Convex Hull Aided Registration Method (CHARM)

Jingfan Fan, Jian Yang, Yitian Zhao, Danni Ai, Yonghuai Liu, Ge Wang, and Yongtian Wang

Abstract—Non-rigid registration finds many applications such as photogrammetry, motion tracking, model retrieval, and object recognition. In this paper we propose a novel convex hull aided registration method (CHARM) to match two point sets subject to a non-rigid transformation. First, two convex hulls are extracted from the source and target respectively. Then, all points of the point sets are projected onto the reference plane through each triangular facet of the hulls. From these projections, invariant features are extracted and matched optimally. The matched feature point pairs are mapped back onto the triangular facets of the convex hulls to remove outliers that are outside any relevant triangular facet. The rigid transformation from the source to the target is robustly estimated by the random sample consensus (RANSAC) scheme through minimizing the distance between the matched feature point pairs. Finally, these feature points are utilized as the control points to achieve non-rigid deformation in the form of thin-plate spline of the entire source point set towards the target one. The experimental results based on both synthetic and real data show that the proposed algorithm outperforms several state-of-the-art ones with respect to sampling, rotational angle, and data noise. In addition, the proposed CHARM algorithm also shows higher computational efficiency compared to these methods.

主要是用于寻找 correspondence

Index Terms—Point set, convex hull, parallel projection, invariant feature, non-rigid registration

1 Introduction

The rapid development of laser scanning devices in recent years has led to the possibility of acquiring high-quality 3D models of large-scale scenes or interesting objects within minutes. The generation of 3D models and the corresponding manipulation techniques have attracted increasing attention and are being studied worldwide. Registration and alignment are usually the most basic step in the quantitative analysis of the 3D models. A 3D model can be generally represented as a mesh or as a point cloud. In this paper, we assume that it is represented as a point set, due to its wide applicability. The registration technique aims to determine the most appropriate geometric warping among different models or models with various representations. The registration of 3D models finds numerous applications in photogrammetry, motion tracking, model retrieval, and object recognition.

Registration techniques can be classified into two main categories based on transformation type: rigid and non-rigid registration. Rigid registration mainly intends to estimate scaling, translation, rotation, and/or shearing that bring the two point sets into the best possible alignment; whereas non-rigid registration is designed to estimate local or global non-regular transformations. Given that non-rigid deformation is generally irregularly distributed over entire models, performing first rigid registration and then non-rigid registration are usually helpful for achieving more accurate results. If the initial transformation between two models is considerably large, then the iterative process of non-rigid registration may fall into a local minimum and fail to accurately estimate the underlying transformation [1]. Rigid transformation can be generally represented using the affine transformation matrix, whereas non-rigid transformation can be represented by the thin-plate spline (TPS) and the B-spline [2], [3]. For the registration technique, the definition of dissimilarity/error between models is one of the most important aspects in determining the registration accuracy and efficiency. The sum of the Euclidean distances between the corresponding points between the models is the most widely used dissimilarity measure for the task. However, random noise generated by the acquisition devices usually affect the registration robustness of the techniques. At present, research on non-rigid registration techniques mainly focuses on feature extraction, feature matching, and structure representation to improve the robustness of the non-rigid registration of overlapping models or point clouds.

We address these challenges by introducing a novel Convex Hull Aided Registration Method (CHARM) for the registration of overlapping discrete 3D point sets. In this paper, the registration problem is first transformed into a feature extraction and matching problem using the following steps: (i) each point set is represented by a convex hull with many triangles, (ii) all the points are projected onto the plane through a triangle in the convex hull, yielding a set of coplanar points, (iii) these coplanar points are transformed into

- *J. Fan, J. Yang, Y. Zhao, D. Ai, and Y. Wang are with the Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China. E-mail: jyang@bit.edu.cn.*
- *Y. Liu is with Department of Computer Science, Aberystwyth University, Ceredigion SY23 3DB, United Kingdom.*
- *G. Wang is with Department of Biomedical Engineering, Rensselaer Polytechnic Institute, NY 12180.*

Manuscript received 2 Mar. 2016; revised 21 July 2016; accepted 15 Aug. 2016. Date of publication 31 Aug. 2016; date of current version 2 Aug. 2017. Recommended for acceptance by M. Botsch.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TVCG.2016.2602858

the xy plane and shifted to the centroid of that triangle, (iv) a virtual grey image is generated where the grey level or intensity represents the sum of the inverse depth of the points in a pixel, (v) the scale-invariant feature transform (SIFT) features are extracted from and matched between the virtual images of the source and target point sets. The correspondences are refined by back-projecting them onto the convex hulls of the 3D models and rejecting those outside relevant triangles. (vi) the refined correspondences are finally utilized to estimate the rigid transformation of the two models based on the random sample consensus (RANSAC) [4]. (vii) after the source model has been subject to the estimated rigid transformation, the refined point matches are utilized as control points to undergo the non-rigid deformation of the source model towards the target one. The main contributions of this paper are summarized as follows.

- We propose a novel CHARM algorithm for the registration of two 3D models with rigid and non-rigid deformations. It is highly robust to noise, sampling, partial overlap, and non-rigid deformation because it is based on a two-step procedure for rigid coarse registration in the first step and non-rigid fine registration in the second.
- We propose a projection invariant feature CH-PIF over discrete point sets that does not require the mesh or texture information of the model. This feature can be applied for model or point set registration, motion tracking, and feature analysis, among others.
- The proposed method is computationally efficient because the number of triangles on the convex hull is significantly smaller than that in the original model.

The rest of this paper is organized as follows. In Section 2, we review the existing main 3D model registration methods. In Section 3, we detail our registration method. In Section 4, we validate our proposed algorithm and compare it with several state-of-the-art ones. Finally, we conclude the paper in Section 5.

2 Related Works

Numerous methods have been developed in the past two decades for the registration of 3D overlapping point sets and mesh models. The most popular method is the iterative closest point (ICP) algorithm, which was proposed by Besl and Mckay [1] in 1992. In this algorithm, the dissimilarity metric of two models is defined as the sum of the euclidean distances between all the closest point correspondences in the two models. ICP achieves an optimal registration by iteratively refining the transformation for the minimization of the dissimilarity metric between the two models in the least squares sense. Thus, ICP is simple and effective for models subject to relatively small transformations.

However, ICP is highly dependent on the initialization of the underlying transformation and the geometries that the two models represent. It can be easily trapped in a local minimum. The expectation maximization (EM)-ICP algorithm [5], [6] extends the ICP algorithm by applying EM principles to first estimate the weights of tentative matches, and then estimate the underlying transformation in the weighted least squares (WLS) sense and repeats these two steps until convergence. EM-ICP provides better robustness and higher accuracy but less computational efficiency than the original ICP algorithm. The anisotropic and inhomogeneous localization errors of EM-ICP are generally assumed as the main causes for inaccurate registration. This assumption led to the development of an anisotropic ICP (A-ICP) algorithm [7], which integrated anisotropic weights for tentative correspondence refinement. Penalizing Closest Point Sharing (PCPS) method [8] assumes that those points in one shape that select the same closest point in another as their tentative correspondents should be penalized, and realizes overlapping free form shape registration in the WLS sense. Gaussian mixture models [9] assume that a model can be represented statistically as the weighted sum of the Gaussian responses of all the points in 3D space. Model registration can then be reformulated as an alignment of different Gaussian mixtures by minimizing their discrepancies. According to the GMM registration framework, coherent point drift (CPD) [10] forces the GMM centroids to move coherently as a group at similar velocities. The group preserves the topological structure of the point sets. In order to achieve rigid registration, the CPD algorithm reparametrizes the GMM centroids with rigid transformations. Non-rigid registration is achieved by regularizing the displacement field and then using variational calculus to estimate the optimal underlying transformation.

After having constructed effective dissimilarity/error measurements, the aforementioned methods generally perform a global optimization on the entire models for the best transformation parameters. Registration accuracy, which is usually affected by the iteration number, is gradually refined during the iterative process. If the initial transformation is inaccurate, then these methods can easily be trapped in a local minimum. A variety of algorithms aim to register models via feature extraction and matching because of their wide applicability to various datasets. The four-point congruent set (4PCS) approach is proposed by extracting a coplanar four-point set and calculating and matching their cross ratios as the geometric feature to optimize the rigid transformation [11]. However, 4PCS is an arbitrary coplanar point set whose definition may be challenging and the combination of all the points may be large. The spin image-based algorithm [12] generates a descriptor from the 2D histogram of neighboring points spinned along the normal at each surface point. The spin descriptor is then utilized to match the 3D models because it exhibits rotation invariance within the 2D images. However, the registration accuracy of these methods is usually affected by the geometry and point distribution of the models. If a significant difference in the point distribution exists between the two models, then the coarse transformation that was estimated from the point matches may fail to initialize ICP, for example, for refinement. The SIFT descriptor has been shown to outperform other descriptors, such as shape context [13], [14] and geometric descriptors [15], for invariant feature description, which leads to the 3D SIFT descriptor [16] for action recognition over video or 3D imagery. The 3D SIFT descriptor calculates the sub-histograms of oriented points to encode time and space information, which proves to be highly powerful for the spatiotemporal information representation of video or 3D data sets. The viewpoint-invariant patch [17] is a

