Performance Analysis for Stable Mobile Robot Navigation Solutions

Chris Scrapper, Raj Madhavan, and Stephen Balakirsky

Intelligent Systems Division
National Institute of Standards and Technology
Gaithersburg, MD 20899, U.S.A.

ABSTRACT

Robot navigation in complex, dynamic and unstructured environments demands robust mapping and localization solutions. One of the most popular methods in recent years has been the use of scan-matching schemes where temporally correlated sensor data sets are registered for obtaining a Simultaneous Localization and Mapping (SLAM) navigation solution. The primary bottleneck of such scan-matching schemes is correspondence determination, i.e. associating a feature (structure) in one dataset to its counterpart in the other. Outliers, occlusions, and sensor noise complicate the determination of reliable correspondences. This paper describes testing scenarios being developed at NIST to analyze the performance of scan-matching algorithms. This analysis is critical for the development of practical SLAM algorithms in various application domains where sensor payload, wheel slippage, and power constraints impose severe restrictions. We will present results using a high-fidelity simulation testbed, the Unified System for Automation and Robot Simulation (USARSim).

Keywords: navigation solutions, scan-matching, ICP, performance analysis, performance evaluation

1. INTRODUCTION

As mobile robots become more ubiquitous, their utility will rely on the ability of the robotic system to safely operate in dynamic, unstructured environments. These systems will need to explore new environments, generate maps that identify obstacles and hazards through exploration, and use these maps to safely navigate to any location. They will also need the ability to intelligently adapt to momentary changes in the environment. Central to the realization of this vision of mobile robots is the system's ability to develop a stable ***navigation solution***, which we define as *the ability of the system to sense the environment, create internal representations of its environment, and estimate pose (where pose consists of position and orientation) with respect to a fixed coordinate frame*.

The primary focus of the work presented in this paper is to bring together an amorphous research community to define standardized test methods for the quantitative evaluation of navigation solutions in dynamic, unstructured environments. Currently, there is no way to quantitatively measure the performance of navigation solutions against user-defined requirements. Additionally, no consensus exists on what objective evaluation procedures need to be followed to deduce the performance of navigation solutions in different domains. The lack of reproducible and repeatable test methods precludes researchers from working towards a common goal. It prevents the communication and inter-comparison of the results, which prevents researchers from leveraging previous work and inhibits technology transfer from the "drawing board" to the field.

Commercial equipment and materials are identified in this paper in order to adequately specify certain procedures. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Further author information: (Send correspondence to C.S.)
C.S.: E-mail: chris.scrapper@nist.gov, Telephone: +1 301-975-4592
R.M.: E-mail: raj.madhavan@nist.gov, Telephone: +1 301-975-2865; R & D Staff Member, Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, U.S.A.
S.B.: E-mail: stephen.balakirsky@nist.gov, Telephone: +1 301-975-4791.

Some researchers have recognized the need for quantitative evaluation of navigation solutions and are attempting to address it through several programs. For example, the Robotics Data Set Repository (Radish)[1] provides a collection of standard robotics data sets. The OpenSLAM repository[2] contains collections of source codes of various Simultaneous Localization and Mapping (SLAM) algorithms. While a step in the right direction, they do not address objective performance evaluation and replication of algorithms is not straightforward.

At the National Institute of Standards and Technology (NIST), we are attempting to develop tools and standardized test methods to classify the performance characteristics of navigation solutions that facilitate the inter-comparison of experimental results. The development of a *de facto* standard testbed for evaluation of navigation solutions will provide a baseline for comparison and the means to target specific aspects of the navigation solution, allowing researchers to assess the performance of various systems in different scenarios and environmental conditions.

Commonly, characterizing the performance of navigation solutions is based on qualitative analysis (i.e. visual inspection) of the robot-generated maps. While this type of analysis provides some indication of the overall performance, it does not allow researchers to understand what errors a specific system is prone to and how these errors impact the overall performance of the system. Identifying *performance singularities*, or the point where the system fails to be well-behaved, is essential for developers to understand the impact of errors on the overall performance. *Performance singularity identification and testing* provides real-time meta-data that allows developers to understand the impact of singularities on the overall performance of a system. This approach to characterizing the performance of navigation solutions is divided into two parts. *Performance evaluation* uses ground truth to evaluate the navigation solution at the system level. By comparing output with ground truth, researchers are able to discover irregularities in the navigation solutions and identify areas where a performance singularity has occurred. *Performance analysis* examines a navigation solution at the algorithmic level to provide insight to the cause and repercussions of the performance singularity.

The remainder of this paper is dedicated to the performance analysis of an range image registration technique that is commonly used in the development of a stable navigation solution. The primary focus of this paper is the development of testing scenarios that enable developers rapidly compare resutls and understand the performance characteristics of a specific implementation. Section 2 provides an overview of navigation solutions and describes common approaches used to formulate them. Section 3 details the two variations of the range image registration technique that were developed as examples to test the testing scenarios. Section 4 illustrates the performance analysis of these techniques in three different environmental scenarios. Section 5 concludes the paper by summarizing the findings and outlining continuing work.

2. NAVIGATION SOLUTIONS

Many approaches to the development of navigation solutions have been proposed or implemented, some with greater success than others. The capabilities and limitations of each of these approaches vary significantly based on the requirements of the end-user, the operational domain, and the limitations of the onboard sensor suite. Understanding the strengths and weaknesses of each of these approaches in essential in producing a stable navigation solution. Methods for developing a stable navigation solution are based on sensors that can be broadly classified into two approaches, exteroceptive and proprioceptive.

