Novel correspondence-based approach for consistent human skeleton extraction

Kang Wang[1] · Abdul Razzaq[1] · Zhongke Wu[1] ·
Feng Tian[2] · Sajid Ali[1] · Taorui Jia[1] · Xingce Wang[1] ·
Mingquan Zhou[1]

Abstract This paper presents a novel base-points-driven shape correspondence (BSC) approach to extract skeletons of articulated objects from 3D mesh shapes. The skeleton extraction based on BSC approach is more accurate than the traditional direct skeleton extraction methods. Since 3D shapes provide more geometric information, BSC offers the consistent information between the source shape and the target shapes. In this paper, we first extract the skeleton from a template shape such as the source shape automatically. Then, the skeletons of the target shapes of different poses are generated based on the correspondence relationship with source shape. The accuracy of the proposed method is demonstrated by presenting a comprehensive performance evaluation on multiple benchmark datasets. The results of the proposed approach can be applied to various applications such as skeleton-driven animation, shape segmentation and human motion analysis.

Keywords Shape correspondence · Heat Kernel signature · Mesh contraction · Skeleton extraction

1 Introduction

The 3D shape correspondence is to find a meaningful relation between pair of shapes and it has a wide range of applications in a variety of domains, such as geometric modeling and processing [17], shape recognition [29, 38], shape registration and shape retrieval [27, 41]. The representation of variations of human shapes has been an important topic in computer

---

✉ Zhongke Wu
zwu@bnu.edu.cn

[1] Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage, Beijing Normal University, Beijing, People's Republic of China

[2] School of Design, Engineering & Computing, Bournemouth University, Bournemouth, UK

vision and graphics fields. Establishing a significant relationship between shape correspondences of similar model with different poses is a challenging task for several reasons. Firstly, among the different types of poses from the same person, there are many non-rigid deformations [15]. Secondly, the human motion capture system could capture 3D postures through multiple cameras environment [5], but we need the meaningful relationship of shape points among distinctive pose shapes.

The shape correspondence is also closely related to the extraction of the curve-skeleton of shape. An articulated 3D model skeleton provides an intuitive abstraction for both geometrical and topological shape of the objects. An extracted 1D curve-skeleton is an effective representation of the model, which facilitates manipulating and understanding the shape of model [18]. The skeletal representation of the model has become very popular and find wide range of applications in shape analysis [26], surface reconstruction [40], character skinning through skeletal rigging [25], skeleton-driven character animation [7], object matching and shape retrieval [12, 22] and shape deformation through skeleton [45]. With the capability of creating a three-dimensional representation of human models, the extraction of curve-skeletons for these models has become a fundamental problem in many applications.

A number of methods can be found to extract a skeleton from 3D shape [4, 23]. Majority of these methods are geometry-based, which have hybrid approach combining topology and geometry-based techniques [44]. An approach which has the combination of capabilities such as robust and high accuracy to extract a skeleton for different poses of same model is highly required. The extracted skeletons by using these methods are often not satisfactory due to undesirable redundant branches, complexity in joints hierarchy and manual approximations during extraction. Technically, in skeleton extraction of the shape, an extracted skeleton is well-centered, well-define joint hierarchy joints. The accurate joint identification that matches the real position in every pose of the shape is a challenging task in computer graphics and interactive applications. A number of commercial 3D packages (such as Maya™, Blender, etc.) have been developed for creating automatic skeleton for the object. However, the generation of skeleton by using these packages is time-consuming, labor-intensive and the results depend on user's skills.

In this paper, we propose and develop an integrated framework of consistent human skeleton extraction based on the base-points-driven shape correspondence (BSC). The source and target shapes are topologically consistent, with different postures and semantically similar to each other. BSC establishes the corresponding relationship between two shapes using local maximum Scale-invariant Heat Kernel signatures as base points to embed the original shapes into a Euclidian space. In this space, we first compute the similarities between points on source and target shapes, and then build the correspondence, based on which the skeleton can be extracted.

To extract the skeleton of a source shape, at first, the source shape is contracted through Laplacian-based mesh contraction method. A 1D curve-skeleton of the contracted shape is then obtained through topological thinning. The final joint-based skeleton shape is produced by applying geometric refinements on extracted 1D curve-skeleton of the source shape. The consistent skeleton of the target shape which possesses different poses of the source shape is then generated based on correspondence relationship between these shapes. The obtained target shape skeleton has the similar topology of the source skeleton such as an equal number of joints, approximately identical joints' positions and hierarchy structure.

The rest of this paper is organized as follows. In Section II, we review previous works on shape correspondence and human skeleton extraction. The details of our proposed BSC are

given in Section III. Firstly, we describe the model downsampling method and how to find the base points. Then this section focuses on how to select the base points and identify the similar points on the source and target shapes. In the end of this section, some results of correspondence are presented. Section IV describes the extraction of the source shape skeleton and generation of target shape skeleton through our shape correspondence approach and compare it with the direct skeleton extracting method. Our experimental results conducted on two benchmark datasets are presented in Section V, followed by conclusion and future work in Section VI.

2 Related work

The extraction of a model's skeleton should give enough information for the overall structure, while maintaining a certain level of details for the model. The pros and cons of various skeleton extraction methods and skeleton properties have been well documented in [13].

In the literature, various methods have been intensively studied for extracting skeletons from 3D shapes. Some methods concentrate on the pairwise skeleton correspondence [51]. Majority of the solutions focused only on the 1D curve-skeleton extraction from 3D objects which describes the abstract structure of the model [20] although some techniques try to compute a joint-based skeleton of the model [28]. Although, the partitioning of the 3D shape can assist in the creation of skeleton and the skeleton of any shape may infer a segmentation of the shape. Lior et al. [37] proposed an approach for decomposing mesh into meaningful parts based on shape diameter function. In their work, Shape diameter function (SDF) was also guided a curve-skeleton by using consistent partition of the given mesh. However, focus of their method is to partition the mesh and its created skeletons are only curve-skeletons rather than joint-based skeletons. Furthermore, skeleton generating by using segmentation algorithms are often lead to expensive computation and resulting skeletons look star-shape skeletons due to over-segmentation and jagged boundaries between segments. The skeleton extracted by these methods generates satisfactory results for single input model. However, they suffer some limitations such as mismatching the location of skeletal joints with the real position of the model, failure to preserve the topology of the model, etc. For automatic extraction methods, the extracted joints hierarchy increases the complexity and fewer unwanted branches could be created. The generated skeletons through these methods also need to tune through skeleton pruning algorithms for every model.

The use of silhouettes as input to recognize human actions [6, 10] is based on the representation of the contour points from human silhouette. But the silhouette is obtained previously by extraction techniques, e.g., background subtraction and mostly from image. And many previous works use edges and silhouettes as pose descriptors [14, 49]. These methods use only image features and are not suitable for a strong dynamic model of human motion. The use of the same model for skeleton extraction and motion capture, depends on a one-to-one correspondence between estimated and ground truth joints [36].

