2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

A Hough Transform Based Scan Registration Strategy for Mobile Robotic Mapping*

Bo Sun*, Weiwei Kong†*, Junhao Xiao† and Jianwei Zhang*
* Department of Informatics, University of Hamburg
Email: {bosun, zhang}@informatik.uni-hamburg.de
† College of Mechatronic Engineering and Automation, National University of Defense Technology
Email: kongweiwei@nudt.edu.cn, junhao.xiao@ieee.org

***Abstract*— The scan registration is the cornerstone to Mobile Robotic Mapping, and the majority of existing global registration methods are dependent on specific features. This paper presents a global feature-less scan registration strategy based on the ground surface, which is extremely common in Mobile Robotic Mapping scenarios. The 3D rotation is decoupled from 3D translation by transforming the input scans into the Hough domain, wherein Phase Only Matched Filtering (POMF) is adopted for the partially overlapped signal registration. No particular features in the input data are prerequisite to our algorithm. The algorithm is validated by the challenging scans captured by our custom-built platform and a public dataset. The result illustrates the reliability of this algorithm to align feature-less, partially overlapped and noisy scans.**

I. Introduction

Acquiring the spatial models of physical environments is one of the fundamental issues in building truly autonomous mobile robots, since map is essential to the subsequent tasks, such as path planning and robot navigation. Range sensor readings are extensively accepted to create a precise map by efficient scan registration techniques. In the context of 3D mapping, the scan registration methods could be categorized into *local* alignment methods and *global* alignment methods, according to an initial estimate is required or not. In practice, the local alignment methods are usually taken as a refinement step of some global alignment methods, which only give a coarse solution.

A. Local registration methods

The two most well-known local registration methods are Iterative Closest Point (ICP) [1][2][3] and Three-Dimensional Normal Distributions Transform (3D-NDT) [4][5][6]. The performance of ICP and NDT in the 3D mapping field has been evaluated in [7], wherein the relative advantages and disadvantages of the two algorithms were proposed. The local alignment methods have three distinct drawbacks:

1) requirement of initial guesses;
2) easy to get trapped in local minima;
3) runtime would vary considerably for different scan pairs, even scan pairs with same size.

* This work was supported by China Scholarship Council (CSC) and the EC Seventh Framework Program theme FP7-ICT-2011-7, grant agreement no. 287752.

B. Global registration methods

On account of the shortcomings of the local alignment methods, global methods that take the global appearance of 3D scan into account are increasingly popular. The majority of global registration methods are *feature-based* techniques, which make use of explicit feature correspondences in the environment, such as Spin Image [8], Fast Point Feature Histograms (FPFH) [9], Depth-Interpolated Local Image Features (DIFT) [10], Signature of Histograms of Orientation (SHOT) [11], Shape Context (SC) descriptor [12], Normal Aligned Radial Feature (NARF) [13] and planar patches [14].

Recently, a number of global *feature-less* registration algorithms have been developed based on global descriptors of the 3D point cloud. The transformation could be solved by aligning the corresponding global descriptors. A fully automatic registration method based on Extended Gaussian Images (EGIs) is presented in [15]. This EGIs-based algorithm makes use of the spherical harmonic transform to correlate two EGIs in the Fourier domain. This algorithm is reported to be capable to deal with the registration problem wherein the two input scans have arbitrarily large displacement and very little overlap. However, it has two blind-spots: (1) it is only feasible to smooth surface, but causes problems at discontinuities; (2) it fails to deal with the scans contain spherical objects, which could lead to constant histograms and less informative EGIs. In the Hough Scan Matched (HSM3D) registration method [16], the Hough/Rough Transform is used to decouple the 3D rotation information from the 3D translation. Although the HSM3D is feature-less in theory, the first part of rotation determination is based on the local minima of Hough Spectrum and the translation recovery is merely dependent on the Hough Transform with several selected direction, so the HSM3D is prone to fall down. To improve the stability of HSM3D, a set of candidate solutions are produced, hence the HSM3D is computationally expensive and an extra evaluation step is necessary to pick out the correct solution. A more recent frequency-based approach SRMR [17] resamples the spectral magnitude of 3D Fast Fourier Transform (FFT) calculated on discrete Cartesian grids of the 3D range data to detach the 3D rotation from 3D translation. This method works only within a restricted range of roll and pitch changes between scans.

Our group has presented an odometry-free scan matching

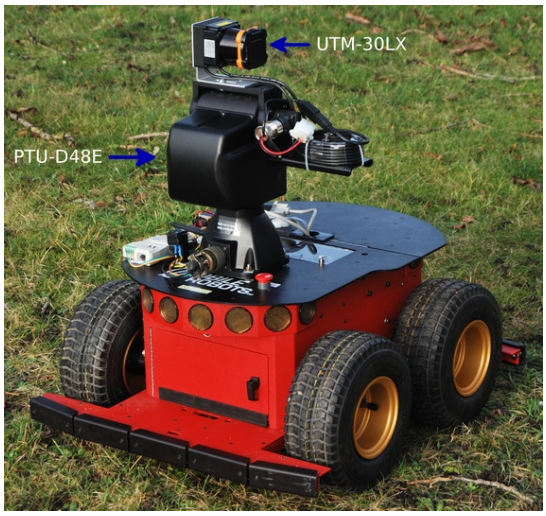


Fig. 1. The custom-built 3D perception platform consists of three main parts, namely a 2D laser range finder (Hokuyo UTM-30LX), a pan-tilt unit (FLIR PTU-D48E) and a mobile robot (Pioneer 3-AT). The FLIR PTU-D48E unit yaw the 2D LRF to produce a proper 3D LRF.

method based on area attributed planar patches [14], but this method works only in the urban area and fails in plane-less scenarios.

