A Hybrid Probabilistic and Point Set Registration Approach for Fusion of 3D Occupancy Grid Maps

Conference Paper · September 2016

DOI: 10.13140/RG.2.2.11515.52006

CITATIONS
0

READS
78

4 authors, including:

Yufeng Yue
Nanyang Technological University
1 PUBLICATION 0 CITATIONS

Danwei Wang
Nanyang Technological University
372 PUBLICATIONS 5,408 CITATIONS

Namal Senarathne
Nanyang Technological University
12 PUBLICATIONS 14 CITATIONS
A Hybrid Probabilistic and Point Set Registration Approach for Fusion of 3D Occupancy Grid Maps

Yufeng Yue∗, Danwei Wang†, P.G.C.N. Senaratne†, and Diluka Moratuwage†
∗School of Electrical and Electronic Engineering, †ST Engineering-NTU Corporate Lab,
Nanyang Technological University, Singapore
Email: yyue001@e.ntu.edu.sg

Abstract—One of the major challenges in multi-robot exploration is to fuse the partial maps generated by individual robots into a consistent global map. We address 3D volumetric map fusion by extending the well known iterative closest point (ICP) algorithm to include probabilistic distance and surface information. In addition, the relative transformation is evaluated based on Mahalanobis distance and map dissimilarities are integrated using relative entropy filter. The efficiency of the proposed algorithm is evaluated using maps generated from both simulated and real environments and is shown to generate more consistent global maps.

Index Terms—map fusion, multi robot, octomap, occupancy iterative closest point, relative entropy filter

I. INTRODUCTION

With the continuing development in the field of mobile robotics and simultaneous localization and mapping (SLAM), in particular, the use of mobile robots to autonomously explore and map an environment using multiple robots has caught the attention of many robotics researchers and practitioners. Coordination among multiple robots [1] enable them to perform difficult tasks in challenging environments more efficiently and reliably, such as search and rescue missions.

One of the major challenges in such missions is to fuse the data generated by individual robots, which can be performed at different levels. Under the condition of limited communication and memory, a compact map generated by compressing raw sensor data is preferred. Hence, it’s vital to fuse the partial maps generated by individual robots into a more consistent global map, which is further compounded with 3D volumetric maps where the map fusion algorithms have to process large amounts of 3D map data.

As will be described in section II, the research in map fusion has not been studied extensively, especially for 3D volumetric maps. Majority of the existing approaches focus on feature matching and topology matching for 2D maps. As the availability of high-quality 3D sensors increase, Octomap based on Octree data structure has been widely used to model the environment with a full probabilistic 3D model [2].

In order to get an accurate and systematic map fusion algorithm, we set two requirements:

a) Relative Transformation: Since we are working on the compressed map level, we should make use of all the characteristics of the map to compute an accurate transformation. In addition, point-wise correspondence should be provided for probability fusion.

b) Probability Fusion: Many of the existing researches pay much attention to compute relative transformation, however place less emphasis on how to fuse the maps. Due to different view and SLAM error, individual maps will have varying probability even for the same object. To fuse the maps into a global one, dissimilarities between individual maps should be measured to fuse the maps and decrease the total uncertainty.

This paper addresses the above requirements and presents a novel method with the following contributions:

Firstly, probabilistic and surface information are incorporated in the registration algorithm, which we call Occupancy Iterative Closet Point (OICP). OICP combines the strength of ICP with the associated probabilistic and surface information of 3D occupancy map, which tends to match voxels with similar occupancy probability.

Secondly, an environment measurement model (EMM) [3] is applied to decide whether to accept or reject the computed relative transformation. Then a relative entropy filter is proposed to fuse the map by combing occupancy probability from both maps which also decreases the overall global map uncertainty.

The paper is organized as follows: Section II reviews related literatures in map fusion, Section III introduces the Occupancy Iterative Closet Point algorithm. Section IV details the evaluation of the computed relative transformation and the fusion of probabilities. Section V presents the outcome of the experiments conducted using simulated and real data. Section VI concludes the paper and discusses the future work.

II. RELATED WORK

In this section we review different approaches proposed in the literature to fuse the local maps into a single global map. The SLAM issue will not be discussed in this context.