Base point means prominent feature point in this paper. We need them to be consistent on shapes with isometric deformations. Reuter and Peinecke [31] have proved that the Laplace–Beltrami eigenvalues as isometry-invariant shape descriptors could be used to recognize the isospectral models dubbed as "Shape-DNA". By the decomposition of the Laplace–Beltrami operator, Reuter got the isometry-invariant signature of a manifold from the eigenfunction corresponding to the first few smallest non-zero eigenvalues. But the "Shape-DNA" is a global

feature. The heat diffusion on a manifold is the diffusion process whose infinitesimal generator is the Laplace–Beltrami operator[11]. Coifman and Lafon [11, 24] have introduced the diffusion processes on manifold which formed a base for the recent studies about the diffusion geometry in shape analysis. Sun and Ovsjanikov [39] proposed an intrinsic, multi-scale and robust shape descriptor, Heat Kernel Signature (HKS). This descriptor is based on the physical processes of heat propagation on a shape and also related to the diffusion geometry proposed by Lafon. It was obtained through the heat kernel of different time interval. HKS is efficiently computable and provides a multi-scale way to capture information about neighborhoods of a given point [39]. In this approach, when time period is small, the HKS takes more information from the close neighbors and catches the local features.

In shape correspondence, isometry is a very important clue for finding the final result. Fortunately, with the development of 3D data acquisition technology, the acquisition of parameterization-free shapes becomes popular and the data are almost in isometric. The BSC method in this paper is based on the invariance of geodesic distance with isometric deformations. To obtain the correspondence directly between two sets is an NP-hard problem. Many researchers have proposed various methods to simplify it and construct some models based on some descriptors [8, 46] and metric structure of shape [1, 19, 48]. H. Zhang et al. [50] proposed a deformation view of shape correspondence and proposed a deformation-driven method. Yusuf Sahillioğlu et al. [32] proposed a greedy optimization on the isometry cost model, and then took different algorithms in [33–35] to detect the correspondence. This approach has solved the correspondence problem with greedy optimization, but it needs a good initialization in order to avoid getting stuck with local maxima. Instead, our approach is able to avoid this by solving a minimum cost max flow problem.

3 Base-points-driven shape correspondence

In this section we first review the downsampling method, Farthest Points Sampling (FPS) [16]. How to represent these down sampled points and eliminate the effects of non-rigid deformations between them are major challenges for obtaining the plausible result of shape correspondence. Thereafter, we introduce our improved correspondence method BSC.

3.1 Downsampling methods

In order to improve the efficiency of the algorithm, a limited number of points have been required under a certain magnitude. A downsampling algorithm FPS has been used to decimate the points. FPS provides almost evenly spaced sampling, the next sample is placed in the center of the largest empty disk on surface, or circle on the plane for 2D cases [16]. In other words, the next sample is placed at a point that is farthest from the previous samples. To this effect, each sample point is as far as possible from other points and at the same time as close as possible.

In the FPS procedure, the points at the anchors (e.g., the feet and the hands of human or animal shapes) will be selected prior to other points on the shape. Various anchor points of each isometric shape will be corresponded and selected firstly by FPS. It obtains a well-separated covering net of the shape (Fig. 1). By taking the advantage of FPS, we have obtained subsets of the shape for our shape correspondence method.
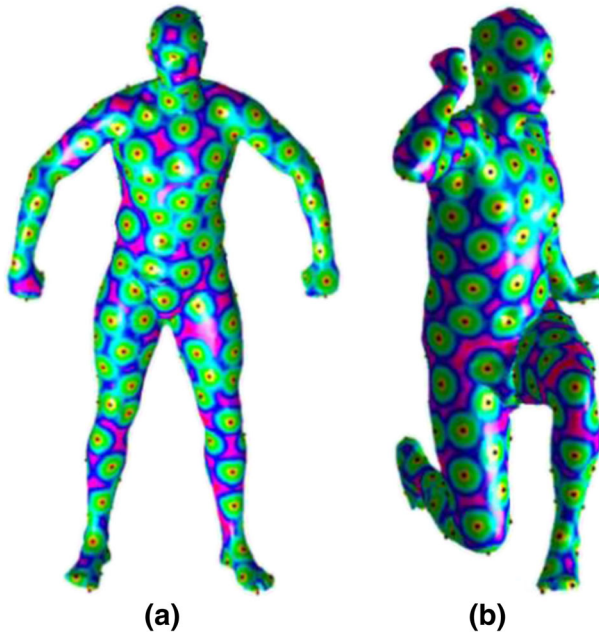
Fig. 1 Results of the FPS downsampling on two different shapes with 200 points

3.2 New shape representation

For a 3D surface $\mathbf{S}(u,v)$, which is a 2D manifold surface, a geodesic metric $d_{geo}(x,y)$ is defined on it, where x, y are two points on the surface. There are three or more points, which are not on a same geodesic path, are selected as the base points: p1, p2, p3. Then arbitrary point p0 on the surface will have an exact location. The $d_{geo}(p_0,p_i)$ has been computed, where $i=1,2,3$. These three geodesic distance values have been taken to be the new coordinates for all the points on the surface $P0(d_{geo}(p0, p1), d_{geo}(p0, p2), d_{geo}(p0, p3))$ in a 3D Euclidian space.

After calculating a point p, we compute the geodesic distances from this point, which forms a field on the surface, through the geodesic function $f(q)=d_{geo}(q,p)$. Every point only has one value. All the same value points are on a close curve, similar to the circular contour on the plane in Euclid space.

In the Euclidean geometry, it has well known three points, which are not on the same curve, can decide one and only one plane. This can be described by the intersection point of three circles, which takes the base points as the centers and the distances from the base points as the radius respectively as shown in Fig. 2. For each point of the surface, we can find the unique group of circles as the coordinate system.

We extend this method to the manifold surface with the geodesic metric. The geodesic distance is selected to replace the Euclid distance, and geodesic contour, which is a closed curve, and used to replace the circle. It is illustrated in Fig. 3.

In Fig. 3a, the geodesic contour is some close curves and there is no intersection point between any two geodesic contours because of a unique distance between any point and the source point. If there are two geodesic contours intersect at one point M, all the points on the
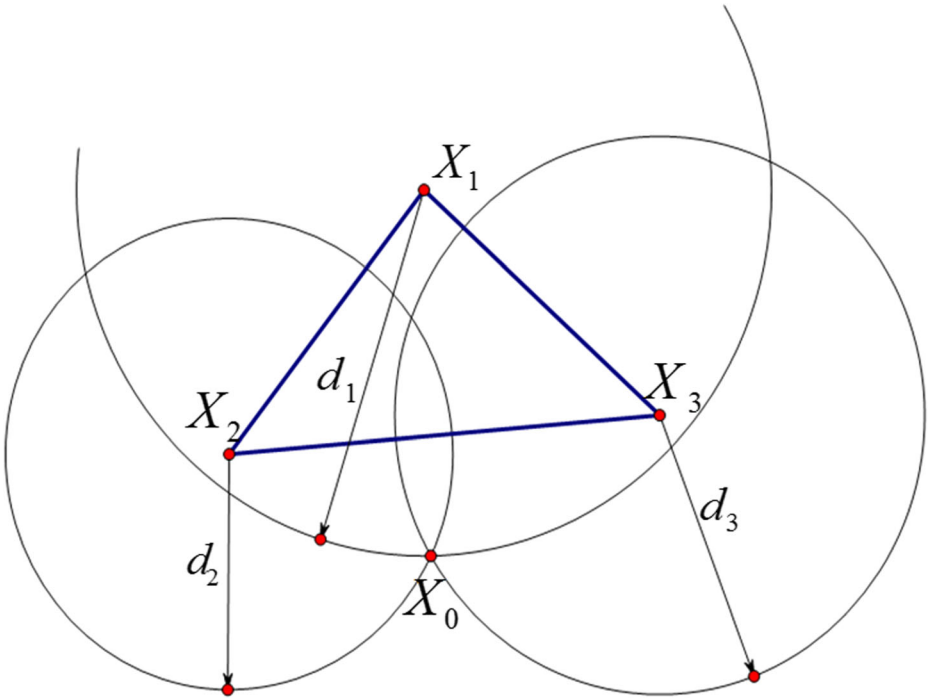
Fig. 2 Any point can be represented by the intersect point of three circles centered by three base points