Dead-reckoning is a widely used method for pose estimation that serves as the backbone for many navigation solutions. It is based on simple mathematical principles that estimate the current pose based on a previous pose. This method "advances" the pose estimate by recursively integrating motion, measured through a proprioceptive sensor, to compute a new heading and the distance traveled. Dead-reckoning is favorable because it provides a simple, cost-effective solution that is self-contained and is capable of computing pose estimates at a high frequency. The major drawback to dead-reckoning is two-fold: 1) because the system is self-contained, *systematic* and *non-systematic* dead-reckoning errors[3] are hard to eliminate and 2) the recursive nature of the algorithms allows errors to propagate and accumulate in an unbounded manner, thus causing the navigation solution to diverge.

Many navigation solutions use a landmark-based approach.[4–6] This approach geometrically computes an estimate based on the recognition of distinct features, occurring naturally or artificially placed, in the environment. The factors contributing to the successful performance and integrity of these methods is the reliable acquisition and extraction of features from sensory data and the ability to efficiently recognize and associate features with some navigational map.[4] While these methods, in general, provide an accurate pose estimate, they require either engineering the environment to provide an adequate set of features, or efficient recognition of features to use as landmarks.[7] In addition, these methods often rely on geometric primitives or models of the environment that are not guaranteed to exist in all environments.

In lieu of the landmark-based approaches, the iconic approaches attempt to utilize whatever sensor data is available to compute a navigation solution by working directly with raw sensor data. This eliminates the need to decide what constitutes a feature by minimizing the discrepancies between the raw sensor data and a model of the environment. Using a maximum likelihood alignment to find the best fit between two sets of data points, this method is capable of providing a computationally efficient pose estimate in complex, unstructured environments. Examples of iconic-based methods in the literature are found in Bailey,[8] Nieto,[9] and Gonzales.[10]

In the late 1980s, Smith *et al.*[11] introduced a new approach in the development of navigation solutions that relied on the correlation of spatial relationships between the vehicle's pose and features in the environment. Later formalized by Leonard and Durrant-Whyte,[4] Simultaneous Localization and Mapping uses statistical methods to fuse high-frequency predictions of vehicle maneuvers with low-frequency observation of the external environment to bound the errors in the pose estimate. Over the past decade, many implementations of SLAM have utilized an iconic approach for the observation model that is based on a range image-registration technique known as scan-matching.[12,13]

3. SCAN-MATCHING ALGORITHM

Many environments in which mobile robots are currently operating do not guarantee static landmarks or the presence of geometric primitives that can be used by navigation solutions. For such environments, an iconic approach that uses direct correlation of unprocessed data to formulate a navigation solution has been employed. This technique, referred in the literature as scan-matching, eliminates the need to define feature models and avoids misclassification due to imperfect sensor models.

We employ a variation of a 3D *fine range image registration method*, known as Iterative Closest Point (ICP) algorithm.[14] The basic scan-matching technique uses point-to-point correspondences in consecutive sets of data points obtained from a laser rangefinder (scans) to compute relative pose estimates. In its simplest form, this navigation solution computes these pose estimates using a maximum likelihood alignment to find the best fit between two sets of data points as shown below:

1. For each point in data set **D**, compute its nearest neighbor in data set **M**.
2. Compute the incremental transformation consisting of a rotation and translation, (**R**,**T**) based on correspondences obtained in step 1.
3. Apply the incremental transformation from step 2 to **D**.
4. If relative changes in **R** and **T** are less than a predetermined threshold or a tolerable number of iterations is exceeded, terminate. Else go to step 1.

Recent improvements in the search strategy for finding data associations between the two sets of data has made this technique a computationally efficient way of generating navigation solutions in environments with minimal structure.[15] However, shortcomings in the basic ICP algorithm can lead to erroneous pose estimates, jeopardizing the integrity of the basic scan-matching technique.[8] In order to improve the convergence characteristics of the basic scan-matching algorithm and to yield more accurate results, we have expanded previous efforts by Bailey[8] and Zhang[16] to develop an *adaptive scan-matching algorithm* that addresses these shortcomings. Understanding the shortcomings of a particular algorithm can provide insight into the performance singularities that might arise in the navigation solutions and will provide insight into how to overcome these singularities. Below we identify two shortcomings of the basic scan-matching algorithm and how the implementation of the adaptive scan-matching algorithm differs from the basic scan-matching algorithm. In Section 4, we will analyze the performance characteristics of these two variants of the scan-matching algorithm.

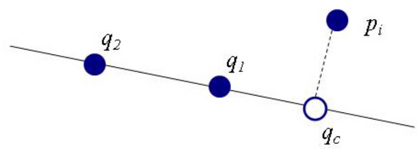

Figure 1: Pseudo point matching is a variant of a point-to-plane data association technique that is an alternative to point-to-point data association.

3.1 Pseudo-Point Matching

Point-to-point data association used by the basic scan-matching algorithm treats scan data as a set of discrete locations. This association of the data points is used to derive a transformation between the successive scans to estimate relative movement of the vehicle. However, the data contained in a scan usually represents a surface and not a set of discrete locations. This can lead to spurious point matching and can cause errors in the pose estimate. Pseudo point matching is a point-to-line data association technique that approximates the real distance between a point and a line[16] (shown in Figure 1). We establish correspondence with a virtual point, q_c, that is the closest point on the line defined by the nearest neighbors found in the model data set, namely q_1 and q_2. The virtual point is derived using the relation:

$$q_c = q_1 + \frac{(p_i - q_1)\dot{(}q_2 - q_1)}{\|q_2 - q_1\|^2}(q_2 - q_1)$$

3.2 Adaptive Thresholding

The least-square objective function used in the basic scan-matching algorithm has no means to tackle uncertainties inherent in sensor data and to evaluate the validity of correspondences. This means that all correspondences between the data points are equally weighted. In order to deal with spurious points/false matches and to account for occlusions and outliers, we modify and weight the least squares objective function such as:

$$min_{(\mathbf{R},\mathbf{T})} \sum_i w_i \; ||\mathbf{M}_i - (\mathbf{R}\mathbf{D}_i + \mathbf{T})\,||^2 \tag{1}$$