Actually, our algorithm is partially inspired by the plane-based registration methods which use the Hough Transform for planes extraction and Hough Transform based feature-less registration method HSM3D. The technique presented in this paper projects the scan data into Hough domain as well as the HSM3D; in this way, the 3D rotation is separated from 3D translation. With regard to rotation recovering, the rotation between two scans could be decomposed into one rotation paralleling two ground surfaces and another rotation compensating the roll about the normal vectors of ground surfaces. After the rotation information is recovered, the translation is quite easy to determine. Both the rotation determination and translation recovering are solved by Phase Only Matched Filter (POMF). The novel algorithm has been experimentally validated on a public dataset and the challenging scans captured by our custom-built platform, as shown in Fig. 1. The technique presented in this paper has the following characteristics:

- ***feature-less***: this method does not depend on specific features.
- ***noise-immune***: this algorithm uses the whole Hough spectrum, so it could resist the noise effectively.
- ***partially overlapped***: the POMF is adopted to effectually deal with the partially overlapped scans.

The rest of the paper is organized as follows: the mathematical fundamentals involve our algorithm are presented in section II; section III describes the algorithm in all subsequent steps; section IV proposes a brief algorithm analysis and several tips in the algorithm implementation; experiments and results with the challenging data are presented in section V; and finally section VI concludes the paper and points out the future work.

II. Mathematical Preliminary

A. Hough Transform Descriptor (HTD)

The Hough Transform is an extensively accepted method for detecting parametrized objects [18], and is widely used in plane-based registration methods. Whereas the plane extraction is inevitable error-prone and the final goal is scan registration rather than plane detection, we are inspired to use Hough Transform to align the scans directly by regarding the Hough Transform of the scan as a global descriptor.

The 3D Hough Transform maps the point cloud into Hough Space defined by (θ, ϕ, ρ), such that each point in the Hough Space corresponds to one plane in $\mathbb{R}^3$. In addition, θ stands for the angle between the normal vector of the plane and xy plane, ϕ is the angle between the projection of the normal vector on the xy-plane and x axis, and ρ represents the distance of the plane to origin. Although the normal Hough Transform method takes pixel/voxel images as the input, it is not a requisite and a set of unorganized points in $\mathbb{R}^3$ could also be used as the input of Hough Transform just like the HSM3D.

The Hough Transform descriptor (HTD) of the 3D point cloud could be defined as follows:

$$HTD(\theta, \phi, \rho) = \sum_{i=1}^{N} \delta(< p_i, \vec{n} > -\rho) \tag{1}$$

$$\vec{n} = [\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta] \tag{2}$$

where p_i is the i_{th} point and N is the number of points. In practical implementation, the Hough Space is divided into discrete cells, and the accumulators related to cells increase with respect to the scores computed by (1). For each point p_i, all the cells getting in touch with its Hough Transform should increase. An exhaustive interpretation of the Hough Transform techniques used in scan registration is available in [19].

B. Translational Invariant Descriptor (TID)

The HTD has two significant properties which qualify it to be a global descriptor for scan registration. Firstly, the rigid transformation between two scans corresponds to the transformation of their HTDs, as (3) shows, where $R \in SO(3)$, $t \in \mathbb{R}^3$ and $HTD\ |_X$ means the HTD of scan X. Note that the translation between HTDs is dependent on direction, which means $\hat{t} \neq t$. This property converts the scan registration problem to determining the transformation between their corresponding HTDs.

$$HTD\ |_{(R \bullet S+t)} = R \bullet HTD\ |_S + \hat{t} \tag{3}$$

Secondly, the (θ, ϕ) parameters of HTD are related to the rotation of the scan, while the ρ parameter is correlated with the translation. In other words, the HTD of the rotated duplication of a scan maintains the ρ invariable, while the HTD of the translated duplication leaves (θ, ϕ) alone. This property could be used to decouple the 6DOF transformation into 3DOF rotation and 3DOF translation.

Integrating the $\| \bullet \|_2$ of HTD with regard to ρ, the HTD could be mapped onto a spherical surface $\mathbb{S}^2$ and then obtain the translational invariant of HTD. Denote the translational invariant as TID, then

$$TID(\bar{\theta}, \bar{\phi}) = \sum_{\rho} \|HTD(\bar{\theta}, \bar{\phi}, \rho)\|_2; \quad (\bar{\theta}, \bar{\phi}) \in \mathbb{S}^2 \tag{4}$$

$$TID\,|_{(R \bullet S+t)} = R \bullet TID\,|_S \tag{5}$$

In this way, the rotational alignment could be achieved in the TID domain regardless of the translation. In this paper, the rotation is first determined in the TID domain regardless of ρ, and then the 3D translation is solved based on the determined rotation matrix.