two contours have the same geodesic distance. In Fig. 3b, considering a point P on the outer contour, the geodesic path from the source S to point P pass through the area, is between the two contours. When it enters into this area, there will be a intersect point Q within the inner contour. However, the points P and Q have the same distance. So there is no intersecting point between any two geodesic contours.
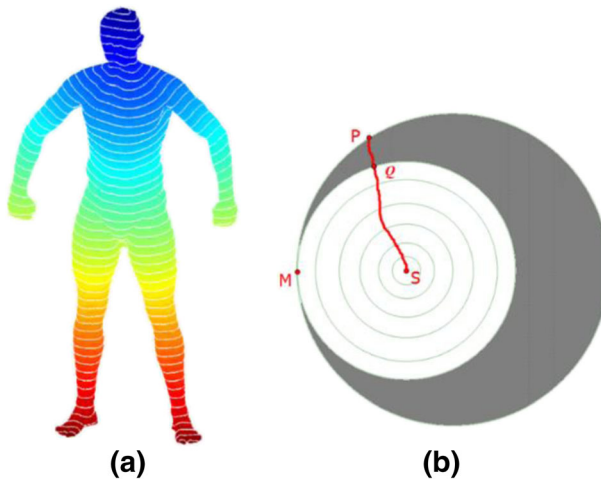
Fig. 3 Example of geodesic. **a** Geodesic contours from a source point on a human pose shape. **b** Illustration on why these contours do not intersect when the two geodesic contour intersect at point M

Three base points are selected, all the points on the surface can be defined by geodesic distances between base points. Under this situation, if base points of two shapes are given in order respectively, a sequence of distance values can be used to represent a point on the surfaces. Because geodesic distances are unique, the sequence will also be unique for one point. It's very important to define these consistent base points on different pose shapes. The following section presents our approach to find the same base points on the source and target shapes using an intrinsic point feature Scale-invariant Heat Kernel Signature (SIHKS) [9].

3.3 Heat Kernel and Heat Kernel signature

3D shapes representations of human posture shapes are convenient in applications such as rendering and visualization. But it is not suitable, at least in a direct way, for many other situations including human motion comparison, motion correspondence, shape estimation, and recognition. In these applications, the pose shapes are considered to be similar if there exist rigid or isometric transformations between them.

Considering the non-rigid deformations of these posture models, this makes it difficult to find their correspondence. Therefore we need a method to eliminate the influence of the deformations. So the HKS as the intrinsic features is chosen to find the base points.

Let M be a compact Riemannian manifold without boundary, the amount of heat at a point $p \in M$ at time t is defined by $u(p,t) : M \times \mathfrak{R}^+ \to \mathfrak{R}^+$. So at time 0, the heat at every point could be represented as the function $f : M \to \mathfrak{R}^+$. Moreover the diffusion of heat on M is governed by the heat equation as follows:

$$(\Delta_M + \frac{\partial}{\partial t})u(p,t) = 0 \tag{1}$$

Here Δ_M is a Laplace-Beltrami operator of M. If the heat distribution f at time t is given by the heat operator H_t, then the following equations are satisfied:

$$\lim_{t \to 0} H_t f = f, \text{ and } u(p,t) = H_t f(p)$$

And these two operators have the relation $H_t = e^{-t\Delta_M}$.

The heat kernel is a function as described in the following equation:

$$k_t(p,q) : \mathfrak{R}^+ \times M \times M \to \mathfrak{R}^+ \tag{2}$$

It satisfies $H_t f(p) = \int_M k_t(p,q) f(q) dq$ for all $p \in M$ and measures the amount of heat transferred from p to point q in time t. According to the eigen-decompositon of Laplace-Beltrami operator and the relation of Δ_M and H_t, the spectral expansion of heat kernel on any compact manifold M has the following form:

$$k_t(p,q) = \sum_{i=0}^{\infty} e^{-t\lambda_i} \phi_i(p) \phi_i(q) \tag{3}$$

Here λ_i and φ_i are i-th eigenvalue and its corresponding eigenfunction of Laplace-Beltrami operator respectively. From Eq. (3), we can see that, the heat kernel is symmetric $k_t(p,q) = k_t(q, p)$. The HKS is also isometry-invariant for the shapes with isometric transformations. Inversely if two shapes have the same HKS, then they are isometric.

$HKS(p) : \mathfrak{R}^+ \rightarrow \mathfrak{R}, HKS(p,t) = k_t(p,p)$ is exactly the Heat Kernel Signature defined on the point p of the manifold M. It could be represented as:

$$k_t(p,p) = \sum_{i=0}^{\infty} e^{-t\lambda_i} \phi_i(p)^2 \tag{4}$$

As a local shape descriptor HKS has also many properties such as multi-scale property especially sampled at a finite set of time $t_1, \cdots, t_n$:

$$HKS(p) = (k_{t_1}(p,p), k_{t_2}(p,p), \cdots k_{t_n}(p,p)) \tag{5}$$

In order to deal with global and local scaling transformations, we need a scale-invariant method which could be achieved by [9, 30, 43]. In this paper, the Scale-invariant Heat Kernel Signature has been employed for defining the base points. According to Bronstein [9], shape scaling factors could be removed by the logarithmically sampling, discrete derivative and discrete-time Fourier transforms. This approach created scale-invariant feature descriptors, and the SIHKS extend the heat kernel signature to deal with global and local scaling transformations. By means of the scale and pose invariant properties of SIHKS, we could find the consistent base points on two shapes with isometric deformations by the local extreme of the SIHKS. Figure 4 shows some results of the base points (marked with pink balls) on the same person with different poses. These points are taken from the local maximum SIHKS in the range of a fixed geodesic distance on shapes.

