The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Landmark rating and selection according to localization coverage: Addressing the challenge of lifelong operation of SLAM in service robots

Siegfried Hochdorfer and Christian Schlegel
University of Applied Sciences Ulm
Computer Science Department, Prittwitzstr. 10, D-89075 Ulm
email: {hochdorfer, schlegel}@hs-ulm.de

Abstract— **Acting in everyday-life environments is still a great challenge in service robotics. Although algorithms and solutions already exist for many relevant subproblems, in particular the aspect of robustness and suitability for everyday use has been neglected so far very often. Robustness and suitability for everyday use are features affecting not only the overall system design but have impact on each single algorithm of each component.**

Although an overwhelming amount of work is available to address the SLAM problem, the challenge of applying a SLAM algorithm over the whole lifecycle of a service robot, perhaps even in different environments, has not been brought into focus very often. An obvious problem to be solved is the continuously growing number of landmarks. A lifelong running SLAM approach requires means to select landmarks such that they best cover the working environment given bounded SLAM resources like the maximum number of manageable landmarks.

This paper proposes a novel solution for selecting appropriate landmarks to limit the number of landmarks. The idea is to quantify the contribution of a landmark to the ability of the robot to localize itself in its working environment. Thus, the core contribution is to base the landmark selection process upon the landmarks' coverage of the working environment.

Real-world experiments on a P3DX-platform with a bearing-only SLAM approach and an omnicam confirm that the addressed question and the proposed first approach might be another step towards the overall goal of suitability for everyday use.

I. INTRODUCTION

Service robots are expected to fulfill tasks in different environments out-of-the-box. Users of service robots cannot be expected to be skilled robotics programmers. Thus, there is a huge demand on appropriate man-machine interfaces and on the ability of service robots to learn about relevant information about their deployment environment by themselves. Of course, there is a fundamental conflict between adaptability and pre-given parameter spaces defining the overall learning space on different time scales. However, various and manifold tasks in complex and varying environments will depend on the robot's ability to perform adaptations and on its ability to integrate new information about its environment.

A fundamental component in nearly every service robot is the SLAM (simultaneous localization and mapping) mechanism. Very often, the SLAM approach is used to build a map of the environment in a deployment or initialization step. Afterwards, the SLAM mechanism often is at least parameterized such that correction steps are still performed for the now known landmarks and the robot pose but no new landmarks are introduced anymore. Thus, the selected landmarks are normally not adapted in case environmental changes require so.

Of course, there is no fundamental reason why one should not let the SLAM mechanism run within a certain environment over the whole lifespan of the robot. However, each newly recognized landmark would then be added to the state vector which results in a growth of the size of the state vector without upper bound. In case of bounded resources, one thus needs a mechanism to get most out of given resources. The SLAM problem thus needs to be extended such that one selects those landmarks that, for example, ensure a certain localization quality within the working environment of the service robot given the maximum number of landmarks.

In particular, the quality of a landmark position alone is *not* a suitable measure to select appropriate landmarks. Rather, we need to represent the positions from which a landmark can be observed and used to localize a robot. Then, we can develop an approach to select those landmarks that provide an appropriate localization quality and whose observation regions cover the working environment while still not exceeding the maximum number of landmarks.

In that approach it isn't relevant where the position of a landmark is. We think that it is relevant from which position it can be observed by the robot. Thus, we quantify the benefit of a landmark in terms of observability regions and not in terms of its pose uncertainty. The rationale behind this approach is the observation that moving close to a landmark position does not yet make sure that it can be reobserved. However, if the robot moves into the region where it observed the landmark previously, the chance of reobservation is much higher. Thus, in case of limited resources, one should select and keep those landmarks that support relocalization in the whole working space and not only those landmarks that are somehow, for example uniformly, spread over the working space.

