

Fully Automatic and Robust 3D Modeling for Range Scan Data of Complex 3D Objects

Jungjae Yim and Guoliang Fan^(✉)

School of Electrical and Computer Engineering, Oklahoma State University,
Stillwater, OK 74078, USA

{jungjae,guoliang.fan}@okstate.edu

Abstract. 3D surface registration of two or more range scans is an important step in building a complete 3D model of an object. When the overlaps between multi-view scans are insufficient, good initial alignment is necessary that usually requires some prior assumption such as pre-defined initial camera configuration or the use of landmarks. Specifically, this paper addresses the problem of registering two or more range scans captured from complex 3D objects which have small overlaps. The proposed technique is based on the integration of a new Partial Artificial Heat Kernel Signature (PA-HKS) and a Modified Multi-view Iterative Contour Coherence (MM-ICC) algorithm. This unique combination allows us to handle multi-view range scan data with large out-of-plane rotation and with limited overlaps between every two adjacent views. The experimental results on several complex 3D objects show the effectiveness and robustness of the proposed approach.

1 Introduction

3D modeling has long been one of the most important research topics in the field of computer vision and pattern recognition. Recently, there are many new techniques for complex 3D object modeling [1,2]. Moreover, with recent technology advancement, many inexpensive 3D sensors [3,4] were developed, which triggers easy access to 3D depth data and allows various applications of using 3D point sets such as 3D modeling, pose estimation, and 3D object recognition. Among many, registration of the range scans to generate 3D models has drawn a great attention and many successful registration techniques have been proposed which usually require sufficient overlap across views and good camera initialization [5,6]. However, registration for partially overlapping data without any initialization still remains a challenge in this field. In this paper, we address this issue by two steps. The first is to support initial alignment of multi-view scans by using a new mesh-based feature. The second is to improve multi-view contour coherence by considering both self-occlusion and visibility of the predicted contours across views. Our goal is aimed at automatic and robust 3D registration from partial overlapping 3D point sets without any camera initialization.

The Heat kernel signature algorithm is generally applied to generate highly localized features from a 3D mesh model for pose estimation or object

recognition [7–9]. This technique shows a successful result exclusively on full 3D mesh data and 2.5D range data without self-occlusion. Because of self-occlusion of 3D data, the accuracy of feature extraction would deteriorate mainly due to the holes on the 3D surface. Accordingly, we develop an improved heat kernel feature in order to accommodate partially overlapped range scan data captured from different view angles. As the first step that was inspired by [7, 8], we propose a new partial artificial Heat kernel signature (PA-HKS) and use it to achieve coarse alignment of the multiple range scan data. As the second step, the original Multi-view iterative contour coherence (M-ICC) algorithm [10] is enhanced by introducing an additional pruning step to remove erroneous contours, leading a modified M-ICC (MM-ICC) algorithm that is more robust and accurate. The original M-ICC algorithm requires initialization of the range scans by pre-defined camera configuration. In conjunction with PA-HKS based alignment, the proposed MM-ICC algorithm is able to self-initialize multi-view camera configuration to support fully automatic and robust registration and to reconstruct a detailed 3D model based on the corresponding contours from multi-view scans.

2 Related Work

In this section, we provide a brief overview of the background of this research, which involves two separate topics: Direct (Appearance-based) registration and Feature-based registration.

Direct (Appearance-Based) Registration. There have been a great deal of research on point set registration focusing on where most pixels agree. Among many, the classic iterative closest point (ICP) algorithm [5] and its variants [11] provide an effective way to register the points and reconstruct the model in a stable and effective manner. However, ICP requires good initialization and is also time consuming mainly because it has to identify the closest point pairs. The Coherence point drift (CPD) algorithm, which is another powerful method of point set registration, and the modified CPD algorithms [12] allow us to get more robust and stable outcomes in presence of noise and outliers. CPD can produce very robust results both in the rigid and non-rigid point sets. However, CPD also entails good initialization to find correct correspondences. Moreover, the M-ICC algorithm in [10] aims to perform registration based on the multi-view contour coherence. These algorithms above have proven effective and accurate under reasonable initialization. However, achieving accurate and robust registration without initialization still remains as a major challenge in point set registration.