3.4 Identification of similar points

In this subsection, we describe the method to compute the correspondence of points on the source and target shape. The base points are selected, and they have been matched to each
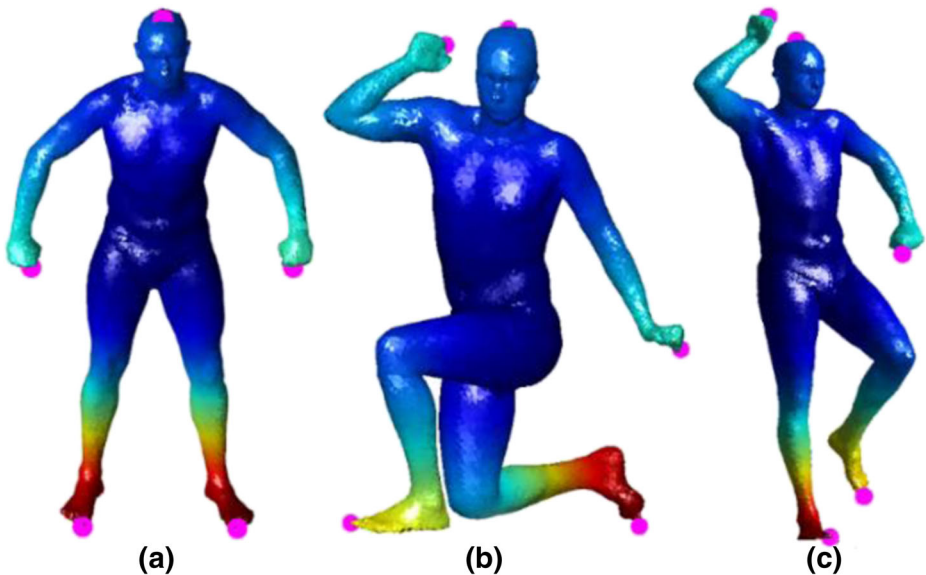
Fig. 4 Example of base points on different shape poses. **a** Base points (*pink balls*) on a standing pose. **b** Base points on the squatting pose. **c** Base points on the dancing pose

other. Now that the geodesics from these base points are computed, we take the order of the base points to be the order of new coordinate in Euclidian embedding space. The geodesic distances matrix M_1 and M_2 have been computed by using the Fast marching algorithm [21].

After embedding the shape into Euclidean space through setting the matched and ordered points of original shapes as base points, two new shapes are achieved, which have the same location and the same surface, as seen in Fig. 5. The correspondence problem then becomes to find the closest pairs of points between two point sets.

Let's denote the $M_1(V_1,F_1)$ and $M_2(V_2,F_2)$ as the source shape and the target shape respectively, where $V_i(i=1,2)$ is the set of new points, which can be written as $\mathbf{v}(v_1,v_2,v_3,\cdots,v_k)$, k is the number of the base points. F_i is the face from the original shape, i.e., the topology of the original and the new shape are same. Now we just find the nearest points between the two sets. The greedy strategy can be taken to find the nearest point from M_2 for each point in M_1. However, it just meets the local optima. Our aim is to find a map $\Psi(v_1):v_1 \in M_1 \rightarrow v_2 \in M_2$, which is given below:

$$\Psi(v) = \underset{v_i \in M_1, \Psi(v_i) \in M_2}{\mathrm{argmin}} \sum_{i=1}^{N} L_2(v_i, \Psi(v_i)) \qquad (6)$$

Here $L_2(v_i, \Psi(v_i))$ is the L_2 distance and N is the number of M_1. Now this problem changes to an optimization problem, which could be solved by the following minimum-cost maximum flow problem. The minimum-cost flow [2] is an important and typical problem in the field of graph theory, and is the core of network optimization problems.

This problem can be described as follows. Giving a flow network, that is, a directed graph $G=(V,E)$, with source point $s \in V$ and a target point $t \in V$, each edge $e(i,j) \in E$ has an associated cost a_{ij} that denotes the cost per unit flow on that edge. A flow $f_{ij} \geq 0$ and a capacity $c_{ij} > 0$ are associated with every edge. We can send a flow f from s to t.
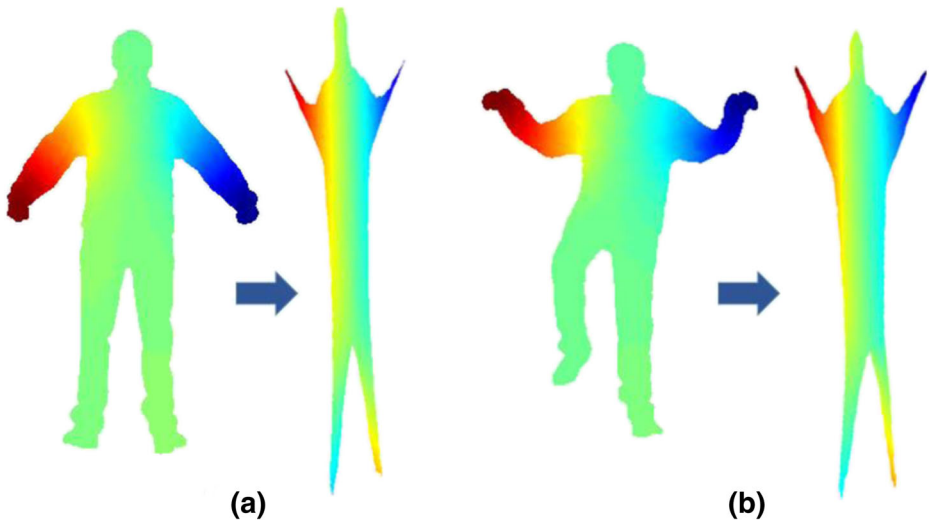
Fig. 5 Embed two new shapes into the same Euclidean space. **a** Standing pose and its representation. **b** Crane pose and its representation

The minimum-cost flow problem is an optimization model formulated as follows:

$$Minimize \sum_{(i,j) \in E} a_{ij} f_{ij} \tag{7}$$

Subject to $f_{ij} \leq c_{ij}, f_{ij} - f_{ji}, \sum_{j \in V} f_{ij} = 0$ for all $i \neq s, i \neq t$ and $\sum_{j \in V} f_{sj} = \sum_{i \in V} f_{it} = d$.

Based on this, if the point set, except s and t, can be separated into two sets with the edges, we call this graph is a bipartite graph. Then with a weight to every edge, we get the minimum weight graph, which will be used in our method to solve the correspondence problem.

3.5 Correspondence results

Some of our correspondence results are illustrated in Fig. 6. We represent the correspondence between each pair of shapes by the lines. The distances are computed between each correspondence points in k (in our experiment $k=5$) dimensional Euclidian space. If k is larger the process is more robust. But it will be more complicate for the base points to be corresponded. With enough base points, nevertheless we will reduce the correspondence error and the symmetry ambiguity. After that we take the average of these distances as the measure standard to test the quality of the result.

Figure 6 also shows the sparse and dense correspondence results. Sometimes the coarse correspondence may lead to less computation time, but this will result in reduced accuracy for skeleton extraction. So the dense correspondence is more suitable for our case.

4 Skeleton extraction through shape correspondence

After the establishment of the correspondence between target and source shapes, we want to extract the target shape skeleton by using the source one. The process of extracting target shape skeleton consists of two steps. At first, a skeleton has been extracted from one of the correspondence shape (source shape). Then a similar plausible skeleton is generated of target shape based on the source shape skeleton. The obtained target skeleton maintains equal number of joints and preserves similar hierarchy of the source skeleton joints.

