New point matching algorithm for panoramic reflectance images

Zhizhong Kang*a, Sisi Zlatanova

a Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands
b OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology Jaffalaan 9, 2628 BX Delft, The Netherlands

ABSTRACT

Much attention is paid to registration of terrestrial point clouds nowadays. Research is carried out towards improved efficiency and automation of the registration process. The most important part of registration is finding correspondence. The panoramic reflectance images are generated according to the angular coordinates and reflectance value of each 3D point of 360° full scans. Since it is similar to a black and white photo, it is possible to implement image matching on this kind of images. Therefore, this paper reports a new corresponding point matching algorithm for panoramic reflectance images. Firstly SIFT (Scale Invariant Feature Transform) method is employed for extracting distinctive invariant features from panoramic images that can be used to perform reliable matching between different views of an object or scene. The correspondences are identified by finding the nearest neighbors of each keypoint form the first image among those in the second image afterwards. The rigid geometric invariance derived from point cloud is used to prune false correspondences. Finally, an iterative process is employed to include more new matches for transformation parameters computation until the computation accuracy reaches predefined accuracy threshold. The approach is tested with panoramic reflectance images (indoor and outdoor scenes) acquired by the laser scanner FARO LS 880.

Keywords: Point matching, panoramic, reflectance image, scale invariant feature transform, Delaunay triangulation, point cloud, registration

1. INTRODUCTION

Presently, laser scanning techniques are used in numerous areas, such as object modelling, 3D object recognition, 3D map construction, and simultaneous localization and map building. One of the largest problems in processing of laser scans is the registration of different point clouds. Due to limited field of view, usually a number of scans have to be captured from different viewpoints to be able to cover completely the object surface. As well known, single scans obtained from different scanner positions are registered to a local coordinate frame defined by the instrument. Therefore the scans must be transformed into a common coordinate frame for data processing. This process is known as registration. Actually point cloud registration determines the transformation parameters bringing one data set into alignment with the other. The transformation parameters are computed by finding correspondences between different data sets representing the same shape from different viewpoints. Since the size of point clouds is usually pretty large, finding the best correspondence is a hard task. Commercial software typically uses separately scanned markers to help the identification corresponding points. Some vendors (e.g. Leica) have implemented algorithms (e.g. ICP (Besl et al, 1992)) allowing registering without markers but still the corresponding points have to be selected manually.

Presently, great effort is given to approaches based on segmentation of laser scan points and consequent matching of extracted features (e.g. Bae and Lichti, 2004; Mian et al., 2004). Features are derived from the point clouds and matched in a semi-automatic or automatic way. In general, feature-based methods might face problems processing large point clouds.

* Z.Kang@tudelft.nl; phone 31 15 278-8338; fax 31 15 278-2348
When the size of point clouds becomes huge, e.g. scans for outdoor scenes, the computation time for point cloud segmentation increases remarkably, which may require specific hardware. Moreover, point cloud registration based on feature-based methods may fail in cities where many planar patches are extracted (Dold and Brenner, 2006).

The approach presented in this paper is inspired by new developments in laser scan technology, i.e. a combination of geometric and radiometric sensors. In the last several years, many scanners have been equipped with image sensors. The 3D information captured by the laser scanner instrument is complemented with digital image data. Because of the generally higher resolution, optical images offer new possibilities in the discrete processing of point clouds. Several researchers have reported investigations in this area (Roth, 1999; Wyngaerd and Gool, 2002; Wendt, 2004; Dold and Brenner, 2006). Roth’s and Wyngaerd and Gool’s methods are similar to ours because they also use feature points based on texture. The difference is that Roth uses only the geometry of 3D triangles for matching and Wyngaerd and Gool use color texture information to drive the matching.

360° full scans are practically made to reduce the number of stations for scan and register in a high efficient way. As a result, the panoramic reflectance images are generated according to the angular coordinates and reflectance value of each 3D point of 360° full scans. It is quite difficult to make any assumptions on the set of possible correspondences for a given feature point as panoramic reflectance images are normally acquired from substantially different viewpoints and moreover the panoramic stereo pair doesn’t simply follow the left-and-right fashion.

This paper presents a new point matching algorithm for panoramic reflectance images. The approach follows three steps: extracting distinctive invariant features, identifying correspondences, pruning false correspondences by rigid geometric invariance. An iterative corresponding process is used to acquire more new matches can be included for transformation parameters computation to reach predefined accuracy threshold. Next section presents a detail description of the approach. Section 3 presents the tests and discusses the results. Section 4 concludes this paper.

