Research Article

Automatic location and semantic labeling of landmarks on 3D human body models

Shan Luo[1], **Qitong Zhang**[1], and **Jieqing Feng**[1] (✉)

Abstract Landmarks on human body models are of great significance for applications such as digital anthropometry and clothing design. The diversity of pose and shape of human body models and the semantic gap make landmarking a challenging problem. In this paper, a learning-based method is proposed to locate landmarks on human body models by analyzing the relationship between geometric descriptors and semantic labels of landmarks. A shape alignment algorithm is proposed to align human body models to break symmetric ambiguity. A symmetry-aware descriptor is proposed based on the structure of the human body models, which is robust to both pose and shape variations in human body models. An AdaBoost regression algorithm is adopted to establish the correspondence between several descriptors and semantic labels of the landmarks. Quantitative and qualitative analyses and comparisons show that the proposed method can obtain more accurate landmarks and distinguish symmetrical landmarks semantically. Additionally, a dataset of landmarked human body models is also provided, containing 271 human body models collected from current human body datasets; each model has 17 landmarks labeled manually.

Keywords 3D human body model; descriptor; landmarking; shape alignment

1 Introduction

A landmark on a model, also referred to as a labeled keypoint, has a specific position and a semantic label [1]. Landmarking is a task to predict both the positions of landmarks and their semantic labels.
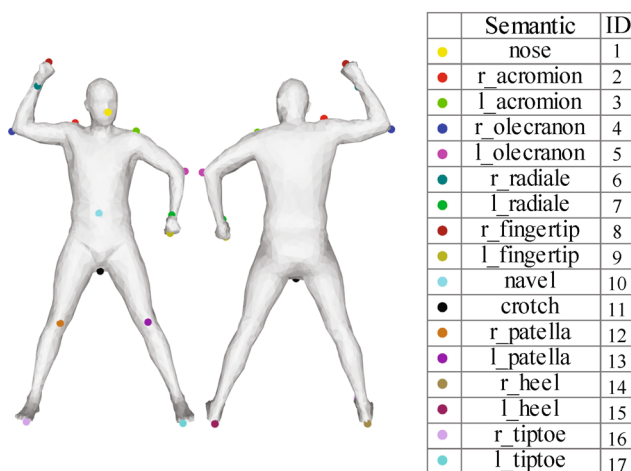
1 State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, China. E-mail: S. Luo, luoshan@zju.edu.cn; Q. Zhang, zhangqitong@zju.edu.cn; J. Feng, jqfeng@cad.zju.edu.cn (✉).

Landmarks on human body models not only play an essential role in many graphics applications, such as shape matching [2], object recognition [3], and surface remeshing [4], but also play a role in human body measurement, clothing design, and healthcare-related applications [5, 6]. However, to label landmarks, one must bridge the semantic gap and have a semantic-level understanding of human body models to identify the positions of the landmarks and attach semantics. Additionally, variations in pose and shape of human body models and their intrinsic symmetry further increase the difficulty of semantic understanding, making landmarking of human body models a challenging problem.

To bridge the semantic gap, the semantics of landmarks can be either transferred from a labeled template model or learned from training data. The former mainly targets shape matching and is limited by the matching accuracy. On the other hand, landmarks can improve the accuracy of shape matching. For learning-based methods, however, few datasets have been released to investigate the landmarking problem on human body models. You et al. [7] collected a keypoint dataset for diverse models in 16 categories, which do not solely cover human body models. Although CAESAR [8] and SHREC'14 Track [9] provide human body model datasets with some landmarks, the human body models in these datasets have similar poses. Therefore, current datasets for human body model landmarking cover shape variations across individuals without considering the diversity of poses, which limits their applications. To solve this problem, we collected 271 human body models with varied shapes and poses and labeled the landmarks manually to form a new keypoint dataset. Considering both anatomical meaning and discrimination of shape features, 17

vertices were selected from a human body model for labeling landmarks. The landmarks and their semantic labels are shown in Fig. 1. Symmetric landmarks, such as the left and right patella are semantically distinguished.

The most common way to label landmarks is to establish correspondences between features and semantic labels. These features can be either handcrafted descriptors or features extracted by deep learning algorithms. Few handcrafted descriptors are robust to simultaneous variations in pose, shape variations between individuals, and the intrinsic symmetry of human body models. Learning-based features [7, 10] may be extracted by treating landmarking as a special segmentation problem, in which each landmark is regarded as an individual class. Since the training data for each class in landmarking are far fewer than those for shape segmentation, the accuracy of these methods is limited. In addition, models for learning-based methods need to be aligned first to extract consistent features and distinguish symmetrical landmarks. However, aligning models manually is tedious and time-consuming, and shape alignment seems to be the most difficult point in the normalization process [11].

As the human body model structure is invariant to changes in shape and pose, it may be possible to locate landmarks via structural information. With the development of shape segmentation methods [12–14], we can easily access structural information through the segments obtained by shape segmentation. In this paper, a new multi-scale descriptor is

defined based on the segments, making it robust to variations in shape and pose of human body models. Moreover, a shape alignment algorithm is adopted to break the symmetric ambiguity of human body models. In this way, the symmetric segments are separated so that the proposed segment-based descriptor is distinguishable for symmetric landmarks. Finally, combining the proposed descriptor and other handcrafted descriptors, an AdaBoost regression model is trained to map these descriptors to landmark labels. To further improve landmarking accuracy, the label is defined as a truncated distance field of the landmark. In summary, the main contributions in this paper are as follows:

- an automatic method to locate symmetry-aware landmarks for 3D human body models with varied shapes and poses, with a keypoint dataset with landmarks labeled manually on 271 human body models,
- a multi-scale descriptor based on segments, which is robust to the variations in shape and pose of human body models and can discriminate symmetric parts, and
- a shape alignment algorithm based on the kinematic skeleton extracted from the segments using existing algorithms, giving human body models a consistent orientation.

2 Related work

Since the descriptors are closely related to the landmarking methods, both descriptors and landmarking methods for human body models will be reviewed in this section.

2.1 Descriptors

A descriptor encodes local or global geometric features of a 3D model. Descriptors can be roughly classified into two types: handcrafted descriptors and learning-based descriptors [2]. Because learning-based descriptors are trained for specific tasks, such descriptors relevant to this paper will be introduced in the landmarking method review.

In earlier years, basic geometric attributes were used as model descriptors, such as 3D coordinates of vertices, which are sensitive to rigid transformations of models. To meet the requirements of applications, more complicated descriptors have been proposed, e.g., Gaussian curvature, spin images [15], 3D

	Semantic	ID
	nose	1
	r_acromion	2
	l_acromion	3
	r_olecranon	4
	l_olecranon	5
	r_radiale	6
	l_radiale	7
	r_fingertip	8
	l_fingertip	9
	navel	10
	crotch	11
	r_patella	12
	l_patella	13
	r_heel	14
	l_heel	15
	r_tiptoe	16
	l_tiptoe	17

Fig. 1 Landmarks defined on human body models and their semantic labels.

shape contexts [16], and shape diameter function (SDF) [17]. These descriptors are robust under rigid transformation but not for non-rigid deformations. To better deal with articulated models that can undergo non-rigid deformations, intrinsic descriptors that are isometric-invariant have been proposed based on geodesic distances or spectral geometry. These descriptors include the global point signature [18], heat kernel signature (HKS) [19], scale-invariant heat kernel signature (SIHKS) [20], wave kernel signature (WKS) [21], etc. Although these descriptors are robust to models of the same person in varied poses, they are not robust to human body models with shape variations between individuals, which are inevitable in practical applications. Moreover, these descriptors cannot distinguish the symmetric parts of human body models since the local or global geometries of the symmetric parts are almost the same. Based on the observation that some orientations, such as the gradient of the average diffusion distance field, are sensitive to local symmetry, some oriented descriptors [22–24] have been proposed to distinguish the symmetric parts of models. However, the orientations adopted by these descriptors are also unstable for non-isometric models, so are not robust for human body models with varied shapes.

In this paper, a new multi-scale descriptor is proposed based on structural information, making it robust to variations in shape and pose of human body models. In addition, the proposed descriptor can discriminate symmetric parts when the adopted structural information is symmetrically distinguishable.

2.2 Landmarking on human body models

Landmarking is a procedure to both predict the positions of landmarks and identify their semantic labels. The semantic labels of corresponding landmarks should be consistent across different models in the same category.

Descriptors play an essential role in early landmarking methods for human body models. Some methods [9, 25][①] locate landmarks directly via similarity of descriptors. Several methods [26, 27] connect landmarks to form a graph, and a Markov network is used to predict the landmarks. In other

methods [9, 28], a classifier is adopted to establish the relationship between several descriptors and landmarks. In this way, effective descriptors can be selected by the classifier automatically. Shu et al. [29] adopted a stacked auto-encoder method to predict points of interest using multiple feature descriptors. Limited by the descriptors, all these methods cannot distinguish symmetric parts, making them applicable only to human body models with known orientations. Wuhrer et al. [26] generalized the landmarking method to human body models in varied poses by mapping the models to canonical forms. However, the best orientation among the canonical forms is chosen manually to avoid symmetric ambiguity.