(a)   **(b)**
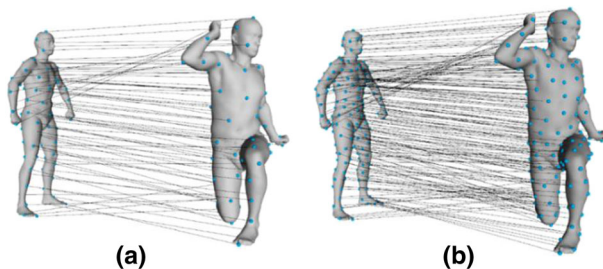
Fig. 6 Result of the correspondence. **a** Sparse result for correspondence. **b** Dense result for correspondence

4.1 Source shape skeleton extraction

The suitable source shape skeleton has been extracted using the method in [40] and then converted curve-skeleton into joint-based skeleton by applying geometric refinements. In source shape skeleton generation, at first the source shape is contracted through Laplacian-based mesh contraction which converts the geometric model into the skeletal shape of the input model. The contracted mesh consists of approximate zero-volume that is visually skeletal representation of that shape. This contraction process preserves the original topology and maintains the real shape of the model. The linear system of vertex positions is iteratively solved during the mesh contraction by using Eq. (8) below:

$$\begin{bmatrix} W_C L \\ W_A \end{bmatrix} V' = \begin{bmatrix} 0 \\ W_A \end{bmatrix} \tag{8}$$

where W_c and W_A are the constraints of contraction and attraction respectively. The contraction and attraction constraints have been minimizing through quadratic energy function by using Eq. (9).

$$\|W_C L V'\|^2 + \sum_i W^2_{A,i} \|v'_t - v_i\|^2 \tag{9}$$

In this equation the first part relates to contraction constraints and the second part describes the attraction constraints. $W_{A,i}$ is the attraction weight for every point i after each iteration. L is the $n \times n$ Laplace operator with elements, V' represents the contracted vertex of the source shape. In order to get more accurate results of the skeleton extraction, we may apply minimum user intervention.

The contracted mesh is then converted into hierarchical joint-based skeleton through topological thinning and geometry refinements. In topological thinning we have applied edge-contraction operation on the contracted mesh to collapse the unnecessary edges until to create the 1D curve-skeleton of the shape. The edge-collapse process is applied repeatedly until all triangles in the contracted mesh are removed and the final curve-skeleton is obtained. The obtained 1D curve-skeleton is then refined into hierarchical joint-based skeleton through geometric refinements. The refinements include: computation of skeleton joints, identification of the root node, establishment of the hierarchy between skeleton joints and creation of the graph of skeleton joints. The extracted joint-based skeleton by implementing topological and geometric refinements on the contracted shape of the source model is given in Fig. 7b.

4.2 Consistent skeleton generation for target shape

The computed skeleton of the source shape is then used to extract the consistent 1D curve-skeleton of the target shape through computed surface correspondence of source to target shape. The joints of the source skeleton are used to calculate the skeleton-vertex correspondence of source shape through the matrix D. D is the shortest distance between every joint of the source skeleton and its shape point described in
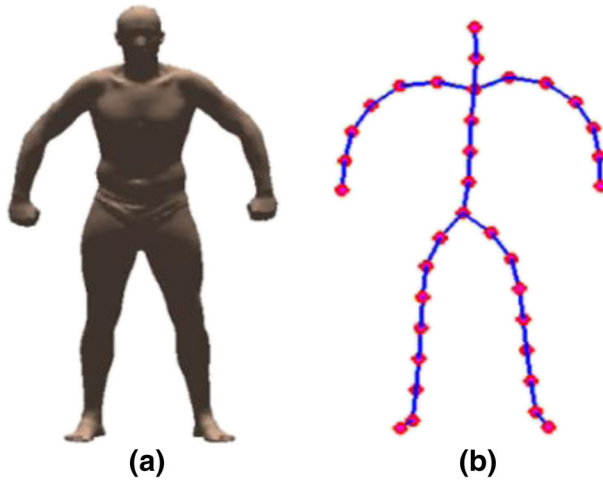
Fig. 7 Extraction of the source shape skeleton. **a** Source shape. **b** Skeleton of the source shape

Eq. (10). Let the source model P and size of the model points is m and the extracted skeleton S consists of n joints. The skeleton to shape surface relations is calculated as:

$$D_{ij} = \left\| P_i S_j \right\|, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n \tag{10}$$

$$\left\| P_i S_j \right\| = \sqrt{\left(P_{i1} - S_{j1}\right)^2 + \left(P_{i2} - S_{j2}\right)^2 + \left(P_{i3} - S_{j3}\right)^2} \tag{11}$$

The closest shape vertex E of source skeleton joints is determined by distance matrix D, $E_i = D_{Ii}$, where I_i is minimum distance of the every skeleton joint from its associated shape vertex.

The projected nodes of the target shape skeleton have been estimated using BSC shape correspondence and skeleton-surface relationship E of the source shape. Our BSC approach computes one to one correspondence between source and target shapes points. Thereafter, the similar points of the target shape are grouped into one group (cluster) based on E. Let the target shape points are V composed of k group corresponding such that $v_1 \cup v_2 \cup \ldots \cup v_k = V$. Let the size of j-th group in target mesh points is $n_j (j = 1, 2 \ldots, k)$. We calculate the center of each group $\bar{v}_j, (j = 1, 2 \ldots, k)$ of the target mesh vertex using Eq. (12) as:

$$\bar{v}_j = \frac{\sum_{i=1}^{n_j} v_{ij}}{n_j}, j = 1, 2, \ldots, k \tag{12}$$

The prototype (center) of each group is computed of target shape that represents the extracted nodes of the target shape skeleton $T = \left\{ \bar{v}_i, \bar{v}_j \cdots, \bar{v}_k \right\}$ as shown in Fig. 8. To connect the target shape nodes, a cyclic connection between nodes has been constructed based on shape correspondence and using rings of share vertices in target shape. To compute the connection between target shape nodes, at first, we compute the connected rings of the shape vertices.
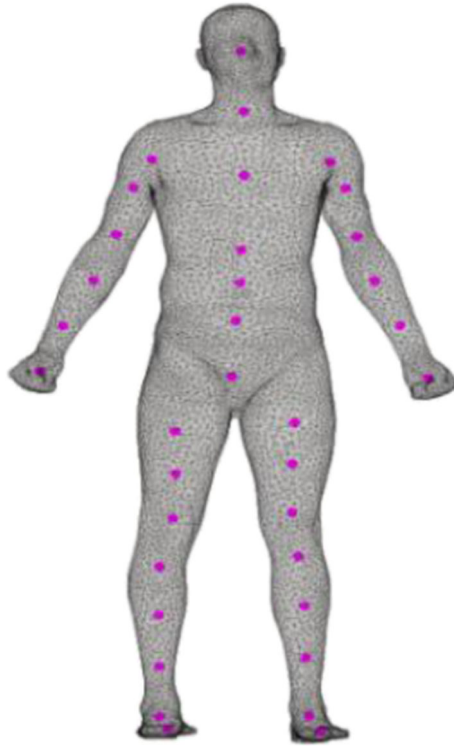
**Fig. 8** The extracted nodes of the target shape

