2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Good Experimental Methodologies for Robotic Mapping: A Proposal

Francesco Amigoni, Simone Gasparini, and Maria Gini

***Abstract*— A way to significantly advance robotic science is to perform experiments that can be replicated by other researchers and be used to compare different methods. This happens rarely in current robotics research. In this paper we present a methodology for performing experimental activities in the area of robotic mapping. The proposed methodology prescribes a number of issues that should be addressed when experimentally validating a mapping method. We present the application of the proposed methodology to a mapping system we have developed.**

I. INTRODUCTION

Efforts are being made to establish standard benchmarks for mobile robots [1], [2]. However, benchmarks only allow to compare different robotic systems by considering them as black boxes, namely by comparing their "external" performance. In our opinion, this is not enough to advance significantly the science of robotics. Beyond, and before, benchmarks, what is needed is to give researchers the ability to repeat and to cross-check the experiments done by others. The repeatability and controllability of experiments are distinctive features of any scientific discipline. However, these issues are not sufficiently addressed by current experimental practices in robotics.

In this paper we propose a methodology for performing experimental activities in the area of robotic mapping [3]. More specifically, we concentrate on mapping methods that operate on segment-based maps. The proposed methodology prescribes a number of issues that, when addressed in the experimental validation of a mapping method, will enable the replication and cross-checking of experiments and the comparison with other methods. We also present the application of the proposed methodology to a specific mapping system we have developed in previous work [4].

The basic idea from which we start is that fully replicable and controllable experiments can enable a widespread comparison of different mapping systems. The methodology presented in this paper is still preliminary. However, we believe it is a first step toward the final goal to foster the good practice of replicating experiments in order to compare different methods and assess their strengths and weaknesses, as usually happens in other scientific disciplines.

The original contributions of this paper, whose nature is more methodological than technical, are (*a*) a critical survey of the current practices in experimental activities for segment-based mapping methods and (*b*) the proposal of a preliminary methodology for performing these experimental activities. To the best of our knowledge, we are not aware of any other similar effort in the area of robotic mapping.

This paper is structured as follows. In the next Section, we overview the current practice in performing experiments for segment-based mapping methods. In Section III we describe the proposed methodology, which we then apply to our mapping system in Section IV.

II. STATE OF THE ART OF EXPERIMENTS FOR ROBOTIC MAPPING SYSTEMS

We concentrate on mapping systems that build maps composed of 2D line segments. Usually, these systems work by incrementally integrating scans acquired in the environment. For this reason, they are usually based on *scan matching* methods. Scans are the result of perception actions performed by robots exploiting their sensors (e.g., laser range scanners). Scans are usually represented by line segments obtained by post-processing the sensor data.

In this Section we overview how the experimental activities aimed at validating scan matching methods have been conducted and reported in the literature. Without attempting to be exhaustive, we selected the major methods for building segment-based maps that have been proposed in the last few years. While algorithms are usually well detailed and explained in published papers (even if making code publicly available should be encouraged), the experimental sections often lack information useful to completely replicate the experiments and evaluate the methods. Table I summarizes the basic information about the experimental activities and whether they are provided by the authors in their papers. We have considered only the features that have an impact on the repeatability of the experimental activities and on the comparison between different mapping methods. In the following we discuss the data reported in the table.

Among the methods we considered, almost all experimental activities are conducted in indoor environments with "private" data, not freely available to other researchers. Only in few cases methods are evaluated using public data, available on web repositories such as [5].

The use of publicly available data can help researchers not only to replicate an experiment but also to evaluate the performance of different methods when applied to the same data. In fact, the publications we examined do not perform any comparison, even qualitative, with other methods, with the exception of [13], [15] that compare their methods with previous methods they have extended or improved.

Information about the size of the mapped environments is a qualitative gauge of a method, since the ability to

F. Amigoni and S. Gasparini are with the Dipartimento di Elettronica e Informazione of the Politecnico di Milano, Milano, Italy, {amigoni,gasparini}@elet.polimi.it.

