Autonomous Exploration with Expectation-Maximization

Jinkun Wang and Brendan Englot

Abstract We consider the problem of autonomous mobile robot exploration in an unknown environment for the purpose of building an accurate feature-based map efficiently. Most literature on this subject is focused on the combination of a variety of utility functions, such as curbing robot pose uncertainty and the entropy of occupancy grid maps. However, the effect of uncertain poses is typically not well incorporated to penalize poor localization, which ultimately leads to an inaccurate map. Instead, we explicitly model unknown landmarks as latent variables, and predict their expected uncertainty, incorporating this into a utility function that is used together with sampling-based motion planning to produce informative and low-uncertainty motion primitives. We propose an iterative expectation-maximization algorithm to perform the planning process driving a robot's step-by-step exploration of an unknown environment. We analyze the performance in simulated experiments, showing that our algorithm maintains the same coverage speed in exploration as competing algorithms, but effectively improves the quality of the resulting map.

1 Introduction

While simultaneous localization and mapping (SLAM) has been applied successfully for state estimation and map-building with data collected passively on sensing platforms, it's still a challenge for an autonomous vehicle to actively explore an unknown environment and manage the quality of its state estimate and map. The capability of autonomous exploration may be especially impactful in scenarios where teleoperation is limited or infeasible due to constrained communication, e.g. in unknown subsea environments with underwater robots.

Jinkun Wang, Brendan Englot - Department of Mechanical Engineering
Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ, 07030.
e-mail: {jwang92, benglot}@stevens.edu

In general, the autonomous exploration problem for mobile robots comprises three stages: 1) the robot identifies candidate locations to explore or paths to follow; 2) a utility function is evaluated for every candidate and the optimal one is selected; 3) the robot executes the optimal action and updates its current knowledge of the environment. While the first task is commonly achieved by enumerating frontier locations or by employing sampling-based methods and the third task is simply performed by feedback controllers, the remaining question is how to properly formulate a utility function that effectively captures the *exploration-exploitation* dilemma, i.e., a balancing of visiting unknown areas to reduce map uncertainty and revisiting known areas to seek better localization.

1.1 Related Work

Without considering the robot's localization uncertainty, the problem has been approached by using Cauchy-Schwarz quadratic mutual information (CSQMI) to reduce computation time [6], by combining global planning with local motion primitives and also refining a trajectory using optimization methods [7], and by exploring on continuous Gaussian process frontier maps [11]. The simplified problem, planning with *a priori* maps, has also been discussed [9], [17],[19] to actively minimize the uncertainty of known landmarks.

Most of the existing research on exploration in unknown environments takes advantage of occupancy grid maps, considering utility functions of map entropy and of the uncertainty of robot poses. An integrated exploration strategy was proposed to combine different utilities [2], [16], e.g. the utility of information gain over the occupancy grid maps, the utility of travel distance to the goal, and the utility of localizability, which incorporates the uncertainty of robot poses and landmarks. However, these methods ignore the correlation between localization and information gain.

High-uncertainty poses are likely to result in inaccurate occupancy grid maps, limiting the usefulness of the information gained by exploring unknown regions. The correlation between localization and information gain was taken into account by integrating over a map's entropy weighted by trajectory probability [20], [22], [23]. The trajectory probability was transformed into particle weight in the particle filter in [20], and the integral was simplified to use only the mean trajectory in [23].

Instead of leveraging the entropy of occupancy grid maps, in this paper, we propose the concept of a *virtual map* comprised of *virtual landmarks*, and acting to minimize the uncertainty of these landmarks. In occupancy grid mapping, grid cells inside the sensor observation cone are less affected by the uncertainty of robot poses, which gives rise to overconfidence in the information gained by exploring unknown maps. In contrast to that, every virtual landmark is deeply connected with robot poses that can observe it in our proposed approach. The proposed metrics contain the uncertainty from potential landmarks that may be observed in the future, and therefore, more accurate landmark positions are obtained in the course of exploring an *a priori* unknown environment.

1.2 Paper Organization

We describe in Section 2 the formal definition of the exploration problem within a feature-based SLAM framework, and the proposed utility function is formulated by considering virtual landmarks. In Section 3, the optimization problem with latent variables is solved through expectation-maximization, where the E-step involves updating the occupancy probabilities (Section 3.3) and the M-step searches for the best path among motion primitives (Section 3.4). Experimental results are presented in Section 4, with conclusions and discussion of future work in Section 5.

2 The Exploration Problem with Feature-Based SLAM

In this paper, we consider the autonomous mobile robot exploration problem in unknown environments, where's the robot's objective guiding exploration is to produce a feature-based map of its surroundings accurately and efficiently. Here we make the following assumptions for simplicity:

1. There are a certain number of static landmarks in the environment which can be used for localization.
2. Landmarks are identifiable without considering the data association problem.
3. The movement of the robot is confined in a limited space, the border of which doesn't have any landmarks.

2.1 The SLAM Framework

We use the *smoothing*-based approach rather than the *filtering*-based approach, adopting incremental smoothing and mapping [14] to repeatedly estimate the entire robot trajectory. The benefit of that will be discussed in Section 3.

Let a mobile robot's motion model be defined as

$$\mathbf{x}_i = f_i(\mathbf{x}_{i-1}, \mathbf{u}_i) + \mathbf{w}_i, \quad \mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, Q_i), \tag{1}$$

and let the measurement model be defined as

$$\mathbf{z}_{ij} = g_{ij}(\mathbf{x}_i, \mathbf{l}_j) + \mathbf{v}_{ij}, \quad \mathbf{v}_{ij} \sim \mathcal{N}(\mathbf{0}, R_{ij}), \tag{2}$$

where we assume the data association between $\mathbf{x}_i, \mathbf{l}_j$ is known.