With the development of deep learning algorithms, some methods have been proposed to locate landmarks via features extracted through deep learning. Wang et al. [2] introduced a triplet convolution neural network (CNN) to learn descriptors that can discriminate landmarks. However, these descriptors aim to improve the accuracy of matching rather than landmarking. Xi et al. [30] located landmarks by rendering 3D human body models into 2D images for use in deep CNNs. Although this method addresses the data imperfections of 3D models and validates its effectiveness on human body models in a standard pose, landmarking results on human body models in varied poses cannot be guaranteed because of the diversity of the rendered images. Other methods [7, 10, 31] treat landmarking as a special segmentation problem, where each landmark is regarded as an individual class. Since the number of training data in each class in landmarking are far fewer than those in shape segmentation, it is more difficult to extract effective features than to perform shape segmentation. As shown in the results [7], these methods face significant difficulties in locating consistent landmarks.

Another way to locate landmarks is to transfer landmarks from a labeled template model to a target model. Several shape matching methods can be used to locate the landmarks by establishing correspondences between the template model and the target model, such as iterative closest point and thin plate spline robust point matching adopted in Ref. [9], and coherent point drift adopted by Zhou and Hao [32]. In these methods, the accuracy of the landmarks directly depends on the shape

---

① Ref. [9] reports on the SHREC'2014 track on automatic location of landmarks, and introduces six landmarking methods.

matching accuracy. To bridge the non-isometric gap, a template model very similar to the target model is selected from the training data. However, limited by the accuracy of shape matching, the accuracy of landmarks located by these methods is still restricted. In recent years, significant progress has been made in shape matching [23, 33, 34]. The accuracy of landmarks based on shape matching may benefit from these new shape matching methods.

In this paper, a learning-based landmarking method is proposed for human body models with arbitrary shapes and poses. First, several descriptors are defined and extracted. An AdaBoost regression algorithm is then adopted to locate the landmarks by selecting effective features from the multiple descriptors.
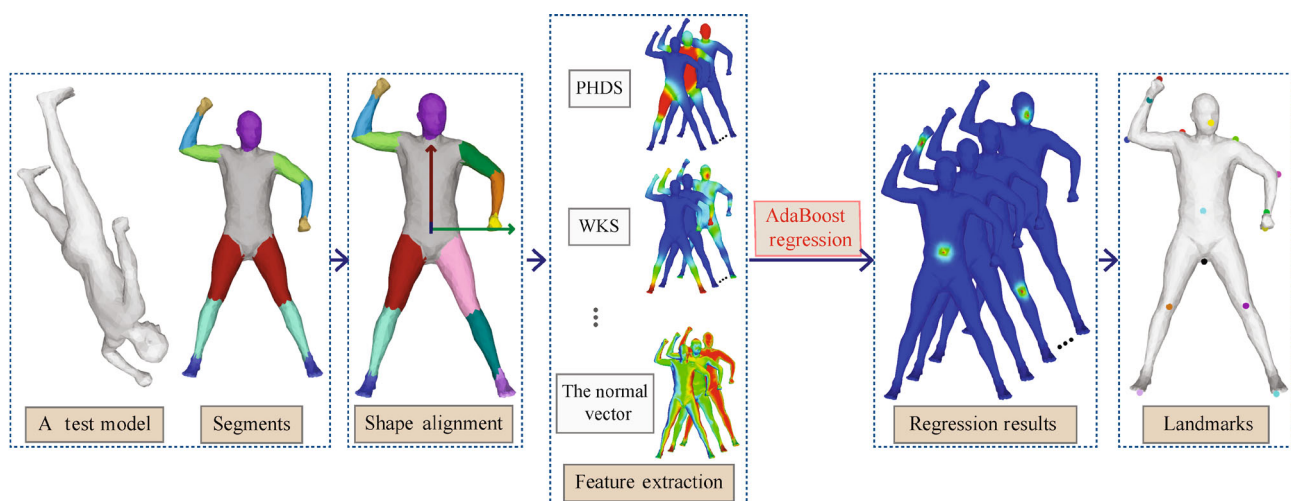
3   Method

3.1   Outline

In this paper, a novel landmarking method for human body models having varied shapes and poses is proposed based on structural information. The input to our method is a triangular mesh of a human body model and the corresponding segments obtained by learning-based shape segmentation methods. Although the segments are semantically labeled, symmetric segments have the same label, as shown in the segments in Fig. 2. Regarding these segments, the symmetric ambiguity is first broken; then, a better way is proposed to enrich structural information by transforming the segments into a

multi-scale descriptor and extending the descriptor to shape analysis.

The proposed landmarking method consists of three steps: shape alignment, feature extraction, and AdaBoost regression. First, a shape alignment algorithm is used to transform the human body models into a consistent orientation. In this way, symmetric segments with the same labels can be distinguished, and some descriptors, such as the normal vector, can be computed consistently. Then, a new multi-scale descriptor is used based on heat diffusion for the symmetry-aware segments, called the *part heat diffusion signature* (PHDS), which is robust to variations in human body models and discriminating with regard to symmetric parts. Additionally, some handcrafted descriptors, such as the HKS and WKS, are also extracted. Combining all these descriptors, an AdaBoost regression algorithm is adopted to establish the correspondence between these descriptors and the probability distributions of the landmarks. The vertices with maximal probability values in the regression results are marked as landmarks. The pipeline of the proposed method is shown in Fig. 2.

3.2   Shape alignment

Due to the various poses and intrinsic symmetry of human body models, it is challenging to define a consistent orientation to align human body models. In recent years, some methods [33, 35] have been proposed to distinguish symmetric parts of human body models with the help of kinematic skeletons.
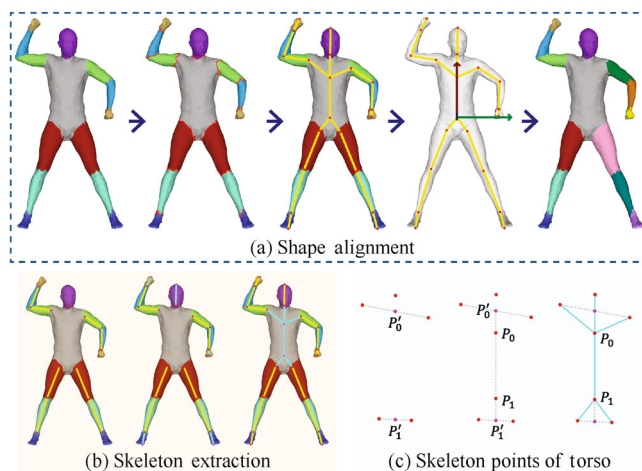

Fig. 2   Pipeline of the proposed method. The color scale in feature extraction and regression results varies from blue to red to indicate values from 0 to 1.

Based on the observation that the feet of human body models always face forward in the standard pose, these methods transmit the orientation information along the kinematic skeleton to the torso and then correctly distinguish the symmetric parts of human body models. However, the kinematic skeleton extraction steps of these methods are complex and time-consuming, whether template model fitting [33] or template skeleton embedding [35]. With the development of learning-based shape segmentation methods, we have easy access to the segments of human body models. This enables us to propose a simpler algorithm to extract kinematic skeletons directly from the segments, and then the orientation transmitting algorithm provided by Luo and Feng [35] is adopted to break the symmetric ambiguity via the extracted kinematic skeleton.

The steps of the proposed shape alignment algorithm are shown in Fig. 3(a). First, the boundary curves between the segments are extracted, and their center points are adopted as the skeleton points. Then, the skeleton points are connected using skeleton bones based on the structures of the segments and are further processed to form a kinematic skeleton. Once the kinematic skeleton has been extracted, the algorithm proposed by Luo and Feng [35] is adopted to establish a local coordinate system on the human body model. Thus, human body models may be consistently oriented by aligning the local coordinate systems, and the symmetric segments are distinguished as well.

Since the skeleton points extracted from the boundary curves may be insufficient to form a complete kinematic skeleton, further steps are taken to form the kinematic skeleton, as shown in Fig. 3(b). For each leaf segment that has only one neighbor, the farthest vertex on the segment to the corresponding skeleton point is defined as a new skeleton point, and a skeleton bone is obtained by connecting these two skeleton points, as shown by the blue lines of the middle model in Fig. 3(b). For the torso segment, two additional skeleton points are extracted to make the kinematic skeletons closer to the medial axis of the human body models. As shown in Fig. 3(c), two auxiliary points P_0' and P_1' are defined as the mid-points of the shoulder skeleton points and the hip skeleton points, respectively. Then, an offset along the line $P_0'P_1'$ is used to gain two skeleton points P_0 and P_1 which satisfy the condition $\|P_0'P_0\| = \|P_1'P_1\| = \|P_0'P_1'\|/5$. In this way, a kinematic skeleton is extracted directly from the segments. Although skeleton points P_0 and P_1 are defined without considering geometric features, they do not affect the establishment of the coordinate system in which only the direction of skeleton bone P_0P_1 is helpful.

To establish the local coordinate system, the algorithm provided by Luo and Feng [35] rotates the leg bones and foot bones of the kinematic skeleton to the standard pose first. An example of this step is shown in Fig. 4. By assuming that the feet face forward in the standard pose and the chest skeleton point (P_0) is above the waist skeleton point (P_1), a local coordinate system may be established to distinguish the symmetry of the human body model, as seen in the fourth model in Fig. 3(a).