The representation of maps can be roughly categorized to feature maps or occupancy grid maps. For feature-based maps, a Delaunay triangulation based topology network to match the features is introduced in [4]. In [5], probabilistic hypothesis density (PHD) SLAM is employed to fuse maps with unknown initial transformation. Although feature-based matching performs effectively in some cases, it leads to loss of information. Then researchers have given more attention to occupancy map fusion. In 2D gird maps, voronoi diagram is generated to represent occupancy map on a higher level, then a deterministic transformation is calculated by extracting
the spectral information [6]. In [7], individual maps are transformed into Hough space, and common areas are found by making use of the properties of Hough space. A variety of standard topology map attributes are evaluated to compute the subgraph-isomorphisms in [8]. A probabilistic edge matching approach is introduced to incorporate the uncertainties of robot pose and transformation in [9].

A. Computing Relative Transformation

The fusion of 3D volumetric maps generated using Octomap is first proposed in [10], where an Octomap is converted to a point cloud by extracting the centers of voxels. Then it utilizes ICP [11] to calculate the relative transformation by minimizing point-wise Euclidean distance [12]. However, the foremost property of Octomap is not combined in registration, which is the occupancy probabilities of individual voxels.

As shown in Fig.1, the raw sensor data generated from simulated environment are highly compressed to occupied voxels. What’s more, due to sensor noise, the raw data will have certain error around the object surface. Hence another property for Octomap is that the surface is more smooth compared to raw sensor data. Although point-to-plane ICP [13] can’t provide voxel-wise correspondence, it should be able to achieve better accuracy by minimizing sum of distance from points to the closest tangent plane and is explored in our proposed approach.

A recent paper [14] reviews the 3D scan matching methods up to date, concluding that Normal Distribution Transform (NDT) [15] provides most accurate results compared to others. NDT models the point cloud into a collection of Gaussian distributions with covariance and mean, and minimizes the sum of $L2$ norm. Since in our case the points are evenly distributed with a distance of the size of a voxel, the Gaussian distribution may not represent the environment explicitly compared to rich sensor data, so the error metric loses its original superiority.

In [16], it is proven that a point-to-point error metric of registration algorithm combining invariant features will converge if the weight parameter is non-increasing.

B. Probability Fusion

In registration algorithm, there is no guarantee that a global minimum will be found. A wrong transformation would lead to an inaccurate global map, which in turn is detrimental for robot exploration. So a model to verify the transformation is essential. In [3], an environment measurement model is proposed to verify the accuracy of transformation by calculating the percentage of inliers. In this paper we apply this model to evaluate the accuracy of relative transformation computed by OICP algorithm.

When a robust and accurate transformation is achieved, the key issue is to fuse the probabilistic information in partial maps into a global and consistent map. In [12], the maps are fused on voxel-wise level by directly adding up the probabilities. Shannon entropy is used to measure the uncertainty of map in robot exploration in [17]. In [9], the paper introduces an entropy filter to reject fusion leads to a higher entropy, but is over confident on single map without taking both maps into consideration. So a relative entropy filter is adopted in this paper to fuse occupancy probability consistently.

III. OCCUPANCY ITERATIVE CLOSEST POINT

In this section, Occupancy iterative closest point (OICP) is presented to incorporate occupancy probability and surface information of 3D occupancy map, which computes the relative transformation between partial maps under the condition of given Octomap and initial position estimation.

A. OICP algorithm

Fig. 2 illustrates an example of Octomap and the associated point cloud extracted from the map. Given an Octomap $M$, we define the point cloud $m$ extracted from Octomap as $m = m_0, m_1, m_2, \cdots, m_i, \cdots, m_n$, where $(m_x^i, m_y^i, m_z^i, L_{m_i}, d_{m_i}, l_{m_i})$ is a tuple that includes the position of the extracted voxel center $m_i = (m_x^i, m_y^i, m_z^i)$, the occupancy probability log-likelihood $L_{m_i}$, voxel size $d_{m_i}$ and Octree level $l_{m_i}$.

The OICP can be divided into two stages. For the first stage, point-to-plane ICP is applied to make use of the smooth surface information of Octomap using Eq.(1), where $\eta$ is the surface normal of tangent plane.

$$f = \arg \min \sum_{i=1}^{n} \| (R m_i + t - n_j) \cdot \eta \|^2$$  \hspace{1cm} (1)

In the second stage, voxel occupancy probability is incorporated in order to achieve accurate transformation with point-wise correspondence. We utilize the weighted linear combination to combine positional distance and feature distance into a single point-wise distance metric as proposed in [16].

$$d(p, q) = d_e(p, q) + \alpha^2 d_f(p, q)$$  \hspace{1cm} (2)
The point-wise distance $d(p,q)$ is described by Euclidean distance $d_e(p,q)$ and feature distance $d_f(p,q)$. The weight parameter $\alpha^2$ is defined to balance the contribution between positions and features.