M. Gini is with the Department of Computer Science and Engineering of the University of Minnesota, Minneapolis, USA, gini@cs.umn.edu.

method	public data	comparisons	env. size	# line segments	displacement	ground-truth	evaluation	parameter values	loop	proc. time
[4]	yes [5]	no	inferable	yes	yes	no	visual	yes	yes	yes
[6]	no	no	inferable	n.a.	n.a.	no	visual	yes	yes	n.a.
[7]	no	no	no	n.a.	yes	no	visual	n.a.	yes	n.a.
[8]	no	no	yes	n.a.	n.a.	yes	pose estimate	n.a.	no	n.a.
[9]	no		inferable	yes	inferable	no	visual	n.a.	no	yes
[10]	no	no	inferable	yes	inferable	simulated	visual and pose estimate	n.a.	yes	n.a.
[11]	no	no	yes	n.a.	yes	simulated	visual and pose estimate	n.a.	yes	n.a.
[12]	no	no	yes	n.a.	n.a.	no	visual and pose estimate	n.a.	no	yes
[13]	no	yes	inferable	n.a.	inferable	simulated	visual	n.a.	yes	n.a.
[14]	yes [5]	no	no	n.a.	inferable	no	visual	n.a.	yes	yes
[15]	no	yes	yes	yes	n.a.	yes	numerically w.r.t. ground-truth map	yes	yes	yes
[16]	no	no	inferable	yes	n.a.	yes	visual	yes	no	n.a.
[17]	no	no	inferable	n.a.	n.a.	no	visual	n.a.	no	n.a.
[18]	no	no	no	n.a.	n.a.	no	visual	n.a.	no	n.a.
[19]	no	no	no	n.a.	yes	no	pose estimate	n.a.	no	n.a.
[20]	no	no	inferable	n.a.	n.a.	no	visual	n.a.	yes	n.a.
[21]	no	no	no	n.a.	n.a.	yes	visual	n.a.	yes	yes
[22]	no	no	yes	n.a.	inferable	no	visual	n.a.	yes	yes
[23]	no	no	no	yes	n.a.	no	visual and pose estimate	n.a.	no	yes
[24]	no	no	yes	yes	n.a.	no	visual	n.a.	no	yes

TABLE I

FEATURES REPORTED IN EXPERIMENTAL ACTIVITIES FOR SOME SCAN MATCHING METHODS

cover large environments suggests that the method is reliable and does not diverge even after many iterations. The size is usually explicitly reported or it is inferable from the figures. Only in few cases this information is missing. The number of line segments that constitute the final map is another important information since it is related to the cost of storing the maps produced by the method. Only few authors provide this kind of information. The displacement between scans is also useful to compare the effectiveness of the methods. Some methods are (sometimes implicitly) based on the assumption of (very) small displacements between scans (e.g., [20]), which implies a simpler approach to the scan matching problem but the use of a larger number of scans to cover the environment. Conversely, larger displacements between scans require fewer scans but present more challenging problems in matching the scans [4]. Despite its importance, information about displacement is absent in many publications; in some cases it is inferable according to the length of the path covered by the robot and the number of scans [9], [10], [22], or from the robot positions marked in the maps [13], or from the information contained in the repository if public data are used [4], [14].

The obtained final maps are usually evaluated visually with respect to the real environment or to a ground-truth map, if available [8], [15], [16], [21]. Only [15] evaluates numerically the likeness of the obtained map by considering the geometrical distance between the map and a ground-truth map. Some authors [10], [11], [13] use simulated data to better compare the results with respect to an artificial known environment that represents the ground-truth. In some cases the final map is evaluated by calculating the error on the robot pose; for example, by calculating the error between the starting and ending poses of the robot after following a closed path or by comparing the pose of the robot with some ground-truth landmarks [15]. Note that multirobot mapping systems (e.g., [14]) usually need to evaluate the errors on the mutual estimated pose of the robots. However, we do not take into consideration this issue in this paper.

