Assessing the behavior of the Iterative Closest Point algorithm in complex environments

David Landry (111 040 494)
March 23, 2017

Supervised by Prof. P. Giguère
and co-supervised by Prof. F. Pomerleau

1 Introduction

When thinking about robots of today, one might imagine a robotic arm in a car factory. This kind of fixed, specialized robots is successfully being used to improve production efficiency in different domains. However a new type of robot is emerging, one that is free to move in its environment. The RHINO robot is a good example [1]. These are not easy to design, as they rely on precise and robust localization. A robot needs to know where it stands with respect to the environment if we want it to act upon its surroundings in a meaningful way. Although many efforts have been made in that direction in the past, localizing a robot in a given environment is still a challenging research problem.

Various sensing modalities are used for localization. Two main examples are cameras and lidar sensors. Cameras capture information about the existing light in the environment and try to localize against it. They are passive sensors in that sense. Lidar sensors, however, are active. They emit light in the environment and capture the emitted signal as it comes back to the sensor. While cameras are good at detecting and recognizing landmarks, lidars are less sensitive to changes in lighting conditions [2].

The need for accurate localization against a map is outlined in the Teach-and-Repeat task. In this task there is a preliminary *Teach* phase. During that time the robot registers sensory inputs in a database as it is driven through a trajectory. In a subsequent *Repeat* phase, the robot repeats the original trajectory by comparing the sensory inputs with those it recorded in the database. In that sense, the database can be taught of as a map that the robot localizes against later. Note that this paradigm is sensor-agnostic as it has been successfully implemented using cameras [3] and lidar sensors [4].

This proposal focuses on lidar sensing as a mean of robust localization. First, it outlines the problematic of *uncertainty estimation* with lidar sensors in Section 2. The importance of this research problem is emphasized in Section 2.2. A summary of the existing approaches to solve it is in Section 3. Then, Section 4 describes my objectives with respect to uncertainty estimation of lidar sensing. It also contains a reasonable set of steps to accomplish these objectives.

2 Research Problem

A viable way to localize a robot using the range data from a lidar sensor is the Iterative Closest Point (ICP) algorithm [5]. It was developed simultaneously by Yang *et al.* [6] and Besl *et al.* [7]. It consists in iteratively aligning two point clouds by minimizing the distance between the closest points in each of them. In a first step, each point from the first point cloud is paired with its closest neighbor in the other point cloud. In a second step, a point cloud is moved such that the error between the pairs of points is minimized. Those steps are repeated until convergence. The output is a rigid transformation between the reference frame of both scans.

When minimizing the error between the pairs of points, the error function could simply be the distance between the associated points. This variant is called the point-to-point version of ICP [8, section 2.6.2]. There is also a point-to-plane version. In that case, the error is the distance between the point and the plane estimated from the shape of the other point cloud around its closest neighbor. In both cases, the error minimization is a non-convex combinatorial problem, because it is based on the pairing of closest points.

In the context of mobile robotics, ICP can be used to compute odometry or localization [9]. It is used as odometry when successive scans are matched to estimate the egomotion of a robot. It is used as a localizer when a scan is matched with a larger point cloud to identify where the robot is within the larger point cloud.

2.1 The Iterative Closest Point algorithm uncertainty

One problem that arises when using ICP is the estimation of its uncertainty. By uncertainty, we mean what kind of error is this algorithm likely to commit with respect to the true transform between the two point clouds. This will be the main theme of my proposal, which can be framed as follows:

What is the uncertainty on the result of the ICP algorithm on two given point clouds?

The answer to this question is complex. The behavior of ICP is greatly conditioned by the geometry of the matched scans [9]. The result of this algorithm is also affected by sensor noise and data association problems. It is common to model this uncertainty as gaussian [9], [10]. This assumption is useful to make the computation and representation of the uncertainty tractable. It also makes it easier to integrate ICP in a Kalman Filter [5]. Consequently, the problem can be posed more precisely as the problem of searching for a covariance matrix that would represent the uncertainty.

2.2 Motivation

A better understanding of the uncertainty of ICP would benefit many areas of research. As already mentioned, predicting the covariance matrix of ICP matches would allow us to integrate them in a Kalman Filter [5]. We would in turn be able to improve our state estimation of the robot.