To ensure the segments agree with human body structure, the segments are subjected to a preprocessing step to update their semantic labels.

(a) Shape alignment

(b) Skeleton extraction   (c) Skeleton points of torso

Fig. 3 (a) Steps of the proposed shape alignment algorithm. (b) Kinematic skeleton extraction and construction directly from the segments. (c) Skeleton points defined on the torso.


Fig. 4 Steps in rotating leg and foot bones to the standard pose. Reproduced with permission from Ref. [35], © Springer Science+Business Media, LLC, part of Springer Nature 2020.

When the number of segments with the same label exceeds the target number, such as three foot-segments for a human body model, the labels of the segments with fewer vertices are modified according to geodesic distances to their adjacent segments. In addition, the symmetric thigh segments sometimes are very close so that the two thigh boundary curves near the torso geometrically merge into one curve, as shown in Fig. 5(b). In this case, first, the merged boundary curve should be subdivided into two boundary curves, and then two skeleton points may be extracted. To this end, the center point of the merged boundary curve is calculated first (red point C). Then, the point on the boundary curve closest to the center point is identified and marked as a temporary separation point (purple point A). Another separation point B is selected as the point closest to point A on the other side of the boundary curve. Point A is further refined to be the point closest to point B on the other side of the boundary curve. Having defined the separation points, the boundary curve can be subdivided into two boundary curves, and the center points of the two parts are regarded as the skeleton points, shown as yellow points in Fig. 5.

Figure 6 shows some results of the proposed shape alignment algorithm. Human body models in various orientations are aligned around the hip region, resulting in a consistent orientation.

3.3 Feature extraction

Features play an essential role in locating landmarks. Handcrafted descriptors, such as HKS and WKS, can distinguish vertices to some extent. However, due to limited discrimination of vertices, accurately locating landmarks and distinguishing symmetrical vertices is challenging. Semantic segmentation results of 3D models can narrow the search for landmarks according to semantic information; for example, the nose landmark must appear in the head segment. Nevertheless, the inclusion relation between vertices
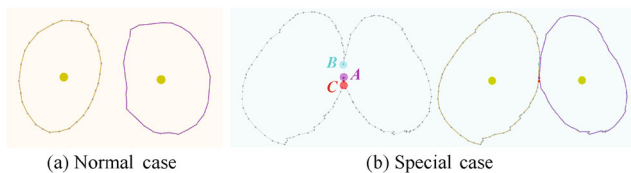
Fig. 5 Skeleton point extraction from the boundary curves between thigh segments and the torso segment.
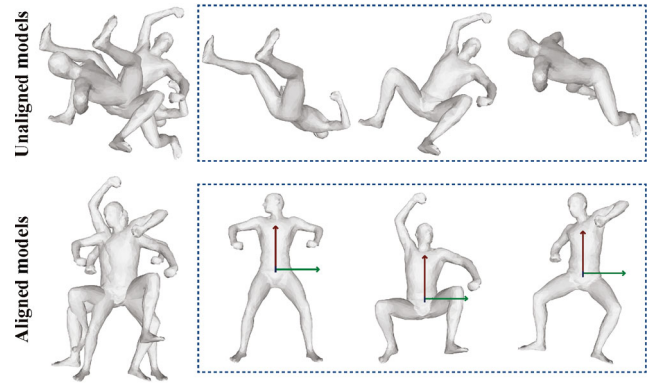
Fig. 6 Results of the shape alignment algorithm based on the segment. Above: models with random initial orientations. Below: aligned models.

and segments is insufficient to locate landmarks accurately.

In this paper, a new multi-scale descriptor called PHDS is proposed to describe the relationship between vertices and segments, which solves a heat equilibrium equation on the surface with different initial heats attached to the segments. In addition to the inclusion relation, the proposed descriptor is also related to the distance between the vertices and the segments, which enriches structural information compared to the segments.

For each segment i, the heat equilibrium equation over the surface can be represented as

$$\partial w^i / \partial t = \Delta w^i + H^i (p^i - w^i) = 0 \qquad (1)$$

where Δ is the discrete cotangent Laplacian matrix, p^i is an indicator vector for the initial heat values of vertices with $p_j^i = 1$ if vertex j belongs to segment i and $p_j^i = 0$ otherwise, and H^i is a diagonal matrix with H_{jj}^i being the heat contribution weight between vertex j and its corresponding segment. In this paper, H_{jj}^i is defined as $H_{jj}^i = (cA_j)/Ad_j^i$, where A is the total area of the model, A_j is the area represented by vertex j, which is one third of the total area of its adjacent triangles, c is a constant value to balance the parameters of Δ and H^i, and d_j^i can be considered as the heat retention ratio of vertex j for the initial heat. The larger d_j^i is, the closer the heat value of vertex j at equilibrium is to its initial heat value. In this paper, d_j^i is defined as follows:

$$d_j^i = \begin{cases} a, & \text{if } j \in \text{segment } i \\ b, & \text{otherwise} \end{cases} \qquad (2)$$

where a and b are constant values. To find the equilibrium parameter w^i, Eq. (1) can also be written:

$$(-\Delta + H^i)w^i = H^i p^i \qquad (3)$$

which can be solved directly by a least squares method.

For each segment i, an equilibrium parameter w^i is calculated by solving Eq. (3). Since a human body model is divided into 14 segments, a 14-dimensional feature vector for each vertex is obtained by taking each segment as the initial heat source in turn: the proposed PHDS is a multi-dimensional descriptor, as shown in each row in Fig. 7. In addition to the initial heat source, the parameters (a, b) also affect the equilibrium parameter w^i. Figure 7 shows the features obtained for three pairs of parameters (a, b). The greater the difference between a and b, the larger the range of heat diffusion on the human body model, so information about segments is transmitted to further vertices. Thus, PHDS is also a multi-scale descriptor with variable parameters (a, b). Moreover, since the features obtained by symmetric segments are arranged in different dimensions, the symmetrical vertices can be distinguished easily according to the features of different dimensions. Thus, the proposed PHDS is discriminative for symmetric parts.

In a broad sense, the proposed PHDS is somewhat similar to the skinning attachment [36]: both are composed of results by taking multiple elements as initial values. The skinning attachment takes each bone as an element, while PHDS takes each segment as an element. However, PHDS does not need complex operations such as skeleton extraction and vertex assignment, making it more effective and feasible. Furthermore, skinning attachment is usually used in shape animation. To the best of our knowledge, our method is the first to extend this type

of descriptor to shape analysis.

Different heat equilibrium parameters w^i can be obtained using different parameters (a, b). There are no best parameters for (a, b) but they should be selected according to task. In this paper, three sets of parameters (a, b) are adopted to form the descriptor PHDS. In this way, heat can be transmitted over different segments so that the source segment can be perceived by vertices at different distances. In addition to PHDS, some other widely used descriptors, such as the HKS, SIHKS, WKS, SDF, curvature, and normal vectors at the vertices, are also adopted to enrich the features. For each vertex on the human body model, a 150-dimensional feature vector is formed using the descriptors above. Details of these descriptors are given in Table 1.

3.4 AdaBoost regression

The essence of landmarking is to automatically establish correspondence between the features and the labels of the landmarks. Since no descriptor can

Table 1 Descriptors used in this paper. Dim indicates number of dimensions. $t_{\min}$ and $t_{\max}$ are determined by analyzing the eigenvalues of the human body models. "−" indicates parameters follow the original paper

Descriptor	Dim	Details
PHDS	42	$(a, b) \in \{(0.1, 0.01),$ $(0.1, 0.005), (0.1, 0.001)\}$
HKS	30	$(t_{\min}, t_{\max}) = (0.003, 0.3)$
SIHKS	30	$(t_{\min}, t_{\max}) = (0.003, 0.3)$
WKS	30	—
Curvature	10	Principal curvatures k_1, k_2, and their 8 combinations introduced in Ref. [1]
Normal	3	Normal vector at each vertex
SDF	5	Cone angle $\in \{40, 50, 60, 70, 80\}$

Fig. 7 Examples of the proposed PHDS, for heat diffusion parameters $(a, b) = (0.1, 0.01)$, $(0.1, 0.005)$, and $(0.1, 0.001)$.

solve the landmarking problem alone, the AdaBoost regression algorithm is adopted to select the effective descriptors from many to locate the landmarks.

First, some training data, including features and labels of the vertices, are needed to train an AdaBoost regression model. If the labels of the vertices are directly defined as 1 (landmark) and 0 (non-landmark), as in previous methods, regardless of whether segmentation or regression is used, it is challenging to learn effective features to locate the landmarks. In this case, the labels for landmarks and neighboring vertices may change suddenly, while the features vary slightly. Therefore, we define the labels in a smooth probability field by constructing a distance field for each landmark. In this way, the labels are more consistent with the features, which is conducive to learning. However, since the values of the probability field range from 0 to 1, the difference of probability values between adjacent vertices becomes smaller with expansion of the probability field distribution, in turn increasing the difficulty of landmark recognition. Thus, a truncation parameter τ is adopted to constrain the range of the probability field, as shown in Fig. 8.

