Information Gain-based Exploration Using Rao-Blackwellized Particle Filters

Cyrill Stachniss[†] Giorgio Grisetti[†‡] Wolfram Burgard[†]

[†] *University of Freiburg, Department of Computer Science, D-79110 Freiburg, Germany*

[‡] *Dipartimento Informatica e Sistemistica, Universitá "La Sapienza", I-00198 Rome, Italy*

***Abstract*— This paper presents an integrated approach to exploration, mapping, and localization. Our algorithm uses a highly efficient Rao-Blackwellized particle filter to represent the posterior about maps and poses. It applies a decision-theoretic framework which simultaneously considers the uncertainty in the map and in the pose of the vehicle to evaluate potential actions. Thereby, it trades off the cost of executing an action with the expected information gain and takes into account possible sensor measurements gathered along the path taken by the robot. We furthermore describe how to utilize the properties of the Rao-Blackwellization to efficiently compute the expected information gain. We present experimental results obtained in the real world and in simulation to demonstrate the effectiveness of our approach.**

I. Introduction

Robots that are able to acquire an accurate model of their environment are regarded as fulfilling a major precondition of truly autonomous mobile vehicles. To learn an environmental model, three problems need to be addressed simultaneously, namely exploration, mapping, and localization.

Recently, Rao-Blackwellized particle filters (RBPF) have been introduced as an effective means for solving the simultaneous localization and mapping (SLAM) problem with occupancy grid maps [4, 19]. The key idea of this technique is to use a particle filter in which each particle represents a potential trajectory. Each particle furthermore carries its own map which is computed based on the associated trajectory.

Whereas a Rao-Blackwellized particle filter for mapping maintains a posterior about the robot's trajectory and the map of the environment, it does not provide information on how to steer the robot through the environment to acquire useful sensor data. As demonstrated in the past [1, 24], the quality of the resulting map constructed by a mobile robot depends on its trajectory during data acquisition. In practice, the major sources of uncertainty about the state of the world are the uncertainty in the robot's pose and the uncertainty resulting from the limited accuracy of the sensor the robot uses to perceive its environment. Therefore, a robot performing an autonomous exploration task should take the uncertainty in the map as well as in its path into account to select an appropriate action.

As a motivating example consider Figure 1. The left image shows a robot that has almost closed a loop. Suppose the vehicle has a high pose uncertainty and has to decide where to go next. Three potential actions are plotted on the map. Action 1 leads the robot to unknown terrain, whereas action 2 performs a loop closure without observing unknown areas. Action 3 does both: after closing the loop it guides the robot to unknown terrain.

Fig. 1. Suppose the robot has a high pose uncertainty and has to decide where to go next. Shown are three opportunities in the left image. Our map and pose entropy-driven exploration system chooses action 3 (as depicted in the right image) since it provides the highest expected utility (see also Figure 7).

Classical exploration approaches, which seek to reduce the amount of unseen area or which only consider the uncertainty in the posterior about the map [2, 14, 23, 26, 27, 29], would choose action 1, since this action guides the robot to the closest location from which information about unknown terrain can be obtained. Active localization approaches, which only estimate the uncertainty in the robots pose [11], would choose either action 2 or 3 to re-localize the vehicle. Our previous loop-closing approach [24] would select action 2 to reduce the entropy in the posterior about potential trajectories. However, the best action to reduce the uncertainty in the posterior about maps and trajectories is action 3. Executing this action yields new sensor information to make the correct data association and close the loop accurately. Additionally, it provides information about terrain so far unknown. As this example shows, exploration approaches should consider both sources of uncertainty to efficiently build accurate maps.

The contribution of this paper is an integrated technique that combines simultaneous localization, mapping, and exploration. In contrast to our previous work [24], in which a heuristic was used to re-traverse loops, the approach presented in this paper is entirely decision-theoretic. Based on the expected uncertainty reduction in the posterior about the trajectory of the robot as well as in the posterior about possible maps, we select the action with the highest expected utility. Thereby, we take into account the sensor information, which is expected to be obtained along the path when carrying out an action, as well as the cost introduced by this action. Real world and simulation experiments show the effectiveness of our

technique.

This paper is organized as follows. After the discussion of related work, we give a brief overview on Rao-Blackwellized mapping in Section III. Section IV and V present our decision-theoretic exploration technique and explain how to compute the expected change in entropy. Section VI describes how the set of possible actions is generated. Finally, Section VII contains experimental results carried out on real robots as well as in simulation.

II. Related Work

In the context of exploration, most of the techniques presented so far focus on generating motion commands that minimize the time needed to cover the whole terrain [2, 14, 26, 29]. Most of these techniques, however, assume that an accurate position estimate is given during exploration. Whaite and Ferrie [27] present an approach that uses the entropy to measure the uncertainty in the geometric structure of objects that are scanned with a laser range sensor. In contrast to the work described here, they use a parametric representation of the objects to be scanned and do not consider the uncertainty in the pose of the sensor. Similar techniques have been applied to mobile robots [23], but such approaches still assume the correct knowledge about the pose of the vehicle. None of the approaches mentioned above take the pose uncertainty into account when selecting the next vantage point. However, there are exploration approaches that have been shown to be robust against uncertainties in the pose estimates [5, 13, 15].

In the area of SLAM, the vast majority of papers have focused on the aspect of state estimation as well as belief representation and update [3, 4, 6, 8, 9, 10, 17, 19, 25]. These techniques, however, are passive and only process incoming sensor data without explicitly generating control commands.

