Automatic Registration of Range Images by Using Shape-Index-Based Artificial Images*

Hu WEI**, Liyan ZHANG**, Shenglan LIU** and Nan YE**

** College of Mechanical and Electronic Engineering, Nanjing University of Aeronautics and Astronautics, 29 Yudao St., Nanjing, Jiangsu, China
E-mail: zhangly@nuaa.edu.cn

Abstract
A novel algorithm is proposed for registration of two partially overlapping range images that are arbitrarily oriented. The algorithm is inspired by the Shape Index (SI) description of 3D geometry and the Scale Invariant Feature Transform (SIFT) algorithm for detecting and matching features between intensity images. For a range image pair to be registered, we first construct an artificial intensity image pair based on SI. Then, the SIFT algorithm is used to extract and match features between the artificial images, and accordingly the 3D feature matches between the range images can be established with ease. A RANSAC procedure, combined with a hysteresis thresholding scheme, is carefully designed to filter out the false 3D feature matches to ensure the robustness. Finally, the transformation is computed according to all correct 3D feature pairs to register the two range images. Experiments and comparisons are included to verify the superior robustness and effectiveness of the proposed approach.

Key words: Range Image, Intensity Image, Registration, Image Matching, Shape Index

1. Introduction

With the features of high accuracy, inexpensiveness and non-intrusiveness, range scanners are the common choice for digitizing 3D objects in many applications, ranging from the automotive industry to the cultural heritage protection. However, to produce a complete surface of the object to be digitized, single scans (i.e., range images) acquired from different scan position and orientation must be transformed into a common coordinate frame. This process is termed as registration. Automatically registering multiple scans without the needs of additional devices has been a hot topic for years. Though being fruitful, the research on this issue is still on the way.

The iterative closest point (ICP) algorithm presented by Besl and Mckay \(^{(1)}\) is a cornerstone for scan registration. It iteratively searches for closest point pairs in two surface patches and optimizes the transformation to minimize the distances between these point pairs. However, since this algorithm implicitly assumes that the closest point pair on different patches corresponds to the same physical point, it only converges toward a local minimal solution if the patches are poorly pre-aligned. Various improvements and variants of the original ICP has been proposed, including the iterative closest compatible point algorithm (ICCP)\(^{(2)}\), which verifies the closest point pairs by additional attributes like color or surface normal. Furthermore, more sophisticated optimization schemes are proposed such as the simulated annealing \(^{(3)}\) and the evolutionary algorithm \(^{(4)}\). Rusinkiewicz and
Levoy [5] and Rodrigues et. al. [6] provided good surveys over the ICP variants. Although these proposed measures can improve the convergence property and the registration accuracy of the original ICP algorithm to some extent, they still do not allow for a registration of completely unaligned surface patches in reasonable time. In practice, one often resorts to interactive operations for an acceptable initial alignment, and then uses the ICP algorithm to obtain more accurate registration. Therefore, how to quickly and robustly estimate an approximate transformation is critical for fully automatic registration of two arbitrarily orientated scans.

A common approach to estimating the approximate transformation is to align a small set of corresponding feature points (i.e., 3D features) in the range images. The feature points typically are extracted and matched by geometry-based local shape descriptors, such as spin image [7], integral volume [8], shape context [9], and multi-scale feature [10]. These methods match at least three pairs of corresponding feature points and compute the rigid transformation by singular value decomposition (SVD) [11] or quaternion method [12]. An extensive review and comparison can be found in Salvi et. al. [13]. The approaches that directly based on geometry feature correspondence primarily differ in their definition of feature points and the way they are matched. A common drawback of these pure geometry approaches is that they rely on the existence of a sufficient number of salient feature points in the geometry.

The detection of feature points was also well established in 2D intensity image applications. Many feature extracting and matching methods for intensity image, as surveyed in [14,15], have shown great applicability and more attractive capabilities than that for 3D geometry. Considering the desirable properties of 2D image feature detection methods and scanning devices commonly capture not only geometry but also texture or light intensities of the scene, it is not surprising that exploiting 2D features for registration of range images is no longer a new idea. Roth [16] and Wyngaerd et. al. [17] use Harris feature detector [18] to extract features from the object’s texture images associated with the range images. Then the range image registration is realized by combining the texture and shape features. Texture-image-aided range image registration has drawn more and more attentions in recent years [19,20]. Some researchers also make effort on adapting algorithms that are originally designed for detecting and matching features of intensity image to range image registration applications [21].