C. Phase Only Matched Filtering (POMF)

Due to its ability to deal with the partially overlapped problem between the signals, the cross-correlation technique is used to determine both the rotation and translation parameters in this paper. However, the value of the standard cross correlation method is heavily dependent on the energy of underlying signals, so it often fails to discriminate the signals which are of different shapes but similar energy. Furthermore, the correlation peaks could be relatively broad depending on the signal structures, which makes it difficult and unreliable to locate the correct displacement between noisy signals. Concerning the scan registration problem, if a disproportionally large value point of HTDs is not present in the overlap area, the standard cross correlation method would fail to recover the rotation information. In order to decrease the impact of such detrimental points and achieve more distinct sharp peaks, the POMF algorithm [20] is used to resolve the signal correlation problem in this paper. The POMF technique has extraordinary significance in transform-based registration methods as regards partially overlapped scans.

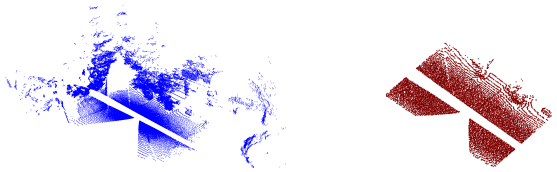
The POMF decouples the local signal energy from the signal structures based on the fact that two shifted signals carry the shift information within the phase of their Fourier spectrum. Let $f_1(x, y, z)$ and $f_2(x, y, z)$ be two shifted signals, and $\mathcal{F}_1(u, v, k)$ and $\mathcal{F}_2(u, v, k)$ be their corresponding Fourier spectra. The shift between these two translated signals could be solved by the following equations:

$$S(u, v, k) = \frac{\mathcal{F}_1(u, v, k)^*}{|\mathcal{F}_1(u, v, k)|} \bullet \frac{\mathcal{F}_2(u, v, k)}{|\mathcal{F}_2(u, v, k)|} \tag{6}$$

$$s(x, y, z) = \mathcal{F}^{-1}\{S(u, v, k)\} \tag{7}$$

$$(x_p, y_p, z_p) = \arg\max_{(x,y,z)} s(x, y, z) \tag{8}$$

where $*$ indicates the complex conjugate, and (x_p, y_p, z_p) is the displacement between the two signals. In theory, it could be used in arbitrary dimensional signal registration problems. Ideally, the $s(x, y, z)$ contains a Dirac peak, but the Dirac pulse deteriorates in practical due to the noise. Meanwhile, a number of variants are developed to improve POMF for subpixel accuracy and the subpixel interpolation method proposed in [21] is adopted in this paper.



(a) (b)

Fig. 2. (a) An example scan from `Gazebo Winter` [22]; (b) The corresponding loose ground points extracted by our algorithm.

III. Description of Algorithm

Consider two scans s_1, s_2 with the relationship:

$$s_2 = R \bullet s_1 + t + \varepsilon$$

where $R \in SO(3)$, $t \in \mathbb{R}^3$, ε is the noise, and their Hough Transform descriptors are HTD_1 and HTD_2, their translational invariant descriptors are TID_1 and TID_2. Our algorithm first recovers the R in the TID domain of the two scans; and then applies the obtained R to the HTDs to solve the translation parameters.

A. Estimation of ground surface direction

Few literature involves the *automatic* extraction of the ground surfaces in our review of published research. In [23], two methods are presented to classify ground and non-ground points, and then approximate the ground surface model. But both of the two approaches are computationally expensive. Considering the scan pattern of the scanner, the density of points is greater closer to the scanner, as shown in Fig. 2 (a). We develop a method to estimate the normal vector of the ground surface based on this prior knowledge, which needs less extra calculation.

In our ground surface estimation method, the xy-plane segment closer to the scanner is divided into square grids, the dimension of which is a compromise of accuracy and efficiency. Normally the range of this plane segment could be a quarter of the whole xy-plane. And then the 3D points, whose (x, y) values are within the scope of the plane segment, are projected into the cells according to their (x, y) coordinates. For each cell, the point of the smallest z value is stored as loose ground surface points; in this way, the scan is classified into loose ground points and non-ground points. The loose ground points of the scan depicted in Fig. 2 (a) are shown in Fig. 2 (b). Our method could not be used to segment the precise ground surface, since it extracts not only the points belonging to the ground surface but also the bottoms of objects. However, our method is capable to deal with the slanted surfaces.

Theoretically, computing the TID of the loose ground points, the normal vector of the ground surface is determined by the main peak of TID. After that, the computation of TID could continue by calculating the Hough Transform of the non-ground points; hence only the operation to separate the loose ground points costs extra computation. Whereas in practise, we found that the normal vector determined by the main peak of TID is sometimes correct with its z

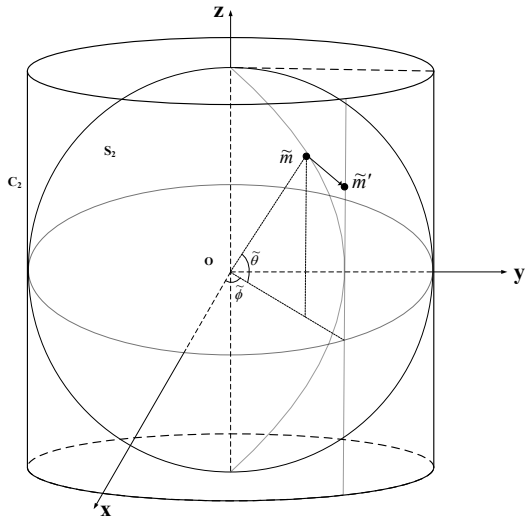


Fig. 3. Mapping from a spherical surface onto its corresponding cylinder surface.

direction but false with its x or y direction. Usually, the second or third large local maximum of TID corresponds to correct direction of ground surface. Therefore, we adopt the RANSAC (*RANdom SAmple Consensus*) based data fitting algorithm to estimate the direction of the ground surface in our code. In addition, we believe this is the same reason why the HSM3D produces several candidate results to improve the robustness.