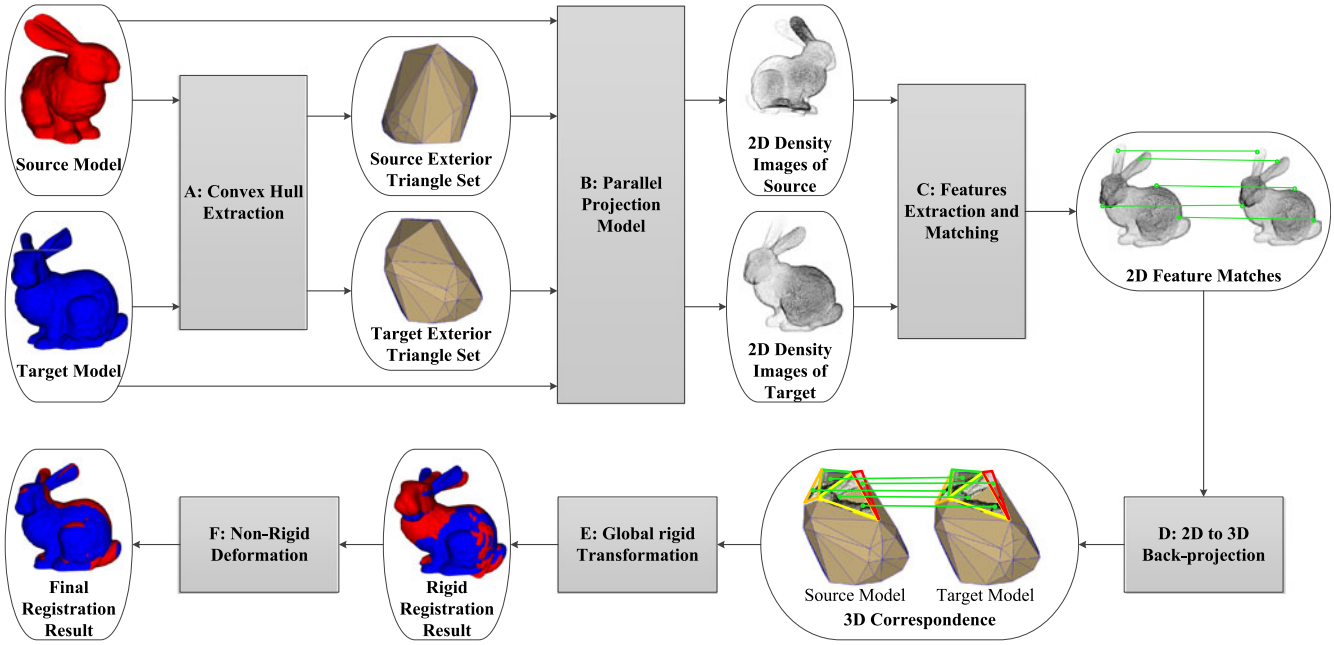

Fig. 1. The flow chart of the proposed CHARM method for the registration of 3D overlapping point sets.

viewpoint-independent local feature for the robust 3D scene alignment and large-scale scene reconstruction. It includes 3D position, patch size, surface normal, texture's dominant orientation and the SIFT descriptor and has an advantage in that the underlying transformation can be estimated from a single point match. MeshDOG [18] is a generalized Gaussian operator that obtains the local extreme of the Laplacian responses of the geometric and photometric properties on the surface of the models. The aforementioned methods improve considerably the effectiveness of 3D model registration. However, feature extraction and matching-based methods achieve model registration by utilizing the geometric information and/or texture information on the surfaces of the 3D models; hence, they generally cannot be utilized to register discrete point sets without mesh information.

Various techniques are also developed to align pairs of models and integrate numerous scans non-rigidly into a complete model. The common non-rigid registration algorithm [19], [20] is based on feature correspondences that are computed using a hierarchical ICP method; feature correspondences usually result in a coarse alignment. The dynamic geometry method [21] attempts to register deforming objects without computing correspondences; alignment is computed by exploiting the underlying smoothness of the space time surface, which results in the property where inter-frame motions remain coherent. Non-rigid registration for partial matching can merge different partial surface data into a complete digital representation [22], [23]. This method is not restricted to part-in-whole matching because of the introduction of the non-linear deformation model instead of matching by explicit prior correspondences. Numerous methods have focused on tracking and reconstructing human motion, including the probabilistic deformable model [24], the dense correspondence method [25], and Markov random field (MRF)-based approaches [26], [27]. Canonical Distortion Coefficients (CDCs) deformation model [26] is proposed for characterizing the deformation of every point on a surface using the distortions along its two principal directions. Then, an objective function is built in the framework of MRF for the search of tentative correspondences. This function minimizes the difference between their feature descriptors and penalizes such a linear transformation between two corresponding triangles whose eigenvalues of the product between its transpose and itself are not in certain ranges. Based on the MRF model, a convex optimization method is proposed in [27] for casting isometric embedding as MRF optimization, and applying linear programming relaxations for efficient global optimization of the correspondences between two given surfaces. These methods implement non-rigid registration by minimizing the difference between the descriptors and (semi-)invariant features among the potential correspondences between the given overlapping 3D surfaces.

3 Our Method

The flowchart of the proposed CHARM method is illustrated in Fig. 1. In this method, the convex hull of the model is first created by extracting the smallest convex set that contains the entire model. A set of 2D images of the model is then generated by parallel projecting the entire model onto each plane that contains a triangle on the convex hull. The salient features of the 2D projection images are effectively extracted and matched because such images effectively preserve the shape and intensity distribution of the entire model. Thereafter, the corresponding features are back-projected onto the surface of the convex hull, and the corresponding relationships between the 3D and 2D features are constructed. Finally, the rigid and non-rigid transformations between these two models can be estimated.

Let X and Y represent the source and target models as point sets with usually different and finite sizes, respectively, where $X, Y \subset \mathbb{R}^3$. The objective of the registration procedure is to estimate the transformation from X to Y, which yields the best possible alignment between the transformed X and target Y models. The transformation may be either rigid or non-rigid.

3.1 Convex Hull Extraction

The principle of parallel projection from 3D to 2D space is to project all the points onto a plane in a specified parallel direction. In each projected 2D image, the outline and point distribution of the entire model are effectively preserved. Hence, the transformation parameters between the source and target models can be effectively estimated by detecting and matching the invariant features in the projection images. To reduce computational time and preserve the convexity of the model, the convex hull is proposed to determine the projection direction of all the points.

The convex hull of a set of 3D points $P \in \mathbb{R}^3$ is the minimum convex combination of these points that contains all the points of the model, which typically represents the outline of the model. In our previous works [28], [29], it is demonstrated that the convex hull of the model exhibits the following properties: (1) The geometric structure of the entire model can be represented by a limited number of vertices and facets on the convex hull. Let n represent the number of points. The convex hull $H \in \mathbb{R}^3$ of the model is a convex polygon that encloses all the points in the space. The convex hull consists of a set of connected triangular facets $F \in \mathbb{R}^3$, and each facet is composed of three vertices $\{f_a, f_b, f_c\}$. The number h of vertices on the convex hull is at most equal to n; hence, the number of facets of the convex hull is at most equal to $2n - 4$ [30]. The number of facets of the convex hull is generally smaller than the extreme values [30]. (2) The convex hull for the model is unique and exhibits the property of invariance to similarity transformation. After translation, rotation, and scaling of the entire model, the extracted convex hull can be deduced to remain identical [15]. However, the extra noise introduced into the model may change the geometric structure of the convex hull. Hence, the key to maintaining the stability of the convex hull is to use effective denoising and smoothing methods [31] to reduce noise before extracting the convex hull. (3) Calculating the convex hull is simple and rapid, and can be achieved effectively by using numerous iterative algorithms [30]. In the current study, the output sensitive-based algorithm [32], [33] was employed to calculate the convex hull of the 3D models. The computational complexity of the convex hull is $O(n \log h)$, where h denotes the number of vertices on the corresponding convex hull. The convex hull is generally invariant to similarity transformations and robust against noise in the geometry. Non-rigid deformation may clearly change the structure of the model, but the geometric structure of the convex hull usually changes less significantly than the surface of the model itself. A local non-rigid transformation may introduce local changes into the facet orientations of the convex hull, which may result in fewer changes in projection views.

In conclusion, the convex hulls of the source and target models exhibit similar geometric structures because of their invariant and unique properties. Hence, the projection images on the plane of the triangular facet of the convex hull include all the points of the entire model. This situation also ensures that the projection images of the corresponding views have the same image intensity distributions. Matching between projection views can significantly improve registration effectiveness.

3.2 Parallel Projection Model

Once the convex hull of a model is obtained, all the triangular facets on the convex hull can be used as projection planes. This projection process is similar to observe a 3D object from a certain viewing direction. The proposed CHARM method utilizes parallel projection principles to project the entire point set onto the triangular facets on the convex hull, and all the projections can be achieved in limited time.

Given a triangular facet F with three vertices $\{f_a, f_b, f_c\}$ on the convex hull, the unit normal vector N of this facet can be calculated as

$$N = \frac{(f_b - f_a) \times (f_c - f_a)}{||(f_b - f_a) \times (f_c - f_a)||}, \tag{1}$$

where $\times$ denotes the cross product of two vectors and $||\cdot||$ denotes the euclidean length of a vector.

The distance from a point p to the plane of the triangular facet F can then be calculated as follows:

$$d = N \cdot (p - f_a), \tag{2}$$

where $\cdot$ denotes the dot product of two vectors.