If the Euclidean distance between a point p_i in one set and its closest point q_i in the other, denoted by $d_i \triangleq d(p_i, q_i)$, is bigger than the maximum tolerable distance threshold $\mathcal{D}_{max}$, then w_i is set to zero in Equation (1). This means that a p_i in one set will not be paired with its counterpart, q_i, since the distance between reasonable pairs can not be very big. The threshold is implemented with respect to two observations: (a) If $\mathcal{D}_{max}$ is too small, then several iterations are required for the algorithm to converge and several good matches will be discarded, and (b) If $\mathcal{D}_{max}$ is too big, then the algorithm may not converge at all since many spurious matches will be included. The adaptive scan-matching algorithm uses statistical analysis of the distance between corresponding points to adaptively set the value of $\mathcal{D}_{max}$.[16]

4. PERFORMANCE ANALYSIS OF TWO SCAN-MATCHING ALGORITHMS

The primary focus of this is paper is not the comparison of scan-matching algorithms to determine which is better. Rather, its focus is the development of standardized test methods and evaluation tools that will facilitate the development of robust scan-matching algorithms and the inter-comparison of results. For illustrative purposes we will use the two scan-matching variants discussed in Section 3, namely the basic scan-matching algorithm and the adaptive scan-matching algorithm, to analyze the performance characteristics of the two different scan-matching algorithms. In the ensuing subsections we will first describe the development of the testing scenarios used to test the resilience to correspondence errors. This will be followed by the performance analysis of the two scan-matching variants in each of these scenarios with an emphasis on vehicle speed.

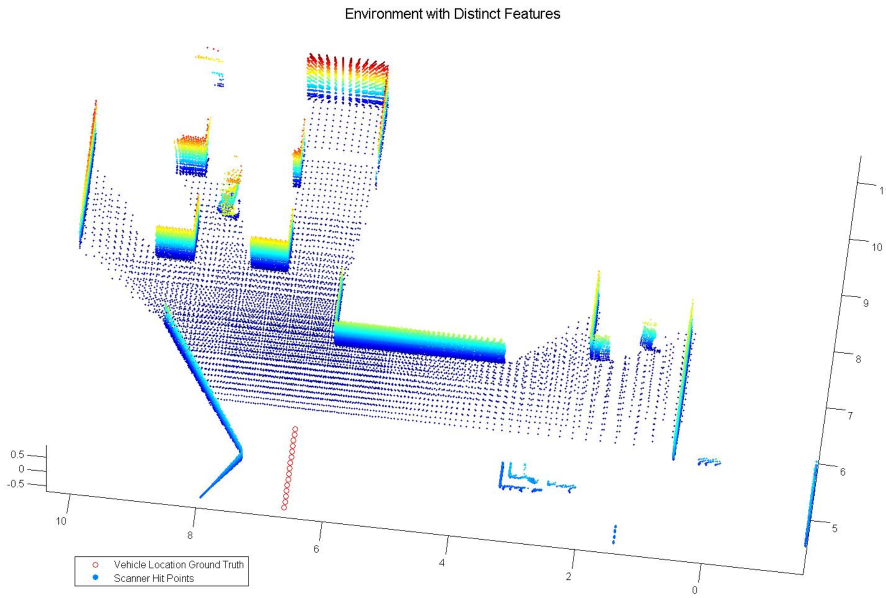

Figure 2: This figure provides a 3D visualization of the testing scenario that provides an environment with distinct features. It uses the ground truth in the reference data set to compute the sensor hit points and plot the locations where the scans were logged. Scanner hit points are colorized based on height.

4.1 Testing Scenarios

When analyzing the performance of navigation solutions, it is important to develop repeatable and reproducible testing scenarios that target specific aspects or shortcomings of the algorithm. The inability to reliably determine valid correspondences, i.e., associating a feature in one scan with its counterpart in another scan, is a performance singularity that jeopardizes the integrity of most navigation solutions that rely on exteroceptive sensors. Outliers, occlusions, and sensor noise are factors that can complicate the determination of valid correspondences. Developing testing scenarios that isolate these factors enables developers address how each factor impacts the overall performance of the system.

As a part of an effort at NIST to define standardized test methods for evaluating navigation solutions, we have developed three testing scenarios that test the robustness of correspondence determination in scan-matching algorithms. Each test scenario is composed of a reference data set that simulates linear motion of the vehicle*. 3D scan data and the ground truth location of the vehicle is captured at 10 cm intervals along a straight line trajectory. The scan data has a horizontal field-of-view of 180°and vertical field-of-view of 20°, both with an angular resolution of 1°. This produces 3D scans that may contain 3801 hit points. Using different combinations of the reference scans will enable developers to test how linear displacement, a function of vehicle speed and the data rate of the sensor, affects the overall performance of scan-matching algorithms.

Environments with Distinct Features

Environments with distinct features serve as the best-case scenario, where the scan-matching algorithms should perform optimally. As seen in Figure 2, this scenario contains a closed set of distinct features that are visible in every scan in the data set. This produces unique patterns in the resulting scans that enable the algorithm to disambiguate features and increases the likelihood of determining valid correspondences. It also produces

*The testing scenarios and referenced data sets have been developed in simulation using Unified System for Robots and Automation (USARSim). For more information on USARSim refer to the website http://sourceforge.net/projects/usarsim.