The texture-image-aided approaches are suitable for the objects which are covered with rich texture, however, the original texture of many objects may be indistinguishable or effected by shading conditions and views of observation.

In this paper, we concentrate on automatic registration of two arbitrarily orientated scans. A new range images registration algorithm is presented, in which two artificial intensity images are generated from each range image pair first, and then feature matching between the artificial image pair is conducted to realize the pair-wise registration. The artificial intensity image pair actually acts as an intermediate of the registration process. The way used for generating the artificial images, as well as the idea for combining a RANSAC procedure and a hysteresis thresholding scheme to filter out false matches, discriminate the proposed method from others. We investigate its applicability for range image registration, and verify its superior robustness and effectiveness by experiments and comparisons.

The rest of the paper is organized as follows. The method for generating the artificial image is given in Section 2. Section 3 describes how to robustly register range image pair based on the artificial image. Registration experiments on both synthetic and real range images are presented in Section 4. The last section concludes the paper.

2. From Range Image to Artificial Intensity Image
A range image is typically organized in a 2.5D structure, i.e., a 2D grid arranged regularly like a chessboard, together with a depth value at each lattice. For the range images captured by the scanners to be registered, there are probably some lattices (i.e., pixels) having no corresponding depth value. If a pixel in the range image corresponds to a depth value, it is called valid pixel. A valid pixel in range image represents a point in 3D space. We use the step discontinuity constrained algorithm [22] to triangulate the range images and then estimate the curvature at each vertex of the discrete mesh surface as suggested by Dong [23].

Then we compute the smallest rectangular bounding box containing all the valid pixels. The artificial intensity image to be created is in the same size as the box. There is a natural one-to-one mapping between the valid pixels in the artificial intensity image and the valid pixels in the range image (i.e., points in 3D space). The next step is to set the intensity for every pixel of the artificial image.

An intuitive way is to compute some local properties of the 3D points in the range image, and then convert these properties into the corresponding intensities. The first requirement for the potential local properties is that they must be invariant to rigid transformation, and the second one is that the derived artificial image should contain rich information so as to facilitate the subsequent image matching. There are quite a few possible local properties of range image that could be used for computing the intensities, such as the principal curvature, Gaussian curvature, the mean curvature, the curvedness [24] and Laplacian smoothing feature [25].

Although all above properties are invariant to rigid transformation, and offer some insight into the intrinsic local shape, these properties must be linearly mapped into intensity interval [0,255] by the following formula

\[ I(v_i) = 255 \times \frac{F(v_i) - \min\{F(v_j)\}_{j=1}^{n}}{\max\{F(v_j)\}_{j=1}^{n} - \min\{F(v_j)\}_{j=1}^{n}} \]  

(1)

where \( F(v_i) \) and \( I(v_i) \) denote the property value of vertex \( v_i \) and its corresponding intensity. Since the \( \max\{F(v_j)\}_{j=1}^{n} \) or \( \min\{F(v_j)\}_{j=1}^{n} \) are usually different for two range images, Eq.(1) could map the same property value in the two range images to different intensities. This probably results in poor matching between the two artificial images. In addition, the above mentioned properties mainly represent some curved extent of the local shape, with no care of the local shape type.

Instead, we introduce the “shape index” to the artificial image generation and verify its superiority for range image registration applications. The “shape index”, originally put forward by Koenderink [24] and modified by Dorai and Jain [26] is defined as

\[ S(v) = \frac{1}{2} \cdot \frac{1}{\pi} \cdot \tan^{-1} \frac{k_1(v) + k_2(v)}{k_1(v) - k_2(v)} \]  

(2)

where \( k_1 \) and \( k_2 \) are the two principal curvatures.

Every distinct surface shape corresponds to a unique value of shape index (except for planar surfaces). The shape index classifies a surface into nine types and their locations on the shape index scale are shown in Fig.1. The shape index captures the intuitive notion of ‘local’ shape of a surface and provides a continuous gradation between salient shapes, such as convex, saddle, and concave. Thus, it has good ability to describe subtle shape variations.

Figure1. Nine well-known shape types and their locations on the shape index scale.
With Eq. (2), all shape indexes are mapped into the interval \([0,1]\). Therefore, for every valid pixel corresponding to point \(v\) in the range image, we directly set \([255 \cdot S(v)]\) as its intensity, where \([\cdot]\) denotes the rounding operation, while the intensities of invalid pixels are all set as 255. This setting ensures the same property values always correspond to the same intensities. We denote the intensity image generated in this way as SIBAI (Shape Index Based Artificial Image).