Nearest neighbor search (NNS) algorithm has used to find the closest points in the shape for computing one-rings of mesh point neighbors. Numerous solutions in computational geometry have been proposed for NNS problems such as linear search which calculates the distance from a query to every data point, local sensitive hashing (make buckets of data based on a distance metric), greedy search (construct graph of the points from query to its neighborhood). We use the space-partitioning algorithm for NNS based on K-d tree data structure [47]. K-d tree algorithm is based on iteratively division of the search space into two regions. The K-nearest neighbor of mesh points is computed by using k-d nearest neighbors search. At first, we compute an approximate neighborhood of target shape V_i by calculating its k nearest neighbors $N_k(V_i)$. The approximate neighbors are projected on a tangent plane defined by their principal components analysis. A planar Delaunay triangulation is constructed and defines one-ring neighbors of mesh points which are used to build the connection between target mesh nodes. The parameter required to define the number of nearest neighbors is $k=0.012$. We use this parameter to estimate the tangential plane and bounding minimum and maximum values (neighbors) in the range of [8:30]. A similar value of k has been used throughout experiments in all shapes for computing the rings of approximate neighbors of the mesh points. In cyclic connection, the value of connected nodes with its neighbor nodes is equal to 1. The generated joint-skeleton of the target shape that is approximately similar to source shape skeleton present in Fig. 9b. In addition, the obtained skeleton from the target shape also generates skeleton joints to surface points mapping that can be directly applicable to mesh skinning deformation.
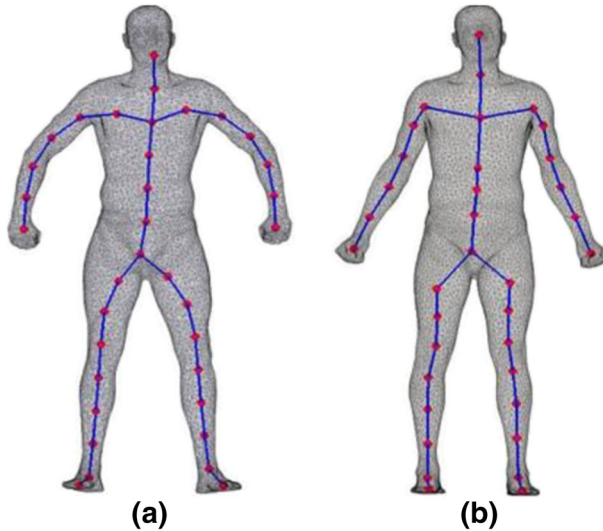
Fig. 9 Consistent skeleton extraction based on source shape skeleton. **a** Source shape, **b** Consistent skeleton of target shape

5 Results and discussions

Although, the extraction of a 1D curve-skeleton through mesh contraction gives a satisfactory skeleton for single shape, but it does not always generate a robust skeleton for every pose of the same shape. In Fig. 10 we compare the extracted skeleton of the target shape skeleton from two approaches: one is the mesh contraction by applying topological and geometric refinements and the other is BSC. In Fig. 10a it can be seen that the target shape skeleton in some extent is off center and has the sharp bend of the bones. A post processing is needed to correct the position of the skeletal joints. On the contrary, our result shown in Fig. 10b is much better. The resulting skeleton remains consistent over the poses change of the source shape and generates analogue skeletons of the target shape as shown in Fig. 9.
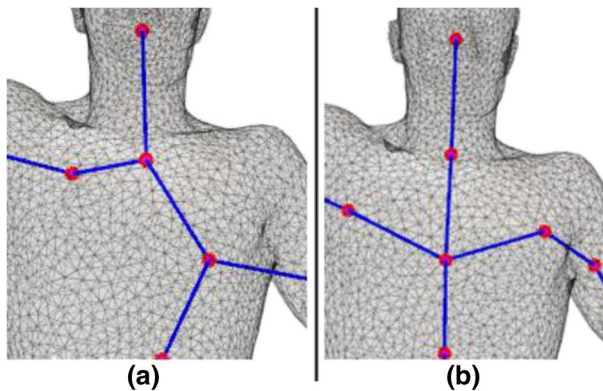
Fig. 10 Comparison of our method and mesh contraction. **a** The result of mesh contraction. **b** Result of our method

Figure 11 gives more results of the skeleton extraction and comparisons of our approach to the direct skeleton extraction of the target shape. We test our approach on two different datasets: Vlasic et al. [42] and SCAPE: Shape Completion and Animation of People [3]. In both datasets each single person performs multiple motions. The shapes in first column of Fig. 11 are the source shapes with their extracted skeleton. The source shape skeletons are extracted through geometric contraction, topological and geometric refinements. Automatic generated skeletons of the target shapes by using BSC are present in second and fourth columns of Fig. 11. The obtained skeleton of the target shape is consistent with the source shape skeleton. The skeletons of the target shapes have an equal number of joints as the source shape. They are well-centered and have similar hierarchy of the source shape skeleton joints.

Majority of the skeletonization methods may generate different skeletons for different postures of the same shape. As shown in Fig. 11, the direct skeletons extraction approach obtains different skeletons of the target shape with different poses. Those skeletons have unequal number of joints, different joints hierarchy and do not match satisfactory position of joints with its original shape. Based on BSC, our approach is able to extract consistent skeletons of target shapes which are different poses of the source shape. And our approach does not require any topological, geometrical and embedding refinement for every poses of same shape.
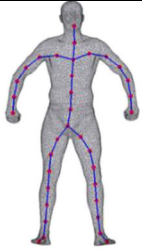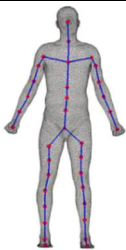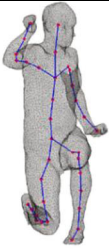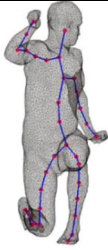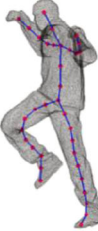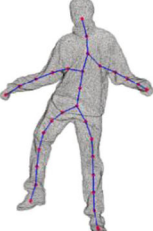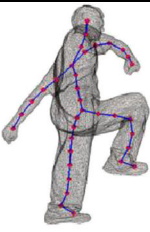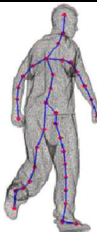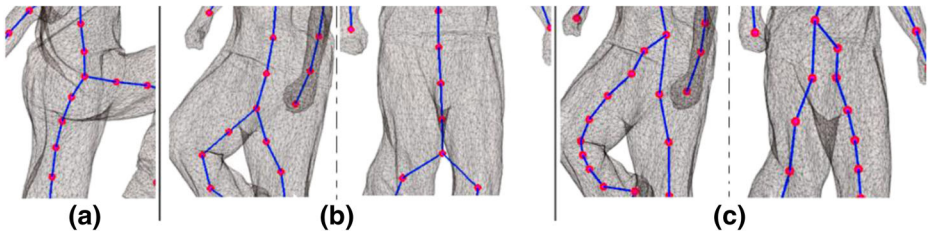

Source shapes' skeleton	Target shapes' skeleton	Direct skeleton extraction result	Target shapes' skeleton	Direct skeleton extraction result

Fig. 11 Comparisons with other results

Fig. 12 Consistency of skeleton point position. **a** Source skeleton, **b** the result of direct skeleton extraction method, **c** the result of our method

There is no need to set approximation parameters for extracting a skeleton of same shape with different poses. The results in Fig. 12 demonstrate the robustness of our approach. The source skeleton joints Fig. 12a and our result Fig. 12c are consistent despite of the position of the model's pants. However, the result Fig. 12b obtained through the direct skeleton extraction is not accurate due to the pants' position in the target pose is lower than that in the source shape. In the implementation, our extraction of skeleton approach is performed in MATLAB. Therefore, it is not yet optimized for speed. The resolution of the mesh in [42] and SCAPE [3] have 20 and 125 K polygons respectively. The extraction result will be more precise for meshes with high resolution, but it will take slightly more time to establish the correspondence.