An improved grasp on the uncertainty of ICP would also allow the construction of better maps for Teach-and-Repeat. In particular, it would enable us to be more careful in areas that are more difficult for ICP. My previous paper [11] is a good step in that direction. It can be considered as a map curating algorithm. It changes the level of detail in a Teach-and-Repeat map according to the uncertainty a robot has when localizing against it. The level of detail is increased in sections that have high uncertainty, and decreased in sections where the level of precision is satisfactory. However, it remains difficult to detect the areas that are uncertain for ICP in the absence of a reliable and quick algorithm to estimate the covariance matrix.

3 Related Work

As of today, a hand-tuned estimate of the covariance matrix (i.e., a covariance that is independent of any sensor input) is still a common solution to this problem [12]. Finding a good covariance matrix requires a lot of experience and trial and error. It is also task-specific. The chosen matrix has to be large enough to accommodate the most difficult match encountered during a task. However, it also has to be small enough to estimate the state of the robot to a satisfactory degree.

The covariance matrix can also be estimated through sampling. This is referred to as the "brute-force" method. Unfortunately, the computational cost of ICP makes sampling many scan matches prohibitive. Thus, estimating the covariance through sampling cannot be done online [9].

Mehra [13], Aghili *et al.* [14] use ICP inside a special form of the Kalman Filter (called Adaptative Kalman Filter). The Adaptative Kalman Filter is a form of Kalman Filter where the covariance matrix of the readings is modified in function of the gap between the predictions and the measurements. The caveat of their approach is that it estimates the covariance of the ICP through purely statistical means. It is reactive, instead of being predictive. Their algorithm is thus still sensitive to outliers and needs to reject matches that seem unlikely.

Censi [9] devised a way to estimate the covariance of ICP by analyzing the underlying error function. There are indications that this estimation is optimistic in real-world situations [12], although more work is needed to prove that observation conclusively. This method was validated on simulated 2D datasets. Bonnabel *et al.* [10] provided a theoretical proof of this approach, but also emphasized the need for the underlying error function to be smooth. Consequently the covariance estimation formula from [9] should not be used with the point-to-point variant of ICP. Effectively, the error function produced by the point-to-point variant is not smooth. Also, this algorithm makes the assumption that ICP always converges in the attraction region of the true solution [9]. However, it is well known that ICP may converge in a local minimum before reaching the true transform between both point clouds [10]. Finally, Manoj *et al.* [15] extended the closed-form approach to 3D. They did not add any assumptions doing so. It means that this technique also makes the true convergence assumption. The indication from Bonnabel *et al.* [10] that this technique should only be applied on the point-to-plan version of ICP also holds.

Afterwards, Vega-Brown *et al.* [16] introduced a data-driven approach to the problem called Covariance Estimation and Learning through Likelihood Optimization (CELLO). They used pre-existing sensory data to predict the covariance of a localization algorithm via machine-learning. It is more general than the previously presented approaches as it is applicable to various sensory modalities. This algorithm requires a predictor function that maps the sensory input in a smaller prediction space. The output of the predictor should capture properties that are important to covariance prediction. One caveat of this approach is the need for ground truth data to train the covariance prediction algorithm. It was used for state estimation in a real-world situation using visual sensors in [17]. To the best of our knowledge, it was not validated on 3D lidar data.

Finally, Vega-Brown *et al.* [12] extended CELLO to the situation where no ground truth is available. It was done using the expectation maximization (EM) algorithm. This approach was validated for lidar on real-world 2D dataset. It was compared with the approach of Censi [9]. Their results suggest that the covariance of this approach is too optimistic. Again, we believe that it has not been tested against 3D lidar data.

A summary of these six methods is shown in Table 1. In this table, strong convergence assumption means that ICP is assumed to always converge to the global minimum. Also, note that the fixed estimate is considered to need pre-existing data because of its trial and error nature. While this pre-existing data does not need to contain precise ground-truth information, it needs to contain enough information about the ICP error so that a human can tune the covariance. Finally, "predictive" means that the algorithm does not need to encounter a difficult match online before it can increase the covariance to more appropriate value.