Fig. 2 shows a range image captured from an angel model in front view \([28]\) and the corresponding artificial image based on SIBAI. Compared to its origin texture image shown in Fig. 2c, the artificial image can not only give an intuitively natural representation but also be independent on the lighting and shading. As shown in Fig. 2b, the texture of the artificial image is rich and clear and captures the intuitive notion of local shape particularly well.

To further evaluate the performance of SIBAI, we compared it with three different artificial images, which are based on mean curvature, curvedness and Laplacian smoothing respectively, as shown in Fig. 3. Compared to the other three ones, the gray histogram of SIBAI in Fig. 3d distributes in a more moderate dynamic range, which means a richer texture. From observations on the four artificial images, we can also see that SIBAI contains much information and reveals more details of the model than the others. This advantage proves very helpful for subsequent 2D-feature extraction.
Figure 3. The artificial images and corresponding gray histograms. From (3a) to (3d), the artificial images are generated by using mean curvature, curvedness, Laplacian smoothing and shape index respectively.

3. Range Image Registration Based on SIBAI

3.1 Matching 2D-Features between the SIBAIs

For two partially overlapping range images, there must be two greatly similar regions in the two SIBAIs, since the intensities of the SIBAIs represent the intrinsic local geometry of the corresponding range image. However, because of the different scanning angles or depth of field for taking the range images, rotation, scale changes and/or affine transformation between the similar image regions inevitably exist. Among the algorithms for detecting and matching 2D image features, the SIFT [29] algorithm has been verified in Refs.[15]and [30] to be the most robust local feature detector and descriptor, invariant to scale, rotation, brightness and affine transforms. Therefore we employ the SIFT algorithm to extract and match 2D-feature points from the two SIBAIs.

Fig.4a shows two range images captured from the angel model in two views (degree 0 and 20 respectively). The two SIBAIs generated are shown in Fig.4b. Figure4c illustrates the matching feature pairs between the SIBAIs output by the SIFT algorithm, where the 2D-feature points are located at the centers of the dark crosses and the matching pairs are linked with solid lines.

According to the one-to-one map (as described in Section 2) between the range image and the artificial image, we can obtain the corresponding 3D-feature pairs between the range images from the matching 2D-feature pairs, except for those with at least one 2D-feature not corresponding to any 3D point in the range image. At first, we select the matching 2D-feature pairs ( $p_i^1, p_j^1$ ) in two SIBAIs, where $p_i^1$ and $p_j^1$ denote matching 2D-feature points in the first and second SIBAI respectively. Then, we determine whether the corresponding 3D points of $p_i^1$ and $p_j^1$ are all in the range images. If they are, we denote the corresponding 3D points as a pair ( $P_i^1, P_j^1$ ). In the angel example, 48 3D-feature pairs are obtained out of the 50 2D-features, as shown in Fig. 4d.

3.2 The RANSAC Procedure
Although the SIFT method provides relatively good matching results, false matches are nevertheless possible, as shown clearly in Fig. 4d. Since the registration is sensitive to such false correspondences, we apply an additional procedure based on the RANSAC (Random Sample Consensus) approach [31] to filter out the false 3D-feature pairs. In each RANSAC sample, three pairs are sampled randomly from all the matching pairs and a transformation is computed through Singular Value Decomposition (SVD) [11], then its support, i.e., the set of all pairs conforming to the transformation, is generated. More specifically, if the Euclidean distance between a 3D-feature pair after the transformation (also called “deviation” in the following text) is smaller than a predefined threshold (\(thr_{mat}\)), the corresponding 3D-feature pair are added into the support set. One way for determining the true transformation is selecting the one with the most elements in its support set. However, every three pairs must be exhaustively sampled in this way, which sometimes means a heavy computational burden. Considering that in our problem, the support rate of a good RANSAC sample is dramatically different from that of a RANSAC sample with obvious wrong match(es), we can simply reject a transformation if the cardinal number of its support set is less than a given percentage \(thr_{sup}\) of all the match pairs output by the SIFT algorithm. When the support rate of a transformation exceeds \(thr_{sup}\), the RANSAC procedure terminates. We denote the support set after the RANSAC procedure as \(S\). More discussions on \(thr_{mat}\) and \(thr_{sup}\) are included in the following section.