The parameters and their values are vital information in order to replicate the experimental activity. Nevertheless only few authors report the values at least of the main parameters involved. Besides allowing repeatability, the values of the parameters and hints about how they influence the behavior and the performance of a method are helpful to evaluate the method itself. Only [4], [15] show the behavior of the system as the values of the parameters change. This enables a more thorough understanding of the proposed method, in addition to a better assessment of its reliability and robustness.

Closing a loop path is a challenging problem that a scan matching method must solve [3]. Accordingly, only few papers do not show any experiment with a loop path.

The processing time for building the map is useful in order to evaluate whether the method can run in real-time or not. This information is not provided by all the authors; few authors report the average time taken between two integrations [4], [9], [12], [23], [24], whereas some other authors say only that their methods work in real time without any other information [14], [15], [21], [22].

III. THE PROPOSED METHODOLOGY: DESCRIPTION

From the discussion of the previous Section, we can see clearly the heterogeneity and incompleteness of the experimental results provided to validate the proposed mapping systems. In order to contribute to overcome this situation, in this Section we propose a methodology for the experimental activities.

In order for experiments to be *repeatable*, the following issues have to be addressed when presenting experimental activities:

- The mapping system has to be applied to publicly available data.
- The values of the parameters should be indicated.
- Some experiments in which the mapping system does not perform well should be shown. This can help other researchers to better understand the system.

In order to *evaluate* and to *compare* the different methods, the following issues have to be addressed when presenting experimental activities:

- All the data about the produced maps should be clearly indicated. These include: the dimensions of the mapped environment, the number of line segments, the time required to build the map, and the displacements between the scans.
- The behavior of the mapping system for different values of the parameters should be shown. This allows one to evaluate the robustness of the system, for example to find out the ranges of values of parameters for which the system works well.
- The map produced following a closed loop path in the environment should be shown, in order to evaluate the ability of the method not to "diverge".
- When a ground-truth map is available (this is not always the case), it should be used to assess the quality of the produced map, by evaluating its distance from the ground-truth map (e.g., according to the Hausdorff metric).

With the availability of the above information, other researchers can evaluate the "soundness" of a proposed mapping method. For example, if the map of a $5\,\mathrm{m}$-long empty corridor contains $1,000$ line segments, then it is evident that the mapping system that produced the map is not "sound". However, apart from some extreme cases, evaluating the "soundness" of a mapping method may not be easy: this issue deserves more investigation in future research.

With the availability of the above information, other researchers can compare different mapping methods. When data are publicly available, the mapping systems can be applied to the same data (possibly, after some pre-processing) and the resulting maps can be compared (this could be facilitated by making the resulting maps publicly available as well). The comparison can be done according to several dimensions, including: time, space (e.g., number of line segments), and quality (e.g., distance from a ground-truth map). These dimensions are contrasting with each other; for example, a method that produces very accurate maps could require more time and produce more line segments than a method that produces less accurate maps. A criterion to compare different mapping systems should be based on an application-dependent combination of the above dimensions. This issue deserves more investigation in future research.

Many of the above issues are quite obvious but, as we have seen, only rarely all of them are addressed in published papers. These issues are by no means exhaustive and others can be easily added (e.g., the average length of line segments). However, we think that they are a good starting point for the definition of a complete and rigorous experimental methodology for robotic mapping.

Although we have presented it for segment-based maps, the proposed methodology is more general and could be applied to mapping methods that work with different types of maps. For instance, it could be applied to grid-based mapping systems without any change, except for defining scans appropriately and for substituting the number of line segments with the number of cells. Other issues could also be added, like the dimensions of the cells. Extensions of the methodology should be deeper investigated, since it may not be easy to compare a segment-based mapping system with a grid-based one.