Let P' represent the coplanar projection of P by the parallel projection above. If p' is a point in P' and a projection of p in P, the relationship between p' and p can then be represented as

$$p' = p - d \cdot N. \tag{3}$$

P' is a 3D coplanar point set. The 2D projection point set P_{2d} can be obtained by rotating P' to the $z = 0$ plane. This transformation $T_{3D->2D}$ of P' could be calculated from the normal direction N to the unit vector $(0, 0, 1)$. Afterwards, a 2D grayscale image that corresponds to the inverse depth map of P_{2d} can be generated; the intensity of a pixel is in inverse proportion to the distance d_i of each projected point p_i inside, and can be calculated as follows:

$$I(u, v) = \frac{\sum_{i=1}^{n} \begin{cases} \frac{1}{d_i + 1}, & p_i^{2d} = (u, v) \\ 0, & p_i^{2d} \neq (u, v) \end{cases}}{v_{\max}} \cdot 255, \tag{4}$$

where $I(u, v)$ represents the intensity of pixel (u, v) in the projection image and $v_{\max}$ represents the maximum value of inverse depth accumulation of all the projected points inside a pixel of the projection image. Based on the previous equation, the accumulated inverse depth of the projected points is normalized to the range of [0, 255]. (u, v) are the horizontal and vertical coordinates of the projection image. In this projection procedure, the projected model will be normalized to a bounding box with a size of 200 centred at the projection of the centroid of the triangular facet F. Therefore, the coordinates of the 2D projected points p_i^{2d} will be restricted within the range of $[-100, 100]$. The resolution of the projection image will affect the number of subsequent matches. On one hand, a high resolution leads to a sparse distribution of the projected points, as thus causes difficulty in feature extraction and matching. On the other hand, a low resolution may cause the structure information of the model to be lost. As a result, the resolution size was empirically chosen as 200×200 in the proposed method. Then, we can generate an independent projection image for each triangular facet on the convex hull. Fig. 2

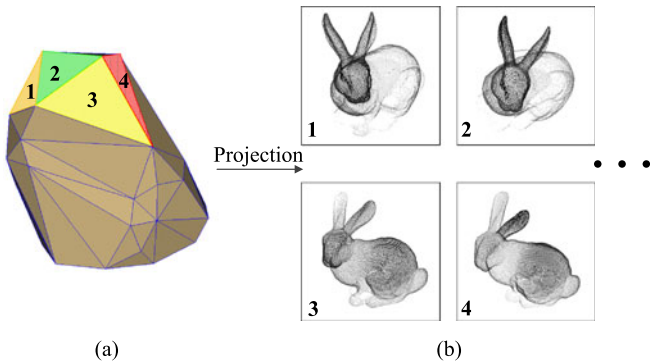

Fig. 2. Schematic diagram of the generated projection images. (a) Triangular facets on the convex hull. (b) The corresponding projection images.

shows four projection images of the bunny data on the corresponding triangular facets of the convex hull. The density and outline of the model can be clearly visualized in the generated images, and the intensity distribution of the projection image is unique for each triangular facet.

3.3 Features Extraction and Matching

Once the projection images from the triangular planes of the convex hull are obtained, they can be used to determine the corresponding point pairs between different models. Salient features are extracted from the images as illustrated in Fig. 3a, and a set of the best-matching candidates are then determined by minimizing the dissimilarity of the intensity or structure distributions of the images. Fig. 3b shows an example of salient feature matching in the projection images. Although the two images are obtained at different projection directions, a certain number of feature points can be identified and matched. Methods for feature extraction and matching have been intensively studied in the past two decades. One of the most extensively used algorithms is the scale invariant feature transform, which was proposed in 1999 by Lowe et al. [34]. The SIFT descriptor extracts a collection of features from a given image in the scale space, each of which is invariant to (slight) changes in translation, rotation, scaling, and intensity. The current study utilizes SIFT to extract and match feature points in the 2D projection images.

Let $D^X_{i,k}$ and $D^Y_{j,l}$ represent the SIFT descriptor of the kth feature point $F^X_{i,k}$ in the ith projection plane for the source model X and the SIFT descriptor of the lth feature point $F^Y_{j,l}$ in the jth projection plane for the target model Y, respectively. The dissimilarity between the ith and jth projection planes is defined as the sum of the minimum Euclidean distances between any feature descriptor in the ith projection image to any feature descriptor in the jth projection image

$$S(i,j) = \sum_k \min_l \left\| D^X_{i,k} - D^Y_{j,l} \right\|, \tag{5}$$

where $\left\| D^X_{i,k} - D^Y_{j,l} \right\|$ is the Euclidean distance of the descriptor vectors $D^X_{i,k}$ and $D^Y_{j,l}$. Then, each projection plane i in X will find the best matched plane $MI(i)$ by minimizing its dissimilarity $S(i,j)$ to all the projection planes j in Y

$$MI(i) = \arg\min_j S(i,j). \tag{6}$$

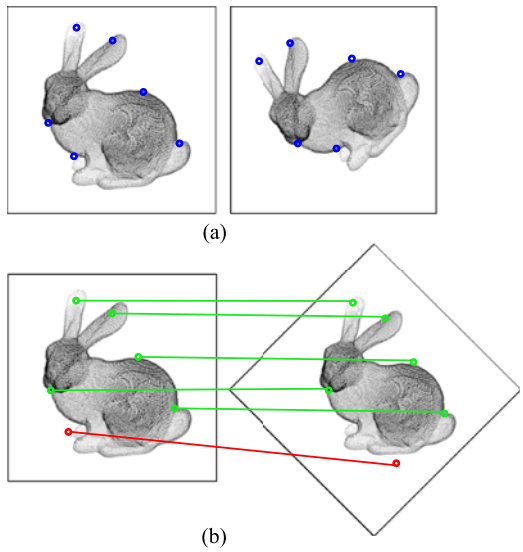

Fig. 3. Feature matching in the virtual projection images. (a) Feature extraction result of the ith image in the source model and the jth image in the target model. (b) Feature matching result of the ith image in the source model and the jth image in the target model.

The best matched point $F^Y_{MI(i),MP(k)}$ of $F^X_{i,k}$ in $F^Y_{MI(i),l}$ is determined through minimizing the difference between its descriptor $D^X_{i,k}$ and any $D^Y_{MI(i),l}$ as

$$\begin{aligned} MP(i,k) &= \arg\min_l \left\| D^X_{i,k} - D^Y_{MI(i),l} \right\|, \\ D(i,k) &= \left\| D^X_{i,k} - D^Y_{MI(i),MP(i,k)} \right\|. \end{aligned} \tag{7}$$

Finally, the best match $q_{c(k)}$ in Y of point p_k in X is determined as: $c(k) = \arg\min_i D(i,k)$.

3.4 2D to 3D Back-Projection

Using the SIFT descriptor, a set of 2D point matches can be established between the projection images of the source and target models. All 2D features are projection points from the 3D space and their original points on the plane of the corresponding convex hull can be restored by the following equation:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ 0 \end{bmatrix} \cdot T_{2D->3D}, \tag{8}$$

where (x,y,z) represent the coordinates of a 3D point p. $T_{2D->3D}$ denotes the transformation from a 2D point in the projection image to a 3D point on the convex hull. $T_{2D->3D}$ can be calculated as the inverse of $T_{3D->2D}$.

Using the back-projection procedure, all points can be mapped into the triangular planes on the convex hull. The back-projection points may lie outside the triangular facets, and the matching process should ensure that the two back-projection points lie inside the regions of the respective trianglar facets. Hence, this principle can be utilized to remove the outliers of the corresponding point pairs. By using this refinement method, the wrong matches introduced by repetitive structures, imaging noise, non-rigid warping, and different resolutions can be effectively removed.

Fig. 4 shows an example of the corresponding matching results on the convex hulls, where the green points and lines

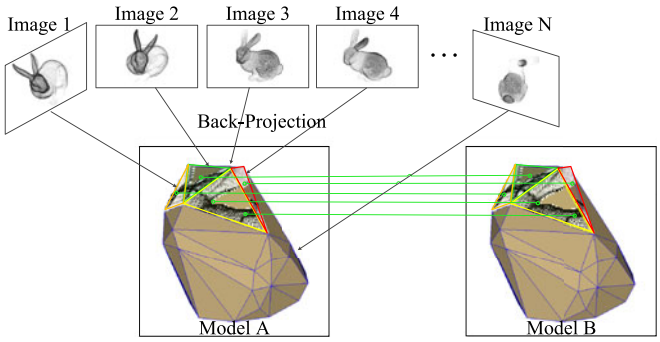

Fig. 4. Schematic diagram of feature point back-projection on the convex hull.

represent the matched feature points and their matching relationships, respectively. The matched feature points are relatively regularly distributed on the triangular facets of the convex hull that would help warp the models.

3.5 Similarity Alignment

After the obtaining of the matched points on the convex hulls, the underlying transformation is estimated by determining the least square minimization of the distances between these points [35]. Let $P = \{p_i\}_{i=1}^{n}$ and $Q = \{q_i\}_{i=1}^{n}$ represent the obtained corresponding point pairs of the source and target models, respectively. Their relationship can then be represented as

$$q_i \simeq T_{3D} \cdot p_i = s \cdot R \cdot p_i + t, \tag{9}$$

where T_{3D} is the transformation matrix between the source and target models. R and t are a 3×3 rotation matrix and a 3×1 translation vector, respectively, and s is the scaling factor. The rotation matrix R is parameterized with three rotation angles, namely, θ_x, θ_y, and θ_z around the x, y, and z axis respectively. In this case, T_{3D} is represented with seven parameters altogether: $(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z, s)$. To determine the least square solution to T_{3D}, the following objective function is built that minimizes the sum of the squares of the euclidean distances between the corresponding points:

$$T_{3D}^{*} = \arg\min_{R,t,s} \sum_{i=1}^{n} \|q_i - (s \cdot R \cdot p_i + t)\|^2, \tag{10}$$

where T_{3D}^{*} is the estimated optimal transformation matrix between the source and target models in the least squares sense [35]. However, if any wrong point match is present, then the estimated transformation matrix will not be accurate. Hence, we employed the random sample consensus scheme [4] for the robust estimation of the best transformation between the source and target models.