To flexibly constrain the range of local regions around the landmarks, the distance field is defined as a Gaussian kernel function with respect to geodesic distance to the landmark. The truncation parameter τ is set as a geodesic distance to control local regions. The labels of the landmark i are defined as

$$g_j = \begin{cases} \exp(-k d_{ij}/\tau), & \text{if } d_{ij} \leqslant \tau \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where d_{ij} is the geodesic distance between vertex j and landmark i, and k is a constant set to 3.

Once the features and the labels are ready, the AdaBoost regression algorithm analyzes the relationship between the features and the labels to obtain a trained model, in which effective descriptors

(a) Distance fields   (b) Distance fields with a truncation parameter
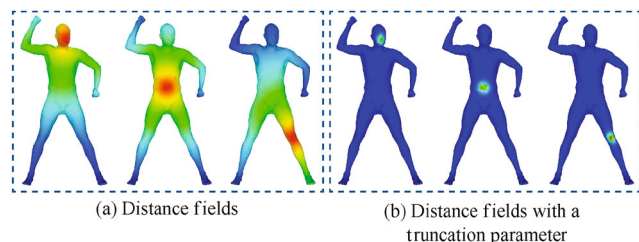
Fig. 8 (a) Examples of labels defined by distance fields. (b) Examples of labels defined by a distance field with a truncation parameter of 0.05.

are selected to locate landmarks. For a test human body model, the predicted labels are obtained by the features and the trained model; vertices with maximal predicted label values are selected as landmarks.

4 Experiments

In this section, the experimental setting for testing the proposed landmarking method is introduced first. Then, the proposed method is experimentally validated and compared with previous methods qualitatively and quantitatively. Limitations of the proposed method are presented at the end. The proposed method was implemented on a desktop PC with an Intel Core i9-10900X CPU and 128 GB RAM.

4.1 Experimental setting

The experimental setting is considered from three points of view: the new dataset, the parameters used, and the evaluation criteria for landmarks.

To validate the proposed landmarking method, some human body models are selected from three datasets: the SCAPE dataset [37], the MPI FAUST dataset [38], and the SPRING dataset [39]; they were labeled manually to provide a dataset for human body model landmarking. The SCAPE dataset contains 71 human body models with various poses of the same person. The MPI FAUST dataset provides 100 human body models of 10 people in 10 different poses. Since the pose of the human body models in the SPRING dataset are all in standard pose, only 100 human body models are selected from this dataset, including 50 male human body models and 50 female human body models. Notably, the vertices of the human body models in MPI FAUST are in one-to-one mapping, and the vertices of the human body models in both SCAPE and SPRING are also one-to-one mapped. Therefore, we downsample these human body models to 2048 vertices for two reasons: to reduce the influence of vertex correspondences on landmarking, and to keep consistency with the keypoint dataset proposed by You et al. [7], to expand it. Taking repeatability of poses into consideration, 120 human body models of various shapes and poses were used as training data, 51 human body models with partially similar poses were used for validation, and the remainder were taken as test data. Details of this data division are shown in Table 2.

Many learning-based shape segmentation methods

Table 2   Human body model dataset division

	SCAPE	MPI FAUST	SPRING	Total
Training set	60	40	20	120
Validation set	11	20	20	51
Test set	0	40	60	100

[12–14] have been proposed to obtain labeled segments with the training data provided by Maron et al. [40]. For convenience, the segments used in this paper are the labeled segments from the training data [40]. Since the human body models are downsampled to 2048 vertices, the parameter c is also set to 2048 to balance the parameters of Δ and H^i. Note that $c = 2048$ is also effective for human body models at other resolutions because the heat contribution weight is determined by both the parameter c and the parameters (a, b). The maximum geodesic distance of all human body models is normalized to 1, and the truncation parameter τ is set to 0.05.

Two evaluation criteria are adopted to evaluate the accuracy of the located landmarks: (i) the geodesic error between the located landmark and the ground truth labeled manually, and (ii) the percentage of correct keypoints (PCK), which considers the fraction of landmarks with geodesic error less than a certain threshold ε_{pck}.

4.2   Experimental results and analysis

The proposed method is comprehensively evaluated from the following five considerations: (i) the necessity of the descriptors used in this paper, (ii) the effectiveness of the setting labels in AdaBoost regression, (iii) the strength of the proposed PHDS, (iv) the landmarking results, and (v) the robustness to changes in resolution of the human body models.

First, we validate the utility of the adopted descriptors in human body model landmarking. An ablation study is conducted by disabling the descriptors in turn, thus evaluating whether each descriptor is helpful for landmarking. The geodesic errors of the landmarks are shown in Table 3. On balance, the experiment that takes all descriptors has the best average performance for all landmarks, as shown by the average values in Table 3. Therefore, all descriptors are helpful for locating landmarks on human body models.

The effectiveness of setting labels as a distance field with a truncation parameter is verified next. If the truncation parameter is too small, the labels have fewer non-zero values, which makes it difficult for AdaBoost to extract effective features. However, the possible regions for landmarks increase as

Table 3   Geodesic errors of landmarks with various descriptors for human body models in the test set. "w/o PHDS" indicates an experiment conducted without PHDS, etc. Best results are shown in red

Landmark	Descriptor							
	w/o PHDS	w/o HKS	w/o SIHKS	w/o WKS	w/o Cur	w/o Norm	w/o SDF	All
1	0.0011	0.0020	0.0017	0.0017	0.0007	0.0030	0.0040	0.0009
2	0.0084	0.0082	0.0086	0.0082	0.0092	0.0100	0.0094	0.0097
3	0.0098	0.0099	0.0105	0.0092	0.0097	0.0106	0.0103	0.0095
4	0.0729	0.0099	0.0082	0.0103	0.0096	0.0181	0.0096	0.0082
5	0.1063	0.0096	0.0093	0.0094	0.0086	0.0234	0.0086	0.0094
6	0.0502	0.0050	0.0053	0.0050	0.0055	0.0074	0.0045	0.0048
7	0.0743	0.0079	0.0069	0.0075	0.0070	0.0113	0.0075	0.0070
8	0.1068	0.0004	0.0009	0.0017	0.0003	0.0008	0.0012	0.0016
9	0.1229	0.0011	0.0012	0.0031	0.0019	0.0004	0.0008	0.0010
10	0.0105	0.0091	0.0076	0.0110	0.0086	0.0118	0.0068	0.0075
11	0.0055	0.0054	0.0053	0.0044	0.0043	0.0046	0.0053	0.0044
12	0.1068	0.0113	0.0093	0.0099	0.0085	0.0283	0.0087	0.0089
13	0.1187	0.0102	0.0094	0.0110	0.0098	0.0216	0.0083	0.0081
14	0.1098	0.0039	0.0041	0.0039	0.0041	0.0058	0.0040	0.0042
15	0.0515	0.0034	0.0038	0.0036	0.0042	0.0058	0.0036	0.0038
16	0.1620	0.0010	0.0006	0.0010	0.0007	0.0006	0.0009	0.0005
17	0.0869	0.0008	0.0003	0.0007	0.0002	0.0001	0.0002	0.0004
Average	0.0709	0.0058	0.0055	0.0060	0.0055	0.0096	0.0055	0.0053

the truncation parameter becomes larger, making the landmarks challenging to locate. Therefore, the truncation parameter needs to balance the number of non-zero labels and the size of the local region. Table 4 shows the geodesic errors in landmarks for different truncation parameters, while the labels of the navel landmark for different truncation parameters are shown in Fig. 9. In addition, the commonly used labeling which sets the label to 1 for a landmark and 0 otherwise, is also tested, the experiment being denoted "0–1" in Table 4. Landmarking performs best on average with the truncation parameter $\tau = 0.05$. In addition, compared to the "0−1" experiment and the distance field without truncation parameter ($\tau = 1$), our method achieves better results, showing the effectiveness of setting labels using a truncated distance field.
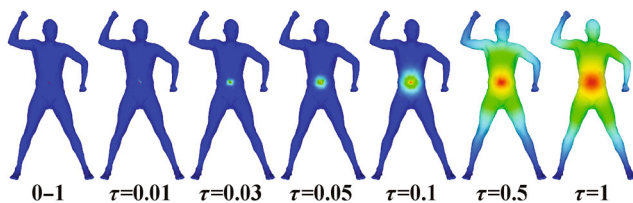
Next, the strength of the proposed PHDS in human body model landmarking is evaluated. The geodesic errors in experiments when disabling the descriptors in turn are shown in Table 3. The geodesic errors increase greatly without PHDS, while in other cases, the geodesic errors are slightly worse. Thus, PHDS has the most significant influence on the accuracy of landmarking. As effective descriptors are selected by AdaBoost during the training process, the importance of descriptors was additionally analyzed for effectiveness, and the results are shown in Fig. 10. The proposed PHDS is of great significance for most landmarks, which further validates the strength of PHDS. Moreover, since human body models are structurally consistent, PHDS, which is based on
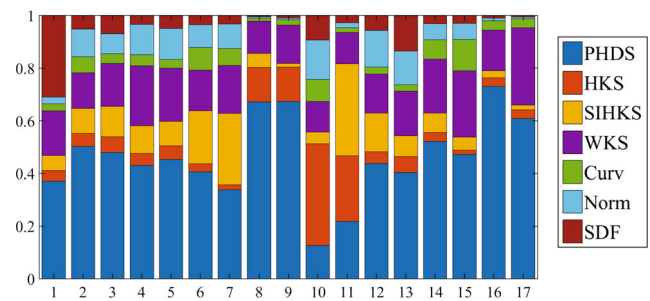

Fig. 9 Labels of the navel landmark for different truncation parameters.