Given measurements $\mathcal{Z} = \{\mathbf{z}_k\}$, we can obtain the best estimate of the entire trajectory $\mathcal{X} = \{\mathbf{x}_i\}$ and observed landmarks $\mathcal{L}_o = \{\mathbf{l}_j\}$,

$$\mathcal{X}^*, \mathcal{L}_o^* | \mathcal{Z} = \underset{\mathcal{X}, \mathcal{L}_o}{\operatorname{argmin}} P(\mathcal{X}, \mathcal{L} | \mathcal{Z}). \tag{3}$$

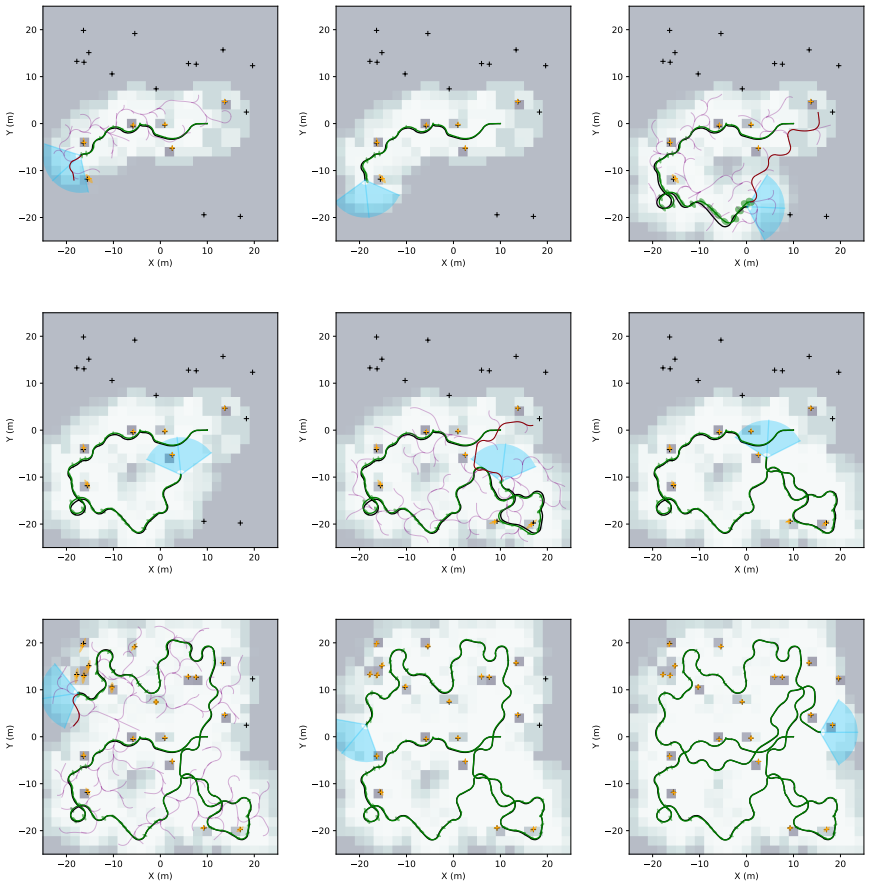


Fig. 1: A few representative steps of planning and execution using our proposed algorithm, over advancing, but non-consecutive time instants during a single exploration process. Virtual landmarks are embedded in the underlying grid cell environment map with occupancy probabilities shown in grayscale color. Error ellipses (2 standard deviations) of observed landmarks (orange) and robot poses (green) are overlaid on top of the corresponding best estimates. An RRT rooted at the current robot pose is constructed (purple). The trajectory from the RRT offering the expected minimum cost is marked by a red line. From top left to bottom right, one figure showing the expected best trajectory is followed by a figure showing its execution, and the final figure simply represents the final step of the exploration task.

The maximum a posteriori (MAP) estimate can be used by maximizing the joint probability, which afterwards leads to a nonlinear least-squares problem. By constructing a graph representation and linearizing nonlinear functions, the marginal distributions and joint marginal distributions, both of which are Gaussian, can be extracted using graphical model-based inference algorithms [14].

2.2 Utility Function

Let $\mathcal{L} = \{\mathbf{l}_i\}$ be the set of landmarks in the environment. We have an estimate for each of them $\mathbf{l}_i \in \mathbb{R}^2$, which is distributed as $\mathcal{N}(\hat{\mathbf{l}}_i, \Sigma_{\mathbf{l}_i})$. The definition of the utility function for the exploration problem is as follows,

$$U = \sum_{\mathbf{l}_i \in \mathcal{L}} \phi(\Sigma_{\mathbf{l}_i}), \tag{4}$$

where $\phi : \Sigma \to \mathbb{R}$ represents the uncertainty criterion for covariance matrices (see Sec. 3.1). In addition to the uncertainty of landmarks, it is beneficial to add a cost-to-go $C(\mathbf{a})$ to favor a shorter path [20]. Thus the optimal action sequence is the one minimizing the uncertainty with a weighted smaller cost:

$$\mathbf{a}^* = \underset{\mathbf{a}}{\operatorname{argmin}}\ U + \alpha C. \tag{5}$$

Since there are landmarks that haven't been observed yet, in most works, the common strategy is to omit the values of unknown landmarks and replace them with the entropy of occupancy grid maps. In this paper, we introduce the concept of *virtual landmarks* and treat the landmarks as latent variables. The objective is then to minimize the uncertainty of possible landmarks that would be observed when following the planned path. This uncertainty is defined

$$U \approx \sum_{\mathbf{v}_k \in \mathcal{V}} \phi(\Sigma_{\mathbf{v}_k}), \tag{6}$$

where $\mathcal{V} = \{\mathbf{v}_k\}$ are *virtual landmarks* (see Fig. 2), which represent all the possible locations of actual landmarks. They won't be incorporated into a robot's SLAM optimization unless they are revealed to be true landmarks, but the potential to reduce their uncertainty is considered throughout the course of autonomous exploration.