For feature distance, the occupancy probability is set as an invariant feature. The probability of voxel $m_i$ given observation $z_{1:t}$ is updated using log-odds notation in Octomap:

$$L(m_i|z_{1:t}) = L(m_i|z_{1:t-1}) + L(m_i|z_t)$$

(3)

Note that log-odds can be easily converted to probabilities using Eq.(4).

$$L(m_i|z_{1:t}) = \log \frac{p(m_i|z_{1:t})}{1 - p(m_i|z_{1:t})}$$

(4)

Combining Euclidean distance and feature distance, occupancy iterative closet point is formulated in Eq.(5). Here the Euclidean distance is defined as the distance between center points of the voxel, and feature distance is represented by the difference in associated occupancy probabilities.

$$f_{oicp} = \sum_{i=1}^{n} \|Rm_i + t - n_j\|^2 + \omega \|p_{m_i} - p_{n_j}\|^2$$

(5)

The OICP iteratively computes the transformation which minimizes the proposed error metric between corresponding points. And the trade off between positional and feature value is determined by weight parameter $\omega$.

B. Uncertainty in occupancy probability

As Octomaps are generated using noisy pose estimates provided by SLAM outputs and noisy sensor data, there is always an inherent level of uncertainty associated with them. When the uncertainty of the map is high, we put less confident in the map. Therefore, Shannon entropy is introduced to measure the uncertainty of the probabilistic voxel map.

$$H(m) = -\sum_i p_{m_i} \log_2 p_{m_i}$$

(6)

When the voxel is occupied certainly with $p_{m_i} = 1$, the entropy achieves minimum with zero. Which means there is no uncertainty. On the contrary, entropy reaches maximum value when the occupancy probability $p_{m_i} = 0.5$.

C. Uncertainty in Positional Value

As shown in Fig. 3, let $d_{TS}$ be the true distance between target point $m_T$ and scene point $n_S$. At each iteration, point $n_S$ iteratively attempts to find its ground truth point $m_T$, however would match to its nearest neighboring point $m_C$ with a distance $d_{CS}$. So, there is always an offset between ground truth distance $d_{TS}$ and the closet point distance $d_{CS}$.

In [16], conditional expectation is applied to model the error due to misalignment in each iteration. The author assumes that the points around the ground truth point $m_T$ are continuously distributed. However, the points are discretized with equal distance $d_m$ in our case.

To model the true distance $d_{TS}$ between point $m_T$ and point $n_S$ using closet point distance $d_{CS}$, we assume the registration error is independent of noise and equal in three directions.

$$E(d_{TS}^2) = \sigma^2_{TS} = \frac{1}{3}\sigma^2_{TSx} = \frac{1}{3}\sigma^2_{TSy} = \frac{1}{3}\sigma^2_{TSz}$$

(7)

We assume the scene point $n_S$ is uniformly distributed on the surface of a sphere with the radius of $d_{TS}$, so we have the probability density function:

$$f(d_{CS}|d_{TS}) = \frac{1}{4\pi d_{TS}^2}$$

(8)

In Eq.(9), the conditional expection of $d_{CS}^2$ given $d_{TS}$ is formulated and calculated by integrating in the spherical coordinate system. $d_{CS}^2$ is computed as the length of the hypotenuse of a right triangle $\triangle_{n_m,m_p}$ as shown in Fig. 3.

$$E(d_{CS}^2|d_{TS}) = \int d_{CS}^2 \cdot f(d_{CS}|d_{TS})ds$$

$$\approx \frac{1}{3} d_{TS}^2 + \frac{2}{3} d_m^2$$

(9)

In conditional expection, we have $E[EX|Y] = E(X)$, so the expectation of $d_{TS}^2$ can be calculated as:

$$E(d_{TS}^2) \approx E(3E(d_{CS}^2|d_{TS}^2) - 2d_m^2) = 3\sigma_{CS}^2 - 2d_m^2$$

(10)

The term $-2d_m^2$ indicates the error caused by discretized point cloud. Since mean square error(MSE) is the global estimate of $d_{CS}$, we have $d_{CS} \approx MSE$. As each dimension is identical, we get $\sigma^2_{TSx} \approx MSE^2 - \frac{2}{3}d_m^2$.