Recently new techniques have been proposed which actively control the robot during SLAM. For example, Makarenko *et al.* [16] as well as Bourgault *et al.* [1] extract landmarks out of laser range scans and use an Extended Kalman Filter (EKF) to solve the SLAM problem. They furthermore introduce an utility function which trades-off the cost of exploring new terrain with the potential reduction of uncertainty by measuring at selected positions. A similar technique has been applied by Sim *et al.* [22], who consider actions to guide the robot back to a known place in order reduce the pose uncertainty of the vehicle. These three techniques differ from the approach presented in this paper in that they rely on the fact that the environment contains landmarks that can be uniquely determined during mapping. In contrast to this, our approach makes no assumptions about distinguishable landmarks and uses raw laser range scans to compute accurate occupancy grid maps.

One disadvantage of feature-based exploration systems is that the underlying environmental models typically do not provide any means to distinguish between known an unknown areas. Therefore, an additional map representation needs to be maintained (like, e.g., an occupancy grid in [1, 16] or a visual map in [22]) to efficiently guide the vehicle. Approaches which do not maintain an additional model to identify unknown areas typically apply strategies in which the robot follows the contours of obstacles [28] or performs wall following combined with random choices at decision points [7].

Duckett *et al.* [5] use relaxation to solve the SLAM problem in their exploration approach. They condense local grid maps into graph nodes and select goal points based on that graph structure, but do not consider the expected change of uncertainty when choosing possible target locations.

There are planning techniques that can compute optimal plans, e.g., by maintaining a belief over possible states of the world and by computing the strategy that is optimal in expectation with respect to that belief. One solution to this is the Partially Observable Markov Decision Process, also known as POMDP [12]. The major disadvantage of the POMDP is its extensive computational cost and most solutions are not applicable to scenarios with more than around one thousand states [20]. Since we reason about a high-dimensional state estimation problem, we have to be content with approximative solutions that rely on strong assumptions. Essentially, our approach can be regarded as an approximation of the POMDP with a one step look-ahead.

Compared to the approaches discussed above, the novelty of the work reported here is that our algorithm for acquiring grid maps simultaneously considers the trajectory and map uncertainty when selecting an appropriate action. We furthermore approximate the information gathered by the robot when it executes an action. Our approach also considers different types of actions, namely so-called exploration actions, which guide the robot to unknown areas and place re-visiting actions, which allow the robot to reliably close loops and this way reduce its pose uncertainty.

III. Rao-Blackwellized Mapping

According to Murphy [19], the key idea of solving the SLAM problem with a Rao-Blackwellized particle filter is to estimate a posterior $p(x_{1:t} \mid z_{1:t}, u_{0:t-1})$ about potential trajectories $x_{1:t}$ of the robot, given its observations $z_{1:t}$ and its odometry measurements $u_{0:t-1}$. This distribution is then used to compute a posterior over maps and trajectories:

$$\begin{aligned} p(x_{1:t}, m \mid z_{1:t}, u_{0:t-1}) = \\ p(m \mid x_{1:t}, z_{1:t}) p(x_{1:t} \mid z_{1:t}, u_{0:t-1}) \end{aligned} \tag{1}$$

This equation can be solved efficiently since the quantity $p(m \mid x_{1:t}, z_{1:t})$ can be computed analytically once $x_{1:t}$ and $z_{1:t}$ are known. To estimate $p(x_{1:t} \mid z_{1:t}, u_{0:t-1})$ over the potential trajectories, Rao-Blackwellized mapping uses a particle filter in which an individual map is associated to each sample. Each of those maps is constructed given the observations $z_{1:t}$ and the trajectory $x_{1:t}$ represented by the corresponding particle. During resampling, the weight ω of each particle is proportional to the likelihood $p(z_t \mid m, x_t)$ of the most recent observation z_t given the map m associated to this particle and its pose x_t.

IV. The Uncertainty of a Rao-Blackwellized Particle Filter for Mapping

The goal of an exploration task is to minimize the uncertainty in the posterior of the robot. The uncertainty can be determined by the entropy H. The entropy of a posterior about two random variables x and y holds

$$\begin{aligned} & H(p(x,y)) \\ &= E_{x,y}[-\log p(x,y)] && (2)\\ &= E_{x,y}[-\log p(x) - \log p(y \mid x))] && (3)\\ &= E_{x,y}[-\log p(x)] + E_{x,y}[-\log p(y \mid x)] && (4)\\ &= H(p(x)) + \int_{x,y} -p(x,y)\log p(y \mid x)\, dx\, dy. && (5)\end{aligned}$$

The integral in Eq. (5) can be transformed as follows:

$$\begin{aligned} & \int_{x,y} -p(x,y)\log p(y \mid x)\, dx\, dy \\ &= \int_{x,y} -p(y \mid x)p(x)\log p(y \mid x)\, dx\, dy && (6)\\ &= \int_x p(x) \int_y -p(y \mid x)\log p(y \mid x)\, dy\, dx && (7)\\ &= \int_x p(x) H(p(y \mid x))\, dx && (8)\end{aligned}$$

Eq. (5) and Eq. (8) can be combined to

$$H(p(x,y)) \quad = \quad H(p(x)) + \int_x p(x) H(p(y \mid x))\, dx. \quad (9)$$

Based on Eq. (9), we can efficiently compute the entropy of a Rao-Blackwellized particle filter for mapping. For a better readability we use d_t instead of $z_{1:t}, u_{0:t-1}$:

$$\begin{aligned} H(p(x_{1:t}, m \mid d_t)) &= H(p(x_{1:t} \mid d_t)) \\ &+ \int_{x_{1:t}} p(x_{1:t} \mid d_t) H(p(m \mid x_{1:t}, d_t))\, dx_{1:t} \quad (10)\end{aligned}$$

Considering that our posterior is represented by a set of weighted particles, we can transform the integral into a sum:

$$\begin{aligned} H(p(m, x_{1:t} \mid d_t)) &\approx H(p(x_{1:t} \mid d_t)) \\ &+ \sum_{i=1}^{\#particles} \omega_t^{[i]} H(p(m^{[i]} \mid x_{1:t}^{[i]}, d_t)) \quad (11)\end{aligned}$$

Here $\omega_t^{[i]}$ is the weight of the i-th particle at time step t.