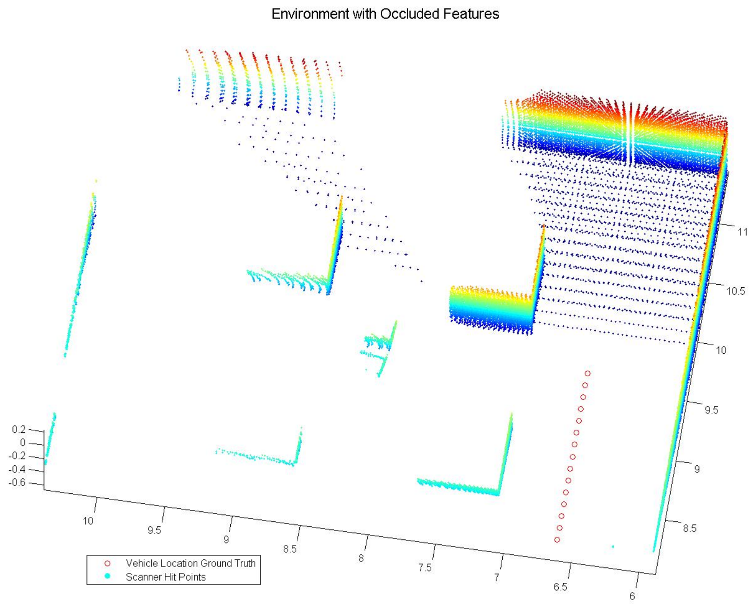

Figure 3: This figure provides a 3D visualization of the testing scenario that provides an environment with occluded features. It uses the ground truth in the reference data set to compute the sensor hit points and plot the locations where the scans were logged. Scanner hit points are colorized based on height.

surfaces in the scan that are perpendicular to the perspective of the sensor allowing the algorithm to produce a more accurate pose estimate.

This scenario limits the complexity in the environment which makes it an ideal testing scenario for two reasons. First, it provides an ideal situation for developers to tune their algorithms. Depending on the expected speed of the vehicle and/or the data rate of the sensor, developers may need to adjust thresholds that are based on distance, as they are in the adaptive scan-matching technique. Second, it provides a baseline for comparison that enables developers to understand how complexities introduced in the other scenarios affect the overall performance of the algorithm.

Environments with Occluded Features

Environments with occluded features are developed to test correspondence determination of scan-matching algorithms. This scenario isolates a problematic situation that occurs frequently in dynamic, unstructured environments, making it an ideal environment for developing mechanisms that detect and handle correspondence errors. As shown in Figure 3, the scenario contains a set of distinct features that may be occluded depending on the sensor's perspective. For example, at the beginning of the scenario the features seen in the middle of the scenario occlude the presence of the wall in the top-left corner. As the vehicle progresses through the scenario, the wall becomes visible producing a situation where consecutive scans may not contain the same set of features, increasing the likelihood of correspondence errors. However, there exists a subset of distinct features that will prevent the solutions from catastrophic failures.

Environments with Minimal Features

Environments with minimal features implement the degenerative case of the scan-matching algorithm. As shown in Figure 4, this scenario is similar to the type of environment that you would expect to encounter in long corridors or tunnels. The two walls on either side of the vehicle run parallel to the perspective of the sensor. This spoofs the algorithm's ability to compute a valid pose estimate by producing a situation where consecutive scans are almost identical, making it appear from the sensor perspective that the vehicle has not moved between

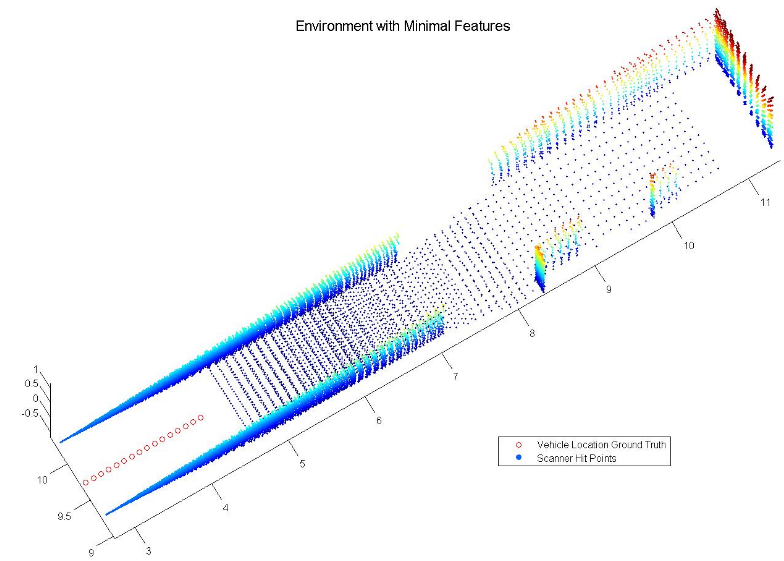

Figure 4: This figure provides a 3D visualization of the testing scenario that provides an environment with minimal features. It uses the ground truth in the reference data set to compute the sensor hit points and plot the locations where the scans were logged. Scanner hit points are colorized based on height.

scans. The only distinct feature allowed in the scenario is the far wall that is perpendicular to the sensor's perspective.

The lack of distinct features will produce catastrophic errors that prevent the scan-matching algorithm from converging to an accurate pose estimate. Understanding how an algorithm performs in the degenerative case is essential when analyzing the performance of any algorithm. It allows us to examine the characteristics symptomatic of the deterioration of the algorithm. This will assist in the development of the meta-level knowledge that will help the system to avoid or overcome performance singularities.

4.2 Performance Analysis & Experimental Results

Performance analysis takes advantage of the ground truth in the reference data sets to measure the error in the pose estimate at each iteration. These errors are plotted to produce a convergence profile. The convergence profile not only shows how well the scan-matching algorithm converges, it elucidates the convergence characteristics, such as the stability of the pose estimate. Other vital information can also be logged that will assist developers in understanding the performance characteristics found in the convergence profile. For instance, the correspondence profiles help to infer how the number of correspondences found at each iteration influences the error in the pose estimate. In the case of the adaptive scan-matching algorithm, the value of $\mathcal{D}_{max}$ reflects the quality of the registration provided by the algorithm. Analysis at this level can provide meta-level knowledge that indicates the error in the system at any given time.