IV. THE PROPOSED METHODOLOGY: APPLICATION

In this Section, we present an example of how to apply the proposed methodology to a mapping system we developed. We first describe our mapping system and then we illustrate some experimental results, according to the proposed methodology.

A. Our Mapping System

The mapping system we use has been extensively described in [4]. Here, we briefly summarize its main features (please refer to the original paper for the details).

The scans and the maps we consider are collections of line segments obtained from 2D laser range data. Our method is exclusively based on the geometrical information contained in the scans to be integrated. In particular, we consider angles between pairs of line segments in the scans as a sort of "geometrical landmarks" on which the matching process is based. We define a INTEGRATE function for integrating two scans into a map. Let's call S_1 and S_2 the two scans and $S_{1,2}$ the resulting map. In our method, two points are considered to coincide when they are closer than POINTDISTANCETOLERANCE and two angles are considered equal when their values differ of less than ANGLEDIFFERENCETOLERANCE. These are the two main parameters of the method. Function INTEGRATE operates in three major steps, as follows:

a) Determine the possible transformations: This step first finds the angles between the line segments in S_1 and between the line segments in S_2 and then finds the possible transformations (namely, the rotations and translations) that superimpose at least one angle α_2 of S_2 to an equal angle α_1 of S_1. Recall that angles between pairs of line segments in a scan are the geometrical landmarks we adopt.

b) Evaluate the transformations: Every transformation found in the previous step needs to be evaluated in order to identify the best one. To determine the goodness of a transformation t we transform S_2 on S_1 (in the reference frame of S_1) according to t (obtaining S_2^t), then we calculate the approximate length of the line segments of S_1 that correspond to (namely, match with) line segments of S_2^t. The value of a transformation is the sum of the lengths of the corresponding line segments that the transformation produces.

c) Apply the best transformation and fuse the line segments: Once the best transformation $\bar{t}$ has been found (i.e., the one with the largest value), this step transforms the second scan S_2 in the reference frame of S_1 according to $\bar{t}$ obtaining $S_2^{\bar{t}}$. The map that constitutes the output of INTEGRATE is then obtained by fusing the line segments of S_1 with the line segments of $S_2^{\bar{t}}$. The main idea behind the fusion of line segments is that a set of matching line segments is substituted by a single polyline, in order to reduce the dimensions of the resulting map.

We have developed methods for integrating a sequence $S_1, S_2, \ldots, S_n$ of n scans by repeatedly calling INTEGRATE. We used one of these methods, called *pivot method*, in the experimental activities presented in this paper (see [4] for details).

B. Experimental Results

Following the proposed methodology, we applied our mapping system to the “stanford-gates1” data set publicly available in the Robotics Data Set Repository (Radish) [5]. (Note that the experiments presented in this Section are complementary to those presented in [4].) The data set is a 30-minute tour through the first floor of the Stanford’s Gates Computer Science Building. The robot used to collect the data is a Pioneer 2DX with a forward-pointing SICK LMS 200. The laser was running at high speed (75 Hz scans) in the 10 mm, $1°$ mode. The data set includes both laser data and odometry data. We considered only laser data (about $115,000$ laser scans). For each scan of the data set, we approximate the points acquired by the laser range scanner by line segments following the approach described in [11] (the distance between a perceived point and its approximating line segment is 25 mm at most). We call scans S_x, where x is the time (in seconds) at which a scan has been acquired, according to the timestamps reported in the data set. The mapping system has been coded in C++ and run on a 2 GHz Pentiunm IV computer with Debian Linux 3.1.

POINTDISTANCETOLERANCE has been set to 10 mm and ANGLEDIFFERENCETOLERANCE to 0.2 rad. We selected these values by performing some integrations of scans randomly taken from the data set. Then, we tested the robustness of our method with respect to variations of these values. To this end, we considered three randomly selected pairs of scans of the stanford-gates1 data set and we applied our scan integration method varying the values of POINTDISTANCETOLERANCE and ANGLEDIFFERENCETOLERANCE. The method has been able to correctly integrate the pairs when the above parameters had values within 3 mm and 13 mm for POINTDISTANCETOLERANCE and 0.16 rad and 1.53 rad for ANGLEDIFFERENCETOLERANCE.