Eq. (11) shows that according to the Rao-Blackwellization, the entropy of the whole system can be divided into two components. Whereas the first term represents the entropy of the posterior about the trajectories of the robot, the second term corresponds to the uncertainty about the map weighted by the likelihood of the corresponding trajectory. Thus, to minimize the robot's overall uncertainty, one needs to reduce the map uncertainty of the individual particles as well as the trajectory uncertainty. In this section, we will describe how we determine both terms in our approach.

Throughout this work, we use occupancy grid maps [18] to model the environment. Note that our technique is not restricted to this kind of representation, it only requires a way to compute the uncertainty for the used map representation. Using occupancy grids, the computation of the map entropy is straightforward. According to the common independence assumption about the cells of such a grid, the entropy of a map m is the sum over the entropy values of all cells. Since each grid cell c is represented by a binary random variable the entropy of m is computed as

$$\begin{aligned} H(m) &= \\ &- \sum_{c \in m} p(c)\log p(c) + (1-p(c))\log(1-p(c)). \quad (12)\end{aligned}$$

Note that the overall entropy calculated for a grid map is not independent from the resolution of the grid. One potential solution to this problem is to weight the entropy of each cell with its size. This approach results in a more or less constant entropy value when refining the grid resolution.

Unfortunately, it is more difficult to compute the uncertainty $H(p(x_{1:t}|d_t))$ of the posterior about the trajectory of the robot, since each pose x_t on the trajectory depends on the previous locations $x_{1:t-1}$. In the context of EKF-based approaches, the pose uncertainty is often calculated by considering only the last pose of the robot, which corresponds to the approximation of $H(p(x_{1:t}|d_t))$ by $H(p(x_t|d_t))$. It is also possible to average over the uncertainty of the different poses along the path as done by Roy *et al.* [21]:

$$H(p(x_{1:t} \mid d_t)) \quad \approx \quad \frac{1}{t}\sum_{t'=1}^{t} H(p(x_{t'} \mid d_t)) \quad (13)$$

Instead one can approximate the posterior about the trajectory by a high-dimensional (length of the trajectory times the dimension of the state vector x_t) Gaussian distribution. The entropy of a n dimensional Gaussian $\mathcal{G}(\mu, \Sigma)$ is computed as

$$H(\mathcal{G}(\mu,\Sigma)) \quad = \quad \log((2\pi e)^{(n/2)}|\Sigma|). \quad (14)$$

Since a finite number of particles is used, the RBPF representation often generates a sparse trajectory posterior for points in time lying further back in the history. Unfortunately, this can lead to a reduced rank of Σ, so that $|\Sigma|$ becomes zero and the entropy $H(\mathcal{G}(\mu,\Sigma))$ approaches minus infinity.

Alternatively, one could consider the individual trajectories represented by the samples as vectors in a high-dimensional state space and compute the entropy of the posterior based on a grid-based discretization. Since the particles typically are extremely sparse, this quantity is in most cases equivalent to or slightly smaller than the logarithm of the number of particles, which is the upper bound for the entropy computed in this way.

Instead we use an approach that is similar to the one proposed by Roy *et al.* [21], who computed the entropy over the trajectory posterior as the average entropy of the pose posteriors over time (see Eq. (13)). To reduce the influence of the dependency between the entropies when the robot gets back to a particular place, we average over the places covered by the vehicle. An example on how the trajectory entropy evolves over time is depicted in the left image of Figure 2.

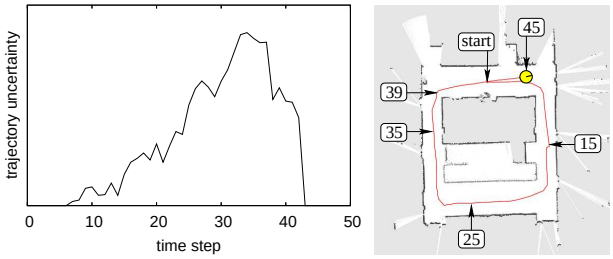

Fig. 2. The trajectory entropy of a robot during a real world experiment. The numbers in the right image illustrate the time step when the robot was at the corresponding location.

V. The Expected Information Gain

To evaluate an action, that guides the robot from its current location to a goal location, we compute the expected information gain, which is the expected change of entropy of our Rao-Blackwellized particle filter when we execute that action. In the last section, we described how to compute the entropy of the robot's world model, whereas in this section we want to estimate the expected entropy after an action has been carried out.

An action a_t generated at time step t is represented by a sequence of relative movements $a_t = \hat{u}_{t:T-1}$ which the robot has to carry out starting from its current pose x_t. During the execution of a_t, it is assumed that the robot obtains a sequence of observations $\hat{z}_{t+1:T}$ at the positions $\hat{x}_{t+1:T}$. In the following, all variables labeled with '^' correspond to values which occur during the execution of an action a_t. For a better readability we replace $\hat{x}_{t+1:T}$ by $\hat{x}$ and $\hat{z}_{t+1:T}$ by $\hat{z}$.