Intuitively, the approximation computes the expected uncertainty of actual landmarks after taking into account the observations collected while following a candidate path. The function provides a trade-off between exploration and localization internally: exploration will decrease the uncertainty of virtual landmarks, which are initialized to have large covariance matrices, but poor-quality localization will lead to higher cost, since the covariance matrices of landmarks are estimated based on the robot poses that can observe them.

3 The EM Exploration Algorithm

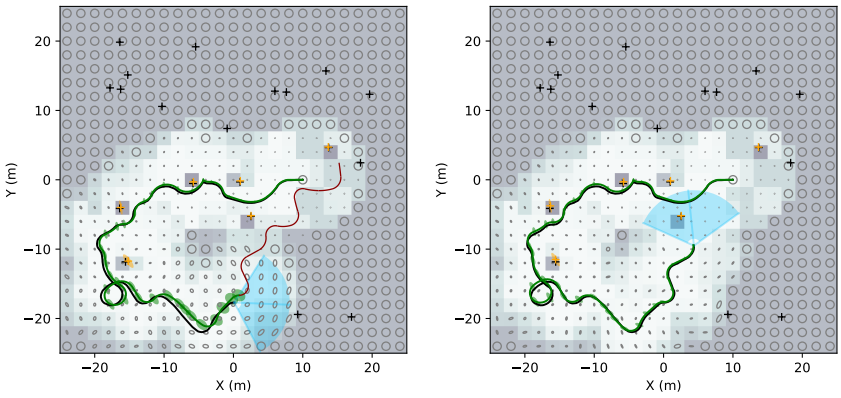


Fig. 2: Gaussian error ellipses (0.5 standard deviations) of virtual landmarks. Left: The uncertainty of virtual landmarks at the bottom of the map isn't reduced significantly after exploration, as the robot hasn't met a loop-closure, which means the possible landmarks observed in the future would have high uncertainty. The proposed algorithm makes a decision to travel upwards in order to revisit a landmark for better localization. Right: The execution leads to an accurate estimate of the trajectory, and the error ellipses of virtual landmarks shrink significantly.

Because of the introduction of latent variables, it is intuitive to approach the exploration problem using an EM algorithm. Such an approach encompasses the update of the probability of a landmark's existence, represented by virtual landmarks, and also the prediction of its uncertainty. Here we show that iteratively the probability can be calculated from an occupancy grid map maintained along the landmarks in the Expectation step, and afterwards it is utilized to evaluate the covariance criteria in the Maximization step. Correspondingly, the E-step provides the expected landmark locations, and the M-step is merely to maximize the utility function, which essentially is an exploration problem with prior knowledge of the environment.

A mathematical formulation is as follows. We wish to maximize the marginal log-likelihood of observations $\mathcal{Z}$ given an action sequence $\mathbf{a}$ with latent variables $\mathcal{V}$:

$$\ln p(\mathcal{Z}|\mathbf{a}) = \ln \Big\{ \sum_{\mathcal{V}} p(\mathcal{Z}, \mathcal{V}|\mathbf{a}) \Big\}. \tag{7}$$

Different from traditional EM applications, the exploration problem involves collecting measurements after each iteration. For that reason, we assume a maximum likelihood measurement model and add measurements of all virtual landmarks into

$\mathcal{Z}$ in Eq. 7. For virtual landmarks beyond the robot's maximum sensing range, we use priors with large uncertainty.

According to the EM algorithm, in the E-step, we use the current $\mathbf{a}^{\text{old}}$ to evaluate the posterior distribution of latent variables $p(\mathcal{V}|\mathcal{Z},\mathbf{a}^{\text{old}})$. In the M-step, we determine $\mathbf{a}^{\text{new}}$ by maximizing the function below:

$$\mathbf{a}^{\text{new}} = \underset{\mathbf{a}}{\operatorname{argmax}} \sum_{\mathcal{V}} p(\mathcal{V}|\mathcal{Z},\mathbf{a}^{\text{old}}) \ln p(\mathcal{Z},\mathcal{V}|\mathbf{a}). \tag{8}$$

The equation above poses a challenge for efficient solution due to the exponential growth of virtual landmark configurations with respect to the number of virtual landmarks. Inspired by classification EM algorithms, an alternative solution would add a classification step (C-step) [5] before the M-step to provide the maximum posterior probability estimate of the virtual landmark configuration,

$$\mathcal{V}^* = \underset{\mathcal{V}}{\operatorname{argmax}}\, p(\mathcal{V}|\mathcal{Z},\mathbf{a}^{\text{old}}). \tag{9}$$

Afterwards, the M-step can be solved given the deterministic variable $\mathcal{V}^*$ as follows,

$$\mathbf{a}^{\text{new}} = \underset{\mathbf{a}}{\operatorname{argmax}} \ln p(\mathcal{Z},\mathcal{V}^*|\mathbf{a}) \tag{10}$$

$$= \underset{\mathbf{a}}{\operatorname{argmin}} \ln \Sigma_{\mathcal{V}^*} = \underset{\mathbf{a}}{\operatorname{argmin}}\, \phi(\Sigma_{\mathcal{V}^*}). \tag{11}$$

Here, we choose to replace the log-likelihood function with the logarithm of the determinant because we assume Gaussian distributions of the virtual landmarks and maximum likelihood observations. We also use ϕ to generalize the log-determinant covariance criterion as shown in Sec. 3.1.

