Sparse Planning Graphs for Information Driven Exploration

Erik Nelson, Vishnu Desaraju, John Yao
The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15217
{enelson, rajeswar, johnyao}@cmu.edu

***Abstract*—Hello, here is some text without a meaning. This text should show, how a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like ≫Huardest gefburn≪. Kjift – Never mind! A blind text like this gives you information about the selected font, how the letters are written and the impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for a special contents, but the length of words should match to the language.**

I. Introduction

Exploration is a key capability that enables robotic vehicles to operate in unknown environments. In this project we develop an active perception policy for robotic exploration. Active perception exploration formulations choose control actions which optimize an information-theoretic objective function such as Shannon's mutual information or entropy [1], [7] over the robot's map, given a sensor measurement model. Other common exploration techniques, such as frontier exploration [], use geometric reasoning to infer explorative paths. While these strategies work well in practice, they operate on a maximum likelihood estimate of the map, and apply heuristics to determine the most uncertain locations in the environment. In contrast, active perception strategies do not utilize geometric or maximum likelihood assumptions, and instead interpret the map as a binary random variable, choosing actions which directly minimize the random variable's uncertainty. Additionally, information-based exploration methods easily extend to 3D configuration spaces, a benefit which is not shared by frontier methods. Julian et al. prove that maximizing mutual information between a robot's map and expected future map naturally yields explorative behaviors [3].

Active perception formulations seek to optimize information-theoretic objectives. While this optimization is real-time for short planning horizons, these metrics are often expensive to compute online, requiring double integration over possible future robot states and measurements, or Monte Carlo sampling from the distribution of measurements. These expensive per-pose computations inhibit online dense evaluation over a configuration space. In this project, we aim to develop an efficient active perception exploration strategy which evaluates the information-theoretic objective in a sparse, but well-chosen set of poses across the configuration space. This strategy evaluation the objective function a limited number of times, while still generating paths that sufficiently explore the space.

To achieve real-time active perception exploration, we use a Rapidly-Exploring Random Tree (RRT) to generate sets of dynamically feasible actions over a finite planning horizon [5]. RRT planners trade trajectory optimality for efficiency, allowing for evaluation of many potential future locations in the configuration space during a single planning step. In addition, RRT planners are anytime, and generate potential trajectories for a pre-specified amount of time before evaluating the most optimal sampled trajectory. Our strategy evaluates each RRT leaf-node using the information-theoretic objective function, and stores the resulting reward in the tree. After planning for a specified amount of time, the maximum reward leaf-node is chosen as the optimal location to visit, and the RRT is traversed to generate a dynamically feasible trajectory to that location.

In addition to the efficiency gains from using an RRT, a recent work by Charrow et al. [2] has proposed the Cauchy-Schwarz Quadratic Mutual Information (CSQMI) as an efficient information-theoretic objective function. CSQMI is theoretically well-motivated: it is derived from Renyi's Quadratic Entropy, a generalization of Shannon's entropy. However, in contrast to Shannon's mutual information (which is derived directly from Shannon's entropy), CSQMI is shown to have superior computational efficiency.

The contribution of this work is an exploration framework to enable online exploration using CSQMI in sparse planning graphs, such as RRTs. We demonstrate an implementation of our approach in simulation, and provide analysis of experiments in which a mobile ground robot must explore an unknown space using a laser scanner range sensor. We discuss the formulation and implementation of the CSQMI metric, RRT, and a controller and Unscented Kalman Filter (UKF) that were developed to enable trajectory tracking and state estimation. Finally, we discuss the implementation of these capabilities on a real ground robot.

This paper is structured in the following manner: Section II gives a brief overview of occupancy grid mapping which is necessary for the CSQMI information metric. Section III details the CSQMI information-

theoretic cost objective and its similarities to Shannon's mutual information. Section IV discusses the measurement model that was chosen to evaluate expected future measurements. Sections V and **??** cover the RRT and UKF formulations used in our implementation. Finally, Sections VII and VIII give results and analysis of our implementation in simulation, and describe future work towards implementing our algorithms on a ground robot.

II. Occupancy Grid Mapping

In order to develop an information-theoretic reward surface, we model the robot's map as a binary random variable. We therefore discretize the space, and represent the map as an occupancy grid - a common environmental representation for robotic mapping.