To compute the information gain of an action, we have to calculate the change of entropy caused by the integration of $\hat{z}, a_t$ into the robot's world model

$$I(\hat{z}, a_t) = H(p(m, x_{1:t} \mid d_t)) - H(p(m, x_{1:t}, \hat{x} \mid d_t, a_t, \hat{z})). \quad (15)$$

Since in general we do not know which measurements the robot will obtain along its path while executing action a_t, we have to integrate over all possible measurement sequences $\hat{z}$ to compute the expected information gain

$$E[I(a_t)] = \int_{\hat{z}} p(\hat{z} \mid a_t, d_t) \cdot I(\hat{z}, a_t)\, d\hat{z}. \quad (16)$$

In the following, we will explain how to approximate $p(\hat{z} \mid a_t, d_t)$ to reason about possible observation sequences $\hat{z}$. The posterior $p(\hat{z} \mid a_t, d_t)$ can be transformed into

$$\begin{aligned} p(\hat{z} \mid a_t, d_t) &= \int_{m, x_{1:t}} p(\hat{z} \mid a_t, m, x_{1:t}, d_t) \cdot p(m, x_{1:t} \mid d_t)\, dm\, dx_{1:t} \quad (17) \\ &= \int_{m, x_{1:t}} p(\hat{z} \mid a_t, m, x_{1:t}, d_t) \cdot p(x_{1:t} \mid d_t) p(m \mid x_{1:t}, d_t)\, dm\, dx_{1:t}. \quad (18) \end{aligned}$$

Eq. (18) is obtained from Eq. (17) by using Eq. (1). If we again assume that our posterior is represented by a set of particles we can rewrite Eq. (18) as follows:

$$p(\hat{z} \mid a_t, d_t) \approx \sum_{i=1}^{\#particles} p(\hat{z} \mid a_t, m^{[i]}, x_{1:t}^{[i]}, d_t) \cdot \omega_t^{[i]} p(m^{[i]} \mid x_{1:t}^{[i]}, d_t) \quad (19)$$

Based on Eq. (19) we can compute $\hat{z}$ for a given action a_t. The factor $p(m^{[i]} \mid x_{1:t}^{[i]}, d_t)$ in Eq. (19) is assumed to be computed analytically due to the assumptions made in the Rao-Blackwellization (see Eq. (1)), namely that we can compute the map $m^{[i]}$ analytically given the positions $x_{1:t}^{[i]}$ as well as the data d_t. We can also approximate the term $p(\hat{z} \mid a_t, d_t)$ of that equation. This can be achieved by performing a ray-casting operation in the map $m^{[i]}$ of the i-th particle to estimate possible observations $\hat{z}$. In other words, the (discrete) posterior about possible observations obtained along the path when executing the action a_t can be computed by ray-casting operations performed in the map of each particle weighted by the likelihood of that particle.

In cases where the ray-casting operation reaches an unknown cell in the map, we have to treat the beam differently. Touching an unknown cell means that we cannot say anything about the beam except that its length will be at least as long as the distance between robot pose and the unknown cell (with a high probability). Since such beams typically have a serious influence on the map uncertainty, we computed statistics about the average change of map entropy introduced by integrating a beam which reaches an unknown cell in the map. Note that in this situation, the change of entropy is approximative proportional to the number of unknown cells covered by that beam. In this way, the system also accounts for unknown areas which are visible from a planned path to any other destination.

This approximation dramatically reduces the amount of potential observations compared to the amount of possible proximity measurements a laser range finder can generate. As we figured out in several experiments, it seems to be a good approximation for robots equipped with a laser range finder.

Despite this approximation, computing the expected information gain based on Eq. (16) is possible but requires a large amount of computational resources. Therefore, we furthermore approximate the posterior in this equation about possible sensory data, by drawing a particle v from the particle set, where each particle is drawn with a probability proportional to its weight. We then use the map associated to v to generate the measurements $\hat{z}$ along the path. This reduces the computational complexity and allows us to run the exploration system on a real robot. Under this simplifying assumption we can rewrite the expected information gain in Eq. (16) by

$$E[I(a_t)] \approx I(\hat{z}(v), a_t). \quad (20)$$

The observation sequence $\hat{z}(v)$ is generated by a ray-casting operation in the map of v. Note that if more computational resources are available this approximation can easily be improved by drawing more (all) particles. This computation can even be parallelized, since there is no interference between the

integration of measurement sequences into different copies of the RBPF.

Now all necessary equations have been introduced to compute the expected information gain $E[I(a_t)]$ for an action a_t. To summarize, $E[I(a_t)]$ describes the expected change of entropy in the Rao-Blackwellized particle filter when executing a_t. To reason about possible observations the robot will obtain along the path, we draw a particle according to its likelihood and perform a ray-casting operation in its map. The expected measurements are then integrated into the filter and the entropies before and after the integration are subtracted.

The complexity of the computation of $E[I(a_t)]$ depends on two quantities. First, the filter needs to be copied to save its current state. This introduces a linear complexity linear in the size of the filter (which in turn depends on the number of particles). The second quantity is the length $l(a_t)$ of the planned path from the current pose of the robot to the desired goal location, because the expected observations along the path are taken into account. The integration of an observation is linear in the number of particles N. This leads to an overall complexity of $O(l(a_t) \cdot N)$ to evaluate an action a_t.