The general idea is summarized in figure 1. We assume that all landmarks in a room are visible from everywhere inside a room. Only considering the uncertainty of a landmark would typically remove landmarks with a high variance. In that example, room 2 would not be covered by landmarks

978-1-4244-3804-4/09/$25.00 ©2009 IEEE

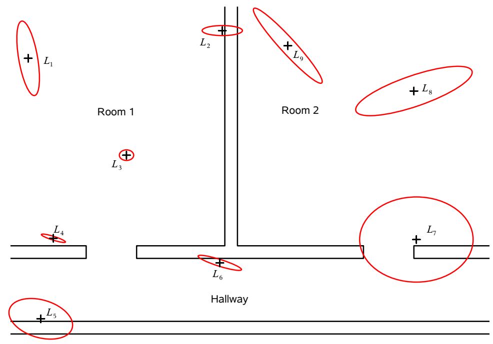

Fig. 1. Landmarks are denoted by L_x and are displayed with a covariance ellipse for illustrating purposes.

anymore. However, one should keep landmarks such that one obtains a certain localization quality in the overall working environment. Thus, it is necessary to keep landmarks such that all rooms are covered with landmarks for relocalization to avoid regions where the robot pose uncertainty becomes too large. Since there are landmarks with a low pose uncertainty in room 1, it is desirable to foremost remove landmarks in room 1 and to keep landmarks (even with higher pose uncertainty) in room 2.

The focus of this paper is of high relevance towards robust and lifelong operating service robot systems. As soon as one allows for lifelong adaptation or learning, one immediately has to face the problem of bounded resources. In case of the SLAM problem that is related to the number of landmarks representing the working environment. Thus, one needs a mechanism to select useful landmarks out of possible landmarks. This process can either be done as offline optimization procedure being invoked from time to time or as continuous selection process during the regular SLAM operation. Of coure, the latter is much more appealing for service robotic applications.

II. RELATED WORK

In the past, some researchers already defined a measure for the quality of a landmark. Maksarov et al established the *Geometric feature track quality evaluation* (GFTQ) [1]. It reflects the probability of the existence of a landmark. The sum of Gaussian probability densities of a so-called established track over the last m steps is divided by the sum of maximum values of the probability density functions over the same m steps. If we want to use GFTQ to quantify the quality of landmarks, we have to store the pair of observation position and landmark measurement for each landmark for the last m steps. This results in tremendous required resources. However, the suitability of the general approach has not been further evaluated.

Another approach is described by Dissanayake in [2]. He recommends to use those landmarks for robot localization which provide the largest information content. The lower the uncertainty of the landmark position estimate, the more benefit is provided by the landmark for robot pose estimation in case of reobservation. Therefore, he recommends to use the sum of the reciprocals of the main diagonal elements of the landmark's covariance matrix. In his experiments he also tested other information measures based on the covariance matrix, such as the Shannon or Fisher information. But their usage had no significant effect on the robot localization accuracy.

In the same work, Dissanayake uses his quality measure to compute the best landmark out of a set of landmarks. First, all landmarks are collected whose state changes in the current step from visible to invisible. From this set only the highest quality landmark is kept and all others are discarded. Thus, the selected highest quality landmark is a single representative for the set of previously visible landmarks. In the example in figure 1, L_3 should remain after leaving room 1 and entering the hallway due to visibility constraints. In case of room 2, L_9 should survive. However, selection of landmark representatives is based on a local set of landmarks and thus depends on the exploration path and the resulting visibility sequence. There is no globally related measure of landmark quality. Nevertheless, this is one of the rare approaches addressing landmark deletion with respect to a landmarks use in terms of observability.

The overall SLAM problem and the impact of the number of landmarks on its algorithmic complexity together with a summary on established approaches for optimized representations or approaches to partition a SLAM problem into separated maps and how to eventually merge them afterwards is described in [3]. However, the aspect of a lifelong running SLAM approach and its challenges are not brought into focus.

III. METHOD

The position of a landmark does not itself give a hint on its usefulness for localizing a robot. In fact, we require to know the poses from which a landmark can be observed to know in which parts of an environment this landmark can be used for localization purposes. Of course, the ability to improve the robot's pose also depends on the variance of the landmark pose estimate. Thus, the observation region of a landmark together with the landmark pose uncertainty form a good starting point for defining a measure for the benefit of a landmark with respect to robot localization.

Thus, the challenge is to determine and to represent the benefit of a landmark with respect to various poses in the working environment of the robot. This requires to calculate or to observe from which positions a certain landmark can be used for localization. Out of a set of landmarks, one than can select those landmarks that cover the working environment and also ensure a certain localization quality given the maximum number of landmarks.