Feature-Based Registration. Feature-based registration methods attempt to find corresponding features in the multiple data sets of different views and to effectively match the extracted features across views. It is a fairly recent research trend that puts an emphasis on detecting interest points on 3D mesh models. Most of the studies along this trend focus on local surface

descriptors [13,14]. A multi-scale approach is commonly employed to analyze the 3D surface at consecutive scales to identify interest points at different levels [15]. A 3D extension of the 2D Harris operator was proposed that is based on the local autocorrelation of images [16,17]. The studies in [17,18] employ the Heat Kernel Signature (HKS) of a 3D mesh model. According to the geometry energy on the vertices, an interest point is selected if it remains as a local maximum of the geometry energy function within several successive scales. Therefore, it is crucial to have the distinctiveness of an interest point for a stable outcome.

Both of the two categories of point set registration reviewed above entail their own strengths and limitations. Accordingly, a number of attempts have been made to combine these two methods to alleviate the constraints and to deliver more robust and accurate outcomes [18].

3 Proposed Approach

The proposed algorithm consists of following two steps: *Coarse Registration* using PA-HKS and *Fine Registration* through MM-ICC. The first involves rough alignment of the multiple partial overlapping range scan based on PA-HKS extracted from each view. The second step is initialized by the first step and performs re-registration by applying via three-round contour pruning and maximizing contour coherence.

3.1 Coarse Registration Based PA-HK

The HKS is grounded based on the Heat kernel and this heat kernel is stable against perturbations of the shape [7]. A HKS formula is derived as $K_t = \sum_{i=0}^{\infty} \exp^{\lambda_i t} \phi_i(x)^2$ where, λ_i and ϕ_i are i^{th} eigenvalue and eigenfunction of Laplace-Beltrami operator. This algorithm uses heat diffusion on the surface of a full 3D model to detect the highly local shape features. Heat diffusion over a longer period of time enables us to spot the summaries of a shape in large neighborhoods whereas the one with a short time period allows the observation of detailed local shape features. Therefore, through heat diffusion, we are able to compare the point signatures at different time intervals and accomplish multi-scale matching between points. The HKS algorithm operates based on the gradient of the surface; thus, it is hard to find the correct feature descriptor in the partial object surfaces under self-occlusion. So, in order to address the problem, we project the range scan data to the plane and connect the projected data with original range data using mesh. This new generated 3D volume mesh is used to apply the HKS algorithm to find the feature descriptor. Although the generated artificial 3D mesh only captures partial geometric information from the 3D model, the extracted PA-HKS features are expected to preserve some local geometric characteristics by which we can produce relatively reliable alignment between the partial surfaces observed from two adjacent views. The idea of PA-HKS is discussed in three steps as follows.

Extracting PA-HKS Keypoints. For each view, we generate a partial artificial 3D mesh data by projecting the range scan data to the back plane. The back and side planar surfaces of the artificial 3D mesh are added to find PA-HKS keypoints on the frontal mesh surface visible in that view. It is possible some keypoints are found along the contour between the back and side surfaces, then we move them to the front mesh surface. Thus, all PA-HKS keypoints are in the original range scan, as shown in Fig. 1(a)–(c).

Grouping PA-HKS Keypoints. There could be multiple local PA-HKS keypoints in the same area among the multiple range scans. In order to ensure one-to-one mapping across views, we group the keypoints by using their 3D Euclidian distance and the similarity of their PA-HKS signatures, as shown in Fig. 1(d). This will result a single local PA-HKS keypoint in those areas that have high energy and rich geometric information. Although those keypoints may not be spatially precise in terms of their locations, they offer important 3D landmarks for initial point set alignment.

Matching PA-HKS Keypoints. Prior to ICP-based alignment, PA-HKS features are used to initialize correspondence between two adjacent views. Given two PA-HKS feature sets from two range scans, \mathbf{S} and \mathbf{T} , we find the correspondence pairs across two views by

$$\mathbf{t}_s = \arg \min_{\mathbf{t} \in \mathbf{T}} d(\mathbf{t}, \mathbf{s}), \forall \mathbf{s} \in \mathbf{S}, \quad (1)$$

where $d(\cdot)$ is a distance function between two PA-HKS features. To ensure robust feature matching, if $d(\mathbf{s}', \mathbf{t}_{s'})$ is larger than a threshold, we declare that $\mathbf{s}' \in \mathbf{S}$ does not have a correspondence in \mathbf{T} and will not be involved in the following ICP step. After PA-HKS based correspondence initialization, we can use ICP to find the initial transformation between the two scans.