B. Rotation Recovery in Feature-rich case

Equation (5) signifies that the analytic solution of R could be achieved if more than two corresponding points were found. The local maxima with relatively large values of HTD are the candidate planes in the physical space. The relationship between local maxima could be used to find the corresponding points of TID. If three corresponding local maxima which are of relatively large TID values were found, the three local maxima could be used to solve the analytic solution of R based on equation (5).

Theoretically, two corresponding points are enough to solve equation (5), but at least three points are required to resist the ambiguity and noise from our practical experience. Admittedly, more corresponding points could refine the closed-form solution through the least square optimization, but three corresponding points are difficult to derive, not to mention more. The reason for that is the TID integrates the candidate planes of the same direction into one point in the TID domain, so the three local maxima of TID stand for candidate planes with three different directions. Moreover, the algorithm requires the maxima that are of relatively large values to prevent the noise. Consequently, the scenarios applicable to this case are uncommon in practice. This motivates us to develop the feature-less registration algorithm.

C. Rotation Recovery in Feature-less case

According to *Euler's rotation theorem*, a rotation could be decomposed into two rotations in 3D space. The rotation R between two scans could be decomposed into R_1 paralleling two ground surfaces and R_2 compensating the roll about the normal vectors of ground surfaces. Projecting this proposition into the TID domain, the rotation R_1 paralleling the two peaks of TIDs stand for normal vectors of ground surfaces, and the rotation R_2 compensating the rotation of TIDs about the two peaks.

Let the g_1 and g_2 be normal vectors of ground surfaces, the R_1 and R_2 could be easily described by the axis-angle representation: 1) the rotation axis of R_1 is $g_1 \times g_2$ and the rotation angle is the angle between g_1 and g_2; 2) the rotation axis of R_2 is g_2 and the rotation angle of R_2 is symbolized by δ. The R_1 could be computed based on g_1 and g_2 directly, so more attention is paid to the R_2 calculation.

There are several methods to determine the rotation angle δ using the TIDs [16]. The algorithm in this paper converts the estimation of δ to a translation recovering problem by projecting the TID defined on $\mathbb{S}^2$ onto its corresponding cylinder surface $\mathbb{C}^2$ whose axis parallels the normal vector of the ground surface, as Fig. 3 outlines. In this way, the radial translation of the $\mathbb{C}^2$ corresponds to the rotation angle δ of $\mathbb{S}^2$, and it could be settled by the cross-correlation techniques, for instance POMF in this paper.

As shown in Fig. 3, let S_2 be a spherical surface and C_2 be its corresponding cylinder surface. Consider a point $\tilde{m}$ on S_2, its the spherical coordinate is $(\tilde{\theta}, \tilde{\phi}, 1)$, and its Cartesian coordinate is $[\cos(\tilde{\theta})\cos(\tilde{\phi}), \cos(\tilde{\theta})\sin(\tilde{\phi}), \sin(\tilde{\theta})]$. And then, re-sampling the Cartesian space by the cylinder coordinate system, the cylindrical coordinate of $\tilde{m}$ could be derived: $[\sin(\tilde{\theta}), \tilde{\phi}, \cos(\tilde{\theta})]$. Finally, map $\tilde{m}$ onto C_2 through normalizing its radial distance to the axis of the cylinder. In such a way, a one-to-one projection $\tilde{m}'$ of $\tilde{m}$ is gained:

$$\begin{aligned}
&\tilde{m} = (\tilde{\theta}, \tilde{\phi}, 1) \\
&\xleftrightarrow{\text{Cartesian resample}} (\cos(\tilde{\theta})\cos(\tilde{\phi}), \cos(\tilde{\theta})\sin(\tilde{\phi}), \sin(\tilde{\theta})) \\
&\xleftrightarrow{\text{Cylindrical resmaple}} (\cos(\tilde{\theta}), \tilde{\phi}, \sin(\tilde{\theta})) \\
&\xRightarrow{\text{radial distance normalization}} \tilde{m}' = (1, \tilde{\phi}, \sin(\tilde{\theta}))
\end{aligned}$$

After mapping the points of S_2 onto C_2, the rotation angle δ of S_2 around the g_2 is converted into the displacement of the cylinder surface along the tangential direction. Fig. 4 proposes the ground-calibrated TIDs projected on cylinder surfaces. It is obvious that there is only tangential shift left after the compensation using R_1. Achieving the cylinder surfaces, the determination of R_2 is converted to a uni-dimensional signal registration issue. In this paper, the two dimensional POMF is used since the shift along the axis could be applied to validate whether the R_1 is solid or not.

D. Translation determination

Supposing the determined rotation R is correct, there is only translation between the two scans after applying the determined R to them.