The standard SLAM approach normally distinguishes an action and a sensing step. The action step comprises the robot motion which introduces further uncertainty into the robot pose estimate. The sensing step either reobserves already known landmarks and thus allows to improve the estimates of landmarks and the robot pose or it detects a previously

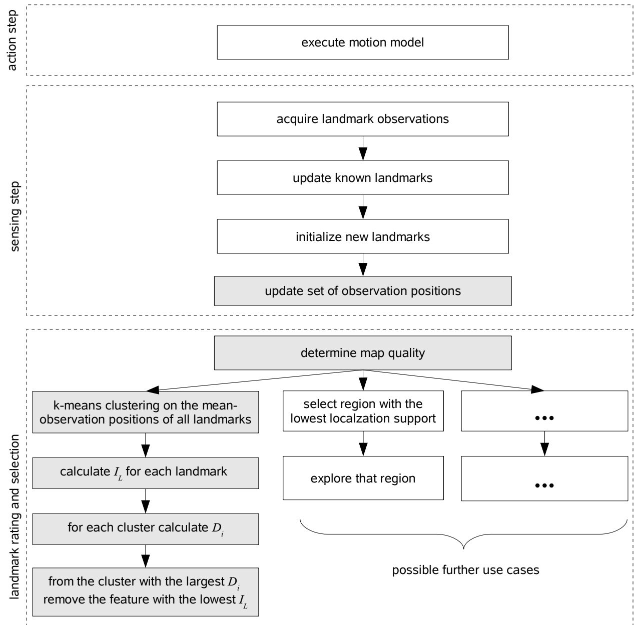

Fig. 2. Extension of the standard SLAM mechanism by functions for landmark rating and selection.

unknown but suitable landmark which can be added to the overall state vector of the SLAM system. Thus, the size of the state vector can grow without any upper bound. In case of limited resources, for example a maximum number of landmarks, a newly discovered landmark rises the question whether one should replace a previously introduced landmark and also which one.

Figure 2 illustrates the principal steps of a SLAM mechanism. The highlighted steps denote the newly added functions for landmark rating and selection. The sensing step keeps track of the set of robot poses from which a landmark has been observed so far. This provides the basis for describing from where in the environment a certain landmark is observable. This information provides the basis for evaluating the benefit of a landmark for localization purposes.

At arbitrary points of time, one can determine the benefits of landmarks for localization purposes. One approach would be to determine the landmark with the lowest impact on reducing the robot pose uncertainty while still ensuring coverage of the working environment. This landmark might then be removed from the SLAM representation of the environment.

A. Simple Approach to Landmark Rating and Selection

In principle, it is possible to store for each landmark every observation position. One approach would be to use a grid map, for example. However, this approach would result in far too huge computational and storage costs. Since we nevertheless require the observability of landmarks, we have to develop a different and much cheaper approach that avoids a fine grained representation of observability.

Thus, we evaluated the overall idea of landmark rating and selection to address the problem of bounded resources by considering restricting the maximum number of landmarks. After starting the SLAM approach, the number of landmarks is growing until reaching the predefined supported maximum. Even in this simple scenario, one from now on has to face the problem of deciding which already included landmark should be replaced by a new one.

B. Update Set of Observation Positions

The set of observation positions defines those robot poses from which the robot actually observed the considered landmark. This set needs to be represented in such a way that it can be used for landmark rating.

In case of the bearing-only SLAM approach [4], [5] used for a first evaluation of the landmark rating and selection process, the observability of the used image features is strongly limited by a minimum and a maximum viewing distance and viewing angle. Thus, the observability region typically looks like a sector of a circular ring. Here, it makes sense to represent the observability region of each landmark by calculating the arithmetic mean $E(X)$ of the observation positions.

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

X_i are the different observation positions and n is the number of observations. Thus, we only need to save the latest average position $E(X)$ and the number of observations. If the robot later re-observes a landmark, the mean can easily be updated by the newer observation position:

$$E(X_{new}) = \frac{(n-1)E(X_{old}) + X_n}{n} \tag{2}$$

C. Clustering of Representatives of Landmark Observability

The clustering step is an intermediate step only. Its purpose is to identify those representatives of landmark observability that cover nearly the same observation region.