As a basis for the core formulation in the following sections, we provide a brief overview of occupancy grid mapping. Occupancy grids are a common and useful We represent the map as an occupancy grid, which consists of a set of cells: $m = \{m^i\}_{i=1}^N$. The probability that an individual cell is occupied is given by $p\left(m^i \mid x_{1:t}, z_{1:t}\right)$, where $x_{1:t}$ denotes the history of states of the vehicle, and $z_{1:t}$ denotes the history of range observations accumulated by the vehicle. We assume that cell occupancy probabilities are independent of one another: $p\left(m \mid x_{1:t}, z_{1:t}\right) = \prod_i p\left(m^i \mid x_{1:t}, z_{1:t}\right)$. For notational simplicity we write the map conditioned on random variables $x_{1:t}$ and $z_{1:t}$ as $p_t\left(m\right) := p\left(m \mid x_{1:t}, z_{1:t}\right)$. Additionally, unobserved grid cells are assigned a uniform prior of being occupied.

We represent the occupancy status of grid cell m^i at time t with a log odds expression

$$l_t := \log \frac{p\left(m^i \mid z_{1:t}\right)}{p\left(\bar{m}^i \mid z_{1:t}\right)} \tag{1}$$

where $\bar{m}^i$ denotes the probability that m^i is unoccupied. When a new observation z_t is obtained, the log odds update is given by

$$l_t = l_{t-1} + \log \frac{p\left(m^i \mid z_t\right)}{p\left(m^i\right)} - \log \frac{p\left(\bar{m}^i \mid z_t\right)}{p\left(\bar{m}^i\right)} \tag{2}$$

where the last two terms represent the inverse sensor model.

III. Information-Theoretic Objective

The goal of active perception exploration is to find a dynamically feasible sequential set of actions over a time interval, $\tau := t+1, \ldots, t+T$, which enable the robot to explore its environment. We use the critera that an explorative action is one which allows the robot to position itself in locations that generate observations which are informative to the robot's map. Under this criteria, choosing the optimal exploration action will maximally reduce the uncertainty in the robot's map. An *action* can be defined as a discrete sequence of states, $x_\tau = [x_{t+1}, \ldots, x_{t+T}]$. While executing an action, the robot will obtain a set of measurements $z_\tau(x_\tau) = [z_{t+1}(x_{t+1}), \ldots, z_{t+T}(x_{t+T})]$ by sensing from the states x_τ. $z_\tau(x_\tau)$ is modeled as a random variable whose distribution is parameterized by a deterministic action, x_τ. In our formulation, we choose to select actions from a library of motion primitives, $\mathcal{X}$, which are generated by a planner. Under these notations, an explorative planner must determine x_τ^*: the action that visits locations which allow the robot to obtain the set of measurements which maximally reduce uncertainty in its current map.

To determine x_τ^*, we follow Charrow et al. [2] and maximize the Cauchy-Schwarz Quadratic Mutual Information (CSQMI) rate between the current map and the expectation of future measurements collected along an action, $\mathrm{I_{CS}}[m; z_\tau | x_\tau]$.

$$x_\tau^* = \underset{x_\tau \in \mathcal{X}^T}{\operatorname{argmax}} \frac{\mathrm{I_{CS}}\left[m; z_\tau(x_\tau)\right]}{R\left(x_\tau\right)} \tag{3}$$

Here, $R : \mathbb{R}^{|x_\tau|} \rightarrow \mathbb{R}^+$ computes the estimated time required to complete action x_τ. We choose to maximize CSQMI as opposed to more common metrics, such as Shannon's mutual information (MI), due to the property that CSQMI can be computed exactly in $\mathcal{O}(N^2)$, and closely in $\mathcal{O}(N)$, when N is the number of occupancy grid cells intersected by all rays in z_τ. In contrast, MI requires averaging many $\mathcal{O}(N^2)$ Monte Carlo samples. The full MI and CSQMI solutions are remarkably similar when computed over an occupancy grid map, further motivating CSQMI's use. CSQMI between the map and collection of measurements from τ is given in Eq. (4).

$$\begin{aligned} &\mathrm{I_{CS}}\left[m; z_\tau\right] = \\ &\quad - \log \frac{\left(\sum \int p(m, z_\tau) p(m) p(z_\tau) dz_\tau\right)^2}{\sum \int p^2(m, z_\tau) dz_\tau \sum \int p^2(m) m^2(z_\tau) dz_\tau} \end{aligned} \tag{4}$$

where sums are over possible enumerations of m, and integrals are over the space of measurements that can be observed during τ. CSQMI is non-negative, and maps two random variables to a real valued quantity which represents the information that one learns about each variable by learning the other. CSQMI is equal to zero i.f.f. its arguments are independent.