3.1 Uncertainty Criteria

In this section, we discuss the choice of uncertainty criteria $\phi(\Sigma_{\mathcal{V}})$ in Eq. 11. In general, there are two commonly used uncertainty criteria in the exploration problem, namely A-Optimality and D-Optimality [4], [19], formulated respectively as:

$$\phi_{\text{A}} = \sum_{i=1}^{n} \lambda_i = \text{tr}(\Sigma_{\mathcal{V}}) \tag{12}$$

$$\phi_{\text{D}} = \prod_{i=1}^{n} \lambda_i = \det(\Sigma_{\mathcal{V}}). \tag{13}$$

Assuming all of the virtual landmarks are independent, the computation of determinant and trace can be written as $\text{tr}(\Sigma_{\mathcal{V}}) = \sum_{\mathbf{v}_k \in \mathcal{V}} \text{tr}(\Sigma_{\mathbf{v}_k})$ and $\det(\Sigma_{\mathcal{V}}) = \prod_{\mathbf{v}_k \in \mathcal{V}} \det(\Sigma_{\mathbf{v}_k})$. As pointed out in [19], we can drive the determinant of the entire covariance matrix to zero by minimizing the eigenvalue of one feature to zero. Specifically, D-

optimality doesn't capture the global information gain and hence we choose A-optimality in the utility function, and also in the evaluation of all of the algorithms examined in this paper. Now we can derive the utility function introduced in Eq. 6.

3.2 Covariance Intersection

To obtain the covariance estimate of a virtual landmark from poses where the robot is able to observe it, it is desirable to add these "virtual measurement factors" into the graphical model and then compute its marginalization. However, two problems arise due to this addition. First, due to the large number of virtual landmarks, the virtual factors increase the burden of optimization. Second, some virtual landmarks will become actual landmarks, providing localization information for the system.

For a virtual landmark $\mathbf{v}_k$, suppose that there are n related robot poses $\check{\mathbf{x}}_k = [\mathbf{x}_1\ \mathbf{x}_2\ ...\ \mathbf{x}_n]^T$, each of which corresponds to a measurement term $\mathbf{z}_{jk} = g_{jk}(\mathbf{x}_j, \mathbf{v}_k) + \mathbf{v}_{jk}$. Then the predicted covariance estimate per an EKF is $H\Sigma_{\check{\mathbf{x}}_k}H^T$, where $H = \text{diag}(H_1, H_2, ..., H_n)$ is the Jacobian matrix with $H_i = \frac{\partial g_{jk}}{\partial \mathbf{x}_j}|_{\hat{\mathbf{x}}_j}$, and $\Sigma_{\check{\mathbf{x}}_k}$ is the marginal covariance of poses observing the virtual landmark. Both [13] and [21] present an efficient algorithm to recover the parts of the covariance matrix associated with non-zero factor entries, which correspond to variables in the same clique of a Bayes tree [14]. However, the condition is unlikely to be satisfied, and the covariance estimate must be approximated to avoid the expensive recovery of full pose uncertainty.

The Covariance Intersection (CI) algorithm [8] was presented to estimate the covariance matrix when the correlation between two sources are unknown. Let $\Sigma_1 = H_1\Sigma_{\mathbf{x}_1}H_1^T$, $\Sigma_2 = H_2\Sigma_{\mathbf{x}_2}H_2^T$ be two independent estimates from the first two poses. Then the CI combines these two estimates to provide an upper bound based on the following equation:

$$\Sigma^{-1} = \omega\Sigma_1^{-1} + (1-\omega)\Sigma_2^{-1}. \tag{14}$$

The weight ω can be optimized to achieve the minimal determinant of Σ as follows:

$$\omega = \frac{2b - ac}{2(a + b - ac)}, \tag{15}$$

where $a = |\Sigma_1|^{-1}$, $b = |\Sigma_2|^{-1}$, and $c = Tr(\Sigma_1\Sigma_2^{-1})$. By iteratively fusing estimates from different poses, we derive an optimal upper bound on the actual covariance in terms of the minimal determinant, given only the block diagonal of $\Sigma_{\check{\mathbf{x}}_k}$.

3.3 Expectation Step

The E-step is used to compute $q_k \triangleq p(\mathbf{v}_k \in \mathcal{L})$, the probability of a potential landmark given the current measurements. By discretizing the map into grid cells, these

Algorithm 1 $\mathcal{Q}^{t+1}, \mathcal{X}^{t+1}, \mathcal{L}_o^{t+1} = Expectation(\mathcal{X}^t, \mathcal{L}_o^t, \mathcal{Z}^t)$

1: $\mathcal{Q}^{t+1} \leftarrow \emptyset$
2: $\mathcal{X}^{t+1}, \mathcal{L}_o^{t+1} \leftarrow Optimize(\mathcal{X}^t, \mathcal{L}_o^t, \mathcal{Z}^t)$ // iSAM2
3: $\{\mathcal{X}^{[n]}\} \leftarrow Sample(\mathcal{X}^{t+1})$
4: **for** $\mathbf{v}_k \in \mathcal{V}$ **do**
5: **for** $\mathcal{X}^{[n]} \in \{\mathcal{X}^{[n]}\}$ **do**
6: $q_k^{t+1} = p(\mathbf{m}_k|\mathcal{X}^{[n]}, \mathcal{Z}^t)$/ #samples
7: **end for**
8: $\mathcal{Q}^{t+1} \leftarrow \mathcal{Q}^{t+1} \cup q_k^{t+1}$
9: **end for**
10: **return** $\mathcal{Q}^{t+1}, \mathcal{X}^{t+1}, \mathcal{L}_o^{t+1}$

probabilities are equivalent to the occupancy probabilities used in occupancy grid maps, that is $q_k = p(\mathbf{m}_k|\mathcal{X}, \mathcal{Z})$ [18].