In the used bearing-only SLAM approach, we apply k-means clustering [6], [7]. The k-means algorithm separates the landmark representatives into n clusters. In first experiments, the number of clusters is set proportional to the number of landmark representatives and is selected empirically at the current state of evaluation. As distance measure for the k-means clustering algorithm, the L_1 distance is used which reduces the computational load compared to a L_2 distance without a relevant effect on the clustering results. The reason is that observation poses are aligned on a trajectory and thus landmark representatives typically also form a chain of observation regions. Figure 3 illustrates an example of spatial clustering of landmark representatives.

Of course, the distance measure for spatial clustering based on k-means does not at all take into account more elaborate details of observability like shape or overlap of observability regions. However, the shape and visibility has already been ignored at the previous step when calculating the landmark representative. The purpose of this step is to identify landmarks that can be observed from a similar position and thus cover the same regions of the environment

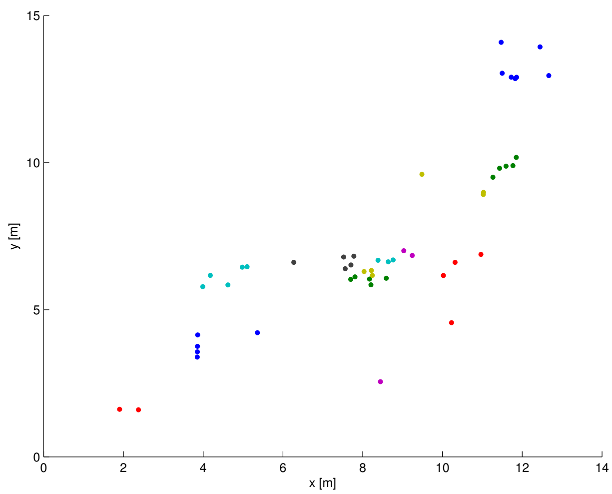

Fig. 3. A cluster comprises several landmark representatives. All representatives belonging to a cluster are drawn with the same color.

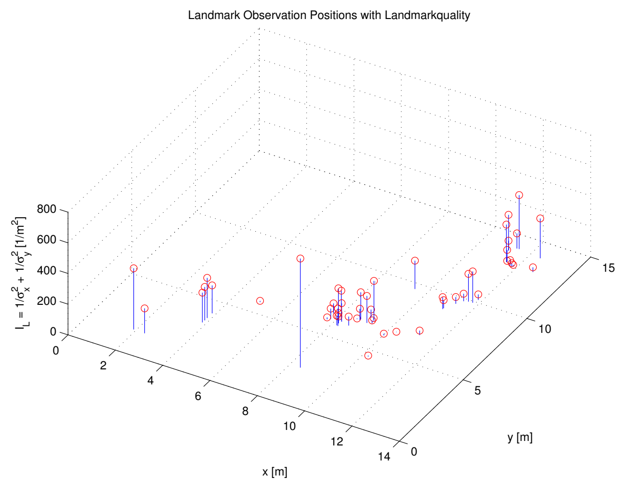

Fig. 4. For all landmarks, the estimated observation positions are plotted with the information content on the z-axis.

with respect to localization support. Therefore, one can plug-in any more advanced approach as soon as more information needs to be considered or exploited.

D. Calculation of Information Content of a Landmark

Landmarks with low position uncertainty provide a high benefit for relocalization of the robot. A simple but efficient measure for the information content of a landmark is provided by its covariance matrix. The information content can be calculated by the sum of the reciprocals of the main diagonal elements of the covariance matrix. It does not consider the correlations of the feature to other features or the vehicle. The information Content measure has also been suggested by Dissanayake in [2].

$$cov(L) = \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix} \tag{3}$$

The information content of a certain landmark is then given by

$$I_L = \frac{1}{\sigma_{xx}^2} + \frac{1}{\sigma_{yy}^2} \tag{4}$$

Figure 4 shows all estimated landmark observation positions. Furthermore, the information content of the landmark is plotted onto the z-axis.