Sometimes, some values for the parameters do not allow to find the correct integration between two scans. Here we report some examples, in order to better clarify the role of the parameters. (In these examples, we do not apply the fusion step of INTEGRATE to the maps.) Fig. 1 shows the integration of scans S_{1238} and S_{1242} for different values of POINTDISTANCETOLERANCE. In this case, a smaller value of POINTDISTANCETOLERANCE allows our method to integrate the two scans with higher accuracy. In general, smaller values of POINTDISTANCETOLERANCE make the method stricter when evaluating if two points coincide, thus achieving an higher accuracy in integrating scans. Fig. 2 shows the integration of scans S_{1536} and S_{1540} for different values of ANGLEDIFFERENCETOLERANCE. In this case, a larger value of ANGLEDIFFERENCETOLERANCE allows our method to integrate the two scans with higher accuracy. In general, larger values of ANGLEDIFFERENCETOLERANCE make the method more tolerant when comparing two angles in order to find a possible transformation between two scans, thus increasing the number of transformations to evaluate.

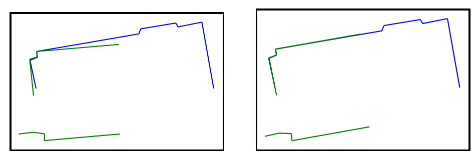

Fig. 1. Integration of scans with POINTDISTANCETOLERANCE = 9 mm (left) and with POINTDISTANCETOLERANCE = 2 mm (right)

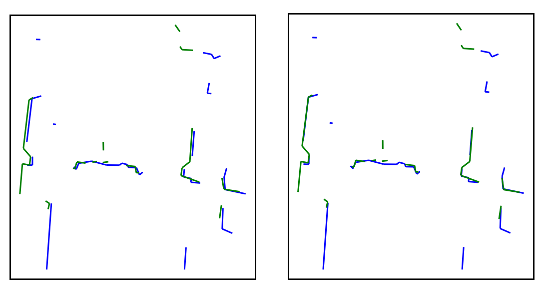

Fig. 2. Integration of scans with ANGLEDIFFERENCETOLERANCE = 0.2 rad (left) and with ANGLEDIFFERENCETOLERANCE = 0.33 rad (right)

We applied our mapping method to build the maps of four large portions of the environment of the stanford-gates1 data set. Fig. 3 shows these maps, while Table II reports some information about them. Finally, we applied our mapping method to integrate these maps to obtain the global map of the environment. Results are shown in Fig. 4 and in the last row of Table II.

The results show that our mapping system is robust and produces maps composed of a limited number of line segments (according to the dimensions of the environment). Note that the majority of the line segments are of little importance (they are due to noise in data acquisition and in scan integration): filtering out those shorter than 300 mm