Figure 4 The workflow for registration of two range images. (4a) Two range images of the angel model. (4b) Two artificial intensity images. (4c) 50 matching 2D-feature pairs based on SIFT. (4d) 48 matching 3D-feature pairs in range images. (4e) 25 correctly matching 3D-feature pairs. (4f) The registration result.
According to the above analysis, while the obvious false 3D-feature pairs with gross error should be filtered out by the RANSAC procedure, the small deviation between the matching pairs after transformation should be tolerated to a certain extent to increase the number of conformal matches. Although three matching pairs can theoretically determine a rigid transformation, the redundancy of more correct matching features is critical for the robustness and accuracy of the registration result. This requires a trade-off between eliminating false matches and keeping reasonable matches. Considering that the resolutions differ with range scanners, we set \( thr_{mat} = m \cdot len_{edg} \), \( thr_{reg} = n \cdot len_{edg} \) (\( m > n \)), where \( len_{edg} \) denotes the average edge length of all triangles in the two range images. The two thresholds \( thr_{mat} \) and \( thr_{reg} \) constitute a hysteresis thresholding scheme, which was expected and has been proved to be helpful for robustness and accuracy of the algorithm. By consulting many experiments on different models, we assign \( m = 3 \), \( n = 2 \). It is easy to understand that the parameter \( thr_{sup} \) is independent of \( thr_{mat} \). Since the support rate of a good RANSAC sample is highly distinct from that of a RANSAC sample with obvious wrong match(es), \( thr_{sup} \) can be safely determined within a wide band. Specifically, given \( m = 3 \), feasible \( thr_{sup} \) is within 35%~65%. We assign \( thr_{sup} = 50\% \) in our experiments. Under above parameter settings, most experiments need no more than 6 RANSAC samples to filter out obvious false candidates, and fewer than 3 iterations to remove the matches with small deviations. In the angel example, 34 matching pairs remain in the support set after the RANSAC filtering, and 25 correct matching 3D-feature pairs survive at last, as shown in Fig.4e. Applied with the final transformation, the two range images are registered into a common reference frame as shown in Fig.4f, which illustrates that the alignment is good enough for further ICP optimization.

4. Experimental Results

4.1 Robustness Test

The robustness of registration mainly relates to the noise and overlapping extent of the two range images. In order to show that the proposed method is robust, we tested it on different range images with noise and pose variations. We take two models for examples: the brain and the valve in the OSU SAMPL database[28].

![Figure 5 The results of matching 3D-features extracted from range images disturbed by noise in different levels. (5a) The matching results of the brain model. (5b) The matching results of the valve model.](image_url)

For every model, we first register two range images with the same pose, yet corrupted with noise in various levels. More specifically, the depth values in the range image is disturbed with \( (0, \sigma) \) Gaussian noise. Let \( L \) denote the diagonal length of the 3D bounding box of the range image and define the noise ratio \( NR = \sigma / L \). Then we compare the correct matching number (CMN) of 3D-features with different NR levels. To verify the
superiority of the SIBAI, we also generate artificial images by using curvedness \cite{24} and Laplacian smoothing \cite{25} respectively for comparison. Figure 5 shows the comparing results. As Fig. 5a illustrated, our method extracts much more correct matches than the other two methods do given the same noise level. And when NR exceeds 0.006, the other two methods find nearly no correct matches, while our method still works until NR closes to 0.01. The comparison result of the valve model shown in Fig. 5b also illustrates that our method generally outperforms the other two methods.

![Figure 5](image)

Figure 5 The results of matching 3D-features extracted from the reference and the test range images. The reference range image is captured from pose degree 0, while the test range images are captured from different rotation angles. (a) The matching results of the brain model. (b) The matching results of the valve model.

In addition, we register two range images with different rotation angles (RA). Set the range image captured with pose of degree 0 as the reference range image (RRI), and the others captured from different rotation angles as the test range image (TRI). We count the CMN extracted between the RRI and the TRI, with increasing RA. For the brain test, we increase RA by 3.6 degree each time, and for the valve test, we increase 2 degree each time. Figure 6 shows the CMNs constructed between the RRI and the TRIs with different RAs. Our method is compared with other three methods in this test. Since the origin textures of the brain and valve model are low discriminative, the methods based on origin texture \cite{19,20} fail when RA is more than 20 degree. The registration resulted with Laplacian smoothing \cite{25} are similar to that with origin texture. The method based on curvedness works only when RA is less than 30 degree. In contrast, our method still works when the RA is bigger than 36 degree for the valve model, and 46.8 degree for the brain model. With different RAs, the corresponding CMN obtained by our method is obviously much more than that of others in most cases.