E. Select Landmark with Lowest Localization Benefit

A landmark with a low information content in a sparsely known region is often more useful than a landmark with a higher information content in a well-known region. Under this assumption, one needs to identify the landmark with the lowest benefit for localization purposes without ignoring spatial coverage. Thus, we consider each cluster of landmarks separately. Within a cluster of landmarks, one would remove the landmark with the lowest information content to have the smallest degradation of localization quality. The idea for selecting the cluster for landmark removal is as follows. The cluster with the largest difference of information content of its landmarks is assumed to cope best with a landmark removal. The rational behind this is the assumption that the bigger the relative difference is the less is the loss of localization support within that particular cluster. This idea takes into account the landmark observability (in terms of spatial clusters) and is fundamentally different from just removing the landmark with the lowest information content from the overall set of landmarks.

The difference of information content D_i of the landmarks within the cluster C_i is calculated as follows with I_L the information content of a landmark.

$$D_i(C_i) = max(I_{L_{0\ldots m}}) - min(I_{L_{0\ldots m}}) \tag{5}$$

The cluster with the maximum difference is determined by $max(D_i)$. If now a landmark from that cluster is removed, there still remains at least another landmark which allows for localization within the spatial region covered by that cluster.

Figure (5) shows the clustered landmark representatives. The landmark with the lowest benefit for localization (at position $p = [12.5, 13.0]$) is marked by a red cross and is to be removed.

In case of an EKF based SLAM, a landmark removal is performed by simply deleting it from the state vector and also removing the appropriate row and column of the corresponding covariance matrix. Other SLAM approaches also allow for removing a landmark by deletion.

IV. RESULTS

A. Experimental Setup

In this section we present the results from a real world experiment. The localization performance is evaluated by comparing the estimated robot positions of the SLAM approach with a maximum number of 50 landmarks with ground truth measurements at 16 different timesteps. The experiment has been performed in our lab, the adjacent hallway and a neighboured room (see figure 6).

We used a Pioneer-3DX platform with a omnicam. The omnicam is a Sony DFW-X710 camera (1024x768, 1/3 inch,

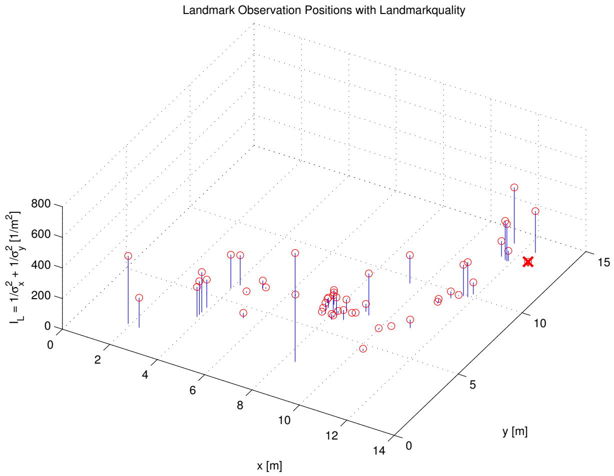

Fig. 5. The landmark with the lowest benefit for localization is marked by a red cross.

Fig. 6. The real world environment used in our experiments. The upper images show the ZAFH laboratory, the lower left image the adjacent hallway and the lower right image a neighboured room in our building (room C12).

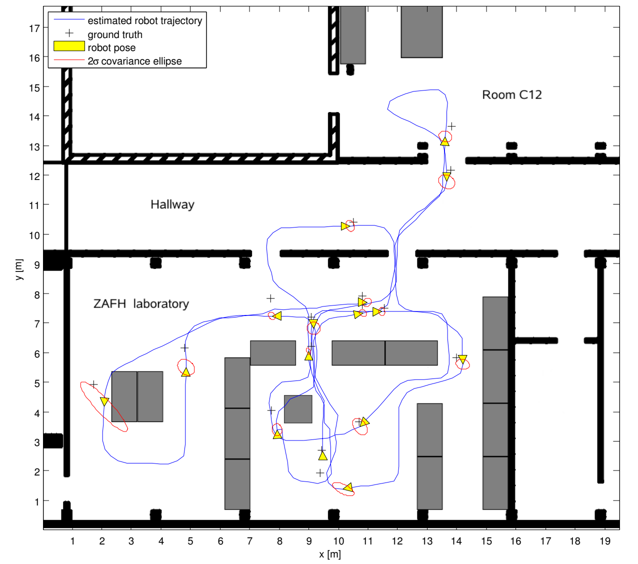

Fig. 7. The robot trajectory (blue line) in the environment. For visualisation purpose we place a floor plan in the background. The grey objects are tables. The yellow triangles represent the robot with the 2σ covariance ellipse (red). For these robot poses the corresponding ground truth positions are marked with a black cross.