Fig. 10 Importance of the descriptors for landmarking during the training process.

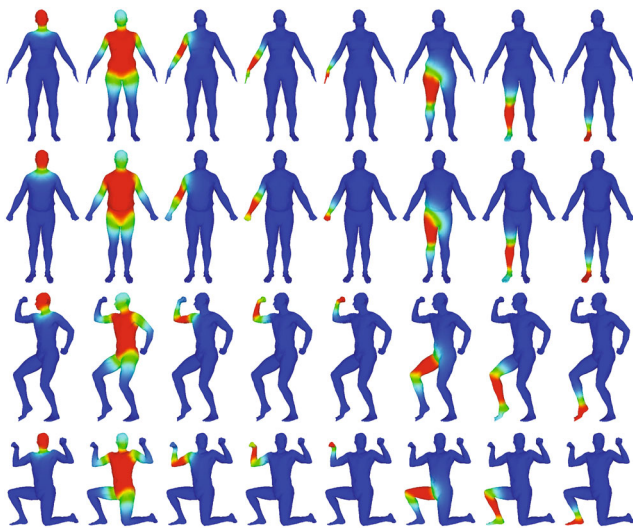**Table 4** Geodesic errors of landmarks with various truncation parameters for the human body models in the test set

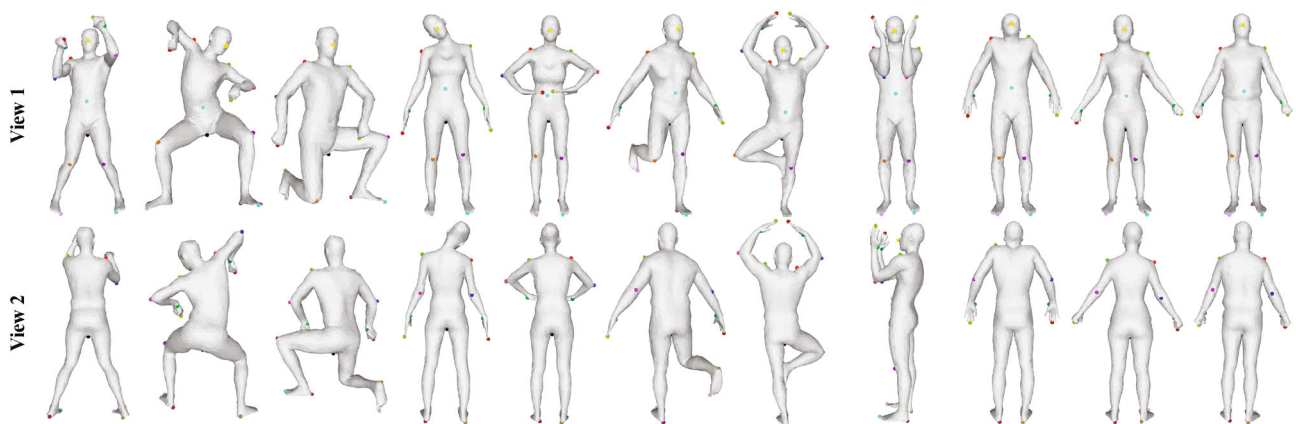| Landmark	Label						
	0–1	$\tau = 0.01$	$\tau = 0.03$	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.5$	$\tau = 1$
1	0.0825	0.0026	0.0010	0.0009	0.0017	0.0020	0.0027
2	0.0573	0.0123	0.0083	0.0097	0.0089	0.0088	0.0101
3	0.1736	0.0197	0.0094	0.0095	0.0090	0.0092	0.0092
4	0.1661	0.0160	0.0097	0.0082	0.0088	0.0089	0.0099
5	0.1616	0.0101	0.0093	0.0094	0.0088	0.0095	0.0104
6	0.0426	0.0060	0.0049	0.0048	0.0045	0.0081	0.0103
7	0.1552	0.0082	0.0068	0.0070	0.0074	0.0081	0.0093
8	0.0191	0.0002	0.0002	0.0016	0.0008	0.0011	0.0017
9	0.0911	0.0017	0.0011	0.0010	0.0011	0.0021	0.0025
10	0.1043	0.0646	0.0136	0.0075	0.0114	0.0072	0.0090
11	0.0571	0.0060	0.0045	0.0044	0.0053	0.0049	0.0059
12	0.2006	0.0352	0.0115	0.0089	0.0106	0.0100	0.0092
13	0.1823	0.0279	0.0104	0.0081	0.0089	0.0092	0.0091
14	0.1490	0.0047	0.0047	0.0042	0.0034	0.0038	0.0045
15	0.1875	0.0040	0.0102	0.0038	0.0037	0.0034	0.0035
16	0.0103	0.0013	0.0007	0.0005	0.0008	0.0006	0.0008
17	0.0086	0.0002	0.0001	0.0004	0.0002	0.0001	0.0002
Average	0.1088	0.0130	0.0063	0.0053	0.0056	0.0057	0.0064

structural information, is also robust to changes in shape and pose, as shown in Fig. 11.

Figure 12 shows the landmarking results of the proposed method on human body models with varied shapes and poses. The landmarks obtained by the proposed method are accurately located and consistent with our perception. Additionally, the symmetric landmarks are semantically distinguished.

Finally, the robustness of the proposed method to the resolution of the human body models is considered. In this paper, landmarks are located via the descriptors of the human body models. Whether the proposed method can be applied to human body models with different resolutions depends on the properties of the descriptors. Figure 13(a) shows some dimensions of the proposed PHDS on human

body models with different resolutions. It can be seen that the PHDS is almost visually the same for these human body models. The other descriptors used in this paper are also robust to changes in resolution of human body models. Figure 13(b) shows the landmarking results for these human body models. The results were obtained by directly applying the AdaBoost regression model trained on human body models with 2048 vertices to other resolution models. We can see that although the resolutions of the test human body models differ from those of the training models, all landmarks for these human body models are labeled in line with our expectations. Thus, the proposed landmarking method is robust to resolution changes in human body models.

The proposed landmarking method was also tested on human body models from the Princeton Segmentation Benchmark [42], where the models are noisier than the training data. The results are shown in Fig. 14, as are segments of these human body models obtained by MeshCNN [13]. Given that MeshCNN is a shape segmentation method for edges on models, we transform the edge segments to face segments, as shown in Fig. 14(b), and then apply them to the landmarking. MeshCNN is trained using models with 752 vertices, a limited quantity that might lead to overly simplified results and detail loss, increasing the difficulty of landmarking. Taking this into account, we also tested the models with the original resolution, in which the segments are obtained from the simplified models through a nearest neighbor algorithm. The segments of human body models at the original resolution are shown in
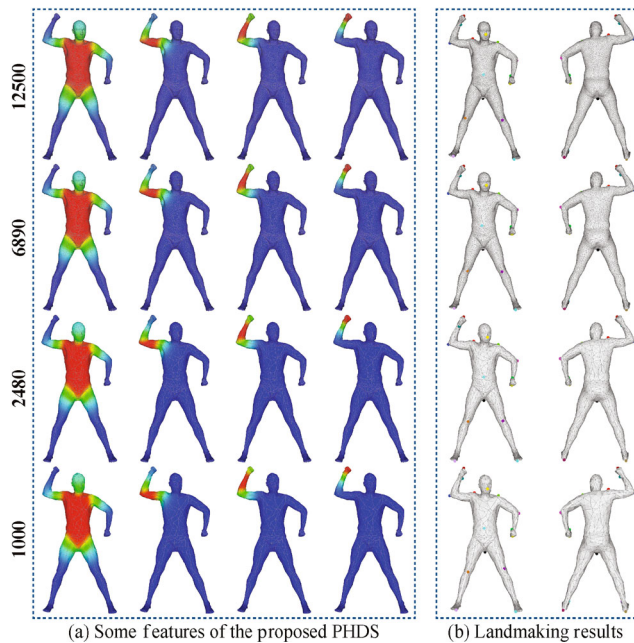

Fig. 11 Some dimensions of the proposed PHDS for human body models with various shapes and poses.


Fig. 12 Landmarking results for the proposed methods for human body models with various shapes and poses. Landmarks with consistent semantic labels are shown in the same color. Left 3 columns: human body models from the validation set. Other columns: models from the test set. Each model is shown from two views.