3.2 Fine Registration Based on MM-ICC

Wang [10] proposed a wide baseline 3D modeling algorithm (called M-ICC) that registers multi-view range scans by maximizing contour coherence between the observed and predicted contours across multiple views. A key idea in the algorithm is to remove incorrect contour points using two step pruning. Here we want to improve contour coherence by refining contour correspondences via a third pruning step. This extra pruning enables us to remove incorrect correspondences of the predicted contours according to the invisibility condition. We name the improved algorithm MM-ICC to differentiate from the original M-ICC one.

Generating Observed R_i and Predicted Range scan data $R_{i \rightarrow j}$. A range scan R_i of view i provides a depth value $R_i(\mathbf{x})$ at each image pixel $\mathbf{x} = (x, y)^T \in \mathbb{R}^2$. A meshed point cloud M_j is generated by range scan R_j and $R_{i \rightarrow j}$ is created by projecting M_j to the i^{th} view. $N_i(\mathbf{x})$ is the surface normal vector of an image pixel \mathbf{x} in R_i .

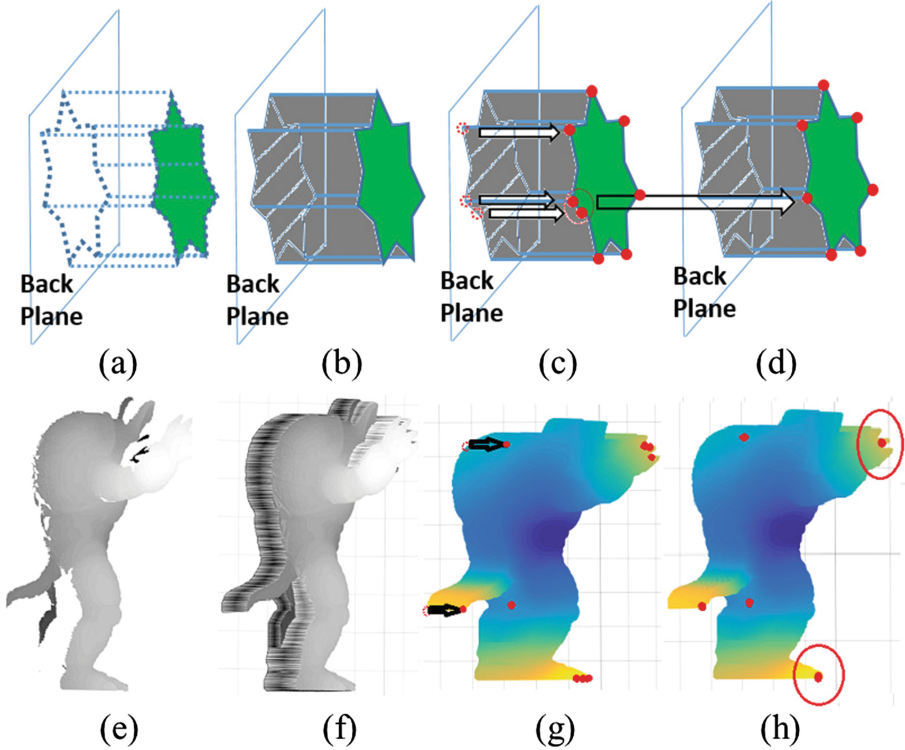


Fig. 1. Illustration of PA-HKS: (a) A range scan in 3D (green); (b) A partial 3D mesh created from (a); (c) PA-HKS features after shifting; (d) PA-HKS features after grouping; (e)–(h) Real object examples corresponding to (a)–(d). (Color figure online)

Extracting Contour Points. The observed range scan data R_i has two different sets of contour points. One set is from the front contours and the other set is from the back contours. The back contours will be set as occlusion ones during the first pruning step. Given pixels belonging to the object in view i as \mathbf{X}_i , the depth of pixels belonging to the background is set by be infinite, i.e., $R_i(\mathbf{x}) = \infty$ for $\mathbf{x} \notin \mathbf{X}_i$. The set of visible contour points \mathcal{C}_i and that of occlusion contour points are distinguished by depth discontinuity of a pixel and its 8 neighboring pixels, $N_{\mathbf{x}}^8$ of range scan as follows:

$$\mathcal{C}_i = \{\mathbf{x} \in \mathbf{X}_i | \exists \mathbf{y} \in N_{\mathbf{x}}^8, R_i(\mathbf{y}) - R_i(\mathbf{x}) > \tau_1\}, \quad (2)$$

$$\mathcal{O}_i = \{\mathbf{x} \in \mathbf{X}_i | \exists \mathbf{y} \in N_{\mathbf{x}}^8, R_i(\mathbf{x}) - R_i(\mathbf{y}) > \tau_2\}, \quad (3)$$

where τ_1 and τ_2 are two thresholds to control the quality of selected PA-HKS features. Depth discontinuity caused by connection of contours and occlusion contours indicates surface holes created by self-occlusion in the range scan data. Those surface holes can remain incorrect contour points which have to be pruned.