![Figure 6](image)

Figure 6 The results of matching 3D-features extracted from the reference and the test range images. The reference range image is captured from pose degree 0, while the test range images are captured from different rotation angles. (a) The matching results of the brain model. (b) The matching results of the valve model.
Figure 7 The results of matching and registration for the range images with rotation transform and noise disturbance simultaneously. The first two rows of (7a) and (7b) respectively are the results of matching 2D-features extracted from SIABIs and 3D features extracted from range images. The last rows in (7a) and (7b) are the final registration results.

Fig.7 shows the matching 2D-features extracted from two SIBAIs, the correct matching 3D-features extracted from the two range images and the aligned results by our method. Both of the two pairs of range images are corrupted with NR=0.002. RAs of the TRIs are 30 and 36 degree respectively. In order to display the matching pairs in the SIBAIs clearly, we hide all dark crosses and only show the matching 2D-feauters linked with solid lines. For this case, the other methods, including origin texture, curvedness, and Laplacian smoothing, all fail, i.e. can not find out any correct matches. All the above experimental results illustrate that our method is most robust to noise and to overlapping extent of the two range images among all the compared methods.

4.2 Effectiveness Test

Figure 8 Examples of matching 3D-features between range image pairs, by using the SIBAIs as intermedia.

In order to further illustrate the effectiveness of the proposed method, we select five models as same as that selected in Ref.[21]. Fig.8 shows the matching results of our method. Compared with the matching results in Ref.[21] (The readers are referred to Fig.21 in Ref.[21]), our results contain much more correct matches for each model. The two range
images of each model in Fig.8 were captured at pose angles -20° and 0° respectively.

The proposed method has been incorporated in the measurement system [32] we developed. Fig.9 shows the matching and registration results of different range images captured by our system. In Fig.9a, one range image is rotated about 180 degree compared to another range image; while in Fig.9b, there is almost only translation between the two range images. In Fig.9c and Fig.9d, the TRIs are translated about 30cm, and rotated about 30 degree with respect to the RRIs. The first two columns respectively are the results of matching 2D-features extracted from SIABIs and the 3D corresponding feature point pairs on the range images. The third column is the registration results of range images by our method. In this experiment, we compare the registration results by our method with the results by the 4PCS method [33], as shown in the last column of Fig.9. To quantitatively evaluate the registration accuracy, we calculated the standard deviations between the registered overlapping regions of the four range image pairs in the commercial software Geomagic Qualify. Table1 illustrates the statistical results. From Fig.9 and Tab.1 we can see that our method has obtained the better registration accuracy than the 4PCS method.

Figure 9 The matching and registration results for the range images captured by our measurement system [32]. The first two columns respectively are the results of matching 2D-features extracted from SIABIs and 3D features extracted from range images. The third column is the registration results of range images by our method, while the last column by
The times used for registering the four range image pairs in Fig.9 by using our method and the 4PCS method respectively are illustrated in Table2, which shows that the algorithm efficiency is largely on the same level.

<table>
<thead>
<tr>
<th>method used</th>
<th>Fig.9a</th>
<th>Fig.9b</th>
<th>Fig.9c</th>
<th>Fig.9d</th>
</tr>
</thead>
<tbody>
<tr>
<td>The proposed method</td>
<td>0.52</td>
<td>3.07</td>
<td>1.22</td>
<td>1.29</td>
</tr>
<tr>
<td>4PCS</td>
<td>1.96</td>
<td>7.06</td>
<td>2.75</td>
<td>2.13</td>
</tr>
</tbody>
</table>

5. Conclusions

In this paper, we have presented an approach to registration of two range images, which are partially overlapping and can be arbitrarily oriented. Our approach is based on constructing a pair of artificial intensity images from the two range images and the SIFT descriptors are used for matching 2D-features between the artificial image pair. By using the RANSAC and the hysteresis thresholding, the correct 3D-feature matches are finally obtained for registration of the two range images.

The proposed method has been compared to some other methods via experiments, the results of which illustrate that our registration approach is significantly more robust and effective than the other compared methods. The good performance should be attributed to the highly descriptive shape-index-based artificial images constructed, as well as the combination of the RANSAC and the hysteresis thresholding.

The limitation of our algorithm is that, analogous to most geometric-based methods, it fails to register surface patches with constant curvature (e.g. planes, cylinders and balls). In these cases, exploiting the original texture (if available) may be helpful.

Acknowledgements

This work is supported by National Science Foundation of China (Grant No.50875130). We thank Dror Aiger, Niloy J.Mitra and Daniel Cohen-Or for opening the source code of 4PCS method, the Signal Analysis and Machine Perception Laboratory at the Ohio State University for sharing the range image database.

References