We represent measurements as B-tuple random variables, such that a measurement z_k^b captured at time $k \in \tau$ contains B beams. Most occupancy grid measurement models assume that cells not intersected by a laser beam are not updated in the map. Let c be the set of cells intersected by laser beam z_k^b, and let $C = |c|$. Then $\mathrm{I_{CS}}[m; z_k^b] = \mathrm{I_{CS}}[c; z_k^b]$. Charrow et al. [2] derive a closed form solution to the CSQMI between a single laser beam and the robot's map.

$$\begin{aligned} \mathrm{I_{CS}}[c; z_k^b] = &\log \sum_{l=0}^{C} w_l \mathcal{N}(0, 2\sigma^2) \\ &+ \Bigg(\log \prod_{i=1}^{C} (o_i^2 + (1 - o_i)^2) \\ &\quad \sum_{j=0}^{C} \sum_{l=0}^{C} p(e_j) p(e_l) \mathcal{N}(\mu_l - \mu_j, 2\sigma^2) \Bigg) \\ &- 2 \log \sum_{j=0}^{C} \sum_{l=0}^{C} p(e_j) w_l \mathcal{N}(\mu_l - \mu_j, 2\sigma^2) \end{aligned} \tag{5}$$

Here, $p(e_j)$ is the probability that the jth cell in the raycast c is the first occupied cell, and $o_i = p(c^i = 1)$, the probability that ith cell in the raycast is occupied. Weights, $w_{l\in[1,C]}$ can be pre-computed over the raycast to avoid duplication, and are evaluated as

$$w_l = p^2(e_l) \prod_{j=l+1}^{C} (o_j^2 + (1 - o_j)^2) \tag{6}$$

By assuming that the resolution of the occupancy grid is greater than the variance of the range sensor, which is typically the case in occupancy grid mapping scenarios, and by assuming that Gaussians approach zero with growth in $|\mu_l - \mu_j|$, one may assert that the inner sum in the double sum terms is only non-zero when l is close to j. In other words, the $\mathcal{O}(C^2)$ double sum terms can be reduced to $\mathcal{O}(C)$ complexity by assuming

$$\begin{aligned} \sum_{j=0}^{C}\sum_{l=0}^{C} \alpha_{j,l} &\approx \sum_{j=0}^{C}\sum_{l=j-\delta}^{j+\delta} \alpha_{j,l} \\ \sum_{j=0}^{C}\sum_{l=0}^{C} \beta_{j,l} &\approx \sum_{j=0}^{C}\sum_{l=j-\delta}^{j+\delta} \beta_{j,l} \end{aligned} \tag{7}$$

where $\delta \ll C$, and $\alpha_{j,l}$ and $\beta_{j,l}$ are the weighted Gaussian terms inside of the double sums in Eq. (5). Our simulator and robot's laser scanners have $\sigma \approx 0.1$ cm, and use occupancy grid sizes of 10.0 cm, allowing us to choose $\delta = 1$ with minimal loss of information.

IV. Measurement Model

The CSQMI of a single beam (Eq. (5)) depends on the ability to estimate $p(e_j)$, the probability that the jth cell in a raycast is the first occupied grid cell. Rather than building a measurement model using the maximum likelihood estimate of z_j like other similar works [2], [8], [3], we compute a discrete distribution over the raycast, where each cell value directly approximates the probability cell c^i is the first occupied cell in c, $p(e_j = c^i)\ \forall c^i \in c$ using the continuous occupancy values in c.