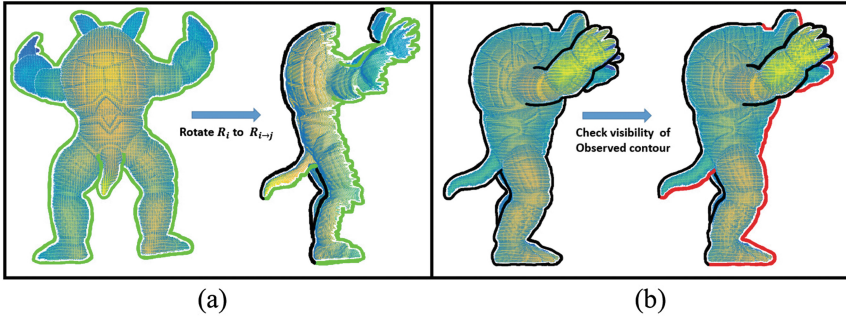


Fig. 2. Illustration of two pruning steps in the original M-ICC algorithm: (a) The first pruning: the green line is \mathcal{C}_i (previous contour points) from observed view i and the black line is $\mathcal{C}_{i \rightarrow j}$ (new contour points) from predicted view j . (b) The second pruning: the black line is \mathcal{C}_j from camera j and the red line shows invisible contour points from view i . (Color figure online)

In this step, \mathcal{C}_i and $\mathcal{C}_{i \rightarrow j}$ are created from R_i and $R_{i \rightarrow j}$ as the sets of observed and predicted contours respectively.

First Pruning: Self-occlusion of Predicted Contour. $\mathcal{C}_{i \rightarrow j}$ contains false contour points which are generated by the boundary points of surface holes created by self-occlusion. Basically, contour points change depending on different view angles in fully covered 3D mesh data. Thus, \mathcal{C}_i and \mathcal{O}_i from the observed view direction should not be one of the members of $\mathcal{C}_{i \rightarrow j}$ from the predicted view direction. In Fig. 2(a), the green color line is a visible contour in previous Camera i direction and the black line is a new contour in j direction after rotated range scan data R_i . This green one should not be a valid contour in camera j direction because the rotated range scan data R_i does not cover the full surface of a target in camera j direction. In Fig. 3(a), the square point corresponds to a contour $\mathcal{C}_{i \rightarrow j}$ from camera view j and the dot point is a contour \mathcal{C}_i from camera view i . If those contours are at the same 3D location, the contour should be pruned, as shown in Fig. 3(a).

$$\mathcal{C}_{i \rightarrow j}^{(1)} = \{\mathbf{X} \in \mathcal{C}_{i \rightarrow j} | \mathcal{C}_{i \rightarrow j}(\mathbf{X}) \cap (\mathcal{O}_i(\mathbf{X}) \cup \mathcal{C}_i(\mathbf{X}))^c\}. \tag{4}$$

Second Pruning: Visibility of Observed Contours. Some contour points in \mathcal{C}_j are not visible from camera location of view i . We prune \mathcal{C}_j based on the visibility of the corresponding contour in view i .

$$\mathcal{C}_{j/i}^{(2)} = \{\mathbf{X} \in \mathcal{C}_j | N_j(\mathbf{X})^T \cdot (\mathcal{O}_{i \rightarrow j} - V_j(\mathbf{X})) > 0\}, \tag{5}$$

where $\mathcal{O}_{i \rightarrow j}$ is the camera location of frame i in camera j and $V_j(\mathbf{X})$ is the back-projection operator which maps \mathbf{X} in frame j to its 3D location.