Table 1: Summary of the different methods to estimate ICP's covariance.

Method	Real-time	Strong convergence assumption	Predictive	Need for pre-existing data	Need for ground-truth	3D
Hand-tuned	x		x	x		x
Brute-force estimate			x			x
Adaptative Kalman Filter	x					x
Closed-form estimate	x	x	x			x
CELLO	x		x	x	x	
CELLO-EM	x		x	x		

As can be seen in Table 1, the closed-form estimate can predict the covariance of 3D ICP matches. However, as mentioned earlier in Section 3, there are indications that this closed-form solution produces optimistic results. A natural reaction to this state of affairs is to extend the CELLO algorithm to the 3D case. To be able to formally compare the 3D version of CELLO and the existing closed form estimate, we also need a well-posed method to evaluate covariance estimation algorithms for ICP.

4 Objectives and Methodology

In light of this information, two main objectives can be defined for my master's. The first objective is to *validate the existing estimations of ICP's covariance experimentally*. In other words, we need a sound framework to evaluate the quality of the provided estimates. To reach that goal, we propose the following steps:

1.1 Review the existing methods of uncertainty estimation and familiarize myself with them. This means reproducing the algorithms of all the common methods for uncertainty estimation. It also means understanding the method of validation used when creating these algorithms, when applicable. In particular, it is important to understand both the algorithms and the methods of validation used by the closed-form estimate [9] and CELLO [16], as I am likely to compare my techniques to these works in the future.

1.2 Create a validation method for ICP covariance prediction that captures the necessary properties.

1.3 Validate the performance of the existing covariance estimation methods on small, toy examples. For example, we could test the validation method on datasets of synthetic point clouds. It is easier to understand the behavior of our validation method if the examples used are easy to comprehend and fully controlled. The validation method can then be adapted if needed.

1.4 Finally, the algorithms should be tested in real-world environments. To do so, I will need datasets that have ground truth information. That way, it will be possible to determine if the compared techniques encompass the reported error.

The second objective is to *improve our prediction of the ICP covariance by using a machine learning based algorithm on a 3D lidar dataset*. More concretely, it implies to do the following:

2.1 Familiarize myself with the CELLO algorithm. Fortunately, this adaptation work is also required in step 1.1.

2.2 Create (or find) a 3D point cloud predictor that captures the information necessary to the execution of the CELLO algorithm. Effectively, the reference papers [12], [16] provide very little detail about the utilized predictor. The overall characteristics of the predictor do not seem to be easily transferable to 3D. Thus, finding a predictor that captures enough information to estimate covariance will be an important step of my master's.

2.3 Collect data that is appropriate to the CELLO algorithm. Effectively, this algorithm requires ground truth information of the localization error. I will have to research for any existing dataset that has this information. Two existing candidates are the *Challenging data sets for point cloud registration algorithms* [18] and the *KITTI* dataset [19]. We could also collect our own dataset for this purpose. This would require equipment that will be described in Section 4.1.

2.4 Use all the tools described above to run the CELLO algorithm on a 3D lidar dataset. The performance of the algorithm could be compared to that of a fixed covariance estimate, and the approach of the closed-form estimate [15]. It would also be good to use the covariance estimation in a real context, by integrating it to a Kalman Filter for example.

Finally, we must not forget the redaction of my master's thesis, which will occupy the last few months of my project.

4.1 Necessary equipment

Some specialized equipment may be required for this project. In the eventuality we need to collect our own dataset, we have access to a Clearpath Husky A200 mobile robot. The Husky is equipped with a Velodyne HDL-32e lidar sensor. The ground truth may be collected using a theodolite. This could be borrowed from the *Dpartement de gomatique*, here at Universit Laval.

5 Schedule

A chronology of how I plan to execute these different objectives is found in Table 2. Note that the number of objectives is optimistic. This is a way to minimize the risk associated with my master's, in case some objectives turn out to be infeasible.

Table 2: Schedule containing the objectives to be accomplished during my master's. The numbers are references to the objectives listed in Section 4.