Generally, occupancy grid maps are constructed based on the most likely hypothesis of the robot path $\mathcal{X}^*$. As a result, the uncertainty of poses is not expressed in grid maps, which, under circumstances with severe drift, would falsely exclude some hypotheses that observe actual landmarks by reducing their probabilities close to zero. Therefore, we adopt the concept of an *expected map* [1] to take a weighted average of occupancy grid maps computed from all path hypotheses. Different from the method in the original paper which relies on samples generated by particle filtering, we employ Monte Carlo sampling to approximate the integral over the continuous space of poses,

$$q_k = \mathbb{E}[p(\mathbf{m}_k|\mathcal{X}, \mathcal{Z})] \tag{16}$$

$$= \int_{\mathcal{X}} p(\mathbf{m}_k|\mathcal{X}, \mathcal{Z}) p(\mathcal{X}, \mathcal{L}_o|\mathcal{Z}) \tag{17}$$

$$\approx \frac{1}{N} \sum_{n=1}^{N} p(\mathbf{m}_k^{[n]}|\mathcal{X}^{[n]}, \mathcal{Z}), \tag{18}$$

where N is the number of samples. The expectation step is performed as shown in Algorithm 1, in which the occupancy probabilities computed from posterior samples (line 3) are averaged (line 4-9) and after that the distributions of virtual landmarks are updated.

3.4 Maximization Step

In the M-step, given the distribution of virtual landmarks, the path candidates are selected and the utility function is evaluated for each of them. The global paths over a long period must take into account two types of actions, exploration actions and place-revisiting actions [20]. Usually exploration actions have destinations near

frontier locations where explored cells meet unknown cells, and place-revisiting actions choose to go back to locations the robot has visited. The prevalence of these particular locations requires us to examine a large amount of free grid cells in order to obtain a near-optimal solution. After identifying destinations, it is also of great benefit to refine the path to a destination rather than take the shortest one [7].

In consideration of these factors, we employ a sampling-based routine for generating action primitives, in which exploration and revisiting actions, possibly with varying paths, are considered simultaneously. The rapidly-exploring random tree (RRT) [15] is a randomized algorithm designed to quickly solve motion planning problems in high-dimensional spaces with constraints such as obstacles and dynamics. RRTs have the property of probabilistic completeness, ensuring that a solution will be found in the limit, if a feasible solution exists. Also due to their randomized construction, a tree might end up with a variety of paths to the same goal region. RRT presents a simple and efficient method for building our action primitives.

An optimal variant of the RRT algorithm, RRT* [12], guarantees asymptotic optimality, converging in the limit toward the shortest path. And RRT* has been successfully applied to active SLAM to generate a robot's action set [23] using two "distance functions", one is the traditional Euclidean travel distance and the other one is cumulative map entropy reduction. As shown in [23], generating candidate paths using the robot's travel distance as a cost metric still provides a sufficient diversity of primitives to achieve a good compromise between computational complexity and entropy reduction. Since our computational task is even more expensive, we also adopt the Euclidean travel distance in the *Nearest* function of RRT.

The M-step is summarized in Algorithm 2, in which the details of building RRTs are omitted for the sake of brevity. The utility function is evaluated for every path $\mathbf{a}$ from the tree root to the leaves. First, the measurements following the path are predicted assuming maximum likelihood observations (line 4). Second, the whole trajectory is re-optimized to incorporate future measurements (line 5). Third, the virtual landmarks are updated using the optimized trajectory (line 6-9). Finally the best action sequence is selected with the minimum utility value, which is computed from the updated virtual landmarks (line 11). To account for the deviation from the nominal trajectory during execution, model predictive control (MPC) performs only the first few steps [10] (line 6-12). One example of the M-step is shown in Fig. 2.

3.5 Algorithm Analysis

The proof that the EM exploration algorithm maximizes the likelihood of measurements is essentially to explain why the function in Eq. 8 is increased after executing action $\mathbf{a}^{\text{new}}$. This will be satisfied if the prior uncertainties of virtual landmarks are large enough to encourage exploration. However, this also implies that the algorithm can fail to explore if it does not have a properly initialized prior - in such a case, a robot may conclude that there isn't a viable path that explores an unknown area with a guarantee of acceptable estimates of the surrounding landmarks.

Algorithm 2 $\mathcal{Z}^t = Maximization(\mathcal{Q}^t, \mathcal{X}^t, \mathcal{L}_o^t)$

Tree $\leftarrow BuildRRT()$
for $\mathbf{a} \in$ Tree **do**
  $\mathcal{V}^{\mathbf{a}} \leftarrow \emptyset$
  $\mathcal{Z}^{\mathbf{a}} \leftarrow Predict(\mathcal{X}^t, \mathbf{a})$
  $\mathcal{X}^{\mathbf{a}}, \mathcal{L}_o^{\mathbf{a}} \leftarrow Optimize(\mathcal{X}^t, \mathcal{L}_o^t, \mathcal{Z}^{\mathbf{a}})$
  for $\mathbf{q}_k^t \in \mathcal{Q}^t$ **do**
    $\Sigma_{\mathbf{v}_k}^{\mathbf{a}} = CI(\check{\mathbf{x}}_k^{\mathbf{a}}, \text{block-diagonal}(\Sigma_{\check{\mathbf{x}}_k}^{\mathbf{a}}))$
    $\mathcal{V}^{\mathbf{a}} \leftarrow \mathcal{V}^{\mathbf{a}} \cup (q_k^t, \Sigma_{\mathbf{v}_k}^{\mathbf{a}})$
  end for
end for
$\mathbf{a}^* = \text{argmin}_{\mathbf{a}} \sum_{\mathbf{v}_k \in \mathcal{V}^{\mathbf{a}}} I_{(q_k \geq p_{\text{occupied}})} \Sigma_{\mathbf{v}_k}$
$\mathcal{Z}^t = Execute(\mathbf{a}^*, \#\text{steps})$
return $\mathcal{Z}^t$

