multi-scale signature based on heat diffusion, Computer Graphics Forum, 28 Wiley Online Library, 2009, pp. 1383–1392.
[28] Y. Fang, M. Sun, K. Ramani, Temperature distribution descriptor for robust 3d shape retrieval, Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on, IEEE, 2011, pp. 9–16.
[29] M.M. Bronstein, I. Kokkinos, Scale-invariant heat kernel signatures for non-rigid shape recognition26 (2)(2010) 1704–1711.
[30] A.M. Bronstein, M.M. Bronstein, L.J. Guibas, M. Ovsjanikov, Shape google: geometric words and expressions for invariant shape retrieval, ACM Trans. Graphics (TOG) 30 (1) (2011) 1.
[31] C. Chu, C. Chen, L. Nolte, G. Zheng, Fully automatic cephalometric x-ray landmark detection using random forest regression and sparse shape composition, submitted to Automatic Cephalometric X-ray Landmark Detection Challenge (2014).
[32] J. Gall, V. Lempitsky, Class-specific hough forests for object detection, Decision forests for computer vision and medical image analysis, Springer, 2013, pp. 143–157.
[33] C. Lindner, S. Thiagarajah, J.M. Wilkinson, G.A. Wallis, T.F. Cootes, arcOGEN Consortium, et al., Accurate fully automatic femur segmentation in pelvic radiographs using regression voting, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2012, pp. 353–360.
[34] A. Criminisi, J. Shotton, D.P. Robertson, E. Konukoglu, Regression forests for efficient anatomy detection and localization in ct studies. MCV, 2010 (2010), pp. 106–117.
[35] H. Huang, P. Mok, Y. Kwok, J. Au, Block pattern generation: from parameterizing human bodies to fit feature-aligned and flattenable 3d garments, Comput. Ind. 63 (7) (2012) 680–691.
[36] I. ISO, 8559: Garment Construction and Anthropometric Surveys-Body Dimensions, ISO: Switzerland, 1989.
[37] P.W. Jones, M. Maggioni, R. Schul, Manifold parametrizations by eigenfunctions of the laplacian and heat kernels, Proc. Natl. Acad. Sci. 105 (6) (2008) 1803–1808.
[38] K. Crane, C. Weischedel, M. Wardetzky, Geodesics in heat: a new approach to computing distance based on heat flow, ACM Trans. Graphics (TOG) 32 (5) (2013) 152.
[39] A.E. Johnson, Spin-Images: A Representation for 3-d Surface Matching, (1997).
[40] F. Tombari, S. Salti, L.D. Stefano, Unique shape context for 3d data description, ACM Workshop on 3d Object Retrieval, (2010), pp. 57–62.
[41] A. Criminisi, J. Shotton, Decision Forests for Computer Vision and Medical Image Analysis, Springer Science & Business Media, 2013.