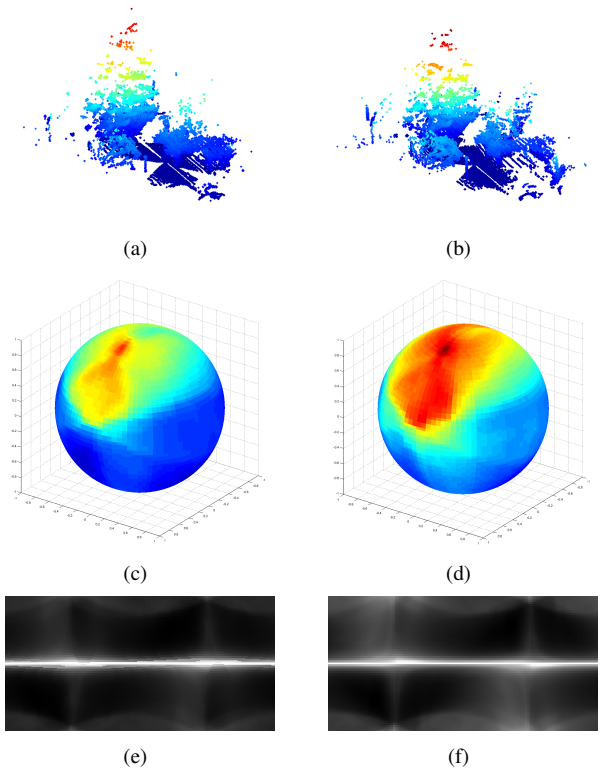


Fig. 4. Examples of Translational Invariant Descriptor (TID): (a) and (b) are the first and fourth scan of dataset `Gazebo Winter` [22]; (c) and (d) are the corresponding TIDs; (e) and (f) are the corresponding ground-calibrated TIDs projected on cylinder surface after R_1 compensation.

The HSM3D algorithm selects several directions and derives the uni-dimensional slices of HTD along the specific directions; afterwards the translation could be calculated on the basis of the uni-dimensional slices. Although this method is feature-less theoretically, it only makes use of several slices of HTD rather than the whole appearance of scans.

The 3D POMF could be adopted to estimate the translation between two shifted scans straightforward, while it is also feasible to integrate the scan along the x, y, z direction and apply the 1D POMF to calculate the translation separately. We use the second method in our algorithm since the 1D POMF is efficient and the 3D POMF is memory-consuming.

Before applying the POMF, it is necessary to rasterize the original 3D surface into volume grids. In general, the way to rasterize the 3D surface is assigning a voxel the value of 1 if it is occupied by the surface, otherwise its value is set to be 0. The Euclidean Distance Transform is applied to 3D surfaces in [24], and then the value at each voxel is given by the negatively exponentiated Euclidean Distance Transform of the point located in it. In our algorithm, we estimate the Gaussian curvature of the 3D surface and assign a voxel the value of the Gaussian curvature on the point it contains. In such a way, the richly structured part of the surface plays a more important role in registration.

Comparing to the rotation determination procedure, the translation recovering is more straightforward and needs less computation.

INPUT: ***Point clouds*** S_1, S_2
OUTPUT: ***The rotation matrix*** R ***and translation vector*** $t = (t_x, t_y, t_z)$
PROCEDURE:

1. ***Loosely extract the points belonging to the ground surfaces and estimate the directions of grounds*** g_1, g_2.
2. ***Compute the HTDs of point clouds*** H_1, H_2 ***and TIDs*** T_1, T_2.
3. ***Find the peaks of*** T_1, T_2 ***and sort the peaks by their TID values.***
4. ***Recover the rotation matrix*** R ***:***
   a. ***If more than three corresponding large peaks of TIDs are available, the*** R ***could be solved easily based on equation (5).***
   b. ***If not, then firstly compute*** R_1 ***paralleling*** g_1, g_2***; and secondly evaluate the rotation angle about*** g_2 ***using POMF to calculate*** R_2***.*** $R = R_2 \bullet R_1$
5. ***Determine translation*** t***:***
   a. ***Rerotate the point clouds according to*** R***.***
   b. ***Rasterize the point clouds into volume grids.***
   c. ***Adopt the 1D POMF to calculate the*** (t_x, t_y, t_z) ***separately.***

Fig. 5. An outline of the registration algorithm proposed in this paper.

Lastly, to conclude this section, the algorithm is outlined in Fig. 5.

IV. Algorithm Analysis and Implementation

A. Accumulator design of Hough Transform

The first accumulator design problem is the discretization of the parameters (θ, ϕ, ρ). The precision of the final solution is closely related to the resolution of accumulators. A trade-off has to be compromised between the coarser discretization which is more noise-immune and efficient but presents low accuracy, and the finer discretization offers more precision but occupies more resource. However, as to the global registration techniques designed for initial crude alignment, the coarser grids are preferred.

The second accumulator design problem is how the accumulator cells correspond to patches of the unit sphere. The classic manner samples the unit sphere uniformly in the spherical longitudinal and azimuthal coordinates, thus the accumulator cells correspond to the patches with equivalent polar and azimuthal angles. This manner is straight and intelligible, but the area of patches are disparate. The patches closer to the equator have larger areas, and the patches closer to the poles are smaller; thus accumulator cells corresponding to patches closer to the equator are correlated to more normal vectors. Another accumulator is presented in [16], and their solution projects the unit sphere onto the smallest cube that contains the sphere. Each face of the cube is discretized regularly. This design is a trade-off between efficiency and manipulability, but the area inequivalent problem is still

unsettled. The accumulator designed in [19] leads to each accumulator cell corresponds to the equal patch area. This design samples the unit sphere uniformly in the longitudinal coordinate, but the azimuthal space of the patch is determined by its longitudinal coordinate to make sure the patches have equal area. Apparently, the azimuthal space between adjacent cells is irregular, and this disorder makes the projection from unit sphere onto cylinder obscure to express. In this paper, the classic manner is adopted, further the values of the accumulator cells are normalized by the area of their corresponding patches on the unit sphere.