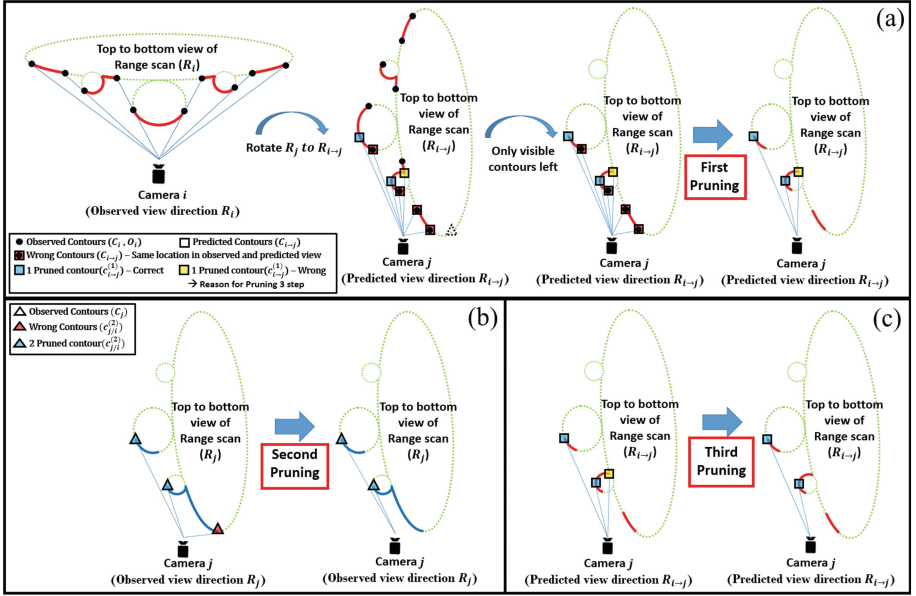


Fig. 3. Illustration of the three-step pruning from a top-down view where each dot represents a contour and each contour correspond to a partial surface: (a) Range Scan R_i , contour C_i (black dots) and occlusion points O_i (black dots) in camera i . (b) notation of Fig. 2. (c) Pruning 1: Range Scan $R_{i \rightarrow j}$, contour points $C_{i \rightarrow j}$ (square) and pruned contour points $C_{i \rightarrow j}^{(1)}$ (blue square and yellow square) in camera j , Some Contours (yellow square) require Third pruning. (d) Pruning 2: Range Scan (R_j) in camera j and pruned contour points ($C_{j/i}^{(2)}$) (Color figure online)

$N_j(\mathbf{x})$ is the surface normal vector of each image pixel in frame j . In Fig. 2(b), the black color line is a visible contour from camera j direction and the red line is an invisible contour in i direction. The red contour should be pruned to find the corresponding contour points only. Only black contours are left after the first and second pruning steps because the black contours have corresponding contours. In Fig. 3(b), the red triangle point presents a contour from camera j and it is invisible from camera i . Those points do not have correspondences in camera i frame thus they should be pruned.

Third Pruning: Visibility of Predicted Contours. As a new step introduced in this work, the third pruning is applied to the results of the first pruning which may still have incorrect contour points. Some contour points $C_{i \rightarrow j}$ generated by $R_{i \rightarrow j}$ are invisible in camera j . However, self-occlusion effect makes the points visible in camera j . Thus, these incorrect points cause errors in the matching step and obstruct accurate registration. In Fig. 3(a) and (c), although the yellow square points should not be visible in camera j , they are still considered as visible contours in the predicted view. Therefore, we perform the third

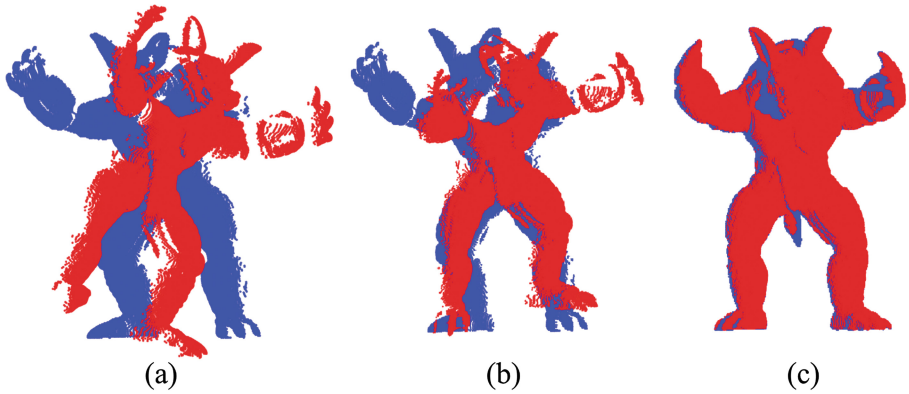


Fig. 4. Illustration of PA-HKS based coarse registration: (a) Initialization of two range scans from the 3D Armadillo model with rotation gap 70° ; (b) Registration results of the M-ICC algorithm (c) Registration results using PA-HKS for ICP.