2. METHODOLOGY

The proposed method consists of three general steps: extracting distinctive invariant features, identifying correspondences, pruning false correspondences by rigid geometric invariance. The last two steps are iterative by using computed transformation parameters between two point clouds behind the panoramic image pair, so that more new matches can be included for transformation parameters computation to reach predefined accuracy threshold. In this paper, the correspondence between image points (pixels) of two overlapping images is called pixel-to-pixel, the correspondence between image points and 3D points of a laser scan is pixel-to-point, and the correspondence between 3D points in two lasers scans is point-to-point correspondence.

The following sections explain in detail the algorithms used in the iterative process.

2.1 Extracting distinctive invariant features

Panoramic reflectance images, as we know, are normally acquired from substantially different viewpoints and moreover the panoramic stereo pair doesn’t simply follow the left-and-right fashion. Therefore, it is quite difficult to make
assumptions on the set of possible correspondences for a given feature point extracted by normal corner detectors. We use SIFT method (Lowe, 2004) to tackle this problem in this paper.

SIFT is a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The features are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images. For details, please reference to Lowe’s paper (Lowe, 2004).

2.2 Identifying correspondence

The invariant descriptor vector for the keypoint is given as a list of 128 integers in range [0,255]. Keypoints from a new image can be matched to those from previous images by simply looking for the descriptor vector with closest Euclidean distance among all vectors from previous images. In this paper, the strategy presented in (Lowe, 2004) is employed to identify matches by finding the 2 nearest neighbors of each keypoint from the first image among those in the second image, and only accepting a match if the distance to the closest neighbor is less than 0.8 of that to the second closest neighbor. The threshold of 0.8 can be adjusted up to select more matches or down to select only the most reliable. Please reference (Lowe, 2004) for the justification behind the determination of threshold of 0.8.

However, this strategy will identify false matches from panoramic reflectance images covering buildings, as building facades are likely to have repetitive patterns. For tackling this problem, the rigid geometric invariance derived from point cloud is used to prune false correspondences.

2.3 Pruning false correspondences by rigid geometric invariance

After the identification of matches, according to the 2D feature points in images, 3D corresponding points are taken from the laser scans on the basis of the known pixel-to-point correspondence.

In the local coordinate systems of different point clouds, Euclidean distance between each two corresponding point pairs is clearly invariant (Fig. 1). Namely, if point \(A\) and \(A'\), \(B\) and \(B'\), \(C\) and \(C'\) are corresponding points respectively, the distances between the points should equal (e.g. \(S_{AB} = S_{A'B'}\)).
It is theoretically possible to verify every two point pairs for distance invariance; however, this process may increase the computation time. To avoid this, we construct Delaunay Triangulated Irregular Network (TIN) and use the relations between the points in the triangles to decide on the distances. The TIN model is selected because of its simplicity and economy. It is also quite efficient alternative to the regular raster of the GRID model. Delaunay triangulation is a proximal method that satisfies the requirement that a circle drawn through the three nodes of a triangle contain no other node (Weisstein, 1999). As constructed for 3D corresponding points, the TIN model is only necessary to be constructed in one scan. Consequently, only those point pairs connected in TIN model will be verified for distance invariance.

The distance invariance error is estimated by error propagation law (e.g. Yu et al., 1989) according to the location error of each two corresponding point pairs. The difference between two distances is computed according to Eq. (1).

\[ Y_{DI} = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2} - \sqrt{(X_{A'} - X_{B'})^2 + (Y_{A'} - Y_{B'})^2 + (Z_{A'} - Z_{B'})^2} \]  

(1)

\(X_i, Y_i, Z_i\) are the 3D coordinates of a point, where \(i\) designates \(A, B, A',\) and \(B'\) respectively;

The location error of point \(i\) is determined by the laser scanner accuracy. As Boehler (Boehler et al., 2003) has pointed out, the laser scanner accuracy depends on many factors: angular accuracy, range accuracy, resolution, edge effects and so on. Among all, angular and range accuracy are mostly used for a laser-scanning instrument. Here we also use them to estimate the location error. If the coordinates of a point \(i\) are computed by a range value \(R_i\), horizontal angle \(\phi_i\), and vertical angle \(\theta_i\), the location accuracy is then determined by angular \(\sigma_\theta\) and \(\sigma_\phi\) and range accuracies \(\sigma_R\), as derived from the following equation:

\[
\begin{align*}
X_i &= R_i \cos \theta_i \cos \phi_i \\
Y_i &= R_i \cos \theta_i \sin \phi_i \\
Z_i &= R_i \sin \theta_i
\end{align*}
\]