Fig. 13 (a) Application of PHDS to human body models with different resolutions. (b) Landmarking results for human body models. Top to bottom: models with 12500, 6890, 2048, and 1000 vertices.
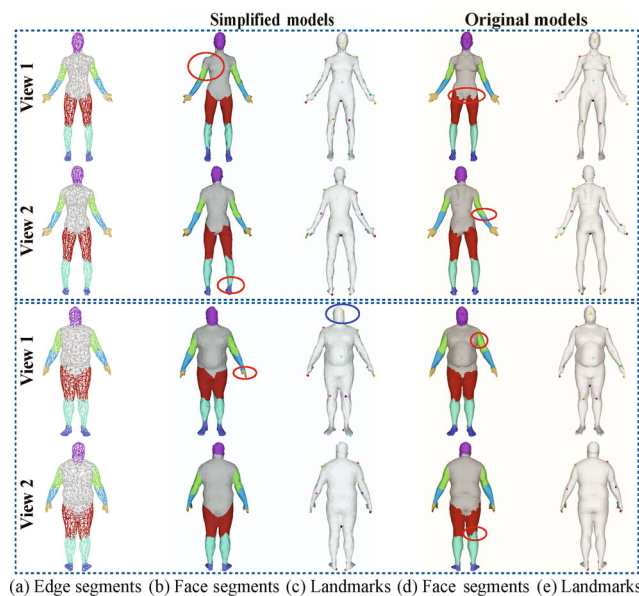
Fig. 14 (a) Edge segments of simplified human body models with 752 vertices from MeshCNN. (b) Face segments of the simplified models. (c) Landmarking results for the simplified human body models. (d) Face segments of the original human body models with 15,700 and 13,703 vertices. (e) Landmarking results for the original human body models.

Fig. 14(d). Although the quality of the segments is not comparable to that in Maron et al. [40], most landmarks are located accurately. Only one nose landmark is labeled at some offset due to the lack of detail. Since the proposed method is robust to the resolution of human body models, better results might

be obtained by combining the landmarking results for human body models with different resolutions, which can be considered in future research.

The speed of the proposed landmarking method is discussed next. Our method contains three steps: shape alignment, feature extraction, and AdaBoost regression. For a human body model with 2048 vertices, it takes approximately 0.48 s to align the human body model. The feature extraction step takes approximately 3.13 s on average to obtain the 150-dimensional feature vector. The AdaBoost regression step costs approximately 12 min for the training step, which contains 120 human body models, and 0.44 s to locate the landmarks on a test human body model. Therefore, for a test human body model, it takes approximately 0.48 s + 3.13 s + 0.44 s = 4.05 s to locate the landmarks by the proposed method.

4.3 Comparisons

In this section, we qualitatively and quantitatively compare the proposed landmarking method with previous landmarking methods, which shows the high accuracy of our method. Additionally, the effectiveness of the proposed shape alignment algorithm is verified.

As described in Section 2, there are two types of methods to locate landmarks: template-based methods, which transfer landmarks from a labeled template model via shape matching, and learning-based methods, which learn the locations and labels of landmarks from the training data. In this paper, we compared the proposed method with two template-based methods, FARM [33] and COCFM [23], and three learning-based methods, SpiderCNN [43], DGCNN [44], and PointConv [45]. FARM is an automatic registration method that deforms a template model to align well with the target, while COCFM is a shape matching method that establishes the correspondence between model vertices. For the three learning-based methods, the landmarking problem is treated as a multi-class classification problem, in which each landmark is associated with its semantic label and all the non-landmarks are assigned a background class label. For these learning-based methods, the models are usually aligned manually first to locate landmarks [7]. Therefore, we test these landmarking methods with both unaligned and aligned human body models. The human body models are given random orientations

to make up the unaligned training data and are processed via the proposed shape alignment algorithm to form the aligned training data.

Table 5 shows the geodesic errors in landmarks located by these methods on the unaligned and aligned human body models. All these landmarking methods achieve better results when aligned human body models are adopted: for template-based methods, although spectral features are adopted to establish correspondence, the skeleton deformation of FARM and the displacement vector used by COCFM are affected by the orientation of the models, which slightly decrease the accuracy of the landmarks when the human body models are unaligned. For learning-based methods, the inconsistent orientations of the unaligned human body models make it challenging to extract effective features, resulting in worse results by a large margin. Although most descriptors used in this paper are invariant to model orientation, the proposed method achieves better results due to the alignment of the normal vectors. Therefore, the proposed shape alignment algorithm could be used as a preprocessing step for other methods to improve their landmarking accuracy. Our method achieves the best performance in terms of landmark accuracy on average, whether unaligned or aligned human body

models are adopted.

Table 6 compares PCK results at a geodesic error threshold $\varepsilon_{\mathrm{pck}} = 0.01$ for aligned human body models. The proposed method outperforms other methods for most landmarks and has better performance on average. Figure 15 shows the PCK curves with geodesic error threshold $\varepsilon_{\mathrm{pck}}$ varying from 0 to 0.1. This shows the significant benefit of our method, with PCK results higher than those of other methods at any geodesic error threshold, which further confirms
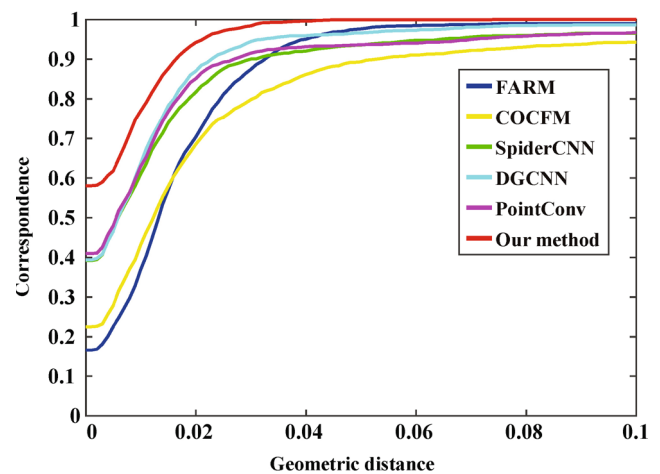
Fig. 15 PCK results for various geodesic error thresholds $\varepsilon_{\mathrm{pck}}$ (0–0.1) for the methods compared.

Table 5 Geodesic errors in landmarks for the human body models in the test set. For FARM, the template model is the SMPL model [41] adopted in the original paper. For COCFM, the template model is a downsampled human body model with standard pose in MPI FAUST