Next we analyze the computational complexity of the algorithm. The computation in Algorithm 1 involves an iSAM2 update, sampling from the robot trajectory, and an occupancy probability update. The complexity of the iSAM2 update generally is bounded by $\mathcal{O}(n^3)$, where n is the number of variables [14]. The most expensive part in the sampling phase is to extract the joint covariance, which depends on the sparsity of the factor matrix. For the case of exploration, we force the robot to perform loop-closure occasionally, and the complexity of covariance recovery could deteriorate to $\mathcal{O}(s^3)$, where s is the size of the dense block in the factor matrix [13]. The update of occupancy probability has the computational complexity $\mathcal{O}(N_{\text{samples}} \cdot N_{\text{vl}})$, in which N_{samples} is the number of samples in Eq. 18, and generally a small set of samples can well approximate the expected map. N_{vl} is the number of virtual landmarks; an appropriate resolution may vary depending on the application.

Algorithm 2 has the computational complexity of $\mathcal{O}(N_{\text{nodes}} + N_{\text{leaves}}(C_1 + C_2))$, where the first term describes the construction of an RRT, and the second term describes the evaluation of our utility function (C_1 for the iSAM2 update and C_2 for the covariance update for virtual landmarks) for each path to a tree leaf. To alleviate costly computation, we propose a simplified process to handle situations without loop-closures. First, EKF propagation is used to predict the covariance instead of using iSAM2. Second, the virtual landmarks are cached in every node, and the covariances of a child node inherit from its parent with a few updates after incorporating the latest robot poses. Overall, by limiting the number of nodes in the RRT, we are able to achieve nearly real-time decision making (see Section 4 for details).

4 Experiments and Results

We analyze the performance of the proposed algorithm in a simulated environment (see Fig. 1 as an example). The simulation employs an environment with point features uniformly distributed in its inner region (the square $[-20\text{m}, 20\text{m}]$ in Fig. 1).

Table 1: Simulation Parameters

Bearing: stddev (deg)	0.5	Translational speed (m/s)	[0.5, 1.0]
Range: stddev (m)	0.002	Rotational speed (rad/s)	[-0.5, 0.5]
Rotation: stddev (deg)	0.2	Simulation step (s)	0.2
Translation: stddev (m)	0.01	Simulation duration (s)	[1.0, 4.0]
Bearing FOV (deg)	120	Safe distance (m)	1.5
Range FOV (m)	[1.0, 8.0]	Number of landmarks	20
Initial position: ([m, m])	[10.0, 0.0]	Initial sigmas ([m, m, deg])	[0.05, 0.05, 0.01]

The robot is equipped with a range sensor with a limited field of view (FOV) that is capable of measuring the relative range and bearing to a landmark. Zero-mean Gaussian noise is added to both measurements. Our mobile robot is configured as a Dubins car, which we assume has constant translational and rotational speed during one simulation period T. The assumed vehicle dynamics are as follows:

$$x_t = x_{t-1} + v_{T_1} dt \cos\theta_{t-1} \tag{19}$$

$$y_t = y_{t-1} + v_{T_1} dt \sin\theta_{t-1} \tag{20}$$

$$\theta_t = \theta_{t-1} + \omega_{T_1} dt, \tag{21}$$

where $v_{T_1} \in [v_{\min}, v_{\max}], \omega_{T_1} \in [\omega_{\min}, \omega_{\max}], t \in [T_{\min}, T_{\max}]$. Similarly, zero-mean Gaussian noise is added to the state propagation equations above. We also place a circumscribed obstacle radius around landmarks to avoid collision. The robot is initialized with low uncertainty to perform the exploration task. The configuration details of our simulation environment are listed in Table 1.

The RRT planner only takes samples within the free space, which is defined by an occupancy probability less than 0.4. To take into account the kinodynamic constraints, we construct a motion library in advance by forward simulating the vehicle with all possible combinations of simulation period, translational and rotational velocity. In the *Steer* function, the connectivity of nodes is checked through searching for the nearest end point in the motion library given a certain radius.

4.1 Comparison

We analyze the performance of our proposed algorithm by comparing it with two variants of entropy-based exploration algorithms.

1. SLAM-OG [2]. The utility function consists of a normalized localization component (SLAM) and a normalized occupancy grid (OG) mapping component:

$$U_{SLAM\_OG} = \alpha \frac{I_{SLAM}(\mathcal{X}, \mathcal{L}|\mathbf{a})}{I_{SLAM\max}} + (1-\alpha)\frac{H_{OG}(\mathbf{m}|\mathbf{a})}{H_{OG\max}}, \tag{22}$$

where $I_{SLAM} \approx \sum_{i=1}^{|\mathcal{L}_o|} \sqrt{\det(\Sigma_{\mathbf{l}_i})}$, $H_{OG} = \sum_{\mathbf{m}_i \in \mathbf{m}} H(\mathbf{m}_i)$.