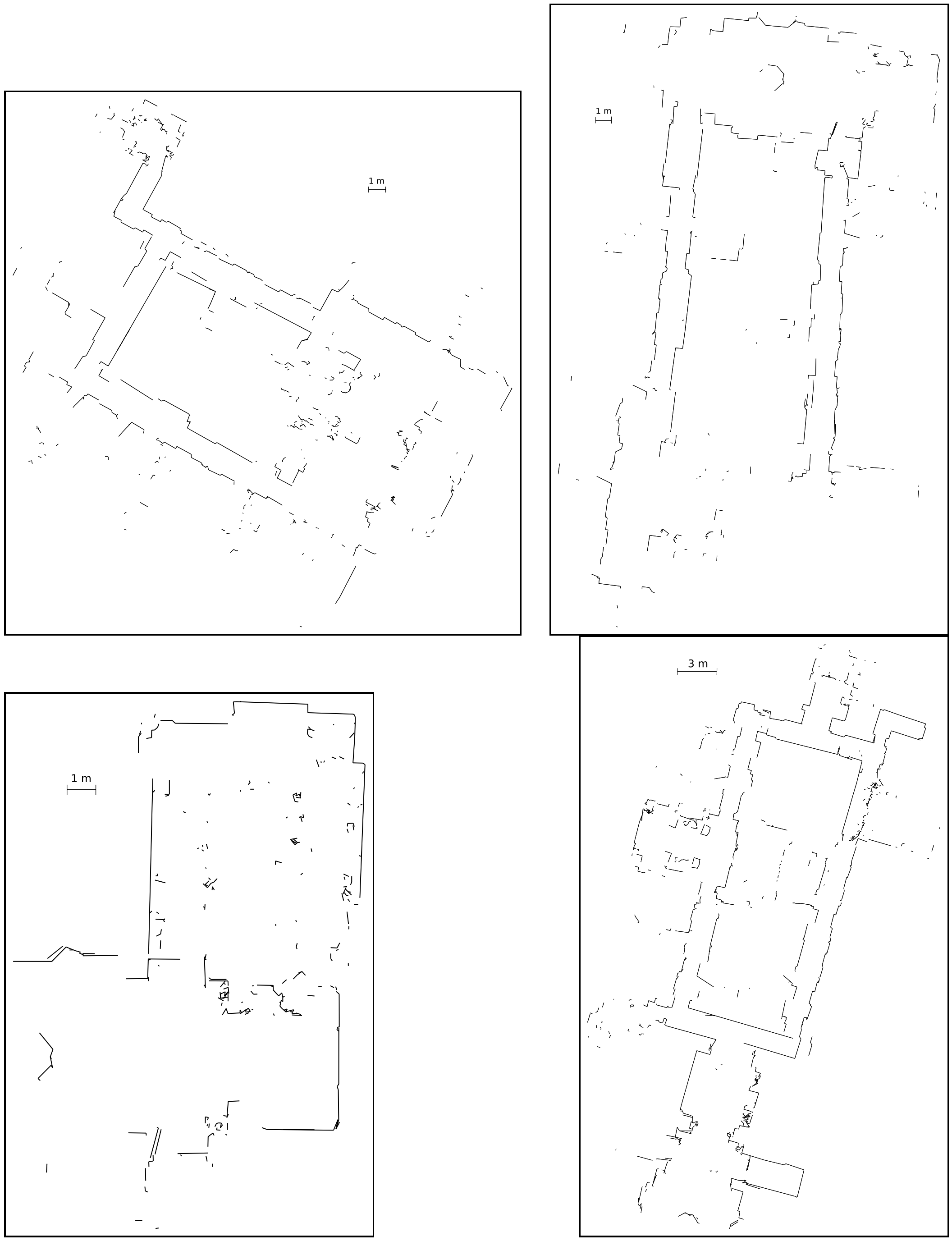

Fig. 3. Maps A (top left), B (top right), C (bottom left), and D (bottom right)

map	# scans	env. size	# line segments	average displacement	proc. time
A	59	20 m $\times$ 15 m	673	0.25 m	12.2 s
B	67	25 m $\times$ 10 m	552	0.25 m	8.9 s
C	31	13 m $\times$ 10 m	296	0.25 m	1.1 s
D	137	33 m $\times$ 15 m	1025	0.25 m	143.5 s
global		45 m $\times$ 50 m	2401		372.5 s

TABLE II

INFORMATION ABOUT MAPS A, B, C, D, AND THE GLOBAL MAP

the number of line segments in the four maps drops to 150, 178, 51, and 273, respectively. The system can also perform in real-time, requiring (on average) less than 1 s to integrate a scan. The error in closing a 50 m loop (upper corridor of the map A in Fig. 3) is about 0.5 m, that is reasonable given the dimensions of the environment. Note that if, differently from our case, a mapping system involves casual choices, mean and variance of closing-loop error should be reported.