B. Inhomogeneity of points cloud

Owing to the scan pattern, the density of the points cloud is inhomogeneous. Usually the ground points closer to the scanner are excessively dense compare to others. Admittedly, the lopsided density is beneficial for ground extraction and the following R_1 calculation. But sometimes, this undue imbalance will overplay the role of the ground and could not fully reveal the structure of objects above the ground, especially when there are sparse constructions in the scene. And the conformation of the objects is crucial to determine the shift of the cylinder and the consequent R_2 computation. To solve this dilemma, the dense points on the ground are used to estimate the direction of the ground, while the points on the loose ground are thrown out of the R_2 determination procedure.

C. Increments of accumulator cells

A physical point represents a sinusoid surface in Hough space, and normally all the accumulator cells related with the sinusoid surface are incremented by $\mathbf{1}$. Due to the inhomogeneity of the points cloud, the planes with equal numbers of points have variable size depending on their distances to the scanner. To solve this ambiguity, the accumulator cells are recommended to be incremented by the area of the points. The area of points could be calculated by its four surrounding closest points. If the scan numbers and the point numbers in the scans are available, the area calculation is straightforward and does not require the computationally expensive nearest neighbour search procedure.

D. Dependency on Ground Surface

From the theoretical derivation of this algorithm, it seems like this method is closely dependent on the continuous ground surface. But that is not the case. When the continuous ground surface is unavailable, for instance two wheels of the robot are standing on the curbside and the other two wheels are standing on the road, our method still solves the registration problem successfully in some cases. That is because our method extracts not only the points on the ground but also the points at the bottom of the objects, and this case gives full play to the RANSAC method. At worst, we could assume that the normal vector of ground surface parallels the z axis and regard the roll and pitch as an undesirable noise, like the algorithm in [17] dose, and then compensate the roll and pitch on the strength of the following refinement step. This is not a real problem for robotic mapping, where there tend to be mainly large changes in yaw and tiny offsets in roll and pitch between subsequent 3D scans. Admittedly, the discontinuous ground surfaces effect the accuracy of the registration result without a doubt.

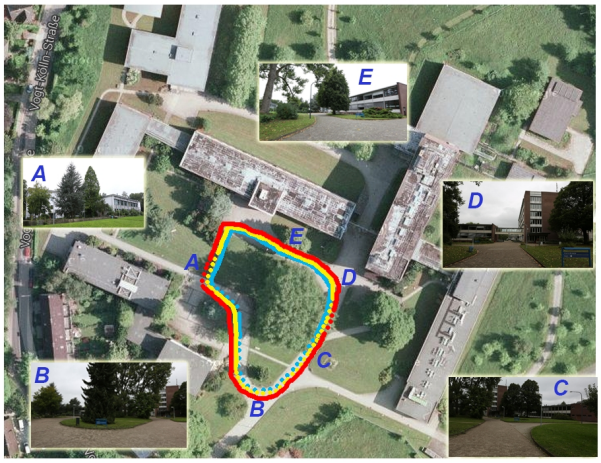


Fig. 6. The overall performance of ICP method, the plane-based method and our algorithm: the inner line (blue) represents the ICP method, the middle line (yellow) corresponds to the plane-based methods, and the outer line (red) stands for our algorithm. Moreover, the the solid lines mean the corresponding algorithm works while the dash lines express failures.

V. Experiments

Our algorithm is a global registration method aiming to prevent the local trapped solutions and resist the wrong alignments caused by less overlap. And this kind of global registration techniques just provide the crude solution, which could be further refined by local registration approaches. So the registration results in the experiments are assessed by visual inspection and the quantitative measure of error is disregarded.

A. Challenging scans captured by our custom-built platform

In this part, we reveal a registration experiment based on the challenging scans captured by our custom-built 3D laser scanner at the Department of Informatics, University of Hamburg, as Fig. 6 shows. Furthermore, Fig. 7 proposes one example scan pair captured at the location **B**. It is seen that one of the trees disappears in another scan as the tree exceeds the range of the LRF after motion of the mobile robot, and the scans are quite noisy since the whole platform is low-cost and they are collected on a rainy day. The large movement and partial overlap would cause the local registration methods to fail. Moreover, the plane-based techniques also do not work well due to lack of enough planes.

As to our method, the ground-calibrated TIDs are projected onto the cylinder, which are shown in Fig. 7 (c)(d), and the logarithms of the normalized TIDs are used to clarify the structure of the cylinder. It is clear that the intensity of the TIDs is quite different because the two scans are