| Land-mark	Unaligned						Aligned					
	FARM	COCFM	SpiderCNN	DGCNN	PointConv	Ours	FARM	COCFM	SpiderCNN	DGCNN	PointConv	Ours
1	0.0062	0.0232	0.4633	0.1441	0.8365	0.0019	0.0028	0.0220	0.0011	0.0036	0.0118	0.0009
2	0.0220	0.0351	0.3965	0.1471	0.2358	0.0095	0.0214	0.0286	0.0263	0.0094	0.0121	0.0097
3	0.0238	0.0272	0.4062	0.1493	0.4128	0.0101	0.0228	0.0285	0.0173	0.0112	0.0088	0.0095
4	0.0155	0.0594	0.6463	0.2918	0.3688	0.0187	0.0160	0.0499	0.0268	0.0139	0.0203	0.0082
5	0.0152	0.0593	0.6386	0.3275	0.4722	0.0254	0.0164	0.0458	0.0346	0.0156	0.0181	0.0094
6	0.0217	0.0613	0.3817	0.3414	0.3545	0.0092	0.0233	0.0617	0.0201	0.0093	0.0247	0.0048
7	0.0289	0.0701	0.4001	0.3286	0.4630	0.0122	0.0294	0.0520	0.0236	0.0301	0.0314	0.0070
8	0.0093	0.0867	0.6082	0.4541	0.4243	0.0014	0.0100	0.0674	0.0256	0.0367	0.0473	0.0016
9	0.0196	0.0924	0.5420	0.3983	0.5376	0.0005	0.0197	0.0549	0.0403	0.0230	0.0501	0.0010
10	0.0224	0.0278	0.1406	0.1472	0.1074	0.0169	0.0225	0.0191	0.0072	0.0064	0.0223	0.0075
11	0.0239	0.0226	0.0918	0.0299	0.1813	0.0055	0.0230	0.0148	0.0161	0.0110	0.0062	0.0044
12	0.0171	0.0419	0.2610	0.1720	0.4934	0.0295	0.0164	0.0240	0.0371	0.0141	0.0173	0.0089
13	0.0156	0.0415	0.2722	0.2047	0.2796	0.0297	0.0154	0.0180	0.0175	0.0167	0.0303	0.0081
14	0.0182	0.0750	0.3817	0.4231	0.8674	0.0066	0.0194	0.0272	0.0302	0.0209	0.0464	0.0042
15	0.0112	0.0755	0.4188	0.4103	0.6258	0.0068	0.0109	0.0232	0.0354	0.0054	0.0104	0.0038
16	0.0290	0.0896	0.6498	0.5256	0.8711	0.0005	0.0304	0.0319	0.0188	0.0171	0.0643	0.0005
17	0.0203	0.0865	0.2755	0.4091	0.9086	0.0002	0.0140	0.0200	0.0024	0.0027	0.0028	0.0004
Average	0.0188	0.0574	0.4103	0.2885	0.4965	0.0109	0.0185	0.0347	0.0224	0.0145	0.0250	0.0053
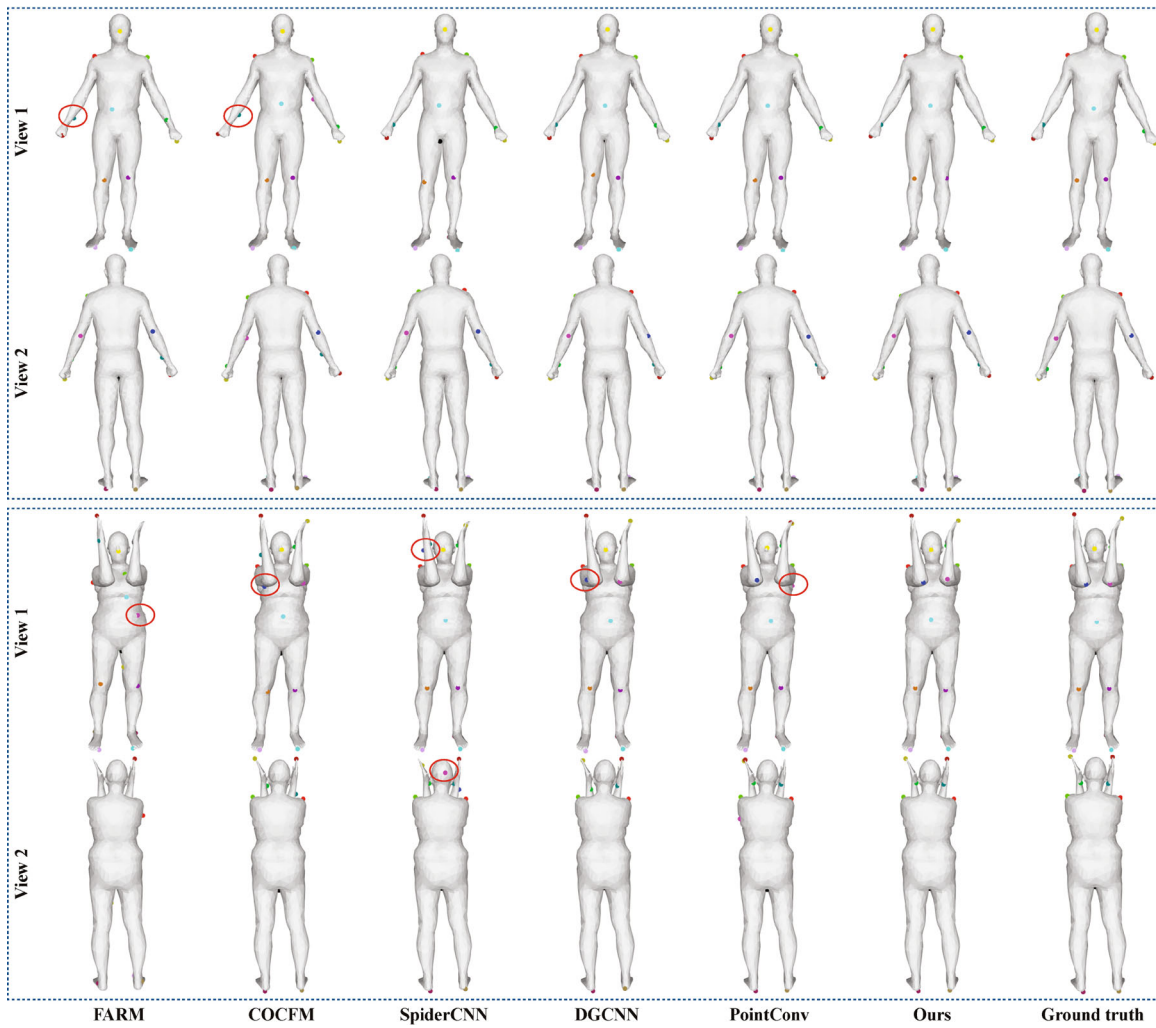
Fig. 16   Comparison of our method and other landmarking methods.

Table 6   PCK results for landmarks with geodesic error threshold $\varepsilon_{\mathrm{pck}} = 0.01$ for human body models in the test set

Landmark	FARM	COCFM	SpiderCNN	DGCNN	PointConv	Ours
1	0.97	0.64	1.00	0.98	0.94	0.98
2	0.08	0.27	0.27	0.54	0.51	0.53
3	0.11	0.20	0.36	0.41	0.54	0.52
4	0.29	0.19	0.25	0.33	0.34	0.59
5	0.40	0.20	0.30	0.39	0.51	0.59
6	0.24	0.04	0.66	0.64	0.51	0.88
7	0.26	0.14	0.49	0.46	0.39	0.69
8	0.59	0.53	0.87	0.83	0.72	0.97
9	0.50	0.59	0.69	0.58	0.38	0.93
10	0.18	0.39	0.68	0.75	0.70	0.68
11	0.19	0.52	0.43	0.56	0.71	0.79
12	0.34	0.34	0.27	0.38	0.43	0.55
13	0.37	0.37	0.44	0.40	0.37	0.58
14	0.32	0.69	0.85	0.82	0.87	0.91
15	0.65	0.77	0.91	0.90	0.97	0.92
16	0.31	0.66	0.95	0.91	0.89	0.98
17	0.48	0.80	0.95	0.94	0.92	0.99
Average	0.3694	0.4318	0.6100	0.6365	0.6294	0.7694

the high accuracy of our method.

Figure 16 shows some landmarking results for different methods. The landmarks located by our method are much more accurate than other landmarking methods and more consistent with the ground truth.

4.4 Limitations

First, because the human body models in the dataset have minimal clothing, limited by the training data, the proposed method is only applicable to human body models with minimal clothing.

Second, the proposed method adopts kinematic skeletons to establish local coordinate systems. As in the method of Luo and Feng [35], if the input human body model does not contain feet, or has such flexibility that the feet can both face backwards simultaneously, then the human body models cannot be aligned correctly, making symmetric segments indistinguishable.

Finally, segments play an important role in the proposed landmarking method. Although the accuracy of current shape segmentation exceeds 90%, some models might still be poorly segmented and thus fail to be accurately landmarked. However, this problem will diminish with further research in shape segmentation or using a combination of landmarking results for models with different resolutions. Nevertheless, for human body models with genus larger than 0 or missing parts, there is no guarantee that shape segmentation methods can obtain correct segments, making the proposed method ineffective.

5 Conclusions

In this paper, we collected and labeled 271 human body models with various shapes and poses to form a new dataset for landmarking. The human body models in this dataset are aligned to a consistent orientation, and 17 landmarks are located manually for each model. A learning-based landmarking method was proposed to map several descriptors to the landmark labels. Since the length of the descriptor is invariant to changes in the number of the vertices, the proposed method is robust to changes in model resolution. In addition to some handcrafted descriptors, a multi-scale descriptor named PHDS was proposed based on structural information, making it robust to variations in human

body models. Further, a shape alignment algorithm was proposed to align human body models to enable PHDS to distinguish symmetric parts. Because of the strength of the proposed PHDS descriptor, a simple AdaBoost regression algorithm can accurately locate the landmarks. Although some steps of the proposed landmarking method are based on existing methods, we use these techniques to solve this challenging problem effectively, with good results.

Since the poses of the human body models vary, it is challenging to select a standard orientation to fully align them. The proposed shape alignment algorithm provides a way to align rough human body models using on local coordinate systems based on the hip, giving the human body models a consistent orientation. In addition, the proposed PHDS is not only symmetry-aware but also applicable to models in the same category, so can be used in applications such as shape morphing and shape correspondence to prevent symmetric ambiguity. In the future, we will focus on landmarking for 3D models in other categories. Various applications, such as shape correspondence and shape morphing, that benefit from the landmarks, will also be further investigated.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Creusot, C.; Pears, N.; Austin, J. A machine-learning approach to keypoint detection and landmarking on 3D meshes. *International Journal of Computer Vision* Vol. 102, Nos. 1–3, 146–179, 2013.

[2] Wang, H.; Guo, J.; Yan, D. M.; Quan, W.; Zhang, X. Learning 3D keypoint descriptors for non-rigid shape matching. In: *Computer Vision – ECCV 2018. Lecture Notes in Computer Science, Vol. 11212.* Ferrari, V.; Hebert, M.; Sminchisescu, C.; Weiss, Y. Eds. Springer Cham, 3–20, 2018.

[3] Guo, Y. L.; Bennamoun, M.; Sohel, F.; Lu, M.; Wan, J. W. 3D object recognition in cluttered scenes with local surface features: A survey. *IEEE Transactions*

on Pattern Analysis and Machine Intelligence Vol. 36, No. 11, 2270–2287, 2014.

[4] Yang, Y.; Fu, X. M.; Chai, S. M.; Xiao, S. W.; Liu, L. G. Volume-enhanced compatible remeshing of 3D models. *IEEE Transactions on Visualization and Computer Graphics* Vol. 25, No. 10, 2999–3010, 2019.

[5] Jones, P. R. M.; Rioux, M. Three-dimensional surface anthropometry: Applications to the human body. *Optics and Lasers in Engineering* Vol. 28, No. 2, 89–117, 1997.

[6] Treleaven, P.; Wells, J. 3D body scanning and healthcare applications. *Computer* Vol. 40, No. 7, 28–34, 2007.

[7] You, Y.; Lou, Y. J.; Li, C. K.; Cheng, Z. J.; Li, L. W.; Ma, L. Z.; Lu, C.; Wang, W. KeypointNet: A large-scale 3D keypoint dataset aggregated from numerous human annotations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 13644–13653, 2020.

[8] Allen, B.; Curless, B.; Popović, Z. The space of human body shapes: Reconstruction and parameterization from range scans. *ACM Transactions on Graphics* Vol. 22, No. 3, 587–594, 2003.