Besides the expected entropy reduction, there is a second quantity the robot should consider when selecting an action. This is the cost to reach the target location. The cost of an action is computed based on the (convolved) occupancy grid map of the most likely particle. Traversing a cell introduces a cost proportional to its occupancy probability.

The expected utility $E[U(a_t)]$ of an action a_t in our exploration system is defined as

$$E[U(a_t)] \quad = \quad E[I(a_t)] - \alpha \cdot cost(a_t). \tag{21}$$

Here α is a weighting factor which trades off the cost with the entropy. This free parameter can be used to trigger the exploration process by adapting the influence of the traveling cost. In our work, we determined α experimentally.

After computing the expected utility for each action under consideration, we select the action a_t^* with the highest expected utility

$$a_t^* = \operatorname*{argmax}_{a_t} E[U(a_t)]. \tag{22}$$

Every time the robot has to make the decision where to go next, it uses Eq. (22) to determine the action a_t^* with the highest expected utility and executes it. As soon as no reachable unknown areas are left, the robot seeks to minimize its pose uncertainty, since the map uncertainty is minimal in that situation. As soon as no action provides an expected improvement in the pose uncertainty, the exploration task is completed.

VI. Computing the Set of Actions

Above we left open how potential actions are generated. One attempt might be to generate a vantage point for each reachable grid cell in the map. Since we reason about observations received along the path, we need to consider all possible trajectories to all reachable grid cells in the map. The number of possible trajectories, however, is huge which makes it intractable to evaluate all of them.

To find appropriate actions to guide a vehicle through the environment, we consider two types of actions, so called *exploration actions* and *place re-visiting actions*. Exploration actions are designed to acquire information about unknown terrain to reduce the map uncertainty. To generate exploration actions, we apply the frontier approach introduced by Yamauchi [29]. For each frontier between known and unknown areas, we generate an action leading the robot from its current pose along the shortest path to that frontier.

Compared to the actions generated from frontiers, the place re-visiting actions do not focus on new terrain acquisition. They guide the robot back to an already known location or perform an active loop-closure. The goal of these actions is to improve the localization of the vehicle, which means to reduce its trajectory uncertainty. In our current implementation, place re-visiting actions are generated based on the trajectory of the robot. Such an action can simply turn the robot around and move it back along its previously taken path. Additionally, we also generate so called loop-closing actions. To determine whether there exists a possibility to close a loop, we would like to refer the reader to a previous work [24] in which we describe how a mobile robot can robustly detect opportunities to actively close a loop.

Given this classification, the actions 1 and 3 depicted in Figure 1 are exploration actions, whereas action 2 is a place re-visiting action performing an active loop-closure.

VII. Experiments

Our approach has been implemented and tested in real world and simulation experiments. The experiments described here are designed to illustrate the benefit of our exploration technique which takes into account the map as well as the trajectory uncertainty to evaluate possible actions.

Our current exploration system uses a highly optimized variant of the original algorithm for mapping with Rao-Blackwellized particle filters, which can handle trajectories with a length of more than one mile. An efficient implementation is necessary due to the online requirement needed for autonomous exploration. Further details can be found in [8].

A. Real World Application

The first experiment is a real world experiment carried out at building 106 at the University of Freiburg using an ActivMedia Pioneer 2 robot equipped with a SICK laser range finder. The exploration run was fully autonomous. The robot started in the lower left room (see Figure 3 (a)). The robot moved through the neighboring room and entered the corridor. After reaching its target location in the horizontal corridor (Figure 3 (b)), the robot decided to move back to in the lower left room to improve its pose estimate (Figure 3 (c)). The robot then explored the neighboring room and afterward returned to the corridor (Figure 3 (d)). It then approached the lower horizontal corridor and moved around the loop (Figure 3 (e)). In the end, the robot returned to the lower left room and finished the

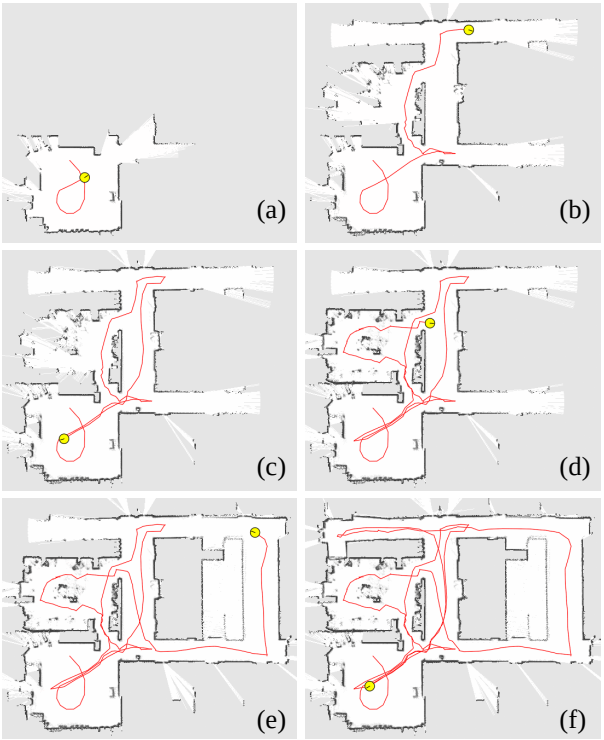

Fig. 3. Six different stages of an autonomous exploration run on the second fbor of building 106 at the University of Freiburg. The map was acquired fully autonomously by our integrated approach.

exploration task. As can be seen from this experiment, as soon as the robot gets too uncertain about its pose, it performs place re-visiting actions or chooses exploration actions which also reduce its pose uncertainty due to the information gathered along the path.