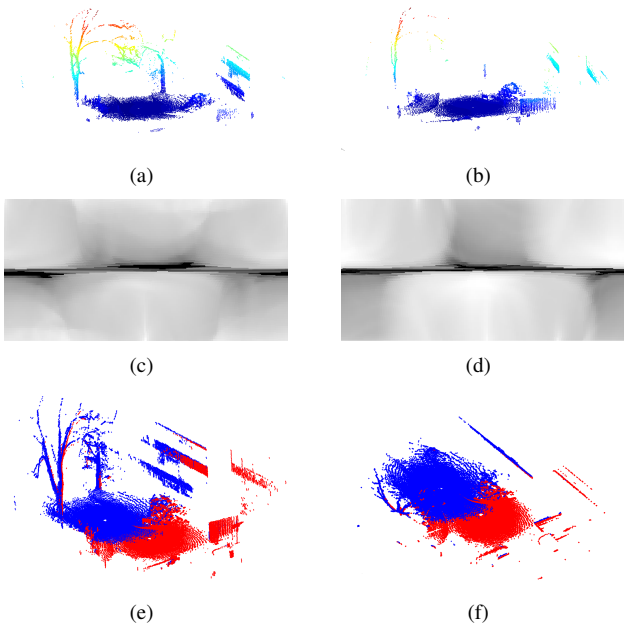


Fig. 7. Examples of the details in the first experiment: (a) and (b) are two example scans captured by our custom-built platform; (c) and (d) are the corresponding ground-calibrated TIDs projected on cylinder; (e) is the side view of the registration result; (f) is the top view of the registration result.

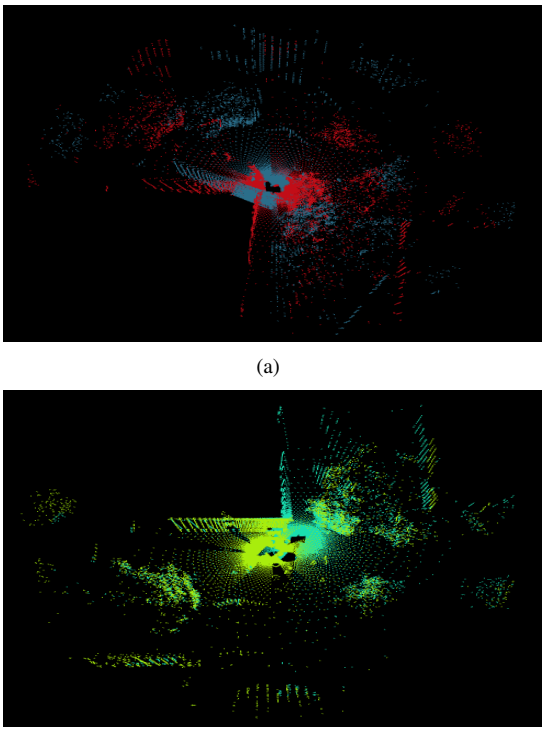


Fig. 8. (a) An example scan pair used in the second experiment; (b) the registration result using our novel method.

partially overlapped. In spite of that, the POMF still detect the shift correctly. Both the side view and the top view of the registration result is presented in Fig. 7 (e)(f). It is shown that there is still little deviation between the two scans, since we set the accumulator of the TID to be $3^\circ \times 3^\circ$. However, the deviation is general and even inevitable for global registration algorithms, and it could be refined by the following local alignment techniques.

We apply the ICP registration method in the Point Cloud Library (*http://pointclouds.org*) and the latest plane-based method [14] which is freely available (*https://github.com/junhaoxiao*) to align the scans. For the overall 46 scans, the ICP method deal with 24 scan pairs successfully and the plane-based method works out 35 pairs of them, while our novel algorithm succeeds in 41 scan pairs. The distribution of the successful cases is shown in Fig. 6, the solid lines mean the algorithms succeeds and the dash lines express fail cases. It can be seen that the ICP method merely performs well in compact environments, and plane-based method succeeds only in the plane-rich surroundings. As to the failure scans of our algorithm, it is because the scenes are mainly composed of symmetric or two similar major parts. Take the scans in location **A** for instance, the reason why our algorithm fails is that there are two buildings and few other objects. In this case, the rotation recovered by our algorithm deviates from the true rotation by 180° around the normal vector of ground surface. This experiment proves that our novel algorithm is capable to handle with noisy, partial overlapped and feature-less scans in practice.

B. Comparison Experiment using Public Dataset

In the second experiment, our novel method is compared to four other state-of-art techniques. First of all, two global methods belonging to different categories are used for this purpose, i.e., the FPFH based method and the plane-based matching method [14]. Furthermore, the two most famous registration methods, ICP and 3D-NDT, are adopted. Though ICP and 3D-NDT are not particularly well suited here since they could not deal with large displacement between scans, they are also included for experimental comparison due to their immense popularity in the context of scan registration.

A publicly available real-world dataset is adopted for comparison. Concretely, the experiment deals with the data set recorded at the Leibniz University by Oliver Wulf (*http://kos.informatik.uni-osnabrueck.de/3Dscans*). The dataset contains 468 3D scans, but we select the scans at intervals of 4 and obtain 117 scans, thus many scan pairs we used have small overlap. Compared to the scans captured by our custom-built platform, this set of scans are more vivid and precise. The eighth and twelfth scans and the registration result of our novel algorithm are depicted in Fig.8.

The ICP, 3D-NDT, FPFH based registration method we applied are in the Point Cloud Library (*http://pointclouds.org*) and the latest plane-based method [14] is freely available on the Internet (*https://github.com/junhaoxiao*). As to the plane-based method, we keep the parameters in the code unchanged. But with regard to the ICP method, we enlarge the maximum distance between two correspondences, since the ICP usually fails to find enough correspondences under

TABLE I

THE OVERALL PERFORMANCE OF FIVE REGISTRATION STRATEGIES IN THE SECOND EXPERIMENT

	Successful scans	Successful ratio with respect to the overall 117 scans
ICP	27	23.1%
3D-NDT	50	42.7%
FPFH	42	35.9%
Plane-based	62	52.9%
Our algorithm	98	83.8%

the inherent parameters for this dataset. While with regard to the 3D-NDT and FPFH method, we increase the leaf size of the filter, because the inherent parameters are too small for the input dataset and the integer indices would overflow on our laptop.