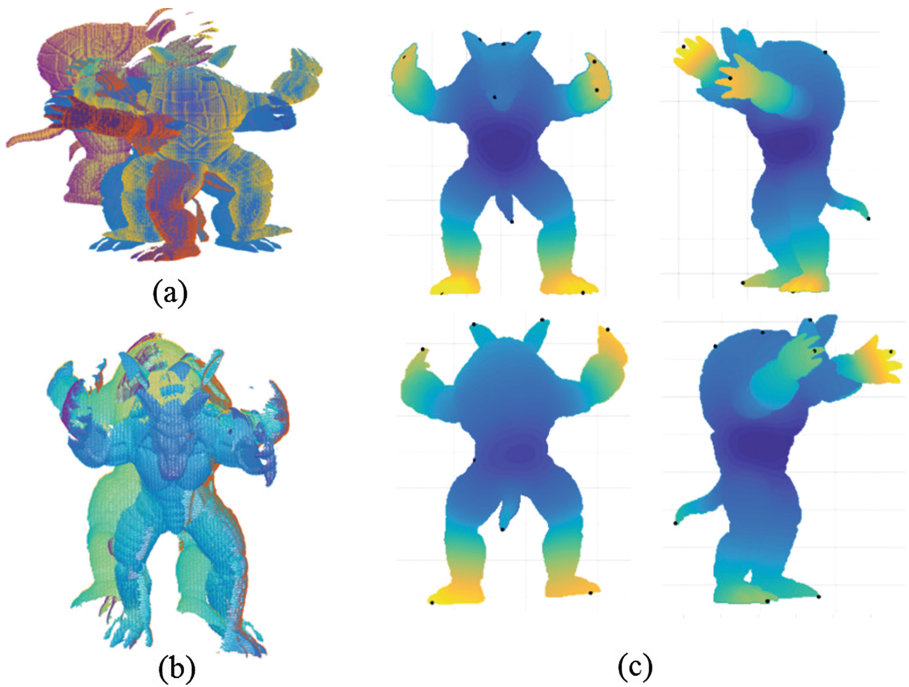


Fig. 5. (a) Four Range scans of the armadillo in the 3D space; (b) PA-HKS features extracted from of four range scans (black dots) (c) Result of the PA-HKS based initial alignment using ICP.

Moreover, four different range scans extracted from the 3D Stanford armadillo model is used to verify the effectiveness of the proposed algorithm compared with the previous M-ICC algorithm. In Fig. 5(a), the four range scans are visualized together in the 3D space without any previously known initialization. All range scan data face to the same one camera direction in the unknown initialization setting. The proposed PA-HKS method extracted corresponding features from the partially overlapped range scan data based on Heat kernel signature(HKS) descriptor. In Fig. 4(b), the heat distribution is indicated by different colors of surface. The warm colors around the PA-HKS features indicate the high variance of the local geometry and similar distribution of feature points in the multi-view scans implies their consistency and reliability. The initial alignment result given in Fig. 4(c) confirms the usefulness of PA-HKS for coarse registration.

4.2 MM-ICC for Refined Registration

We compare M-ICC and MM-ICC based on the same view initialization obtained from PA-HKS based coarse registration. In Fig. 6, the left column represents the results from M-ICC under seven different settings listed in Table 2, while the right column shows the results from MM-ICC. Table 2 shows the RMS errors under seven settings. It is shown that M-ICC could result in successful results in the case of two simple views (with little occlusion), as shown Fig. 6 (the 1st and 4th rows) where registration error is relatively small (Table 2 (setting (i,iv))). However, in more challenging conditions, such as complex views which have more self-occlusion areas (setting ii, v) or large rotation angles among multi-view range scans (setting iii), M-ICC was incapable of yielding accurate results, as in Fig. 6 (the 2nd, 3rd and 5th row). The RMS results in Table 2 clearly show the effectiveness, accuracy and robustness of the proposed MM-ICC algorithm. Moreover, in Fig. 6 (the 6th and 7th rows), two of four pair-wise registrations from M-ICC show successful outcomes whereas the other two has significant mis-match. On the contrary, the results of MM-ICC show stable and accurate matching outcomes for all four pair-wise registration. Overall, the proposed MM-ICC algorithm is proven to be robust and accurate for complex 3D objects under large gaps of multi-view scans and significant occlusion in each scan.