B. Decision Process

The next experiment is designed to show how the robot chooses actions to reduce its pose uncertainty as well as its map uncertainty. Figure 4 depicts parts of a simulated exploration task performed in a map acquired at Sieg Hall, University of Washington. Each row depicts a decision step of the robot during autonomous exploration. In the first step, shown in the first row, the robot has almost closed the loop. It had to decide if it is better to move through the loop again or to focus on exploring the horizontal corridor. In this situation, the robot moved to target point 1 and actively closed the loop, since this provided the highest expected utility (see right plot in the first row of Figure 4). Target location 1 had the highest expected utility because the robot expected a chance to re-localize itself by closing the loop and to observe parts of the unknown areas close to the planned trajectory. Therefore, this actions provided an expected reduction of map and trajectory uncertainty. In the second decision, the robot focused on acquiring new terrain and approached the horizontal corridor, since target location 6 had the highest expected utility. The same happened in the third decision step, shown in the last row of this figure. Moving back through the known areas of the loop provided less expected entropy reduction and therefore the robot continued exploring the horizontal corridor (target location 5).

Figure 5 shows the map after reaching target location 5 from the last decision step. To visualize the change of entropy over time, the right plot shows the evolution of the map as well as the pose uncertainty. The labels in the left image show the time steps in which the robot was at the corresponding location. As can be seen, the entropy stayed more or less constant in the beginning, since the map uncertainty decreased, whereas the pose uncertainty increased. After closing the loop at around time step 45, the pose uncertainty dropped so that the overall uncertainty was also reduced. Moving through known areas between time step 50 and 90 did not provide a lot of new information and did not change the entropy that much. As soon as the robot entered especially the wide part of the horizontal corridor, the overall uncertainty dropped again due to the serious reduction of map uncertainty compared to the moderate increase of pose uncertainty.

C. Comparison to Previous Approaches

The third experiment addresses the decision problem of the motivating example presented in the introduction of this paper. It shows how our approach chooses the actions which lead to the highest uncertainty reduction in the posterior about poses and maps compared to previous techniques. As can be seen in Figure 7, the robot has almost closed the loop. Suppose the robot has a high pose uncertainty and considers three potential actions to approach different target locations (see left image of Figure 7). Action 1 is a new terrain acquisition action, whereas action 2 performs a loop closure. Action 3 leads the robot to unknown terrain while simultaneously closing the loop. Since action 3 combines a loop-closure with new terrain acquisition, it provides the highest expected utility (see right image of Figure 7). Therefore, our approach chooses this target point. This is an advantage compared to other approaches which seek to actively close loops in an heuristic way [24]. Such a technique would choose action 2 to reduce the pose uncertainty of the vehicle. Classical exploration approaches, which only take into account the map uncertainty or guide the robot to the closest unknown area [2, 14, 23, 26, 27, 29] would select action 1. Even an active localization technique which seeks to reduce the pose uncertainty of the vehicle [11] would choose either action 2 or 3 (with a 50% chance each).

D. Corridor Exploration

The last experiment was performed at building 79 at the University of Freiburg and is depicted in Figure 6. The environment has a long corridor and contains no loop. According to a very short sensor range used in this experiment, it was hard for the robot keep track of its own position. As can be seen, our approach leads to an interesting intuitive behavior. Due to the large uncertainty in the pose of the vehicle, the robot chooses several times actions which guide it back to a well-known place (which is the starting location in this experiment) to reduce its pose uncertainty.