2. OG [1], [20]. The utility function computes the map entropy weighted by the likelihood of robot poses:

$$U_{OG} = \int_{\mathcal{X}} H(\mathbf{m}|\mathcal{X}, \mathcal{Z}) p(\mathcal{X}, \mathcal{L}_o|\mathcal{Z}) \tag{23}$$

$$\approx \frac{1}{N} \sum_{n=1}^{N} H(\mathbf{m}^{[n]}|\mathcal{X}^{[n]}, \mathcal{Z}). \tag{24}$$

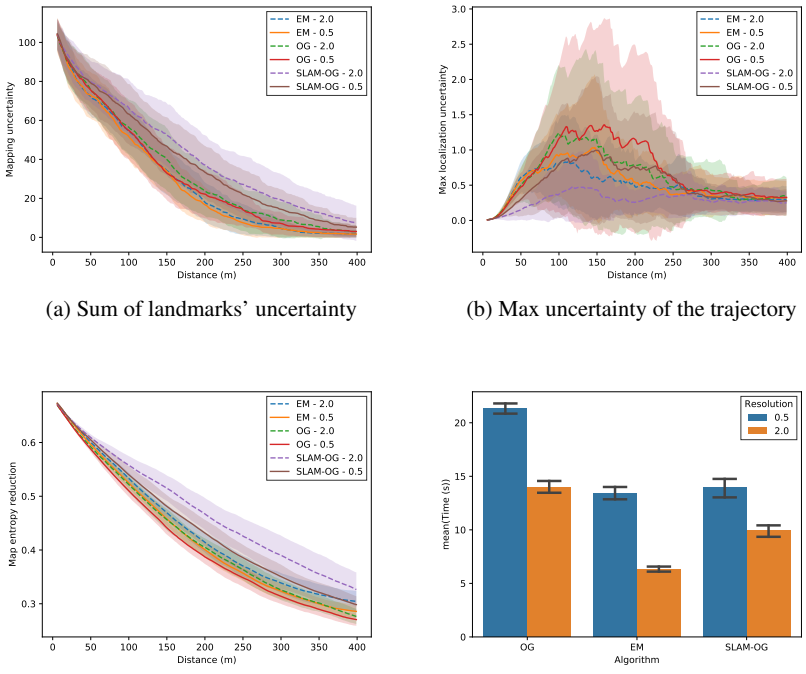
In the first utility function (SLAM-OG), the uncertainty of continuous pose variables is expressed using differential entropy, unlike occupancy grid maps with a discrete probability distribution. As a consequence, the scale of pose uncertainty is much smaller than that of map entropy [3]. The second utility function (OG) prioritizes grid cells that lie inside the robot's current sensor observation cone. In addition, the potential impact of loop closures on previously observed portions of the map (and their accuracy) is ignored, due to the utility function's emphasis on the entropy of occupancy maps.

4.2 Results

The simulated experiments were conducted such that all the parameters are the same for all algorithms except for the weight of the distance cost. Here we used a distance weight with a constant rate of decay with respect to the area discovered, or $\alpha = w_{\max} \frac{\text{\#free cells}}{\text{\#total cells}}$. For each algorithm, the maximum weight $w_{\max}$ was chosen to achieve the best performance through exhaustive search. Landmarks were uniformly sampled in the environment and 50 environments were generated. We also tested all algorithms using two levels of resolution for the virtual landmarks and occupancy grids: $2.0(m)$ and $0.5(m)$ to see the effect of using a high resolution map. One example of exploration using the EM algorithm with a low resolution is demonstrated in Fig. 1. The performance of the algorithms was averaged among the 50 experiments performed, as listed in Fig 3.

Fig. 3a shows the uncertainty reduction of landmarks, where the uncertainty is computed by $\sum_{\mathbf{l} \in \mathcal{L}} \text{tr}(\Sigma_{\mathbf{l}})$, in which unobserved landmarks have initial estimates $\sigma_0 = 3.0$. On one hand, this metric takes into account localization and mapping uncertainty through observed landmarks; on the other hand, it also measures the exploration rate by incorporating unknown landmarks. Fig. 3b shows the evolution of localization error by evaluating $\max_{\mathbf{x} \in \mathcal{X}} \text{tr}(\Sigma_{\mathbf{x}})$. We also compare different algorithms with respect to average map entropy ($\sum H(\mathbf{m}_i)/N_{\mathbf{m}}$) reduction, shown in Fig. 3c. The planning time is shown in Fig. 3d under Ubuntu 14.04 on an i7-6950X CPU, although these Python-based implementations are capable of futher optimization.

We have the following observations from the comparisons. The OG algorithm, weighing map entropy by localization uncertainty, has a comparable exploration rate with the proposed algorithm (EM) in the beginning (Fig. 3a, 3c). However,



(a) Sum of landmarks' uncertainty

(b) Max uncertainty of the trajectory

(c) Average entropy reduction

(d) Planning time

Fig. 3: The results of 50 exploration trials with the same randomly initialized landmarks for every algorithm.

landmark uncertainty is reduced at a lower rate when most of the landmarks are observed (Fig. 3a). In addition, by exploring with the EM algorithm, we end up with more accurate feature-based maps - their curves are closer to zero in the final stage of Fig. 3a. The SLAM-OG algorithm places constant relative weights on SLAM and OG uncertainty, and thus its exploration is the most conservative, but it has the lowest localization error as shown in Fig. 3b. In contrast, the EM algorithm is able to maintain low trajectory uncertainty and it achieves a superior exploration rate until landmark uncertainty is close to zero. More importantly, the planning time of the EM algorithm is much cheaper because it doesn't need to calculate the map entropy for every candidate by sampling from the posterior trajectory distributions (Eq. 24). Overall, by using a finer resolution, the performance of all algorithms improves. However, for the EM algorithm, the small gains in performance do not appear to merit the sacrifice in computation time.

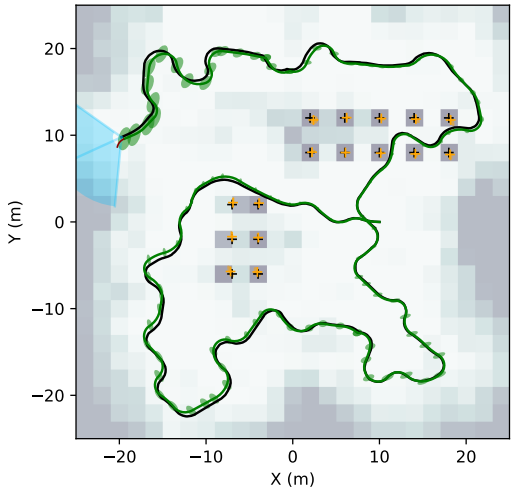


Fig. 4: An exploration example in a structured environment.