We will now analyze the performance of the adaptive scan-matching algorithm, discussed in Section 3, by using the basic scan-matching algorithm as a baseline for evaluating the effects of the pseudo-matching and adaptive thresholding. In each testing scenario, described in Section 4.1, the algorithm will be subjected to several runs to measure how vehicle speed and/or data rate of the sensor impacts the algorithm's ability to compute a single pose estimate in environments with varying degrees of complexity. The linear motion of the vehicle is simulated by using different combinations of scans in the referenced data set along a straight line trajectory. The first scan in each of the referenced data sets will be used as the model data, $\mathbf{M}$, for the algorithm, or starting location. For each of the runs, the algorithm will iterate through the referenced data set, using those scans as the observed

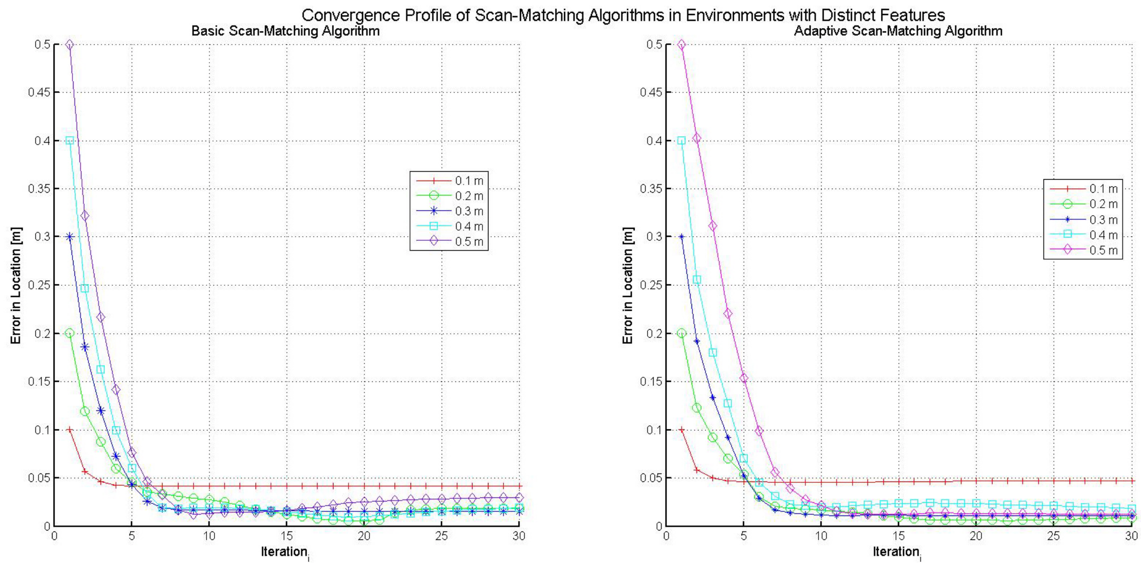

Figure 5: Convergence profile of the scan-matching variants in environments with distinct features. The convergence profiles use the ground truth in the reference data set to compute the error in the pose estimate at each iteration of the scan-matching algorithm. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

data, **D**, or the end location. This will produce linear displacements of 10 cm, 20 cm, 30, 40 cm, and 50 cm along the straight line trajectory in each of the scenarios that will be logged and analyzed below.

In Environments with Distinct Features

The convergence profiles of the two scan-matching techniques in environments with distinct features, shown in Figure 5, shows both approaches were able to rapidly converge to an accurate pose estimation (within 5 cm of ground truth) in under 10 iterations. However, it is noteworthy to point out three residual artifacts that are present in the convergence profile. First, in the 10 cm linear displacement run, the pose estimate computed by both algorithms is less accurate than their counterparts at larger displacements. Second, perturbations in the pose estimate of the basic scan-matching technique suggests the adaptive scan-matching technique produces a more stable solution. Finally, the pose estimate in the basic scan-matching algorithm appears to converge quicker than the adaptive scan-matching technique in the run with 50 cm linear displacement.

In order to gain insight into the nature of these artifacts, a close examination of the correspondence and threshold profiles of the adaptive scan-matching technique is needed, as shown in Figure 6. First, the number of correspondences found in the 10 cm of linear displacement run begins to plateau almost immediately and remains fairly constant for the remaining iterations. This differs drastically from the other runs where the number of correspondences is monotonically decreasing leading to more accurate pose estimates. Second, the adaptive scan-matching technique uses statistical analysis of the data to eliminate correspondence errors and improve registration between the data. Looking at the threshold profile, the value of $\mathcal{D}_{max}$ fluctuates until finally converging to a value close to zero. This indicates there is good registration between the points, making the pose estimate more stable. Finally, the adaptive scan-matching uses a threshold based on distance that discards correspondences with large spatial relationships, preventing it from converging as quickly under ideal conditions with no occlusions.

In Environments with Occluded Features

In environments with occluded features the convergence profiles in Figure 7 show that the adaptive scan-matching algorithm provides a more robust algorithm that was able to outperform the basic scan-matching algorithm. In the 50 cm of displacement run, neither of the scan-matching variants were able to converge to an accurate pose