3.6 Non-Rigid Deformation

To estimate non-rigid transformation [19] between X and Y, the thin plate spline [2], [3] is used to represent the deformation constrained by the control points. In 1989, Bookstein first proposed the use of the TPS interpolation as a point-based elastic registration algorithm for medical images. One of the most important attributes of the TPS is its capability to decompose a space transformation into a global similarity transformation and a local non-affine warping component. Under the restriction of the corresponding points, the deformation parameters can be obtained. Hence, the deformation extent of all the points can be effectively manipulated under the combined effects of global and local transformations.

Suppose $C = \{c_i\}_{i=1}^{n_C}$ are the control points extracted from the source model X. Then, the radial basis function of a given point p to a control point c_i is $\varphi(\|p - c_i\|)$, where $\|p - c_i\|$ is the Euclidean distance between p and c_i. The radial basis function $\varphi(r)$ is defined as: $\varphi(r) = -r$. Thus, the point p with coordinates of (x, y, z) can be transformed towards q with coordinates of (x', y', z') using the following TPS equations $TPS(p) = (x', y', z')^T$:

$$\begin{aligned} x' &= \sum_{i=1}^{n} w_i \varphi(\|p - c_i\|) + w_{n+1}x + w_{n+2}y + w_{n+3}z + w_{n+4} \\ y' &= \sum_{i=1}^{n} w_i' \varphi(\|p - c_i\|) + w_{n+1}'x + w_{n+2}'y + w_{n+3}'z + w_{n+4}' \\ z' &= \sum_{i=1}^{n} w_i'' \varphi(\|p - c_i\|) + w_{n+1}''x + w_{n+2}''y + w_{n+3}''z + w_{n+4}'' \end{aligned} \tag{11}$$

where w_i, w_i' and w_i'' are the regression coefficients of the TPS transformation and the superscript T denotes the transpose of a vector. TPS-based model registration can be defined as the optimization for the best coefficients, from which the deformation of the source model X is warped towards the target model Y.

Conventional methods frequently select regular grid points that contain the entire model as the control points, which cannot guarantee the effective deformation of the model with sparse and irregular distributions. The density of the grid that divides the space of all the points is also difficult to determine. Hence, the point matches $P = \{p_i\}_{i=1}^{n}$ and $Q = \{q_i\}_{i=1}^{n}$ extracted from the back-projection procedures above are used as the control points in this paper. The regression coefficient w_i in the TPS-based deformation of the source model X towards the target model Y is estimated from the following objective function:

$$W = \arg\min_{w_i} \sum_{i=1}^{n} \|TPS(p_i) - q_i\|^2. \tag{12}$$

The back-projection points on the convex hull of the entire model are relatively regularly distributed around the model, and they can thus control the deformation of the surface effectively. Furthermore, the limited number of back-projection points considerably reduces the computational complexity. However, a small number of the control points on the convex hulls may limit their capabilities in representing the large scale local deformations inside the point sets.

4 Experimental Results

To evaluate the performance of the proposed algorithm, a set of experiments was designed and carried out based on both simulated data and real models acquired using the Kinect device (Microsoft Corporation, USA). The proposed CHARM method was compared with six state-of-the-art algorithms, including 4PCS [11], CPD [10], A-ICP [7], GMM [9], CDCs-MRF [26] and Li's method [23] for rigid and non-

rigid registration respectively, in terms of noise, sampling rate, and overlapping ratio. The synthetic transformation or noise were generated, unless stated otherwise, using the parameters randomly selected as follows: θ_x, θ_y, and θ_z over the range of $[30, 180$ degree$]$, t_x, t_y, and t_z over the range of $[100, 300]$, s over the range of $[0.8, 1.2]$, and $\sigma = 2$. The algorithm was implemented in C++ under the UNIX environment, and all the experiments were conducted on a relatively low-cost PC with 16 GB RAM and 3.4 GHz Intel CPU.

In this paper, the following parameters are used to measure the performance of the proposed method: the root-mean-square (RMS) error, inlier ratio, and failure or successful registration. The RMS error is calculated as the square root of the average of the squared Euclidean distances of all corresponding point pairs between the source and target point sets. Such correspondences are available for synthetic data, and are established as the closest points between the transformed source and target models for other data. Inliers are such correspondences whose errors are smaller than 5. Inlier ratio is defined as the ratio in percentage between the number of inliers and the total number of the correspondences established. A successful registration is defined as such a registration whose RMS error is smaller than 10. To quantify the registration accuracy, all the models were normalized into a bounding box with a size of 100 before registration. In this case, the registration error is a relative error measurement. Furthermore, two types of noises, including modeling and processing Gaussian noises, were added to the simulated data respectively in the experiments. The modeling Gaussian noise was used to simulate the noise in the process of data acquisition and was such white Gaussian noise that was randomly generated independent translations and added to the three coordinates of each point. The processing Gaussian noise was used to simulate the noise in the process of data processing and was such white Gaussian noise that was generated by randomly translating the three coordinates of the vertices on the convex hull. The magnitudes of the modeling and processing noises were controlled by their standard deviations of σ_G and σ_{TPS}. As the translations to the three coordinates of the points in space were independent and controlled by the standard deviation of Gaussian distribution, the resulting perturbations of the two models can be effectively controlled.

4.1 Feature Matching Performance

The accuracy of feature matching is the most important aspect in demonstrating the effectiveness of the proposed CHARM registration method. In this section, the accuracy of feature extraction and matching is investigated. Fig. 5 shows the matching results of the proposed CHARM method using different shapes for the model projection. The bunny and lion data sets were downloaded from the Stanford [36] and Aim [37] data repositories respectively. Figs. 5 A1 and 5B1 show the initial target data sets, whereas Figs. 5 A2 and 5B2 show the data sets that were randomly transformed from Figs. 5A1 and 5B1 respectively by a similarity transformation. Both of these two models were corrupted by the modeling Gaussian noise with a standard deviation of $\sigma_G = 2.0$. Figs. 5A3 and 5B3 illustrate the matching relationships of 63 and 38 tentative point matches (TPMs) on the extracted convex hulls. All the connecting green lines are parallel, and they indicate that these TPMs are correct. In contrast, Figs. 5A6 and 5B6 illustrate the established TPMs on the spheres, and the numbers of these TPMs are 37 and 29 respectively, which are fewer than those on the convex hulls. Because the convex hull of a point cloud will be more accurate in enclosing the point cloud than the bounding sphere, the projection images on the convex hulls can encode more information of the models and thus enable more TPMs to be found, and achieve better feature matching results. The registration results demonstrate the proposed CHARM method is highly robust to the noise. The smaller registration errors in the color error maps indicate that the convex hull is more suitable to be used as the projection planes than a sphere.

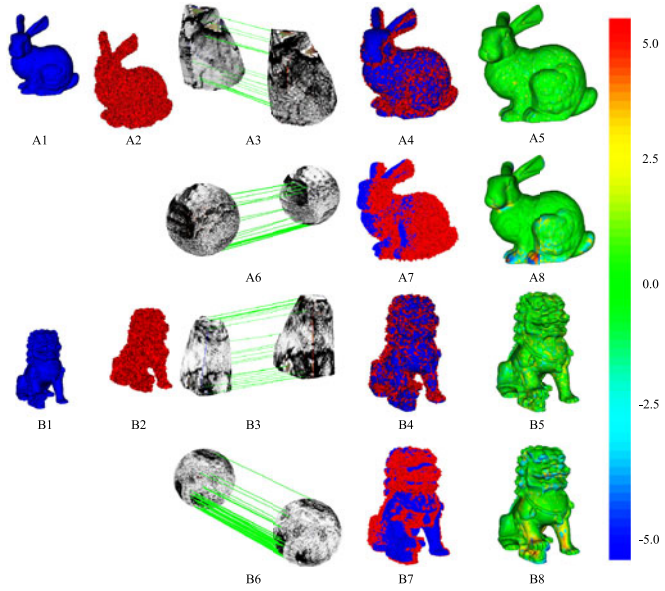

Fig. 5. Results of the proposed CHARM method with different shapes used for point projection. (A1 and B1) Target models; (A2 and B2) models randomly transformed and corrupted by the modeling Gaussian noise; (A3 and B3) feature matching on the convex hulls; (A6 and B6) feature matching on the spheres; (A4, A7, B4 and B7) results of the rigid registration; (A5, A8, B5 and B8) error maps of the rigid registration results.

Table 1 shows the intermediate results for the registration of the bunny and lion data sets. It can be seen that the more the noise added, the fewer the established TPMs, the fewer the inliers, the smaller the inlier ratio, and the larger the registration RMS error. When the standard deviation of the modeling Gaussian noise increased from 1.0 to 4.0, the number of 2D matches in the projection images decreased from 2,387 to 1,253 for the bunny data and from 1,847 to 1,064 for the lion data. Once the 2D matches were back-projected into the 3D space, the number of the TPMs considerably decreases due to the removal of the wrong ones. With the noise level increasing from 1.0 to 4.0, the inlier ratio for the bunny data is 100, 98.4, and 89.2 percent; whereas that for the lion data is 98.1, 92.1, and 87.5 percent respectively. It is worth noting that by using the convex hull as projected planes, the number of inliers and the inlier ratio are much larger than those using a sphere. These remarkable results demonstrate that the convex hull is more suitable for point projection and the proposed CHARM method is robust to different levels of noise.