(2)

In general, \(\sigma_R\), \(\sigma_\theta\) and \(\sigma_\phi\) can be considered as constant per a laser scanner. Laser scanners for distances up to 100 m show about the same range accuracy for any instrument (Boehler et al., 2003). As a result, range accuracy can be considered as an invariant for the whole point cloud because scanned targets of terrestrial laser scanner are usually within 100m. Three times of error of distance invariance is chosen as threshold to determine the correct correspondence in our approach, i.e.:

\[ |Y_{DI}| < 3\sigma_{DI} \]

(3)
Where, $\sigma_{DI}$ is the distance variance error which is computed by Eq. (1) with respect to the error propagation. $\sigma_{DI}$ is related to each of the two corresponding pairs, therefore the threshold chosen here is self-adaptive instead of a constant. If the above condition is satisfied, those two point pairs are considered as corresponding.

### 2.4 Iterative corresponding process

As mentioned earlier, it is likely to identify false matches from panoramic reflectance images covering buildings using only invariant descriptor vector for the keypoint. Those false matches are pruned in previous section by rigid geometric invariance. In this section, we discuss how to find more correct matches by iterative corresponding process.

After pruning false correspondences, only correct matches are kept. A least-square adjustment, based on correct matches, computes the six transformation parameters (defining rotation and translation) between two point clouds behind the panoramic reflectance image pair. Using the transformation parameters computed in a previous iteration, the correspondences in the image pair can be better predicted which results in increased number matched points. The new matches are included in the computation of new transformation parameters. The iterative process continues until the transformation parameters reach predefined accuracy threshold.

#### 2.4.1 Computation of transformation parameters

As well know, single scans from different scan positions are registered in a local coordinate frame defined by the instrument. Using corresponding points detected at previous step, it is possible to compute transformation parameters between deferent coordinate frames and thus register the two point clouds.

The least-square parameter adjustment for absolute orientation in photogrammetry (e.g. Wang, 1990; Mikhail et al., 2001) is used to solve least-square optimized values of transformation parameters. Iterative process is implemented to acquire higher accuracy because error equations have been linearised.

It should be noticed that after the outlier detection, the wrong matched points are removed and the transformation parameters are computed only with the correct ones. However, the outlier detection may remove many points and the transformation parameters will be determined from very few points. Therefore, these parameters cannot be considered final. To be able to improve the transformation parameters, more points appropriate for matching have to be found. The candidate points are searched amongst the keypoints already extracted in section 2.1. Therefore an iterative process is implemented.

#### 2.4.2 Corresponding point prediction

Using the initial transformation parameters, the position of corresponding points in one image (The right one in this paper) can be predicted based on the extracted feature points in the other (The left one in this paper). As mentioned above, all the points on the left image, extracted by the feature point extraction algorithm are used in the iterative process. As presented earlier, based on image coordinate ($x$, $y$) of feature point in the left image, we can acquire the coordinate ($X$, $Y$, $Z$) of corresponding 3D point of left scan. Using the initial transformation parameters, the coordinate ($X'$, $Y'$, $Z'$) in right scan can be calculated from ($X$, $Y$, $Z$). The image coordinates ($x'$, $y'$) corresponding to ($X'$, $Y'$, $Z'$) are certainly the expected position of corresponding point in right image. Thereafter, a certain region centered at ($x'$, $y'$) is determined for searching exact corresponding point.

Each iteration accordingly consists of four steps, i.e. corresponding point prediction using transformation parameters computed from previous iteration, identifying correspondences, pruning false matches and transformation parameters...
computation. This iterative process ensures matching of larger number of points and reasonable distribution of corresponding point, which leads to improved values of the transformation parameters. The iterative process continues until the RMS error of transformation parameters computation satisfies a given threshold. This threshold is in the range of millimeter and is determined with respect to the range accuracy of the scanner.