[9] Giachetti, A.; Mazzi, E.; Piscitelli, F.; Aono, M.; Hamza, A. B.; Bonis, T.; Claes, P.; Godil, A.; Li, C.; Ovsjanikov, M.; et al. SHREC'14 track: Automatic location of landmarks used in manual anthropometry. In: *Eurographics Workshop on 3D Object Retrieval (2014)*. Bustos, B.; Tabia, H.; Vandeborre, J. P.; Veltkamp, R. Eds. The Eurographics Association, 2014.

[10] Sung, M.; Su, H.; Yu, R.; Guibas, L. Deep functional dictionaries: Learning consistent semantic structures on 3D models from functions. In: Proceedings of the 32nd Conference on Neural Information Processing Systems, 2018.

[11] Chaouch, M.; Verroust-Blondet, A. Alignment of 3D models. *Graphical Models* Vol. 71, No. 2, 63–76, 2009.

[12] Haim, N.; Segol, N.; Ben-Hamu, H.; Maron, H.; Lipman, Y. Surface networks via general covers. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 632–641, 2019.

[13] Hanocka, R.; Hertz, A.; Fish, N.; Giryes, R.; Fleishman, S.; Cohen-Or, D. MeshCNN: A network with an edge. *ACM Transactions on Graphics* Vol. 38, No. 4, Article No. 90, 2019.

[14] Wiersma, R.; Eisemann, E.; Hildebrandt, K. CNNs on surfaces using rotation-equivariant features. *ACM Transactions on Graphics* Vol. 39, No. 4, Article No. 92, 2020.

[15] Johnson, A. E.; Hebert, M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 21, No. 5, 433–449, 1999.

[16] Frome, A.; Huber, D.; Kolluri, R.; Bülow, T.; Malik, J. Recognizing objects in range data using regional point descriptors. In: *Computer Vision - ECCV 2004. Lecture Notes in Computer Science, Vol. 3023*. Pajdla, T.; Matas, J. Eds. Springer Berlin Heidelberg, 224–237, 2004.

[17] Shapira, L.; Shamir, A.; Cohen-Or, D. Consistent mesh partitioning and skeletonisation using the shape diameter function. *The Visual Computer* Vol. 24, No. 4, 249–259, 2008.

[18] Rustamov, R. M. Laplace–Beltrami eigenfunctions for deformation invariant shape representation. In: Proceedings of the 5th Eurographics Symposium on Geometry Processing, 225–233, 2007.

[19] Sun, J.; Ovsjanikov, M.; Guibas, L. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum* Vol. 28, No. 5, 1383–1392, 2009.

[20] Bronstein, M. M.; Kokkinos, I. Scale-invariant heat kernel signatures for non-rigid shape recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1704–1711, 2010.

[21] Aubry, M.; Schlickewei, U.; Cremers, D. The wave kernel signature: A quantum mechanical approach to shape analysis. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, 1626–1633, 2011.

[22] Meng, W.; Yi, F. Local diffusion map signature for symmetry-aware non-rigid shape correspondence. In: Proceedings of the 24th ACM International Conference on Multimedia, 526–530, 2016.

[23] Ren, J.; Poulenard, A.; Wonka, P.; Ovsjanikov, M. Continuous and orientation-preserving correspondences via functional maps. *ACM Transactions on Graphics* Vol. 37, No. 6, Article No. 248, 2018.

[24] Wang, Y. Q.; Guo, J. W.; Yan, D. M.; Wang, K.; Zhang, X. P. A robust local spectral descriptor for matching non-rigid shapes with incompatible shape structures. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6224–6233, 2019.

[25] Li, Y.; Zhong, Y. Q. Automatic detecting anthropometric landmarks based on spin image. *Textile Research Journal* Vol. 82, No. 6, 622–632, 2012.

[26] Wuhrer, S.; Azouz, Z. B.; Shu, C. Semi-automatic prediction of landmarks on human models in varying poses. In: Proceedings of the Canadian Conference on Computer and Robot Vision, 136–142, 2010.

[27] Azouz, Z. B.; Shu, C.; Mantel, A. Automatic locating of anthropometric landmarks on 3D human models.

In: Proceedings of the International Symposium on 3D Data Processing, Visualization, and Transmission, 750–757, 2006.

[28] Lovato, C.; Castellani, U.; Zancanaro, C.; Giachetti, A. Automatic labelling of anatomical landmarks on 3D body scans. *Graphical Models* Vol. 76, No. 6, 648–657, 2014.

[29] Shu, Z. Y.; Xin, S. Q.; Xu, X.; Liu, L. G.; Kavan, L. Detecting 3D points of interest using multiple features and stacked auto-encoder. *IEEE Transactions on Visualization and Computer Graphics* Vol. 25, No. 8, 2583–2596, 2019.

[30] Xi, P. C.; Shu, C.; Goubran, R. Localizing 3-D anatomical landmarks using deep convolutional neural networks. In: Proceedings of the 14th Conference on Computer and Robot Vision, 197–204, 2017.

[31] Yi, L.; Su, H.; Guo, X. W.; Guibas, L. SyncSpecCNN: Synchronized spectral CNN for 3D shape segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 6584–6592, 2017.

[32] Zhou, Z. K.; Hao, S. J. Anatomical landmark detection on 3D human shapes by hierarchically utilizing multiple shape features. *Neurocomputing* Vol. 253, 162–168, 2017.

[33] Marin, R.; Melzi, S.; Rodolà, E.; Castellani, U. FARM: Functional automatic registration method for 3D human bodies. *Computer Graphics Forum* Vol. 39, No. 1, 160–173, 2020.

[34] Guo, J. W.; Wang, H. Y.; Cheng, Z. L.; Zhang, X. P.; Yan, D. M. Learning local shape descriptors for computing non-rigid dense correspondence. *Computational Visual Media* Vol. 6, No. 1, 95–112, 2020.

[35] Luo, S.; Feng, J. Q. Symmetry-aware kinematic skeleton generation of a 3D human body model. *Multimedia Tools and Applications* Vol. 79, Nos. 29–30, 20579–20602, 2020.

[36] Baran, I.; Popović, J. Automatic rigging and animation of 3D characters. *ACM Transactions on Graphics* Vol. 26, No. 3, 72–es, 2007.

[37] Anguelov, D.; Srinivasan, P.; Koller, D.; Thrun, S.; Davis, J. SCAPE: Shape completion and animation of people. *ACM Transactions on Graphics* Vol. 24, No. 3, 408–416, 2005.

[38] Bogo, F.; Romero, J.; Loper, M.; Black, M. J. FAUST: Dataset and evaluation for 3D mesh registration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3794–3801, 2014.

[39] Yang, Y. P.; Yu, Y.; Zhou, Y.; Du, S. D.; Davis, J.; Yang, R. G. Semantic parametric reshaping of human body models. In: Proceedings of the 2nd International Conference on 3D Vision, 41–48, 2014.

[40] Maron, H.; Galun, M.; Aigerman, N.; Trope, M.; Dym, N.; Yumer, E.; Kim, V. G.; Lipman, Y. Convolutional neural networks on surfaces via seamless toric covers. *ACM Transactions on Graphics* Vol. 36, No. 4, Article No. 71, 2017.

[41] Loper, M.; Mahmood, N.; Romero, J.; Pons-Moll, G.; Black, M. J. SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics* Vol. 34, No. 6, Article No. 248, 2015.

[42] Chen, X. B.; Golovinskiy, A.; Funkhouser, T. A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics* Vol. 28, No. 3, Article No. 73, 2009.

[43] Xu, Y. F.; Fan, T. Q.; Xu, M. Y.; Zeng, L.; Qiao, Y. SpiderCNN: Deep learning on point sets with parameterized convolutional filters. In: *Computer Vision – ECCV 2018. Lecture Notes in Computer Science, Vol. 11212.* Ferrari, V.; Hebert, M.; Sminchisescu, C.; Weiss, Y. Eds. Springer Cham, 90–105, 2018.

[44] Wang, Y.; Sun, Y. B.; Liu, Z. W.; Sarma, S. E.; Bronstein, M. M.; Solomon, J. M. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics* Vol. 38, No. 5, Article No. 146, 2019.

[45] Wu, W. X.; Qi, Z. A.; Li, F. X. PointConv: Deep convolutional networks on 3D point clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9613–9622, 2019.

Shan Luo received her B.S. degree in computer science from Jilin University, Changchun, China, in 2015. She is now a Ph.D. candidate in the State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China. Her research interests include human body shape analysis and synthesis.

Qitong Zhang received her B.S. degree in digital media technology from Shandong University, Jinan, China, in 2017. She is now a Ph.D. candidate in the State Key Laboratory of CAD&CG, Zhejiang University. Her fields of interest are multi-view stereo and 3D reconstruction.

Jieqing Feng is a professor in the State Key Laboratory of CAD&CG, Zhejiang University. He received his B.Sc. degree in applied mathematics from the National University of Defense Technology in 1992 and his Ph.D. degree in computer graphics from Zhejiang University in 1997. His research interests include geometric modeling, real-time rendering, stereo vision, and modeling and simulation in solar thermal power.