Hence, we get the final form for the weight as:

$$\omega = k \cdot \frac{\sigma_{TSx}}{H(m)} \approx k \cdot \frac{\sqrt{MSE^2 - \frac{2}{3}d_m^2}}{H(m)}$$

(11)

where $k$ is positive constant factor. As the registration iteratively minimizes the error function, MSE should be non-increasing. Then weight $\omega$ is non-increasing with invarint $H(m)$, and should tend to zero as the OICP converges to a perfect match. When the positional error is large, we should rely more on features. As the two maps get closer, the algorithm tends to rely more on positional value.
IV. TRANSFORMATION EVALUATION AND PROBABILITY FUSION

A. Transformation Evaluation

Although the registration algorithm will achieve convergence, there is no guarantee that a global minimum will be found and there is a high chance that the algorithms would fall into a local minima.

In [3], an environment measurement model (EMM) is proposed to verify the transformation by calculating the Mahalanobis distance between matched points. Then the inliers percentage is set as a threshold to decide whether to accept or reject the transformation. The probability of point \( m_i \) given point \( n_j \) can be written as:

\[
p(m_i|n_j) = N(m_i;n_j, \Sigma_{ij})
\]

When operating on the Octomap representation, we have

\[
\Sigma_i = \frac{1}{p_{m_i}}, \Sigma_j = \frac{1}{p_{n_j}}
\]

Since the covariance is equal in three directions, rotation matrix has no impact on \( \Sigma_i \)

\[
\Sigma_{ij} = R^T \Sigma_i R + \Sigma_j = \frac{1}{p_{m_i}} + \frac{1}{p_{n_j}}
\]

Then for each correlated pairs, Mahalanobis distance can be computed as follows:

\[
f_{m,ij} = -\frac{(Rm_i + t - n_j)^T \cdot (Rm_i + t - n_j)}{2(p_{m_i}^{-1} + p_{n_j}^{-1})}
\]

The inliers are decided by setting a threshold of Mahalanobis distance. Then a confidence level is applied as a criteria for rejecting the wrong estimates by computing the overall fraction of inliers. It is shown in the experiments that the applying of EMM effectcively rejects high erroneous estimates that would largely undermine the quality of the fused map.

B. Relative Entropy Filter

Under sensor noise and motion uncertainty, the same object may have different probabilities in separated maps. Hence its vital to establish a strategy to consider the dissimilarities and fuse them into the global probability map.

Kullback-Leibler(KL) divergence is a commonly used relative entropy filter to measure the difference between two probabilities. The original form is:

\[
SKL(R||S) = KL(R||S) + KL(S||R)
\]

Here we regard the occupancy probability \( p_{m_i} \) and \( p_{n_j} \) as discrete random variables (RVs), and the difference between the two RVs are computed as:

\[
SKL(p_{m_i}||p_{n_j}) = p_{m_i} \log \frac{p_{m_i}}{p_{n_j}} + p_{n_j} \log \frac{p_{n_j}}{p_{m_i}}
\]

\[
SKL(p_{m_i}||p_{n_j})\text{ represents the difference between the two probabilities. If SKL is within a certain threshold, we regard the dissimilarity as random errors and still fuse the corresponding pairs by taking the average } p_{fuse} = \frac{p_{m_i} + p_{n_j}}{2}. \text{ If SKL is beyond the threshold with } p_{m_i} > p_{n_j}, \text{ it means the voxel with lower probability } p_{n_j} \text{ doesn’t have enough information for the environment and will be rejected, then we only trust the voxel with higher probability and let } p_{fuse} = p_{m_i}.
\]

V. EXPERIMENTAL RESULTS

Experiments conducted using both simulated and real data are presented in this section. For simulation, a robot equipped with a simulated 3D depth sensor traversed the environment twice to create two maps in Gazebo simulator [18] using RTAB [19]. A Pioneer 3 AT robot was tele-operated twice in the WKW lobby at Nanyang Technological University to build two maps. The robot was equipped with a Hokuyo Laser range finder for pose estimation using Gmapping [20] and an Asus Xtion PRO for 3D perceptions, as shown in Fig. 4. The individually generated Octomaps were post-processed using C++ and Matlab to generate the fused maps. The size of Octomaps in both experiments was set to be 0.1m.

As mentioned before, we focus our research on the fusion stage with an initial relative transform estimation. The initial relative transformations between the coordinate frames of different robots can be extracted during the deployment phase of the mission if such information is available. Else it can be extracted later on through coordinated rendezvous actions [21]. To simulate the availability of noisy initial relative transformations, we generated two maps with different initial transformations both in position and rotation. Fig. 5 illustrates the point clouds extracted from the generated individual Octomaps with an en erroneous initial transformation. The dimension was about 50m * 30m * 10m for both maps.
A. Registration Results

In this section, we present the performance of OICP compared to ICP, point-to-plane ICP and normal distribution transform (NDT). As a baseline for comparison, standard ICP, point-to-plane ICP and NDT implementations from the Point Cloud Library (PCL) were used [22] and parameters were set according to the environments. The number of iterations for ICP, point-to-plane ICP and NDT is set to be 50. As mentioned before, the point clouds extracted from Octomap has a smoother local planes compared to raw sensor data. For OICP, point-to-plane ICP was performed for 15 iterations as initial registration stage to make use of this surface information, then ICP with occupancy probability formulated in Sec III was performed 35 iterations to get the point-wise correspondence.