6 Conclusion and future work

In this paper, we have proposed an automatic and pose-invariant approach to extract the skeleton of an articulated human model based on BSC. We apply BSC to find the consistency of geometric information between shapes. Based on the corresponding relationship between the source and target shapes, the plausible skeleton of the target shape is extracted. Our approach supports the reusability of the source skeleton from the same model. It does not require any pre-processing to shapes, such as downsampling, mesh contraction, simplification and etc. In the proposed solution, the skeletons of different poses of the target shapes are computed instead of single shape.

The robustness of our proposed method is demonstrated through our tests on two well-known datasets. Experiments have shown that our algorithm produces both effective and accurate skeletons of the target shapes. Compared with the directly extracting method, our approach performs better in term of consistency such as equal number of skeleton joints and similar hierarchy between skeleton joints for a variety of target shapes. Furthermore, the skeleton generated by our approach is efficient because it does not require post-processing such as topological and geometric refinements. In the future we will investigate and study more efficient and robust correspondence method so that the accuracy and consistency of the skeleton extraction can be improved further.

References

1. Aflalo Y, Dubrovina A, Kimmel R (2013) Spectral generalized multi-dimensional scaling. arXiv preprint arXiv:13112187
2. Ahuja RK, Magnanti TL, Orlin JB (1993) Network flows: theory, algorithms, and applications
3. Anguelov D, Srinivasan P, Koller D, Thrun S, Rodgers J, Scape DJ (2005) Shape completion and animation of people. ACM Trans Graph 3:408–416
4. Au OK-C, Tai C-L, Chu H-K, Cohen-Or D, Lee T-Y (2008) Skeleton extraction by mesh contraction. ACM Trans Graph 3:44
5. Bandouch J, Engstler F, Beetz M (2008) Accurate human motion capture using an ergonomics-based anthropometric human model. In: Articulated motion and deformable objects. Springer 248–258
6. Bandouch J, Jenkins OC, Beetz M (2012) A self-training approach for visual tracking and recognition of complex human activity patterns. Int J Comput Vis 99(2):166–189
7. Baran I, Popović J (2007) Automatic rigging and animation of 3d characters. ACM Trans Graph 3:72
8. Berretti S, Werghi N, Del Bimbo A, Pala P (2013) Matching 3D face scans using interest points and local histogram descriptors. Comput Graph 37(5):509–525
9. Bronstein MM, Kokkinos I (2010) Scale-invariant heat kernel signatures for non-rigid shape recognition. IEEE Conf Comput Vis Pattern Recognit. IEEE 1704–1711
10. Chaaraoui AA, Climent-Pérez P, Flórez-Revuelta F (2013) Silhouette-based human action recognition using sequences of key poses. Pattern Recogn Lett 34(15):1799–1807. doi:10.1016/j.patrec.2013.01.021
11. Coifman RR, Lafon S (2006) Diffusion maps. Appl Comput Harmon Anal 21(1):5–30. doi:10.1016/j.acha.2006.04.006
12. Cornea ND, Demirci MF, Silver D, Shokoufandeh A, Dickinson SJ, Kantor PB (2005) 3D object retrieval using many-to-many matching of curve skeletons. In: Shape modeling and applications. Int Conf IEEE 366–371
13. Cornea ND, Silver D, Min P (2007) Curve-skeleton properties, applications, and algorithms. IEEE Trans Vis Comput Graph 13(3):530–548. doi:10.1109/TVCG.2007.1002
14. Deutscher J, Reid I (2005) Articulated body motion capture by stochastic search. Int J Comput Vis 61(2):185–205
15. Dey TK, Li K, Luo C, Ranjan P, Safa I, Wang Y (2010) Persistent heat signature for pose-oblivious matching of incomplete models. Comput Graphics Forum 5:1545–1554, Wiley Online Library
16. Eldar Y, Lindenbaum M, Porat M, Zeevi YY (1997) The farthest point strategy for progressive image sampling. IEEE Trans Image Proc Publ IEEE Signal Proc Soc 6(9):1305–1315. doi:10.1109/83.623193
17. Golovinskiy A, Funkhouser T (2009) Consistent segmentation of 3D models. Comput Graph 33(3):262–269
18. Hahmann S, Belyaev A, Busé L, Elber G, Mourrain B, Rössl C (2008) Shape interrogation. In: Shape analysis and structuring. Springer 1–51
19. Jia T, Wang K, Wu Z, Zhao J, Xu P, Liu C, Zhou M (2014) Isometric shape matching based on the geodesic structure and minimum cost flow. In: Cyberworlds (CW), 2014 International Conference on. 122–129. doi:10.1109/CW.2014.25
20. Jiang W, Xu K, Cheng Z-Q, Martin RR, Dang G (2013) Curve skeleton extraction by coupled graph contraction and surface clustering. Graph Model 75(3):137–148
21. Kimmel R, Sethian JA (1998) Computing geodesic paths on manifolds. Proc Natl Acad Sci 95(15):8431–8435
22. Kin-Chung Au O, Tai CL, Cohen-Or D, Zheng Y, Fu H (2010) Electors voting for fast automatic shape correspondence. Comput Graphics Forum 2:645–654, Wiley Online Library
23. Kravtsov D, Fryazinov O, Adzhiev V, Pasko A, Comninos P (2014) Controlled metamorphosis between skeleton-driven animated polyhedral meshes of arbitrary topologies. Comput Graphics Forum 1:64–72, Wiley Online Library
24. Lafon SS (2004) Diffusion maps and geometric harmonics. Ph.D. Dissertation, Yale University, May 2004
25. Le BH, Deng Z (2014) Robust and accurate skeletal rigging from mesh sequences. ACM Trans Graph 33(4):84
26. Li X, Woon TW, Tan TS, Huang Z Decomposing polygon meshes for interactive applications. In: Proceedings of the 2001 symposium on Interactive 3D graphics, 2001. ACM 35–42
27. Lian Z, Godil A, Bustos B, Daoudi M, Hermans J, Kawamura S, Kurita Y, Lavoué G, Van Nguyen H, Ohbuchi R (2011) SHREC'11 track: shape retrieval on non-rigid 3D watertight meshes. 3DOR 11:79–88
28. Ma J, Choi S (2014) Kinematic skeleton extraction from 3D articulated models. Comput Aided Des 46:221–226
29. Mahmoudi M, Sapiro G (2009) Three-dimensional point cloud recognition via distributions of geometric distances. Graph Model 71(1):22–31