We expect that the information presented in this Section and the fact that we have used publicly available data (following the proposed methodology) will allow other researchers

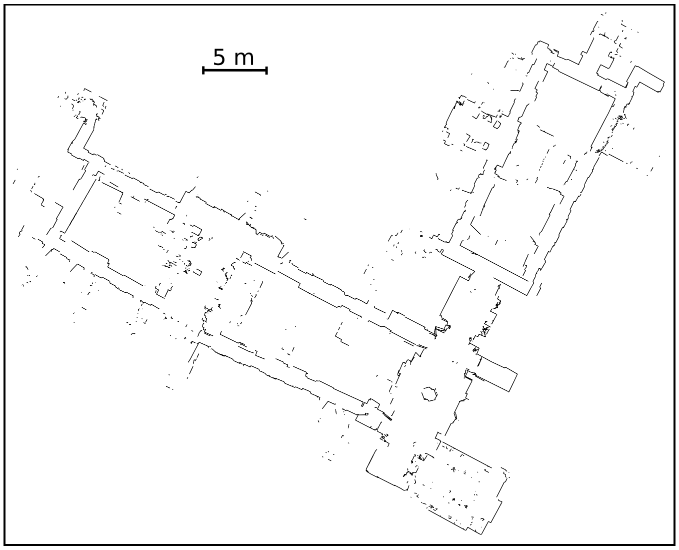

Fig. 4. Global map

to repeat our experiments both for cross-checking them and for comparing other mapping methods to ours.

V. CONCLUSIONS

In this paper, we have presented a preliminary version of a methodology for performing good experimental activities in the field of robot mapping systems. The proposed methodology has been applied to a mapping system to show how it can improve the repeatability and controllability of experimental results and the comparison of different mapping systems.

Future work will address a more complete definition of the proposed methodology. For example, some ways to evaluate the "soundness" and the robustness of a mapping method and to compare different mapping methods are needed. An interesting approach has been adopted in [25], where the authors compare their method with other two methods based on the Iterative Closest Point framework. They match each scan against itself with an initial random displacement. In this way, the ground-truth is known and the accuracy of each method can be evaluated and compared. Although ICP methods work with point-based maps, this experimental approach can be easily redeployed for segment-based maps. Finally, the methodology should be "put at work" by applying it to different mapping systems in order to assess its potential.

VI. ACKNOWLEDGMENTS

The authors would like to thank Igor Giussani for producing some of the experimental results and Brian Gerkey for providing the stanford-gates1 data set.