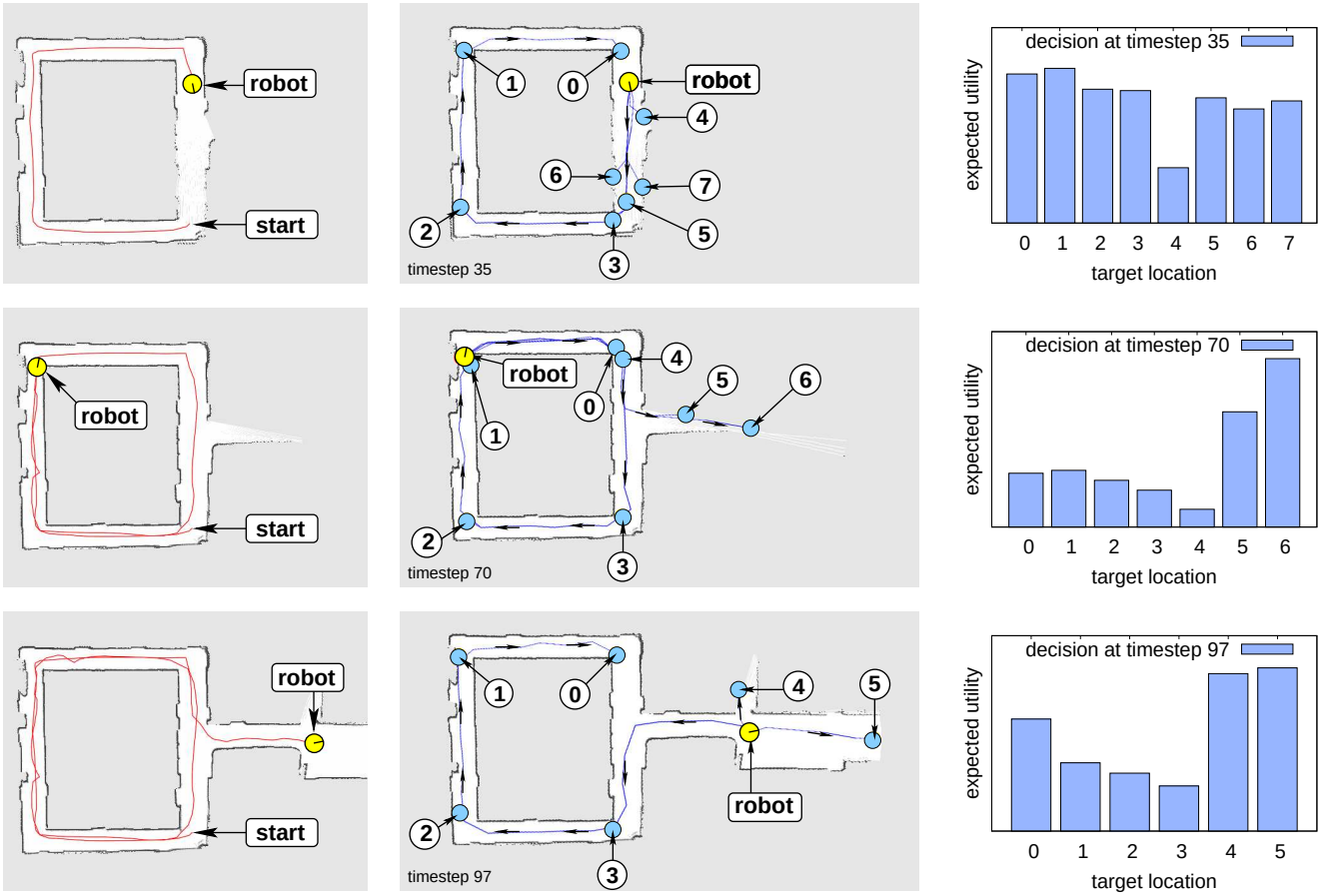

Fig. 4. Three points in time in which the robot had to decide where to move next. The left images depict the trajectory of the robot up the corresponding point in time. The middle images illustrate the best maps and possible actions to be evaluated. The plots on the right-hand side show the expected utility of the individual target locations.

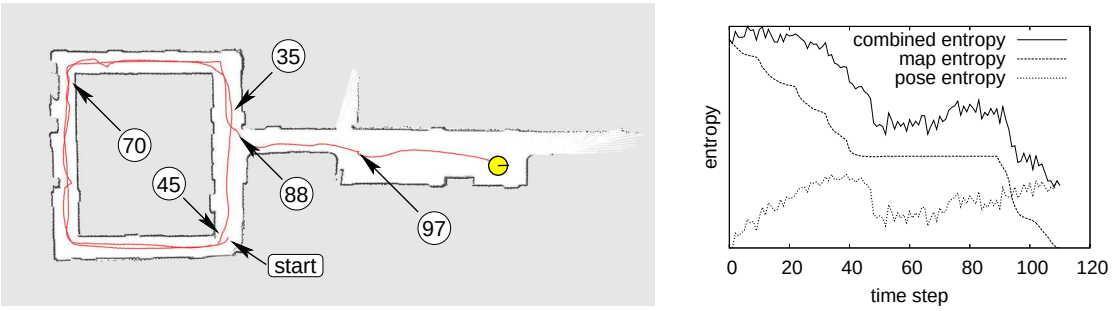

Fig. 5. This figure illustrates the evolution of the entropy during the experiment shown in Figure 4. The marker in the left image correspond to the different points in time when the robot was at the corresponding location. The right plot depicts the entropy during the data acquisition phase. It depicts the map entropy, the pose uncertainty, and the overall (combined) entropy over time.

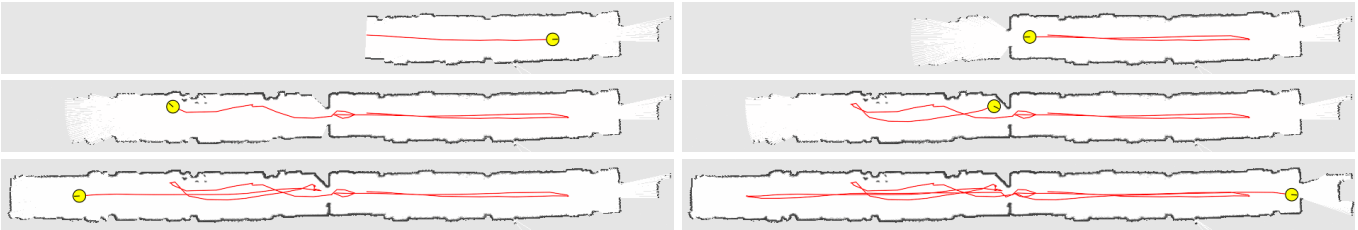

Fig. 6. The exploration of a long corridor. Due to the high pose uncertainty, the exploration system chooses actions which guide the robot on a path close to the starting location in order to re-localize.

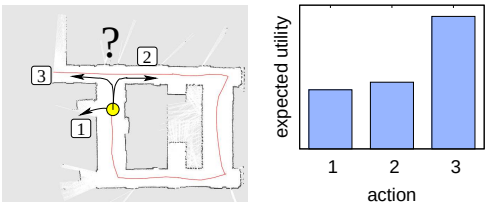

Fig. 7. Deciding where to go next. Shown in the left image are three potential actions and the corresponding expected utilities in the right image.