The success rates of all five methods are proposed in Table 1. It can be seen that our novel technique outperforms other four methods significantly. It is worth to talk about the performance of the FPFH feature based method. In this experimentation, we found that the FPFH method confronts challenges about how to eliminate the mismatch, especially in the large scale scene with numerous similar 3D patches. We conjecture this is why the feature-based registration algorithms are popular in registration of the indoor scans but often fail in the field robotic mapping applications.

VI. CONCLUSION AND FUTURE WORK

The algorithm proposed in this paper decomposes a 6DOF scan matching issue into several uni-dimensional correlation problems, and uses more effective Phase Only Matched Filtering (POMF) instead of the standard cross correlation method to resolve the correlation problems. This approach is a global registration method and does not rely on specific features of the input scans. And compared to the feature-based registration algorithms, the parameters of which should be tuned for different scenarios, this method is less dependent on the parameters. Furthermore, this algorithm is more comprehensible and combines the ground surface and Hough Transform craftily.

In the near future work, we would like to focus on the automatic validation of the registration results. The solid self-check technique is extremely important for robotic intelligence, and is the weakest part of the existing registration methods. The verification procedure of our algorithm is still under-developed, and two simple criteria could be applied. Firstly, the sum of distances between each point in the overlapping region and its closest point in the corresponding scan should be small enough. Secondly, the surface orientation should be consistent, which means that the normals of the overlapped points should be equivalent.

REFERENCES

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

[2] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol. 10, no. 3, pp. 145–155, 1992.

[3] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing ICP variants on real-world data sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, 2013.

[4] E. Takeuchi and T.Tsubouchi, "A 3D scan matching using improved 3D Normal Distributions Transform for mobile robotic mapping," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systerms(IROS)*, Oct. 2006, pp. 3068–3073.

[5] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.

[6] B. Huhle, M. Magnusson, W.Strasser, and A. Lilienthal, "Registration of colored 3D point clouds with a kernel-based extension to the Normal Distributions Transform," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, May 2008, pp. 4025–4030.

[7] M. Magnusson, A. Nchter, C. Lrken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3D registration reliability and speed:a comparison of ICP and NDT," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Kobe,Japan, 2009, pp. 3907–3912.

[8] A. Johnson, "Spin-Images: A representation for 3-D surface matching," Ph.D. dissertation, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.

[9] R. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Kobe,Japan, 2009.

[10] H. Andreasson and A. Lilienthal, "6D scan registration using depth-interpolated local image featues," *Robotics and Autonomous Systems*, vol. 58, no. 2, pp. 157–165, 2010.

[11] F. Tombari, S. Salti, and L. D. Stefano, "Unique signatures of histograms for local surface description," *Proceeding of 11th European Conference on Computer Vision*, pp. 356–369, 2010.

[12] A. Frome, D. Huber, R. Kolluri, T. Blow, and J. Malik, "Recognizing objects in range data using regional point descriptors," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 356–369, May 2004.

[13] B. Steder, R. B. Rusu, K. Konolige, and W. Burgard, "NARF: 3D range image features for object recognition," in *Workshop on Defining and Solving Realistic Perception Problems in Personal Robotics at the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, 2010.

[14] J. Xiao, B. Adler, H. Zhang, and J. Zhang, "Planar segment based threedimensional point cloud registration in outdoor environments," *Journal of Field Robotics*, vol. 30, pp. 552–582, 2013.

[15] A. Makadia, A. P. IV, and K. Daniilidis, "Fully automatic registration of 3D point clouds," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, New York,USA, 2006, pp. 1297–1304.

[16] A. Censi and S. Carpin, "HSM3D:feature-less global 6DOF scanmathing in the Hough/Radon domain," in *Proc. IEEE International Conference on Robotics and Automation(ICRA)*, Kobe,Japan, 2009.

[17] H. Blow and A. Birk, "Spectral 6DOF registration of noisy 3D range data with partial overlap," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, April 2013.

[18] J. Illingworth and J. Kittler, "A survey of the hough transform," *Computer Vision, Graphics, and Image Processing*, vol. 44, no. 1, pp. 87–116, October 1988.

[19] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nchter, "The 3D Hough Transform for plane detection in point clouds-a review and a new accumulator design," *3D research*, vol. 2, no. 2, Mar. 2011.

[20] J. L. Horner and P. D. Gianino, "Phase-only matched filtering," *Applied Optics*, vol. 23, no. 6, pp. 812–816, 1984.

[21] H. Foroosh, J. B.Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, Mar. 2002.

[22] F. Pomerleau, M. Liu, F. Colas, and R. Siegwart, "Challenging data sets for point cloud registration algorithms," *The International Journal of Robotics Research*, vol. 31, no. 14, pp. 1705–1711, Dec. 2012.

[23] R. Rao, A. Konda, D. Opitz, and S. Blundell, "Ground surface extraction from side-scan (vehicular) lidar," in *Proc. MAPPS/ASPRS Fall Conference*, San Antonio,USA, 2006.

[24] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonics representation of 3D shape descriptors," in *Proc. Eurographics Symposium on Geometry Processing*, Aachen, Germany, June 2003.