5 Conclusions and Future Work

In this paper, we proposed the concept of virtual landmarks, which are latent variables representing the possible locations of actual landmarks in the environment. We presented a novel utility function for the autonomous exploration problem in feature-based maps, which essentially computes the covariance criteria at virtual landmarks. The direct modeling of landmarks potentially observed in the future enables more accurate mapping and also a comparable exploration rate with respect to traditional methods. However, the evaluation of the proposed utility function relies on the recovery of a full trajectory, which is expensive for real-time applications. In addition, a more memory-efficient approach to storing and updating virtual landmarks is a subject of ongoing and future work.

Acknowledgements This research has been supported in part by the National Science Foundation, grant number IIS-1551391.

References

1. J. L. Blanco, J. A. Fernandez-Madrigal and J. González. A novel measure of uncertainty for mobile robot slam with Rao-Blackwellized particle filters. *Int. J. Robotics Research*, 27(1): 73–89 (2008)
2. F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky and H. F. Durrant-Whyte. Information based adaptive robotic exploration. *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 540–545 (2002)

3. H. Carrillo, P. Dames, V. Kumar and J. A. Castellanos. Autonomous robotic exploration using occupancy grid maps and graph slam based on shannon and rényi entropy. *Proc. IEEE Int. Conf. on Robotics and Automation*, 487–494 (2015)
4. H. Carrillo, I. Reid and J. A. Castellanos. On the comparison of uncertainty criteria for active SLAM. *Proc. IEEE Int. Conf. on Robotics and Automation*, 2080–2087 (2012)
5. G. Celeux and G. Govaert. A classification EM algorithm for clustering and two stochastic versions. *Computational statistics and Data analysis*, 14(3):315–332 (1992)
6. B. Charrow, S. Liu, V. Kumar and N. Michael. Information-theoretic mapping using Cauchy-Schwarz quadratic mutual information. *Proc. IEEE Int. Conf. on Robotics and Automation*, 4791–4798 (2015)
7. B. Charrow, G. Kahn, S. Patil, S. Liu, K. Goldberg, P. Abbeel, N. Michael and V. Kumar. Information-theoretic planning with trajectory optimization for dense 3D mapping. *Proc. Robotics: Science and Systems*, 2015
8. L. Chen, P. O. Arambel, and R. K. Mehra. Estimation under unknown correlation: covariance intersection revisited. *IEEE Trans. on Automatic Control*, 47(11):1879–1882 (2002)
9. H. J. S. Feder, J. J. Leonard and C. M. Smith. Adaptive mobile robot navigation and mapping. *Int. J. Robotics Research*, 18(7):650–668 (1999)
10. S. Huang, N. M. Kwok, G. Dissanayake, Q. P. Ha and G. Fang. Multi-step look-ahead trajectory planning in SLAM: Possibility and necessity. *Proc. IEEE Int. Conf. on Robotics and Automation*, 1091–1096 (2005)
11. M. G. Jadidi, J. V. Miró, R. Valencia and J. Andrade-Cetto. Exploration on continuous Gaussian process frontier maps. *Proc. IEEE Int. Conf. on Robotics and Automation*, 6077–6082 (2014)
12. S. Karaman and E. Frazzoli. Sampling-based algorithms for optimal motion planning. *Int. J. Robotics Research*, 30(7):846–894 (2011)
13. M. Kaess, and F. Dellaert. Covariance recovery from a square root information matrix for data association. *Robotics and Autonomous Systems*, 57(12):1198–1210 (2009)
14. M. Kaess, H. Johannsson, R. Roberts, V. Ila, J. J. Leonard and F. Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *Int. J. Robotics Research*, 31(2):216–235 (2012)
15. S. M. LaValle. Rapidly-exploring random trees: A new tool for path planning. (1998)
16. A. A. Makarenko, S. B. Williams, F. Bourgault and H. F. Durrant-Whyte. An experiment in integrated exploration. *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 534–539 (2002)
17. R. Martinez-Cantin, N. de Freitas, A. Doucet, and J. A. Castellanos. Active policy learning for robot planning and exploration under uncertainty. *Proc. Robotics: Science and Systems*, 321–328 (2007)
18. H. Moravec and A. Elfes. High resolution maps from wide angle sonar. *Proc. IEEE Int. Conf. on Robotics and Automation*, 116–121 (1985)
19. R. Sim and N. Roy. Global A-optimal robot exploration in slam. *Proc. IEEE Int. Conf. on Robotics and Automation*, 1050–4729 (2005)
20. C. Stachniss and G. Grisetti and W. Burgard. Information gain-based exploration using Rao-Blackwellized particle filters. *Proc. Robotics: Science and Systems*, 65–72 (2005)
21. B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon. Bundle adjustment - a modern synthesis. *Int. Workshop on Vision Algorithms*, 298–372 (1999)
22. R. Valencia, J. V. Miró, G. Dissanayake and J. Andrade-Cetto. Active pose SLAM. *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 1885–1891 (2012)
23. J. Vallvé and J. Andrade-Cetto. Active pose SLAM with RRT*. *Proc. IEEE Int. Conf. on Robotics and Automation*, 2167–2173 (2015)
24. M. Yguel, O. Aycard and C. Laugier. Update policy of dense maps: Efficient algorithms and sparse representation. *Field and Service Robotics*, Springer Berlin/Heidelberg, 23–33 (2008)