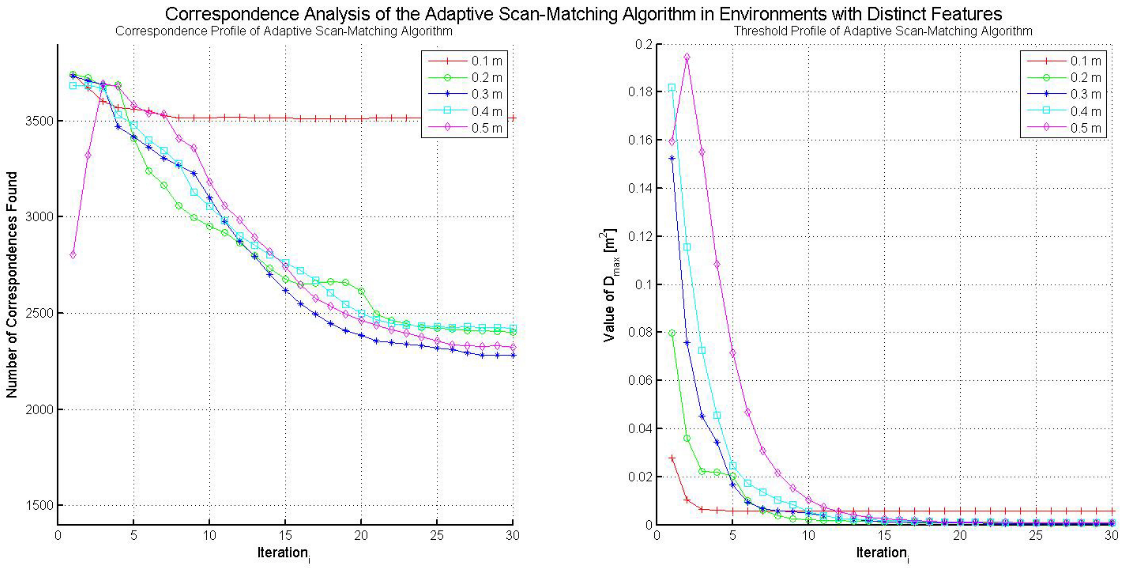

Figure 6: The two figures here show the correspondence profile and the threshold profile of the adaptive scan-matching algorithm in environments with distinct features. The correspondence profile plots the number of correspondences found at each iteration. The threshold profile shows the value of adaptive threshold, $\mathcal{D}_{max}$, at each iteration. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

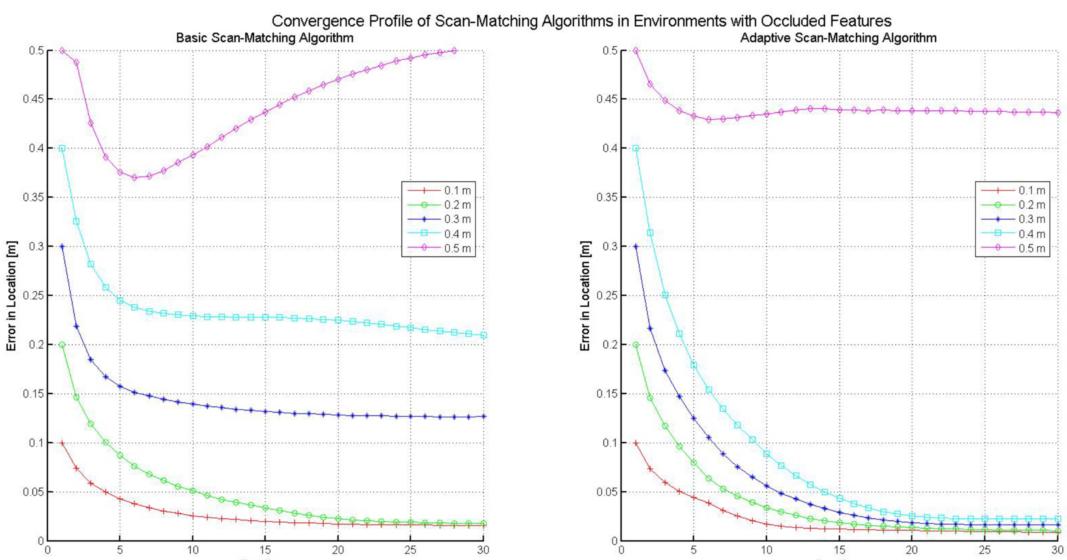

Figure 7: Convergence profile of the scan-matching variants in environments with occluded features. The convergence profiles use the ground truth in the reference data set to compute the error in the pose estimate at each iteration of the scan-matching algorithm. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

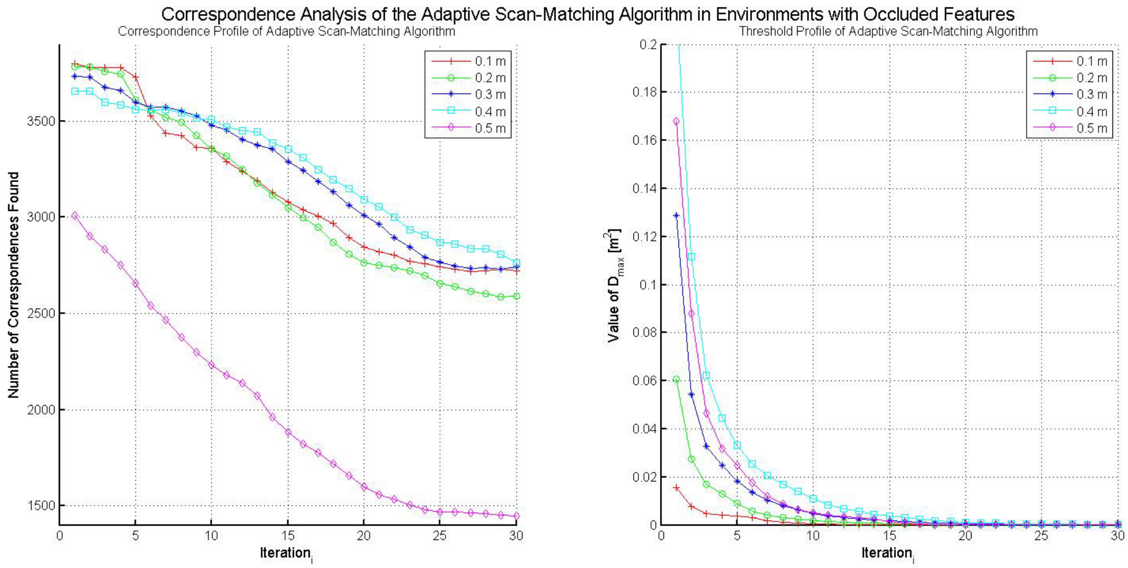

Figure 8: The two figures here show the correspondence profile and the threshold profile of the adaptive scan-matching algorithm in environments with occluded features. The correspondence profile plots the number of correspondences found at each iteration. The threshold profile shows the value of adaptive threshold, $\mathcal{D}_{max}$, at each iteration. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

estimate. However, in this run the adaptive scan-matching algorithm was able to maintain stability, where the basic scan-matching pose estimate diverges.