progressive scan, firewire, YUV color, 15 images/second) with a hyperbolic glass mirror (H3G, Neovision). And for SLAM we use an improved version of the EKF based Bearing-Only SLAM approach with SURF Features [8] as landmarks, as described in [5]. We set the maximum number of landmarks to 50.

The trajectory forms a run of approximately 115m, so we can test whether the robot can handle the *loop closing* with the reduced number of landmarks. During the experiment three loops with a length of approximately 8m, 10m and 14m have to be closed.

The travel distance between two observation positions is approximately 0.3m. The number of clusters is set dynamically to $1/4$ of the number of currently known landmarks.

Due to the lack of GPS in indoor environments it is quite hard to get the ground truth position of the robot. We solve the problem of determining the ground truth position by measuring manually the distance from the robot to two a priori known coordinates in the environment with a Bosch Digital Laser Rangefinder (DLE 150). The accuracy of this method is approximately [0.1m, 0.1m].

B. Localization Performance in case of Restricted Number of Landmarks

We restricted the number of landmarks to 50. The robot was teleoperated while performing SLAM. It started in the ZAFH laboratory and closed the first loop (timestep 30-57) in the laboratory. Afterwards the robot entered the hallway and moved back through a second door into the already seen ZAFH lab, where the second loop (timestep 115-149) was closed. Then the robot moves through the hallway into room C12 and again back to the ZAFH lab. The experiment finished after a last loop (timestep 295-342) with a length of approximately 14m around the meeting table. The trajectory can also be seen in figure 7. Therefore the SLAM approach had to cope with loop closure three times during this experiment. The progression of the robot pose uncertainty (eigen values from the covariance matrix of the x and the y component) is plotted in figure 8. As long as the robot explores unknown regions, the pose uncertainty grows. As soon as it comes back into already known regions, the loop closure reduces the accumulated high uncertainty values. This effect can be seen around the time steps 57, 149 and 342. The loop closure capability has not been influenced by the restricted number of landmarks at all.

Figure 9 shows that the amount of initialized landmarks stays constant after the given maximum number is reached (blue line). In comparison to that the red line shows the increase of the landmark count without restricting the number of landmarks. The reduced set of landmarks covered the

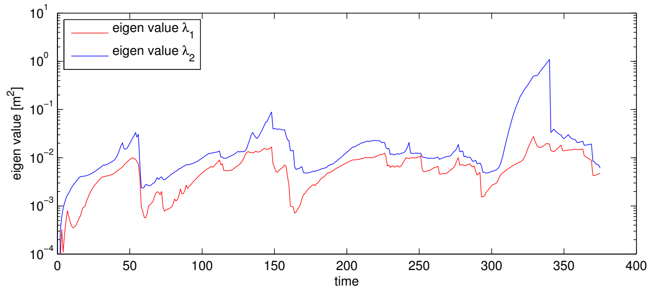

Fig. 8. Eigen values of the robot position covariance matrix during the run with restricted number of landmarks. The y-axis is log-scaled.

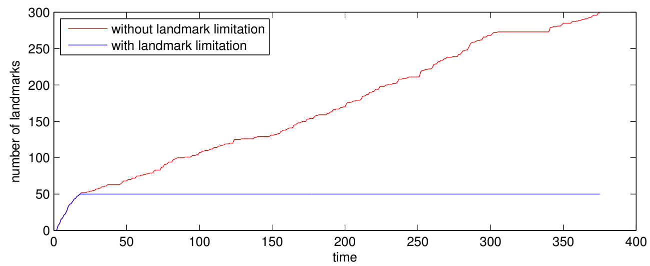

Fig. 9. The really used landmark count is illustrated as blue line and the red line shows the increase of the landmark count without restricting.

explored region in an appropriate manner as can be seen in figure 5.