The comparison of accuracy of registration algorithms are summarized in Fig. 6. The translation error and rotation error are the Euclidean norm of and Euler norm of the difference between the ground-truth and the output.

OICP outperforms NDT, point-to-plane ICP, as well as map registration algorithm ICP used in [12]. Although point-to-plane ICP unable to provide point-wise correspondence, it makes use of the smooth plane and achieves better accuracy compared to default ICP based fusion. However, point-to-plane ICP still has drifts in translation and rotation. OICP combines surface and probability information to compute the point-wise correspondence with similar probability to generate the most accurate relative transformation. We should make special notice to NDT, which produces large errors, especially in translation. The reason might be the Gaussian distribution extracted from discretized sparse point cloud can’t represent the environment explicitly compared to rich sensor data.

As described before, point-to-plane ICP and NDT can’t provide point-wise correspondence. Hence we only discuss ICP and OICP in the following experiments. The registered results are shown in Figs. 7 and 8. A large offset in translation is shown for the result of ICP, and OICP produces approximately a perfect alignment. For both results, rotation errors are considered to be relatively accurate.

B. Transformation evaluation

After computing a transformation, the accuracy is evaluated by calculating the percentage of inliers. Since ICP is deterministic with only coordinates, the covariance of each point is set to be identity matrix \( I_{3 \times 3} \). For map fusion, a strict criteria is set to guarantee the accuracy: the inliers threshold is set as 0.15 Mahalanobis distance and the fraction of rejecting a transformation is set to be below 60%.

Fig. 9 illustrates the comparison of the average Mahalanobis distances between OICP and ICP in the two experiment settings. The box plot shows a box that encloses the middle 50% of the data, with the median value marked by a line and the mean value marked by a small square.

The inlier fraction based on the threshold is shown in Fig. 10. The inlier percentage of ICP is lower than OICP because ICP didn’t converge to perfect alignment. Based on the selected threshold, OICP transformation generates an accurate and reliable map registration result compared to pure ICP based map registration approach in [12].

C. Probabilistic Map Fusion

After determining the correct transformation for map registration, a relative entropy filter is used to achieve map fusion.
by considering the dissimilarities between matched voxels. The relative entropy threshold is set as 0.1 in KL-divergence. In comparison, the fusion strategy in [12] is performed by directly taking the average of probability for all the matched voxels.

In simulated environment, the entropy for map fusion by directly taking the average is 0.16. And the entropy of using relative entropy filter is 0.13. For real experiment in WKW Lobby environment, the entropy of directly taking the average is 0.24, and is 0.2 by using relative entropy filter.

The entropy of WKW Lobby is higher than simulation due to real world noise and uncertainty. The experiment indicates that our map fusion strategy is able to combine probabilities of both maps effectively to further decrease the total uncertainty. Figs. 11 and 12 show the final results of the fused 3D occupancy grid maps of simulation and WKW Lobby.

VI. CONCLUSION AND FUTURE WORK

A probabilistic 3D map fusion algorithm Occupancy iterative closest point(OICP) was proposed in this paper. The OICP algorithm combines occupancy probability in conjunction with surface information to register point cloud extracted from Octomap. The weight between positional error and map uncertainty is controlled by modelling errors using Shannon entropy and positional misalignment. EMM is applied to evaluate the quality of transformation by rejecting low inlier fraction. The final map fusion is then achieved with a relative entropy filter which integrates the measurements of both maps and decreases the uncertainty of the global map. Experiments on simulated and real world environments were conducted and the evaluations confirmed that the proposed OCIP method is able to produce more consistent global 3D occupancy grid maps compared to existing methods in the literatures.

In future work, we hope to extend our work to fuse maps with unknown initial transformation. The availability of a consistent global requires periodic fusion of partial maps from multiple robots. Hence making a decision on when to fuse the maps warrants further research and would also be explored in the future.

ACKNOWLEDGMENT

The research was partially supported by the ST Engineering NTU Corporate Lab through the NRF corporate lab@university scheme.

REFERENCES