TABLE 1
Statistics of Feature Matching Results on the Bunny and Lion Data Sets

Scenes	Convex Hull						Sphere					
	Bunny (Fig.6. a)			Lion (Fig.6. b)			Bunny (Fig.6. a)			Lion (Fig.6. b)		
Gaussian Noise	$\sigma=1.0$	$\sigma=2.0$	$\sigma=4.0$	$\sigma=1.0$	$\sigma=2.0$	$\sigma=4.0$	$\sigma=1.0$	$\sigma=2.0$	$\sigma=4.0$	$\sigma=1.0$	$\sigma=2.0$	$\sigma=4.0$
2D Matches	2387	2030	1253	1847	1687	1064	1274	975	731	1076	943	864
3D Matches	84	63	28	52	38	24	43	37	15	37	29	10
Inliers (error<5)	84	62	25	51	35	21	38	31	10	33	24	4
Inlier Ratio	100%	98.40%	89.20%	98.10%	92.10%	87.50%	87.50%	83.78%	66.67%	89.19%	82.76%	40.00%
Error of Matches after Registration	1.347	1.928	3.124	4.212	3.987	6.413	2.171	2.764	4.521	5.134	5.328	13.147
Successful	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N

4.2 Rigid Registration Results

The registration of free form shapes is usually challenging due to a number of factors: large underlying transformation, different resolutions and point distributions, and small and partial overlap. Hence, in this section we investigate the convergence and robustness of the proposed method to these factors.

Fig. 6 shows the registration results of the car shapes subject to different transformations. The blue car in the first row is the target model, whereas the red cars are the transformations of the target model with the specified rotational angles between 45 and 180 degree. These transformed models were used as source models. The modeling Gaussian noise with a standard deviation of 1 was added to each model. The second row shows the extracted matches on the convex hulls. The numbers of the extracted TPMs are 73, 69, 74, and 70 respectively; no significant difference clearly exists among these numbers. The third row shows the final registration results and the fourth row shows the corresponding error maps. The mean and the largest registration errors of all the point matches are 0.13 and 0.82, respectively. Thus, the CHARM method is highly stable for the registration of data sets subject to random transformations with large rotations.

The registration accuracy of the CHARM method is dependent on the feature matching results of the projection images, whereas the grayscale distribution of the projection image is determined by the depth distribution of the points in the model. Fig. 7 shows the performance of the CHARM method for the registration of the elephant models subject to different rates of decimation [38]. The blue model in the first row is the target model with 5,000 points; the red models in the same row correspond to those with fewer sampled points from 4,000 to 1,000. The second row corresponds to the matching relationship on the surfaces of the convex hulls of the target and transformed source models. The identified point matches decrease with the decrease in the sampled points. For the case of 1,000 sampled points, the number of the identified point matches is clearly limited, which indicates the difficulty in determining the corresponding point pairs. The third row shows the registration results of all the sampled points from 4,000 to 1,000. The CHARM method is highly effective for the cases of 4,000, 3,000 and 2,000 sampled points, in which the registration errors are less than 1 for more than 90 percent of the sampled points. This means that the proposed CHARM method is effective for the registration of the overlapping point sets with a significant difference in point density. However, for the case of 1,000 sampled points, they can hardly represent the true geometry and details of the elephant, leading to large global and local registration errors over the entire model. This is probably as expected. If the density of the points is significantly low and the density differences are considerably large, then the registration accuracy can be improved by interpolating the neighboring points in the two models.

The registration accuracy of the proposed CHARM method is also affected by the size of the overlap between the two models to be registered. Fig. 8 shows the results of the proposed CHARM method for the registration of the fertility model with different sizes of overlap. The blue

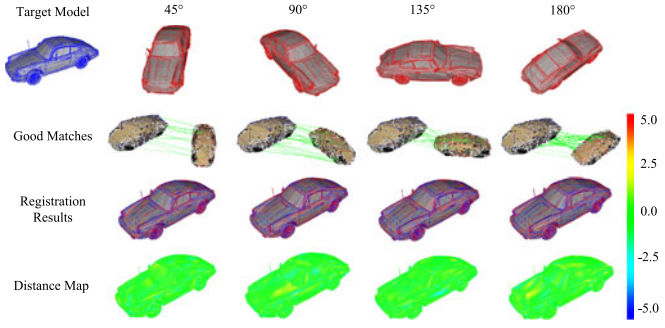

Fig. 6. Results of the proposed CHARM method for the registration of the car shape subject to different transformations with different rotation angles. First row: The target model and the 4 transformed models with different rotational angles. Second row: Matching relationships of points on the convex hulls. Third row: Rigid registration results. Fourth row: Error maps of the rigid registration results.

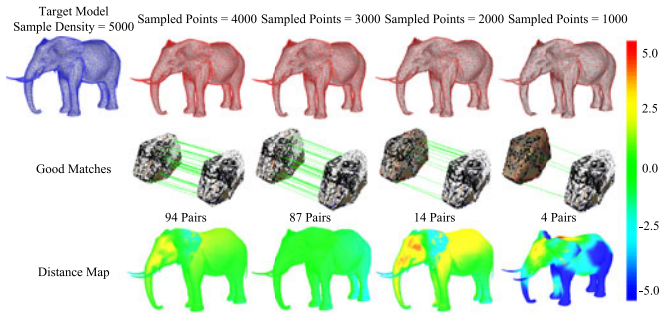

Fig. 7. Results of the proposed CHARM method for the registration of the elephant models with different sampling rates. First row: The target model and its corresponding models with different sampling rates. Second row: Feature matching on the convex hulls. Third row: Error maps of the final registration results.

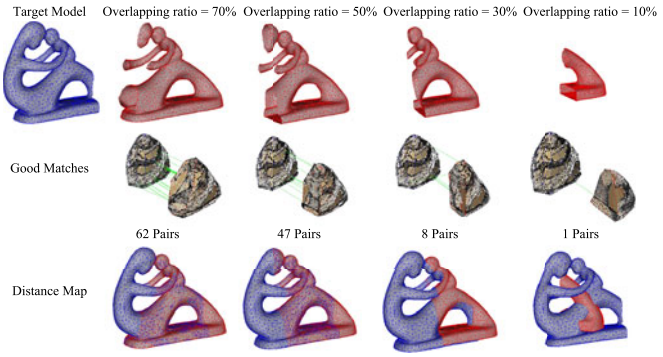

Fig. 8. Results of the proposed CHARM method for the registration of the fertility models with different sizes of overlap. First row: The target model and its corresponding models with different sizes of overlap. Second row: Feature matching on the convex hulls. Third row: Final registration results.

model in the first row is the target model; the red models in the same row correspond to those with an overlapping ratio of 70 to 10 percent with the target model. The second row corresponds to the matching relationship on the surfaces of the convex hulls of the target and transformed source models. The found point matches decrease rapidly with the decrease in overlapping ratio. For the overlapping ratios of 70 and 50 percent, sufficient TPMs were found. For the overlapping ratio of 30 percent, the number of the found TPMs is small. And for the overlapping ratio of as low as 10 percent, there is only one pair of TPM found. The third row shows the final registration results. The CHARM method is highly effective for the registration of models with an overlapping ratio as high as 70 and 50 percent, and still effective for the overlapping ratio as low as 30 percent. These results show that the proposed CHARM method is applicable for the registration of partially overlapping models. However, for the overlapping ratio of 10 percent, the proposed CHARM method usually fails due to too small common areas between the models to be registered, which is also challenging for any other registration method.

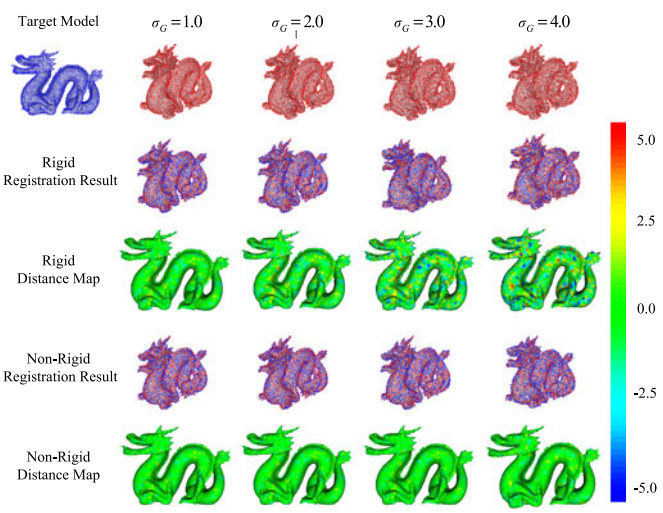

Fig. 9. Results of the proposed CHARM method for the registration of the dragon models subject to different levels of the modeling Gaussian noise. First row: The target model and its corresponding models corrupted with different levels of the modeling noise. Second row: Results of rigid registration. Third row: Error maps of the rigid registration results. Fourth row: Results of non-rigid registration. Fifth row: Error maps of the non-rigid registration results.

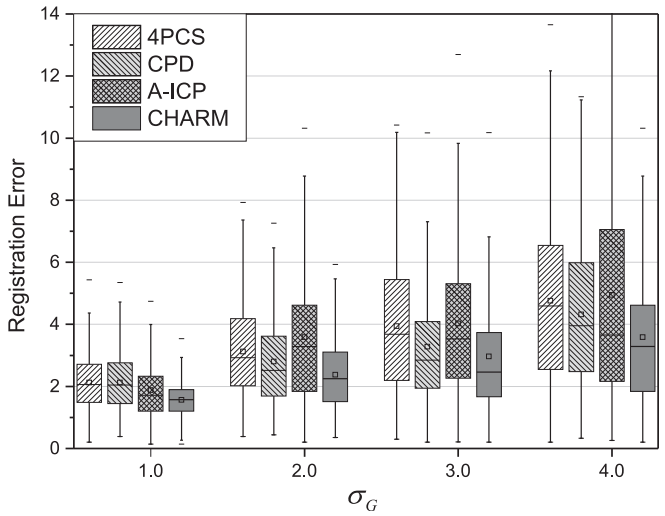

Fig. 10. The registration errors of different algorithms over four levels of the modeling Gaussian noise.