Table I illustrates the euclidean distance between the ground truth measurements and the estimated robot pose for the specified timesteps. The standard deviation of the robot pose is listed in the two right columns of the table. The ground truth positions, the robot pose and the 2σ covariance ellipse can also be seen in figure 7.

V. CONCLUSIONS AND FUTURE WORK

In this paper we introduced a novel approach for landmark rating and selection. It allows to decide on which landmark to remove from a set of already discovered landmarks and to replace it by a newly detected one. This is the key towards avoiding the otherwise ever growing number of landmarks in case a SLAM algorithm is run over the lifetime of a service robot, for example.

The difference compared to existing work is that we not only select the landmark with the lowest information content to be removed. In fact, we argue that the benefit of a landmark for localization purposes needs to take into account its observability regions. This allows to keep track of which landmarks contribute to which parts of the working environment of a service robot and to thus also keep even landmarks with high uncertainty in such regions that would otherwise simply provide no localization support at all.

Of course, the implemented mechanisms to verify the relevance of the question and the overall benefit of such an approach are very rudimentary at the moment. However, even applying such a simple approach already resulted in the ability to adhere to a predefined maximum number of allowed landmarks without loss of localization quality or coverage.

The results of these first experiments exceeded our expectations. In particular, these results were achieved in the demanding setting of using an everyday indoor environment and a bearing-only SLAM approach based on SURF features. Thus, the achieved results show a promising way towards addressing suitability of SLAM algorithms for everyday use where the capability of lifelong adaptation within bounded resources is mandatory.

Future work will focus evaluating further approaches for landmark evaluations. In particular, refining the representation of the observability regions while still being able to handle these online as part of the SLAM approach is necessary. Further, it is neccessary to analyze other algorithms to group the observation positions in a spatial manner.

TABLE I

EUCLIDEAN DISTANCE BETWEEN GROUND TRUTH AND THE ESTIMATED ROBOT POSE

timestep	euclidean distance $[m]$	$\sigma_x[m]$	$\sigma_y[m]$
1	0.000	0,000	0,000
25	0.169	0,052	0,069
52	0.164	0,126	0,141
75	0.202	0,038	0,055
100	0.206	0,076	0,095
125	0.267	0,114	0,087
150	0.905	0,182	0,115
175	0.277	0,063	0,060
200	0.418	0,110	0,093
225	0.423	0,137	0,118
251	0.374	0,110	0,106
281	0.653	0,083	0,102
300	0.575	0,061	0,064
325	0.804	0,399	0,423
350	0.649	0,142	0,136
375	0.233	0,075	0,072

VI. ACKNOWLEDGMENTS

This work has been conducted within the *ZAFH Servicerobotik* (http://www.zafh-servicerobotik.de/). The authors gratefully acknowledge the research grants of state of Baden-Württemberg and the European Union.

REFERENCES

[1] D. Maksarov and H. Durrant-Whyte, "Mobile vehicle navigation in unknown environments: a multiple hypothesis approach," *IEEE Proceedings - Control Theory and Applications*, vol. 142, no. 4, pp. 385–400, Jul 1995.

[2] G. Dissanayake, H. F. Durrant-Whyte, and T. Bailey, "A Computationally Efficient Solution to the Simultaneous Localisation and Map Building (SLAM) Problem," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2000, pp. 1009–1014.

[3] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, September 2005.

[4] T. Bailey, "Constrained initialisation for bearing-only SLAM," *IEEE International Conference on Robotics and Automation (ICRA)*, vol. 2, pp. 1966–1971, Sept. 2003.

[5] C. Schlegel and S. Hochdorfer, *Service Robotics*. ITECH, ISBN 978-3-902613-29-5, 2008, ch. Localization and Mapping for Service Robots: Bearing-Only SLAM with an Omnicam, pp. 253–278.

[6] G. A. F. Seber, *Multivariate Observations*. New York: Wiley, 1984.

[7] H. Spath, *Cluster Dissection and Analysis: Theory, FORTRAN Programs, Examples*. Halsted Press,New York, 226 pp., 1985.

[8] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *9th European Conference on Computer Vision*, Graz Austria, May 2006.