30. Ovsjanikov M, Mérigot Q, Mémoli F, Guibas L (2010) One point isometric matching with the heat kernel. Comput Graphics Forum 29(5):1555–1564. doi:10.1111/j.1467-8659.2010.01764.x
31. Reuter M, Wolter F-E, Peinecke N (2006) Laplace–Beltrami spectra as 'Shape-DNA'of surfaces and solids. Comput Aided Des 38(4):342–366
32. Sahillioğlu Y, Yemez Y (2010) 3d shape correspondence by isometry-driven greedy optimization. IEEE Conf Comput Vis Pattern Recognit. IEEE 453–458
33. Sahillioğlu Y, Yemez Y (2011) Coarse-to-fine combinatorial matching for dense isometric shape correspondence. Comput Graphics Forum 30(5):1461–1470. doi:10.1111/j.1467-8659.2011.02020.x
34. Sahillioğlu Y, Yemez Y (2012) Minimum-distortion isometric shape correspondence using EM algorithm. IEEE Trans Pattern Anal Mach Intell 34(11):2203–2215. doi:10.1109/TPAMI.2012.26
35. Sahillioğlu Y, Yemez Y (2014) Partial 3-D correspondence from shape extremities. Comput Graphics Forum 33(6):63–76. doi:10.1111/cgf.12278
36. Schwarz LA, Mkhitaryan A, Mateus D, Navab N (2012) Human skeleton tracking from depth data using geodesic distances and optical flow. Image Vis Comput 30(3):217–226. doi:10.1016/j.imavis.2011.12.001
37. Shapira L, Shamir A, Cohen-Or D (2008) Consistent mesh partitioning and skeletonisation using the shape diameter function. Vis Comput 24(4):249–259. doi:10.1007/s00371-007-0197-5
38. Smeets D, Hermans J, Vandermeulen D, Suetens P (2012) Isometric deformation invariant 3D shape recognition. Pattern Recog 45(7):2817–2831. doi:10.1016/j.patcog.2012.01.020
39. Sun J, Ovsjanikov M, Guibas LA (2009) Concise and provably informative multi-scale signature based on heat diffusion. Comput Graphics Forum 5:1383–1392, Wiley Online Library
40. Tagliasacchi A, Zhang H, Cohen-Or D (2009) Curve skeleton extraction from incomplete point cloud. ACM Trans Graph 3:71
41. Tam GK, Cheng ZQ, Lai YK, Langbein FC, Liu Y, Marshall D, Martin RR, Sun XF, Rosin PL (2013) Registration of 3D point clouds and meshes: a survey from rigid to nonrigid. IEEE Trans Vis Comput Graph 19(7):1199–1217. doi:10.1109/TVCG.2012.310
42. Vlasic D, Baran I, Matusik W, Popović J (2008) Articulated mesh animation from multi-view silhouettes. ACM Trans Graph 27(3):97
43. Wang K, Wu Z, Xu P, Zhao J, Jia T, Shui W, Ali S, Zhou M (2014) Scale-invariant heat kernel mapping. Cyberworlds, Int Conf 114–121. doi:10.1109/CW.2014.24
44. Wong S-K, Yang J-A, Ho T-C, Chen Y-C, Chuang J-H (2012) A skeleton-based approach for shape correspondence. Paper presented at the Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, Singapore, Singapore
45. Yan HB, Hu SM, Martin RR, Yang YL (2008) Shape deformation using a skeleton to drive simplex transformations. IEEE Trans Vis Comput Graph 14(3):693–706. doi:10.1109/Tvcg.2008.28
46. Yang Y, Günther D, Wuhrer S, Brunton A, Ivrissimtzis I, Seidel H-P, Weinkauf T (2012) Correspondences of persistent feature points on near-isometric surfaces. In: Computer Vision–ECCV 2012. Workshops and Demonstrations. Springer 102–112
47. Yianilos PN (1993) Data structures and algorithms for nearest neighbor search in general metric spaces. Paper presented at the Proceedings of the fourth annual ACM-SIAM Symposium on Discrete algorithms, Austin, Texas, USA
48. Zhang MQ, Li F, Wang XC, Wu ZK, Xin SQ, Lui LM, Shi L, Wang DF, He Y (2014) Automatic registration of vestibular systems with exact landmark correspondence. Graph Model 76:532–541. doi:10.1016/j.gmod.2014.04.010
49. Zhang W, Shang L, Chan AB (2014) A robust likelihood function for 3D human pose tracking
50. Zhang H, Sheffer A, Cohen-or D, Zhou Q, Kaick OV, Tagliasacchi A (2008) Deformation-driven shape correspondence. Comput Graphics Forum 27(5):1431–1439. doi:10.1111/j.1467-8659.2008.01283.x
51. Zheng Q, Sharf A, Tagliasacchi A, Chen B, Zhang H, Sheffer A, Cohen-Or D (2010) Consensus skeleton for Non-rigid space-time registration. Comput Graphics Forum 29(2):635–644. doi:10.1111/j.1467-8659.2009.01633.x

Kang Wang is a Ph.D. candidate in the Institute of Virtual Reality and Visualization, Beijing Normal University, Beijing, China. He received his B.S. degree from School of Computer Science, Qufu Normal University in 2010. His research interests include computer graphics, geometric modeling and differential geometry processing and analysis.

Abdul Razzaq received his M.S(Computer Science) degree from the Islamia University of Bahawalpur, Pakistan. He received M.Sc(Computer Science) degree from the University of Agriculture, Faisalabad, Pakistan. He is currently a Ph.D. candidate in the Institute of Virtual Reality and Visualization, Beijing Normal University, Beijing, China. His research interests include computer graphics, 3D Human motion, and Comptuer Animation.

Zhongke Wu is full professor and a Ph.D. student supervisor in the College of Information Science and Technology, Beijing Normal University, China. Prior to joining in BNU, he worked in the Nanyang Technological University (NTU), Singapore, Institute National de Recherche en Informatique et en Automatique (INRIA) in France, Institute of High Performance Computing (IHPC) and Institute of Software, Chinese Academy of Sciences in China from 1995 to 2006. His research interests are image processing and computer graphics.

Feng Tian received the PhD degree from Xi'an Jiaotong University, China. Currently he is an Associate Professor Bournemouth University (BU), United Kingdom. He was an Assistant Professor in Nanyang Technological University in Singapore before joining BU in 2009. His current research interests include computer graphics, computer animation, augmented reality, image processing.

Sajid Ali received his PhD degree from Beijing Normal University, Beijing, China. He is a Faculty member in department of computer science at University of Education, Lahore, Pakistan. Currently, he is a Postdoctoral candidate in department of computer Science, Beijing Normal University, China. His research interests are computer graphics, mocap 3D data, computer vision and computer animation and Computer Network.

Taorui Jia will received the master's degree from Beijing Normal University, China. He was an undergraduate student in Jilin University from 2003 to 2007. His current research interests include computer graphics, shape correspondence.

Xingce Wang received her Ph.D.degree in computer science from Harbin Engineering University in 2005. And she finished postdoctoral research in Beijing Normal University in 2010. Now she is an associate professor in college of information science and technology, Beijing Normal University. Her current research interests include computer graphic, medical imaging, artificial intelligence and Machine learning.

Mingquan Zhou is a professor and doctoral supervisor at the College of Information Science and Technology, Beijing Normal University and director of Beijing Key Laboratory of Digital Preservation and Virtual Reality for Cultural Heritage. His research interests are computer graphic and virtual reality.