Noise is generally the most common factor that affects the registration accuracy. Fig. 9 shows the results of the CHARM method for the registration of the dragon models under different levels of the modeling Gaussian noise. The blue dragon in the first row is the target model, whereas the red models in the first row are the dragon data sets corrupted by different levels of the modeling Gaussian noise. The second to fifth columns correspond to the modeling noises with standard deviations of 1.0 to 4.0 at steps of 1. Without loss of generality, the red dragons were subject to similarity transformations randomly generated. The degrees of degeneration in the mesh of the dragon become higher with an increase in the standard deviation of the modeling noise. The difference between the degenerated dragon with $\sigma_G = 4.0$ and the target model is considerable. All the models corrupted with the modeling noise were registered with the target model. The second and third rows show the registration results and their error maps, respectively. The mismatching errors increase and are randomly distributed on the surface of the dragon. The fourth and fifth rows show the results of the CHARM method after non-rigid registration. All the local mismatched regions of the dragon in the third row have been accurately registered. The registration error can barely be seen on the error maps of the final registration results.

Fig. 10 shows the comparison of CHARM, 4PCS [11], CPD [10], and A-ICP [7] with respect to different levels of the modeling Gaussian noise. The registration errors increase with its standard deviation, as expected. The CHARM method also achieves the minimum registration errors at all noise levels compared with the other three registration approaches. This is because it is difficult for the 4CPS algorithm to define the coplanar points and use their cross ratios to search for point matches, the CPD algorithm considers the combination of all the points in the two models as potential point matches and the coherence constraint of their movement is not always powerful enough to distinguish the true matches from false ones, and the A-ICP algorithm adopts the greedy closest point strategy for finding potential point matches and the established ones are not always correct. In sharp contrast, the proposed CHARM method detects

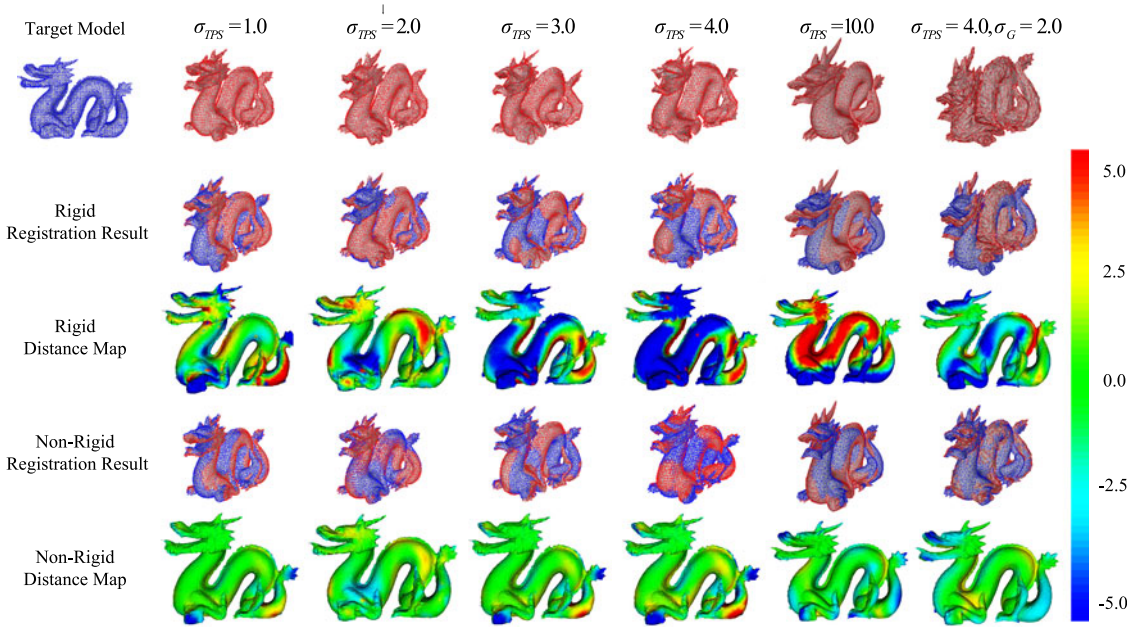

Fig. 11. Results of the proposed CHARM method for the registration of the dragon models subject to different levels of the processing Gaussian noise. First row: The target model and its corresponding models corrupted with different levels of the processing Gaussian noise. Second row: Results of rigid registration. Third row: Error maps of the rigid registration results. Fourth row: Results of non-rigid registration. Fifth row: Error maps of the non-rigid registration results.

distinct points and then describes and matches them, leading more reliable TPMs to be established.

4.3 Non-Rigid Registration Results

Fig. 11 shows the results of the proposed CHARM method for the registration of the dragon models subject to different levels of the processing Gaussian noise. In the first row of Fig. 11, the blue dragon is the target model, whereas the red dragons are the models corrupted with various levels of the processing Gaussian noise. The red dragons are the source models subject to further similarity transformations randomly generated. The second to seventh columns correspond to the transformed dragons with the processing Gaussian noise levels of 1.0, 2.0, 3.0, 4.0, 10.0 and 4.0 combined with a modeling noise ($\sigma_G = 2.0$). The second row shows the alignment of the source and target models by the rigid registration step in the proposed CHARM method, while the third row shows the error maps of the registration results. The transformed source and target models are roughly aligned together for all noise levels; because of the influence of the processing Gaussian noise, a number of local areas exhibit relatively large errors. The fourth and fifth rows show the non-rigid registration results. It can be seen that most of the locally poorly-registered areas in the process of the rigid registration have been brought into more accurate alignment with smaller errors, showing that the local deformations of the transformed source and target models are necessary and effective. The registration results for the processing Gaussian noise levels of 3.0 and 4.0 are inferior to those of 1.0 and 2.0 because several local areas at the tail of the dragon have relatively large errors. The accurate registration results for the processing Gaussian noise levels of 10.0 and 4.0 combined with the modeling Gaussian noise with a standard deviation of 2 show the robust performance of the CHARM method over the models subject to large non-rigid deformation and mixed noise.

In Fig. 12, three non-rigid 3D shape retrieval data with large deformations from SHREC [39] were used to test the robustness of the proposed method. From the book data to the woman data, the magnitude of the non-rigid deformation between the source and target models becomes much larger. The registration results and error maps show that the transformed source and target models have been brought into accurate alignment. In the book and the bird data, the source and target models have been accurately registered, and the maximum error in the deformed regions is less than 2.0. However, in the woman data, due to the different postures of the woman, the magnitude of non-rigid deformation is much larger than those in the previous two data sets. After the non-rigid registration, the source and target models have been still brought into good alignment. But the outline of the transformed source model does not fit that of the target model very well, and the maximum error in some regions of the point matches is larger than 5.0. These results demonstrate that the proposed CHARM

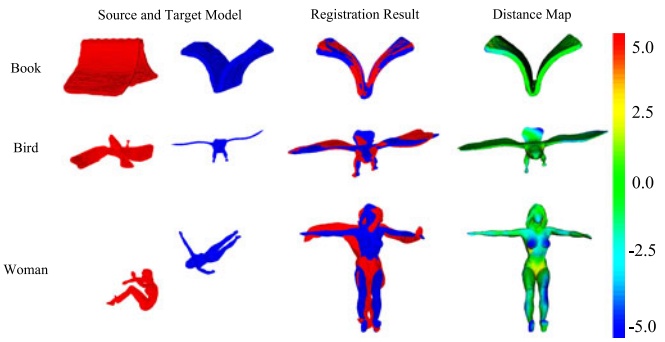

Fig. 12. Results of the proposed CHARM method for the registration of non-rigid 3D shape retrieval data from SHREC. First column: The source and target book, bird and woman models for registration. Second column: The non-rigid registration results. Third column: Error maps of the non-rigid registration results.

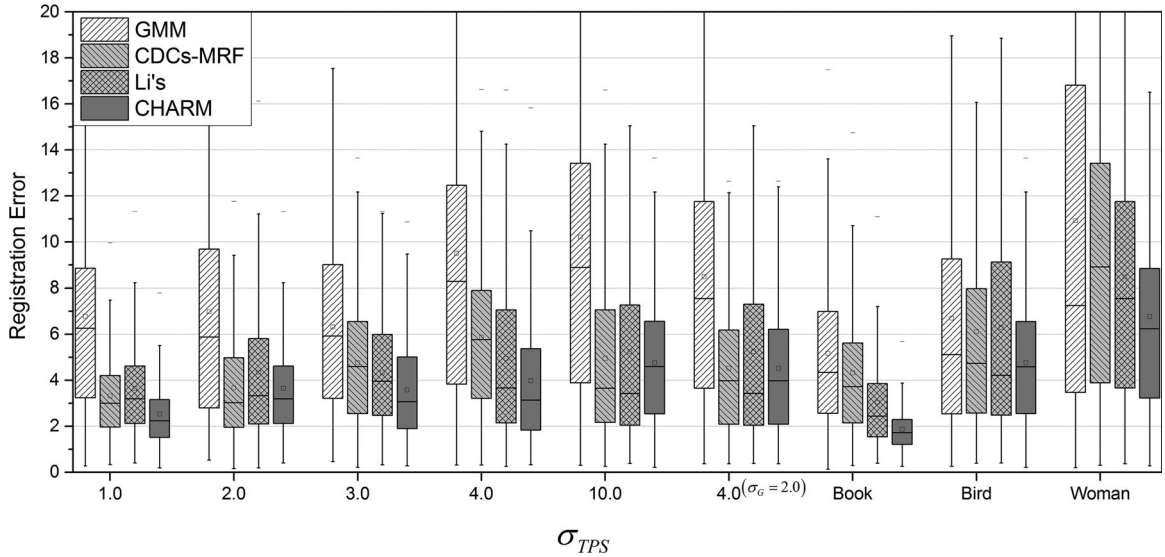

Fig. 13. The errors of different algorithms for the registration of synthetic data subject to six levels of the processing Gaussian noise and three real SHREC data sets.

method is able to register the source and target models even with relatively large non-rigid deformation.