Table 2. RMS errors in results

Different test setting	M-ICC	Proposed
(i) Armadillo, 70°, 2 simple views	1.099	0.290
(ii) Armadillo, 70°, 2 complex views	3.169	0.281
(iii) Armadillo, 90°, 2 complex views	3.289	0.233
(iv) Bunny, 60°, 2 simple views	0.521	0.189
(v) Bunny, 60°, 2 complex views	5.116	0.116
(vi) Bunny, 90°, 4 complex views	Ave. of 4 pairs = 1.595	Ave. of 4 pairs = 0.206
(vii) Armadillo, 90°, 4 complex views	Ave. of 4 pairs = 4.476	Ave. of 4 pairs = 0.279

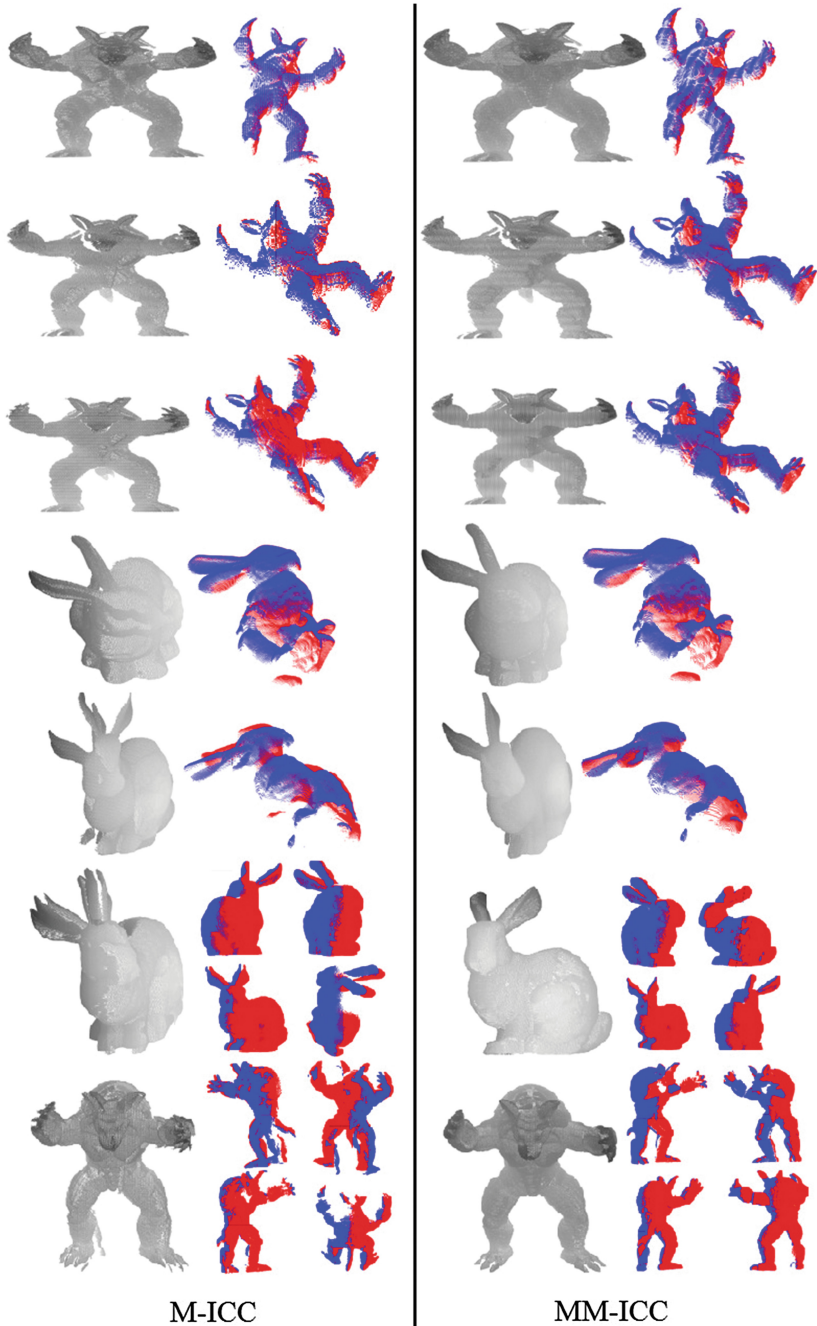


Fig. 6. Comparison between M-ICC and MM-ICC based on two 3D models under seven settings given in Table 2. For each setting, we show the final 3D modeling result (left) and the pair-wise scan matching results of every two adjacent views (right).