In order to understand why the adaptive scan-matching algorithm was able to exhibit superior performance in environments where correspondence errors are more likely, it is important to examine the correspondence and threshold profiles shown in Figure 8. Recall that each scan in the reference data set contains approximately 3801 hit points. The correspondence profile for the adaptive scan-matching technique shows the number of correspondences found at each iteration is monotonically decreasing. This is driven by the value of $\mathcal{D}_{max}$, shown in the threshold profile, which used statistical analysis of the data to dynamically set the value for $\mathcal{D}_{max}$. The statistical analysis of the data causes the value of $\mathcal{D}_{max}$ to converge as the registration between the data improves. It is also important to note two additional observation that indicate that meta-level knowledge in the adaptive scan-matching technique can help recognize the stability of the system. First, the convergence of the threshold profile coincides with the convergence of the pose estimate. Second, the correspondence profile for the 50 cm of displacement run is noticeably different than the profiles for the other run.

In Environments with Minimal Features

Figure 9 shows the effects of the degenerative case, environments that lack distinct features, on the scan-matching algorithms. Even though the basic scan-matching algorithm appears to converge slightly better, both scan-matching variants fail to produce valid pose estimates. The lack of features produces identical scans where the points in each of the resulting scans are located in the same place from the perspective of the sensor. This undermines the ability of both techniques to determine valid correspondences, which impedes their ability to compute a valid pose estimate. Since the majority of the points in the resulting scans are aligned, the adaptive scan-matching algorithm assumes good registration between the data and discards valid correspondences. The basic scan-matching algorithm does not weight the correspondences, so all available information is used. This means that valid correspondences are not discarded, which enables the pose estimate in the basic scan-matching to converge slightly better.

In order to better understand the effects of this scenario on the adaptive scan-matching algorithm, we examine the correspondence and threshold profiles in Figure 10. First, it is important to note that the value of $\mathcal{D}_{max}$

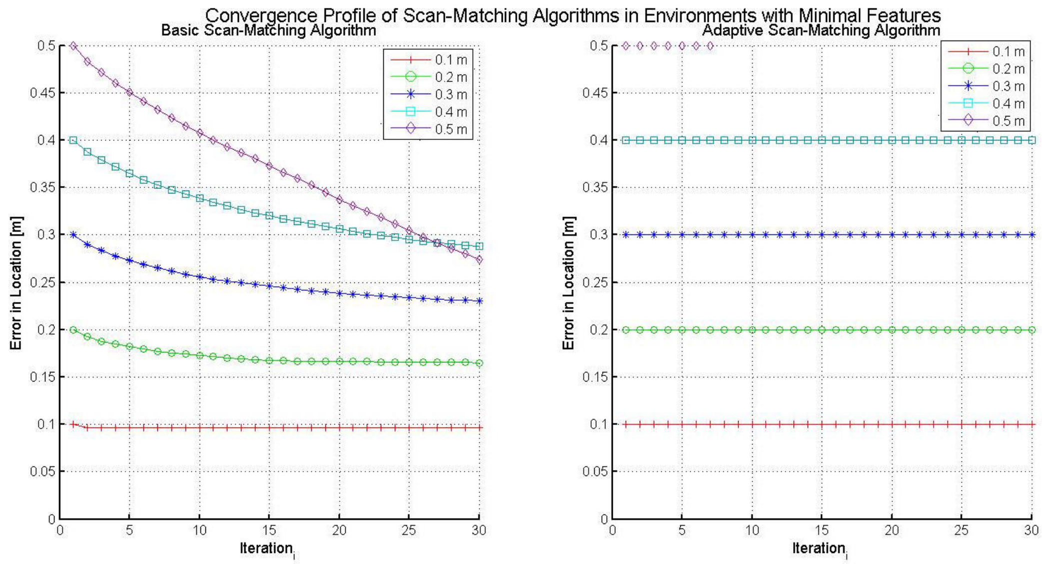

Figure 9: Convergence profile of the scan-matching variants in environments with minimal features. The convergence profiles use the ground truth in the reference data set to compute the error in the pose estimate at each iteration of the scan-matching algorithm. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