Previous experiments compared 4PCS, CPD, A-ICP and the proposed CHARM method for rigid registrations. Fig. 13 shows the performances of the proposed CHARM method, GMM [9], CDCs-MRF [26], and Li's [23] algorithm for the registration of both the synthetic data subject to different levels of the processing Gaussian noise and real data. The mean of the box plots helps relate the registration errors of all the methods with increased magnitudes of non-rigid deformation. The CHARM method achieved again the minimum registration errors over all the data compared with the other three approaches.

To compare with the standard aggregated measures, the proposed CHARM algorithm is evaluated on the FAUST challenge datasets [40] as shown in Fig. 14. The first row to the third row illustrate the 1st, 8th, and 10th case in the challenge, respectively. The left two columns are the scans to be registered. The third column reports the registration errors in centimeters of different correspondences, and the proposed method is compared to all the ground-truth correspondences (sorted by decreasing error). The fourth column shows the error distributions over the vertices in the first scan in different front/back or side views (the blue color represents the minimum error, the red color represents the maximum error). The registration error in the first two pairs of the scans are acceptable, but the error in the last case is relatively large. This is because in one case, the woman stood with two arms raised. In the other case, she stooped and made her arms touch her legs. As a result, the two gestures of the woman are significantly different and the magnitude of deformation between the two models is so large, leading their convex hulls to fail to define properly the problem for the registration of the two models. Such datasets will also be challenging for any other method to register.

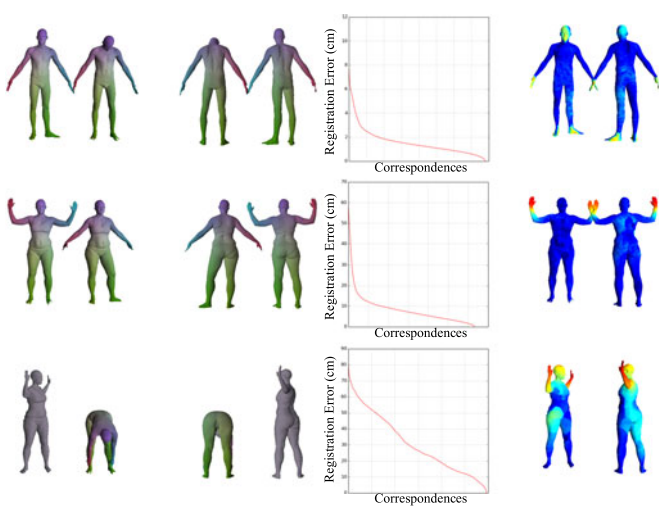

Fig. 14. Results of the proposed CHARM method for the registration of the woman scans subject to different deformations. Left: the scans to be registered, Middle: the error distribution of different correspondences, Right: the error distribution over different vertices in the first scan (visualized from different directions).

4.4 Registration Results on Range Scans

A number of range scans captured by the Kinect imaging device were used to evaluate the performance of the proposed CHARM method. Fig. 15 shows the experimental

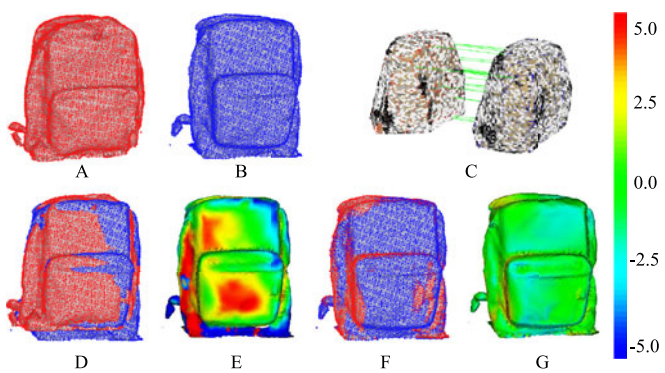

Fig. 15. Results of the proposed CHARM method for the registration of the schoolbag with and without books scanned by a Kinect sensor. A: Source model. B: Target model. C: The feature matching results on the convex hulls. D: Rigid registration results. E: Error map of the rigid registration results. F: Non-rigid registration results. G: Error map of the non-rigid registration results.

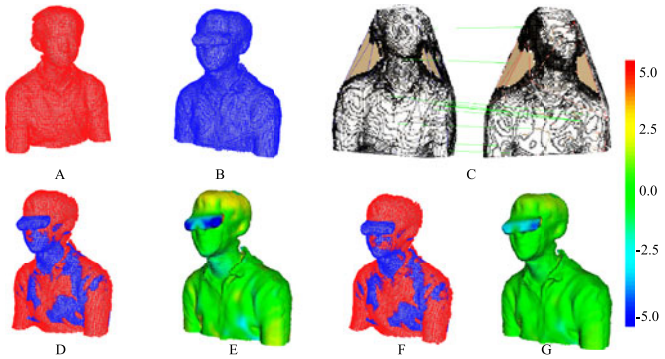

Fig. 16. Results of the proposed CHARM method for the registration of two scans of a man with and without wearing glasses. A: Source model. B: Target model. C: The feature matching results on the convex hulls. D: Rigid registration results. E: Error map of the rigid registration results. F: Non-rigid registration results. G: Error map of the non-rigid registration results.

results of a scanned schoolbag at different levels of fullness. The two scans are accurately registered together, and the largest registration error of the point matches is less than 2.

Fig. 16 shows the upper parts of a human body scanned using the Kinect device. Figs. 16A and 16B show the scanned data sets of a man without and with wearing glasses. The difference between the two scanned data sets can be effectively visualized. Most parts of the models are accurately registered together except several local areas, such as the eyes and top clothes. The mean and largest errors of the point matches are 0.93 and 3.78, respectively. Figs. 16F and 16G show the results and the error map, respectively, after non-rigid registration. The two models are accurately registered; the largest error which is less than 3 is distributed at the border of the glasses, as expected. Thus, the CHARM method is highly effective for the registration of the models with global and local deformations.

3D registration methods are widely applied for the task of object tracking. Fig. 17 demonstrates the performance of the proposed CHARM method for object tracking. In this experiment, a skull model was used to be tracked in a cluttered environment. As the blue scans (captured by Kinect) show, there are many other objects in the scene besides the skull. However, the tracking results show that the source model and target model have been aligned accurately at each position. These results demonstrate that the proposed CHARM method is robust for the rigid registration of the range scans with typical overlap size, and is powerful for object tracking.

4.5 Evaluation of Computational Efficiency

Fig. 18a shows the computational time of four different rigid registration methods: 4PCS, CPD, A-ICP, and the proposed

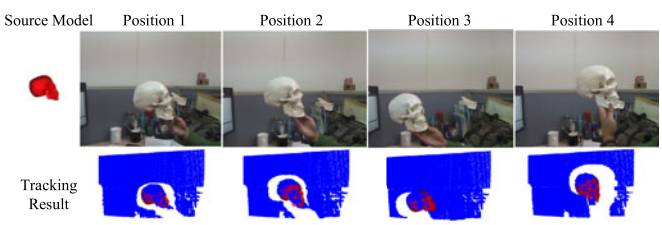

Fig. 17. Results of the proposed CHARM method for object tracking. First row: The source model of the skull and the scenes where the skull sits at different positions. Second row: The tracking results of the source model (red) and the target model (blue).

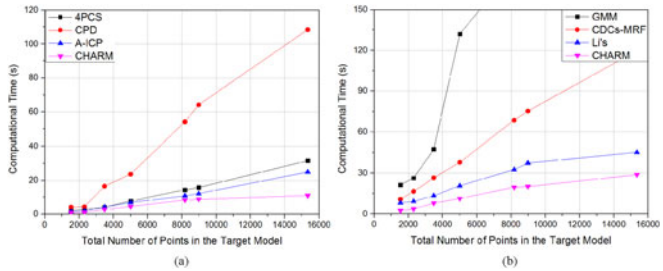

Fig. 18. The computational time in seconds of different rigid and non-rigid registration algorithms.

CHARM method. Fig. 18b shows the computational time of four different non-rigid registration methods: GMM, CDCs-MRF, Li's and CHARM. It can be clearly seen that the proposed CHARM method required less computational time than all the competitors. This is because the existing methods either have to search iteratively for the tentative correspondences and update the estimation of the underlying transformations, or have to consider the combination of all the points in the two models and estimate their weights as tentative correspondences. In contrast, the proposed method established a small set of point matches between the convex hulls of the two models and then use them to constrain the underlying non-rigid deformation. Therefore, the proposed CHARM method exhibits higher computational efficiency for both rigid and non-rigid registrations.