5 Conclusion and Future Work

We have proposed a new 3D modeling algorithm that involves two new techniques to perform coarse-to-fine registration of multi-view range scan data without any initial condition. Specifically, the PA-HKS features extracted from partial artificial mesh models are found useful to roughly align range scan data despite large view gaps and complex object contours. This step makes the proposed algorithm more robust and flexible to handle various multi-view camera settings and 3D objects of complex shapes. The MM-ICC algorithm incorporates an additional pruning step to further remove supposedly invisible points in predicted contours. This third pruning is very effective to assist the second-round re-registration and significantly improves the final results. In the future, we will apply this method to real depth data captured from RGB-D sensors. In addition, we will also make an attempt to generalize the proposed approach for robust modeling of non-rigid objects.

References

1. Allen, B., Curless, B., Popovic, Z.: Articulated body deformation from range scan data. *ACM Trans. Graph.* **21**, 612–619 (2002)
2. Chang, W., Zwicker, M.: Automatic registration for articulated shapes. In: *Proceedings of Eurographics Symposium on Geometry Processing*, vol. 27, pp. 1459–1468 (2002)
3. Rogge, S., Hentschel, C.: A multi-depth camera capture system for point cloud library. In: *Proceedings of IEEE Fourth International Conference on Consumer Electronics*, pp. 50–54 (2014)
4. Zhang, Z.: Microsoft kinect sensor and its effect. *IEEE Multimedia* **19**, 4–10 (2012)
5. Besl, P., McKay, N.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 239–255 (1992)
6. Myronenko, A., Song, X., Carreria-Perpinan, M.: Non-rigid point set registration: Coherent point drift. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 2262–2275 (2010)
7. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: *Proceedings of Eurographics Symposium on Geometry Processing*, vol. 28, pp. 1383–1392 (2009)
8. Brandao, S., Costeria, J., Veloso, M.: The partial view heat kernel descriptor for 3D object representation. In: *Proceedings of ICRA 2014*, vol. 28 (2014)
9. Bronstein, M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1704–1711 (2010)
10. Wang, R., Choi, J., Medioni, G.: 3D modeling from wide baseline range scans using contour coherence. In: *Proceedings of CVPR 2014*, pp. 1–8 (2014)
11. Rusinkiewicz, S., Levoy, M.: Efficient variants of the ICP algorithm. In: *Proceedings of 3rd International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152 (2001)
12. Aiger, D., Mitra, N., Cohen-Cor, D.: 4-points congruent sets for robust pairwise surface registration. *ACM Trans. Graph. (TOG)* **27**(3), 85 (2008)
13. Shah, S., Bennamoun, F., Boussaid, F., El-Sallam, A.: 3D-Div: A novel local surface descriptor for feature matching and pairwise range image registration. In: *2013 IEEE International Conference on Image Processing*, pp. 2934–2938 (2013)

14. Guo, Y., Sohel, F., Bennamoun, M., Lu, M., Wan, J.: TrisI: A distinctive local surface descriptor for 3D modeling and object recognition. In: 2013 International Conference on Computer Graphics Theory and Applications, GRAPP 2013, pp. 86–93 (2013)
15. Tardif, C., Schafer, A., Waehnert, M., Dinse, J., Turner, R., Bazin, P.: Multi-contrast multi-scale surface registration for improved alignment of cortical areas. *Neuroimage* **111**, 107–122 (2015)
16. Sipiran, I., Bustos, B.: A robust 3D interest points detector based on harris operator. In: Proceedings of Eurographics Workshop on 3D Object Retrieval, pp. 7–14 (2010)
17. Dutagaci, H., Cheung, C., Godil, A.: Evaluation of 3D interest point detection techniques via human-generated ground truth. In: Proceedings of Eurographics Workshop on 3D Object Retrieval (2012)
18. Xia, R., Zhao, J., Liu, Y.: A robust feature-based registration method of multi-modal image using phase congruency and coherent point drift. In: SPIE Proceedings: Pattern Recognition and Computer Vision, vol. 8919 (2013)
19. Chetverikov, D., Svirko, D., Stepanov, D.: The trimmed iterative closest point algorithm. In: 16th International Conference on Pattern Recognition, vol. 3, pp. 1–4 (2002)
20. The stanford 3D scanning repository (Retrieved July 2015)