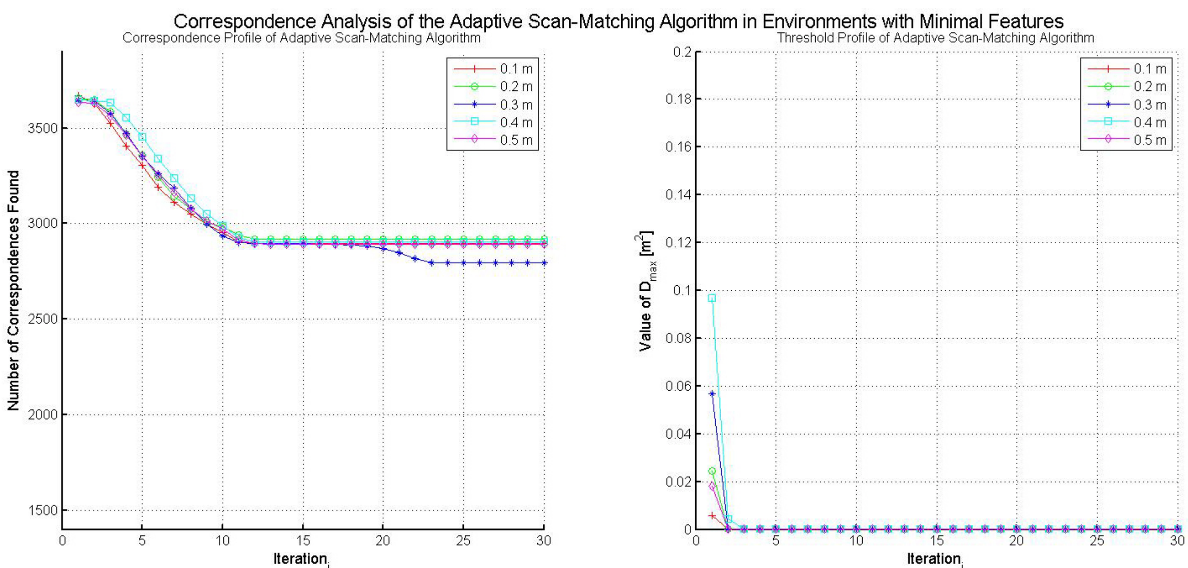

Figure 10: The two figures here show the correspondence profile and the threshold profile of the adaptive scan-matching algorithm in environments with minimal features. The correspondence profile plots the number of correspondences found at each iteration. The threshold profile shows the value of adaptive threshold, $\mathcal{D}_{max}$, at each iteration. Each of the figures shows the results of 5 runs of linear displacements along a straight line trajectory. The linear displacements for each of the runs is depicted in the legend.

in the threshold profile converges instantly. This indicates that the statistical analysis of the data incorrectly assumes there is good correspondence between the data sets. Moreover, the correspondence profile shows the number of correspondences are not monotonically decreasing and actually level out after 10 iterations. This differs drastically from the meta-level knowledge derived in the other testing scenarios where the adaptive scan-matching algorithm was able to compute a valid pose estimate.

5. CONCLUSION AND FUTURE WORK

In this paper we discussed an ongoing effort at NIST to develop standardized test methods and evaluation tools for the quantitative evaluation of navigation solutions. We presented three testing scenarios developed to analyze the performance characteristics of scan-matching algorithms with an emphasis on linear displacement. These testing scenarios have permitted us to explore how correspondence determination can jeopardize the integrity of the techniques. We illustrated how these methods facilitate the inter-comparison of experimental results and how they can be used to explore meta-level knowledge that can help researchers to develop scan-matching techniques that are more resilient to correspondence errors. Developing this meta-level knowledge is essential to recognize when a performance singularity is occurring and for the development of mechanisms that will allow these techniques to overcome and/or avoid performance singularities.

The performance analysis of the two scan-matching techniques has shown that the adaptive scan-matching technique provides a more robust correspondence determination that enables it outperform the basic scan-matching techniques in situations where correspondence errors are more likely. The analysis also showed that the adaptive scan-matching algorithm provides a more stable solution with meta-level knowledge that indicates when performance singularities are occurring. In the future, we would like to evaluate how to apply this knowledge and develop mechanisms to compensate for errors that may be introduced into the scan-matching process.

REFERENCES

[1] "Robotics Data Set Repository (Radish)." http://radish.sourceforge.net.

[2] "OpenSLAM." http://openSLAM.org.

[3] Borenstein, J., Everett, H., and Feng, L., [*Navigating Mobile Robots: Systems and Techniques*] (1996).

[4] Leonard, J. and Durrant-Whyte, H., "Mobile Localization by Tracking Geometric Beacons," in [*IEEE Transactions on Robotics and Automation*], **7**, 376–382 (1991).

[5] Kleiner, A., Prediger, J., and Nebel, B., "RFID Technology-based Exploration and SLAM for Search And Rescue," in [*Proceedings of the IEEE/RSJ International Conference on Robots and Systems*], 4054–4059 (October 2006).

[6] Latecki, L. J. and Lakaemper, R., "Polygonal Approximation of Laser Range Data Based on Perceptual Grouping and EM," in [*IEEE Int. Conf. on Robotics and Automation (ICRA)*], (2006).

[7] Gutmann, J., Burgard, W., Fox, D., and Konolige, K., "An Experimental Comparison of Localization Methods," in [*IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'98)*], (1998).

[8] Bailey, T., *Mobile Robot Localisation and Mapping in Extensive Outdoor Environments*, PhD thesis, University of Sydney (2001).

[9] Nieto, J., Bailey, T., and Nebot, E., "Recursive scan-matching SLAM," *Robotics and Autonomous Systems* **55**(1), 39–49 (2007).

[10] Gonzalez, J., Stentz, A. T., and Ollero, A., "An Iconic Position Estimator for a 2D Laser RangeFinder," in [*Proceedings of IEEE International Conference on Robotics and Automation (ICRA '92)*], **3**, 2646 – 2651 (May 1992).

[11] Smith, R., Self, M., and Cheeseman, P., "Estimating uncertain spatial relationships in robotics," 167–193 (1990).

[12] Wulf, O., Nchter, A., Hertzberg, J., and Wagner, B., "Ground Truth Evaluation of Large Urban 6D SLAM," in [*Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '07)*], (2007).

[13] Thrun, S., Hähnel, D., Ferguson, D., Montemerlo, M., Triebel, R., Burgard, W., Baker, C., Omohundro, Z., Thayer, S., and Whittaker, W., "A System for Volumetric Robotic Mapping of Abandoned Mines," in [*Proc. of the Intl. Conf. on Rob. & Autom.*], (2003).

[14] Besl, P. J. and McKay, N. D., "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992).

[15] Nuchter, A., Lingeman, K., Hertzberg, J., and Surmann, H., "6D SLAM with Approximate Data Association," in [*Proc. Intl. Conf. on Adv. Robotics*], (July 2005).

[16] Zhang, Z., "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Comput. Vision* **13**(2), 119–152 (1994).