Spring 2017	Summer 2017	Fall 2017	Winter 2018	Summer 2018
Internship at rebro Universitet	1.1 1.2 2.1	1.3 1.4 2.3	2.2 2.4	Redaction

6 Conclusion

This document described my project for the rest of my master's at Universit Laval. It introduced the key problematic of covariance estimation for the ICP algorithm. The relevance of this problematic was demonstrated through the examples of state estimation and map curating. Several state of the art techniques were introduced, and their respective caveats was discussed. Finally, this document described my objectives in more details and presented a reasonable sequence of steps to accomplish them.

References

[1] W. Burgard, A. Cremers, and D. Fox, "The interactive museum tour-guide robot," *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial AAAI '98/IAAI '98 intelligence*, pp. 11–18, 1998.

[2] C. McManus, P. Furgale, B. Stenning, and T. D. Barfoot, "Lighting-invariant visual teach and repeat using appearance-based lidar," *Journal of Field Robotics*, vol. 30, no. 2, 2013.

[3] P. Furgale and T. D. Barfoot, "Visual teach and repeat for long-range rover autonomy," *Journal of Field Robotics*, vol. 27, no. 5, pp. 534–560, 2010.

[4] C. Sprunk, G. D. Tipaldi, A. Cherubini, and W. Burgard, "Lidar-based teach-and-repeat of mobile robot trajectories," in *IEEE International Conference on Intelligent Robots and Systems*, Karlsruhe, 2013, pp. 3144–3149.

[5] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 119–152, 1994.

[6] C. Yang and G. Medioni, "Object modelling by registration of multiple range images," in *Image and Vision Computing*, vol. 10, Sacramento, 1992, pp. 145–155.

[7] P. Besl and N. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[8] F. Pomerleau, F. Colas, and R. Siegwart, "A review of point cloud registration algorithms for mobile robotics," *Foundations and Trends in Robotics*, vol. 4, no. 1, pp. 1–104, 2015.

[9] A. Censi, "An accurate closed-form estimate of ICP's covariance," in *IEEE International Conference on Robotics and Automation*, Rome, 2007, pp. 3167–3172.

[10] S. Bonnabel, M. Barczyk, and F. Goulette, “On the covariance of ICP-based scan-matching techniques,” in *Proceedings of the American Control Conference*, Boston, 2016, pp. 5498–5503.

[11] D. Landry and P. Giguère, “Automating node pruning for lidar-based topometric maps in the context of teach-and-repeat,” in *13th Conference on Computer and Robot Vision*, Victoria, 2016, pp. 132–139.

[12] W. Vega-Brown and N. Roy, “CELLO-EM: adaptive sensor models without ground truth,” in *IEEE International Conference on Intelligent Robots and Systems*, Tokyo, 2013, pp. 1907–1914.

[13] R. K. Mehra, “On the identification of variances and adaptive Kalman filtering,” *IEEE Transactions on Automatic Control*, vol. AC-15, no. 2, pp. 175–184, 1970.

[14] F. Aghili and C.-Y. Su, “Robust relative navigation by integration of ICP and Adaptive Kalman Filter using laser scanner and IMU,” *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 4, pp. 2015–2026, 2016.

[15] P. S. Manoj, L. Bingbing, Y. Rui, and W. Lin, “A closed-form estimate of 3D ICP covariance,” in *Proceedings of the 14th IAPR International Conference on Machine Vision Applications*, Tokyo, 2015, pp. 526–529.

[16] W. Vega-Brown, A. Bachrach, A. Bry, J. Kelly, and N. Roy, “CELLO: a fast algorithm for covariance estimation,” in *IEEE International Conference on Robotics and Automation*, Karlsruhe, 2013, pp. 3160–3167.

[17] V. Peretroukhin, L. Clement, M. Giamou, and J. Kelly, “PROBE: predictive robust estimation for visual-inertial navigation,” in *IEEE International Conference on Intelligent Robots and Systems*, Hamburg, 2015, pp. 3668–3675.

[18] 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.

[19] A Geiger, P Lenz, C Stiller, and R Urtasun, “Vision meets robotics: the KITTI dataset,” *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.