VIII. Conclusion

In this paper, we presented an integrated approach which simultaneously addresses mapping, localization, and exploration. We use a decision-theoretic framework for mobile robot exploration in combination with a Rao-Blackwellized particle filter to build accurate grid maps. Our approach considers different types of actions, namely exploration actions forcing terrain acquisition as well as place re-visiting and active loop-closing actions that reduce the robot's pose uncertainty. By estimating the expected entropy of the particle filter after carrying out an action, we are able to determine the action which promises the highest expected uncertainty reduction. Thereby, we take potential measurements gathered along the path into account. The approach has been implemented and tested on real robots and in simulation. As a result, we obtain a robust decision-theoretic exploration algorithm that produces highly accurate grid maps. Practical experiments showed the effectiveness of our action selection technique in comparison to previous approaches.

Acknowledgment

This work has partly been supported by the DFG under contract number SFB/TR-8 (A3), by the Marie Curie program under contract number HPMT-CT-2001-00251, and by the EC under contract number FP6-004250-CoSy. Furthermore, the authors would like to thank Kristian Kersting and Joelle Pineau for their helpful comments.

References

[1] F. Bourgoult, A.A. Makarenko, S.B. Williams, B. Grocholsky, and F. Durrant-Whyte. Information based adaptive robotic exploration. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland, 2002.

[2] W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun. Collaborative multi-robot exploration. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 476–481, 2000.

[3] G. Dissanayake, H. Durrant-Whyte, and T. Bailey. A computationally efficient solution to the simultaneous localisation and map building (SLAM) problem. In *ICRA'2000 Workshop on Mobile Robot Navigation and Mapping*, San Francisco, CA, USA, 2000.

[4] A. Doucet, J.F.G. de Freitas, K. Murphy, and S. Russel. Rao-blackwellized partcile filtering for dynamic bayesian networks. In *Proc. of the Conf. on Uncertainty in Artificial Intelligence (UAI)*, Stanford, CA, USA, 2000.

[5] T. Duckett, S. Marsland, and J. Shapiro. Fast, on-line learning of globally consistent maps. *Autonomous Robots*, 12(3):287 – 300, 2002.

[6] A. Eliazar and R. Parr. DP-SLAM: Fast, robust simultainous localization and mapping without predetermined landmarks. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, Acapulco, Mexico, 2003.

[7] J. Folkesson and H. Christensen. Outdoor exploration and slam using a compressed filter. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2003.

[8] G. Grisetti, C. Stachniss, and W. Burgard. Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 2443–2448, Barcelona, Spain, 2005.

[9] J.-S. Gutmann and K. Konolige. Incremental mapping of large cyclic environments. In *Proc. of the IEEE Int. Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, pages 318–325, Monterey, CA, USA, 1999.

[10] D. Hähnel, W. Burgard, D. Fox, and S. Thrun. An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2003.

[11] L.P. Kaelbling, A.R. Cassandra, and J.A. Kurien. Acting under uncertainty: Discrete Bayesian models for mobile-robot navigation. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 1996.

[12] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra. Planning and acting in partially observable stochastic domains. Technical report, Brown University, 1995.

[13] J. Ko, B. Stewart, D. Fox, K. Konolige, and B. Limketkai. A practical, decision-theoretic approach to multi-robot mapping and exploration. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 3232–3238, 2003.

[14] S. Koenig and C. Tovey. Improved analysis of greedy mapping. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Las Vegas, NV, USA, 2003.

[15] B. Kuipers and Y.T. Byun. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Robotics and Autonomous Systems*, 1981.

[16] A.A. Makarenko, S.B. Williams, F. Bourgoult, and F. Durrant-Whyte. An experiment in integrated exploration. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Lausanne, Switzerland, 2002.

[17] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: A factored solution to simultaneous localization and mapping. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, Edmonton, Canada, 2002.

[18] H.P. Moravec and A.E. Elfes. High resolution maps from wide angle sonar. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 116–121, St. Louis, MO, USA, 1985.

[19] K. Murphy. Bayesian map learning in dynamic environments. In *Neural Info. Proc. Systems (NIPS)*, Denver, CO, USA, 1999.

[20] J. Pineau, G. Gordon, and S. Thrun. Point-based value iteration: An anytime algorithm for pomdps. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, Acapulco, Mexico, 2003.

[21] N. Roy, W. Burgard, D. Fox, and S. Thrun. Coastal navigation – robot motion with uncertainty. In *Proceedings of the AAAI Fall Symposium: Planning with POMDPs*, 1998.

[22] R. Sim, G. Dudek, and N. Roy. Online control policy optimization for minimizing map uncertainty during exploration. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, New Orleans, LA, USA, 2004.

[23] C. Stachniss and W. Burgard. Exploring unknown environments with mobile robots using coverage maps. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, 2003.

[24] C. Stachniss, D. Hähnel, and W. Burgard. Exploration with active loop-closing for FastSLAM. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2004.

[25] S. Thrun. An online mapping algorithm for teams of mobile robots. *Int. Journal of Robotics Research*, 2001.

[26] G. Weiß, C. Wetzler, and E. von Puttkamer. Keeping track of position and orientation of moving indoor systems by correlation of range-finder scans. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 595–601, Munich, Germany, 1994.

[27] P. Whaite and F.P. Ferrie. Autonomous exploration: Driven by uncertainty. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.

[28] F.H. Wullschleger, K.O. Arras, and S.J. Vestli. A flexible exploration framework for map building. In *Proc. of the Third European Workshop on Advanced Mobile Robots*, 1999.

[29] B. Yamauchi. Frontier-based exploration using multiple robots. In *Proc. of the Second International Conference on Autonomous Agents*, pages 47–53, Minneapolis, MN, USA, 1998.