3. RESULTS

The approach is tested with several panoramic reflectance image pairs generated from point clouds (indoor and outdoor scenes) acquired by FARO LS 880. The angular resolution selected for FARO LS 880 is 0.036° in both of horizontal and vertical directions which is a quarter of full resolution the instrument claims. Dataset 1 is acquired for the office environment and Dataset 2 is scanned for outside buildings. The proposed method was implemented in C++. All the tests are performed on a PC with CPU Intel Pentium IV 3 GHZ and 1 GB RAM.

As mentioned previously, the panoramic reflectance images are generated with respect to the angular coordinates and reflectance value of every 3D point of point cloud.

3.1 Indoor data set

In the reflectance images of FARO LS 880 (Fig. 2), the pixel-to-point correspondence is straightforward and corresponding 3D points are readily available in the data file.

As presented in Section 2, SIFT method was used to extract distinctive invariant features from panoramic images and matches were identified from keypoints by looking for the descriptor vector with closest Euclidean distance. 655 corresponding point pairs were identified (Fig.2), however many are false accepted. The rigid geometric invariance derived from point cloud was accordingly used to prune false correspondences. Strict threshold was employed to ensure only correct matches can be remained. As a result, only 99 correct corresponding points (Partially illustrated in Fig.3) were kept against 655 shown in Fig.2. Trying to include more new matches, as presented in Section 2.4, we used an iterative corresponding process to ensure matching of larger number of points and reasonable distribution of corresponding point.
As Fig.4, 676 corresponding point pairs were acquired after iterative process and 99% of them are correct.

The registration of Dataset 1 was implemented with those correct corresponding points. The registration accuracy is 1.1mm after 2 iterations and average distance between corresponding points is 2.7mm as shown in Table 1. Both are the order of millimeter. The whole process of our method cost 5 minutes.

Table 1 Result of proposed method on indoor datasets

<table>
<thead>
<tr>
<th>Proposed method</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$i$</th>
<th>RMS (m)</th>
<th>Max (m)</th>
<th>Min (m)</th>
<th>AVG (m)</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data set 1</td>
<td>11987424</td>
<td>11974976</td>
<td>2</td>
<td>0.0011</td>
<td>0.0351</td>
<td>0.0002</td>
<td>0.0027</td>
<td>5</td>
</tr>
</tbody>
</table>

Please note, all the notations in the tables are the same, i.e. $n_i$ is the total number of points of Dataset 1, $i$ is the number of total iterations. RMS is the accuracy of registration computed from the least-square parameter adjustment. Max, Min and AVG are respectively the maximum, minimum and average distance between 3D corresponding point pairs after registered in a common coordinate frame.

3.2 Outdoor data set
Dataset 2 consists of two point clouds of outside building. As Fig.5, the building facade has repetitive pattern, therefore, few corresponding points on the facade were kept after pruning false matches. By iterative matching process, plenty of correct corresponding point pairs on the facade were identified and the distribution of matches becomes even in the panoramic images (Fig.6). The registration result is listed in Table 2. The RMS is 4.4 mm and average distance between corresponding points is 4.8 mm. The whole process completed in 6 minutes after only 2 iterations.

Table 2 Result of proposed method on outdoor dataset

<table>
<thead>
<tr>
<th>Proposed method</th>
<th>(n_1)</th>
<th>(n_2)</th>
<th>(i)</th>
<th>RMS (m)</th>
<th>Max (m)</th>
<th>Min (m)</th>
<th>AVG (m)</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset 2</td>
<td>16726500 16713375</td>
<td>2</td>
<td>0.0044</td>
<td>0.0430</td>
<td>0.0008</td>
<td>0.0048</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

In this paper, a new point matching algorithm for panoramic reflectance images is presented and tested with several data
sets. The approach follows three general steps: extracting distinctive invariant features, identifying correspondences, pruning false correspondences by rigid geometric invariance. An iterative corresponding process is used to acquire more new matches can be included for transformation parameters computation to reach predefined accuracy threshold.

The point cloud registration implemented by corresponding points matched from panoramic reflectance images is able to acquire the accuracy of millimetre order. It is proven by the experiments that our algorithm is able to work without assuming any prior knowledge of the transformation between these images. To use the presented point matching algorithm there should be sufficient, i.e. at least 20% to 30%, overlap between image pairs. This degree of overlap is not difficult to ensure when collecting panoramic reflectance images.

REFERENCE


ACKNOWLEDGEMENT

This research was supported by the BSIK Project of The Netherlands “Virtual reality for urban planning and safety”.