REFERENCES

[1] EURON, "Survey and inventory of current efforts in comparative robotics research," http://www.robot.uji.es/benchmarks/index.html.
[2] NIST, "Intelligent systems division - performance metrics and test arenas for autonomous mobile robots," http://www.isd.mel.nist.gov/projects/USAR/rules.htm.
[3] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millenium*, G. Lakemeyer and B. Nebel, Eds. Morgan Kaufmann, 2003, pp. 1–35.
[4] F. Amigoni, S. Gasparini, and M. Gini, "Building segment-based maps without pose information," *P IEEE*, vol. 94, no. 7, pp. 1340–1359, 2006.
[5] A. Howard and N. Roy, "The robotics data set repository (radish)," http://www.robot.uji.es/benchmarks/index.html.
[6] E. Brunskill and N. Roy, "SLAM using incremental probabilistic PCA and dimensionality reduction," in *Proc. of the IEEE Int'l Conf. on Robotics and Automation*, 2005, pp. 342–347.
[7] J. Canou, G. Mourioux, C. Novales, and G. Poisson, "A local map building process for a reactive navigation of a mobile robot," in *Proc. of the IEEE Int'l Conf. on Robotics and Automation*, 2004, pp. 4839–4844.
[8] I. Cox, "Blanche: An experiment in guidance and navigation of an autonomous robot vehicle," *IEEE T ROBOTIC AUTOM*, vol. 7, no. 2, pp. 193–204, 1991.
[9] T. Einsele, "Real-time self-localization in unknown indoor environments using a panorama laser range finder," in *Proc. of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, 1997, pp. 697–702.
[10] A. Garulli, A. Giannitrapani, A. Rossi, and A. Vicino, "Simultaneous localization and map building using linear features," in *Proc. of the 2nd European Conf. on Mobile Robots*, 2005.
[11] H. H. Gonzáles-Baños and J. C. Latombe, "Navigation strategies for exploring indoor environments," *INT J ROBOT RES*, vol. 21, no. 10-11, pp. 829–848, 2002.
[12] J. Gonzalez, A. Ollero, and A. Reina, "Map building for a mobile robot equipped with a 2D laser rangefinder," in *Proc. of the IEEE Int'l Conf. on Robotics and Automation*, 1994, pp. 1904–1909.
[13] J. S. Gutmann and C. Schlegel, "Amos: Comparison of scan matching approaches for self-localization in indoor environments," in *Proc. of the Euromicro Workshop on Advanced Mobile Robots*, 1996, pp. 61–67.
[14] A. Howard, G. S. Sukhatme, and M. J. Matarić, "Multirobot simultaneous localization and mapping using manifold representations," *P IEEE*, vol. 94, no. 7, pp. 1360–1369, 2006.
[15] Y. L. Ip, A. B. Rad, K. M. Chow, and Y. K. Wong, "Segment-based map building using enhanced adaptive fuzzy clustering algorithm for mobile robot applications," *J INTELL ROBOT SYST*, vol. 35, pp. 221–245, 2002.
[16] K.-H. Kim and H. S. Cho, "Range and contour fused environment recognition for mobile robot," in *Proc. of the Int'l Conf. on Multisensor Fusion and Integration for Intelligent Systems*, 2001, pp. 183–188.
[17] L. J. Latecki, R. Lakaemper, X. Sun, and D. Wolter, "Building polygonal maps from laser range data," in *Proc. of the Int'l Cognitive Robotics Workshop*, 2004, pp. 56–62.
[18] M. López-Sánchez, F. Esteva, R. L. de Màntaras, C. Sierra, and J. Amat, "Map generation by cooperative low-cost robots in structured unknown environments," *AUTON ROBOT*, vol. 5, pp. 53–61, 1998.
[19] F. Lu and E. Milios, "Gloablly consistent range scan alignment for environment mapping," *AUTON ROBOT*, vol. 4, no. 4, pp. 333–349, 1997.
[20] A. Martignoni III and W. Smart, "Localizing while mapping: A segment approach," in *Proc. of the Eighteen Nat'l Conf. on Artificial Intelligence*, 2002, pp. 959–960.
[21] R. Mazl and L. Preucil, "Building a 2D environment map from laser range-finder data," in *Proc. of the IEEE Intelligent Vehicles Symposium*, Detroit, USA, Oct. 2000, pp. 290–295.
[22] T. Röfer, "Building consistent laser scan map," in *Proc. of the European Workshop On Advanced Mobile Robots*, 2001, pp. 83–90.
[23] X. Zezhong, L. Jilin, and X. Zhiyu, "Scan matching based on CLS relationships," in *Proc. of the IEEE Int'l Conf. on Robotics, Intelligent Systems and Signal Processing*, 2003, pp. 99–104.
[24] L. Zhang and B. Ghosh, "Line segment based map building and localization using 2D laser rangefinder," in *Proc. of the IEEE Int'l Conf. on Robotics and Automation*, 2000, pp. 2538–2543.
[25] J. Minguez, L. Montesano, and F. Lamiraux, "Metric-based iterative closest point scan matching for sensor displacement estimation," *IEEE Transactions on Robotics*, vol. 22, no. 5, pp. 1047–1054, 2006.