5 Conclusion

In this paper, we proposed a novel convex hull based method for the registration of 3D models subject to non-rigid transformations. The proposed CHARM method first projects each point of the model onto the triangular facet of the convex hull under the parallel projection principle; a set of projection images is then generated. The generated images effectively preserve the point distribution and outline of the entire model and can be utilized to search for point matches. This paper employs the extensively used SIFT algorithm for feature extraction and matching in the projection images. Once the corresponding point pairs have been found, all the feature points are back-projected onto the triangular facets of the convex hulls. The matching points outside the convex hulls can be effectively removed because they must lie inside the relevant triangular facets. A set of refined point matches on the convex hull can be eventually obtained; hence, the model registration problem can be first transformed into a minimization of the distances between the extracted matched point pairs on the convex hulls. These point matches are further utilized as the control points to drive the non-rigid deformation of the source model towards the target one.

A series of experiments based on both simulated and scanned data sets has been conducted for the evaluation of the performance of the proposed CHARM method. The experiments were designed by considering feature matching capability, rotational angle, different Gaussian noises, density variations and overlap size. The proposed algorithm was also compared with the state-of-the-art ones, including 4PCS, CPD, A-ICP, GMM, CDCs-MRF, and Li's. The experimental results show that our method is more effective, accurate and efficient for the registration of 3D models subject to rigid or

non-rigid transformations. The proposed CHARM method achieves significantly smaller registration errors and requires considerably less computational time compared with all the selected methods. Further research will focus on developing methods to deal with the registration of the 3D models subject to considerably large deformations and applying the graphics processing unit to further improve the computational efficiency of the proposed CHARM method.

Acknowledgments

This work was supported by the National Basic Research Program of China (2013CB328806), Key Projects in the National Science & Technology Pillar Program (2013BAI01B01), National Hi-Tech Research and Development Program (2015AA043203), and the National Science Foundation Program of China (81430039, 61501030). Jian Yang is the corresponding author.

References

[1] P. Besl and N. McKay, "A method for registration of 3D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[2] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.

[3] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vision Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[4] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[5] S. Granger and X. Pennec, "Multi-scale EM-ICP: A fast and robust approach for surface registration," in *Proc. 7th Eur. Conf. Comput. Vision*, 2002, pp. 418–32.

[6] J. Hermans, D. Smeets, D. Vandermeulen, and P. Suetens, "Robust point set registration using EM-ICP with information-theoretically optimal outlier handling," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, 2011, pp. 2465–2472.

[7] L. Maier-Hein, et al., "Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1520–1532, Aug. 2012.

[8] Y. Liu, "Penalizing closest point sharing for automatic free form shape registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 1058–1064, May 2011.

[9] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.

[10] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.

[11] D. Aiger, N. Mitra, and D. Cohen-Or, "4-points congruent sets for robust pairwise surface registration," *ACM Trans. Graph.*, vol. 27, pp. 85:1–85:10, 2008.

[12] A. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 433–449, May 1999.

[13] G. Mori, S. Belongie, and J. Malik, "Efficient shape matching using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1832–1837, Nov. 2005.

[14] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[15] A. Frome, D. Huber, R. Kolluri, T. Blow, and J. Malik, "Recognizing objects in range data using regional point descriptors," in *Proc. 8th Eur. Conf. Comput. Vision*, 2004, pp. 224–237.

[16] P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional sift descriptor and its application to action recognition," in *Proc. 15th Int. Conf. Multimedia*, 2007, pp. 357–360.

[17] C. Wu, B. Clipp, X. Li, J. M. Frahm, and M. Pollefeys, "3D model matching with Viewpoint-Invariant Patches (VIP)," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, 2008, pp. 1–8.

[18] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface feature detection and description with applications to mesh matching," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, 2009, pp. 373–380.

[19] B. J. Brown and S. Rusinkiewicz, "Non-rigid range-scan alignment using thin-plate splines," in *Proc. 2nd Int. Symp. 3D Data Process. Vis. Transmission*, 2004, pp. 759–765.

[20] B. J. Brown, "Global non-rigid alignment of 3D scans," in *Proc. ACM SIGGRAPH*, 2007, pp. 21:1–21:9.

[21] N. J. Mitra, S. Floery, M. Ovsjanikov, N. Gelfand, L. J. Guibas, and H. Pottmann, "Dynamic geometry registration," in *Proc. 5th Eurographics Symp. Geometry Process.*, 2007, pp. 173–182.

[22] H. Li, R. W. Sumner, and M. Pauly, "Global correspondence optimization for nonrigid registration of depth scans," in *Proc. 6th Eurographics Symp. Geometry Comput. Graph. Forum*, 2008, vol. 27, pp. 1421–1430.

[23] H. Li, B. Adams, L. J. Guibas, and M. Pauly, "Robust single-view geometry and motion reconstruction," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 175:1–175:10, 2009.

[24] C. Cagniart, E. Boyer, and S. Ilic, "Probabilistic deformable surface tracking from multiple videos," in *Proc. 11th Eur. Conf. Computer Vision*, 2010, pp. 326–339.

[25] M. Wand, et al., "Efficient reconstruction of nonrigid shape and motion from real-time 3D scanner data," *ACM Trans. Graph*, vol. 28, no. 2, pp. 15:1–15:15, 2009.

[26] Y. Zeng, C. Wang, X. Gu, D. Samaras, and N. Paragios, "A generic deformation model for dense non-rigid surface registration: A higher-order MRF-based approach," in *Proc. IEEE Conf. Comput. Vision*, 2013, pp. 3360–3367.

[27] Q. Chen and V. Koltun, "Robust nonrigid registration by convex optimization," in *Proc. IEEE Int. Conf. Comput. Vision*, 2015, pp. 2039–2047.

[28] J. Fan, et al., "Convex hull indexed gaussian mixture model (CH-GMM) for 3D point set registration," *Pattern Recognition*, vol. 59, pp. 126–141, 2016, doi:10.1016/j.patcog.2016.02.023.

[29] J. Fan, J. Yang, F. Lu, D. Ai, Y. Zhao, and Y. Wang, "3-points convex hull matching (3PCHM) for fast and robust point set registration," *Neurocomputing*, vol. 194, no. 6, pp. 227–240, 2016.

[30] D. M. Berg, O. Cheong, and V. M. Kreveld, *Computational Geometry: Algorithms and Applications*. New York, NY, USA: Springer-Verlag, 2008.

[31] S. Fleishman, I. Drori, and D. Cohen-Or, "Bilateral mesh denoising," in *Proc. ACM SIGGRAPH*, 2003, pp. 950–953.

[32] K. A. Mohamed and C. Kupich, "An O(n log n) output-sensitive algorithm to detect and resolve conflicts for 1D range filters in router tables," Albert-Ludwigs-Universitt Freiburg, Freiburg, Germany, Tech. Rep. 226, pp. 1–12, 2006.

[33] C. Barber, D. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, 1996.

[34] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[35] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 4, pp. 376–380, Apr. 1991.

[36] Stanford, "The stanford 3d scanning repository." (2010). [Online]. Available: http://graphics.stanford.edu

[37] Aim@Shape, "Aim@shape shape repository v4.0," 2007. [Online]. Available: http://shapes.aimatshape.net

[38] J. Wu and L. Kobbelt, "A stream algorithm for the decimation of massive meshes," in *Proc. Graph. Interface*, 2003, vol. 3, pp. 185–192.

[39] Z. Lian, et al., "Shrec'11 track: Shape retrieval on non-rigid 3D watertight meshes," in *Proc. 4th Eurographics Conf. 3D Object Retrieval*, 2011, vol. 11, pp. 79–88.

[40] F. Bogo, J. Romero, M. Loper, and M. J. Black, "FAUST: Dataset and evaluation for 3D mesh registration," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, Jun. 2014, pp. 3794–3801.

Jingfan Fan received the BSc degree in optical engineering from Beijing Institute of Technology, in 2010. He is currently working toward the PhD degree. His research interests include medical image processing and augmented reality.

Jian Yang received the PhD degree in optical engineering from Beijing Institute of Technology, in 2007. He is currently a professor in the School of Optoelectronics, Beijing Institute of Technology, China. His research interests focuses on medical image processing, augmented reality, and computer vision.

Yitian Zhao received the PhD degree in computer science from Aberystwyth University, United Kingdom, in 2013. He is a lecturer with Beijing Institute of Technology, China. From November of 2013 to October 2014, he worked as a research assistant with the University of Liverpool. His primary research interests lie in 2D/3D image processing, medical image analysis, pattern recognition, computer graphics, and data-base and knowledge base systems.

Danni Ai received the BE and ME degrees from Xi'an Jiaotong University, in 2005 and 2008, respectively, and the DE degree from Ritsumeikan University, Japan, in 2011. She joined Beijing Institute of Technology as a postdoctoral, in 2013. Her research interests include medical image analysis, virtual reality and augmented reality.

Yonghuai Liu received the first PhD degree from Northwestern Polytechnical University, PR China, in 1998, and second PhD degree from the University of Hull, United Kingdom, in 2001. He is a senior lecturer with Aberystwyth University, since 2011. His primary research interests lie in 3D computer vision, pattern recognition, image processing, machine learning, artificial intelligence, and intelligent systems. He is a fellow of the Higher Education Academy of United Kingdom.

Ge Wang received the PhD in ECE, is Clark & Crossan Endowed Chair Professor and Director of Biomedical Imaging Center/Cluster, Rensselaer Polytechnic Institute, Troy, New York, USA. His expertise includes x-ray computed tomography (CT), optical molecular tomography, and multimodality imaging.

Yongtian Wang received the PhD degree in optics from the University of Reading, England, in 1986. He is currently a Yangtze River scholar, and the director of the Optoelectronics and Information Technology Center, Beijing Institute of Technology. His research interests include optical design and CAD, virtual reality (VR) and augmented reality (AR) technologies and applications.