To calculate this distribution, we use a generalization of the geometric distribution, which is commonly used to determine the probability distribution over the number of Bernoulli trials necessary to obtain one success, supported on the set $\mathbb{N}^+$. Since each occupancy grid cell in a raycast contains continuous probabilities $\in [0, 1]$ of the cell's occupancy, we build on the geometric distribution to support independent, but not identically distributed (i.n.i.d) random variables. The probability that a cell c^i is the first cell that terminates the raycast is then the probability that no previous cells in the raycast did.

$$p(e_j = c^i) = o^i \prod_{j=1}^{i-1} (1 - o^j) \tag{8}$$

Although Eq. (8) requires $\mathcal{O}(C^2)$ to be computed for all $c^i \in c$, we were able to derive an efficient recursive formula to compute the distribution $p(e_j)$, which runs in $\mathcal{O}(C)$.

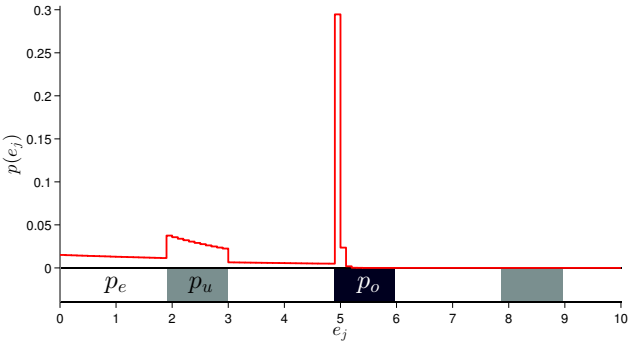


Figure 1: The distribution $p(e_j)$ over a depicted 1-dimensional map with $p_e = 0.01$, $p_o = 0.92$, $p_u = 0.05$.

$$\begin{aligned} p(e_j = c^i) &= o^i \prod_{j=1}^{i-1} (1 - o^j) \\ &= o^i \left(\frac{o^{i-1}}{o^{i-1}}\right) (1 - o^{i-1}) \prod_{j=1}^{i-2} (1 - o^j) \\ &= o^i \frac{1 - o^{i-1}}{o^{i-1}} p(e_j = c^{i-1}) \\ &= o^i \left(\frac{1}{o^{i-1}} - 1\right) p(e_j = c^{i-1}) \end{aligned} \tag{9}$$

The distribution $p(e_j)$ is depicted over a one-dimensional raycast in Fig. 1.

V. Closed Loop RRTs

To make use of this information cost function to guide exploration, we consider a sampling-based planning approach that can evaluate the predicted information gain throughout the environment. In addition, we wish to use the occupancy grid that is being updated online (as described in Sect. **??**) to guide the vehicle around obstacles in the environment. The rapidly-exploring random tree (RRT) algorithm is well suited to planning paths through these types of large environments, and works as follows. The planner starts from the vehicle's current state and samples a point x in the environment. Using the occupancy grid, we can reject samples that lie in cells with a sufficiently high probability of being occupied. If the sample is valid, we find the closest node in the tree of paths (initially just the vehicle state), where closeness is measured in terms of Euclidean distance, and add a new edge to the tree connecting the sample point to the nearest node. Then a new sample is drawn and the process repeats to grow a tree of path segments through the environment. This tree-growing process terminates after a specified time, and the minimum cost path is returned.

Figure **??** shows snapshots of the system planning through the environment while updating the occupancy grid. The edges in the tree are smooth since they are generated by forward simulating the closed-loop vehicle dynamics toward the sample point, resulting

in a variant of the RRT algorithm known as Closed-loop RRT (CL-RRT) [5]. This approach is traditionally used to ensure dynamic feasibility and dense collision checking. However, the forward simulation also means we have full state information for the system at the end of each segment. This allows us to evaluate the predicted information gain at that point and assign a corresponding cost to candidate trajectory.

To change CL-RRT from a goal-directed planner to an exploration-driven planner, we first define the sampling distribution to be a Gaussian centered about the root of the tree, with no bias toward any direction (unlike standard sampling-based planners that will sample the goal some small probability). We also define the cost of each branch segment to just be the information metric computed at its endpoint (as opposed to a more traditional setup where the cost is the total distance traveled from the root plus a cost to go based on an admissible heuristic, such as Euclidean distance to a goal). Finally, since there is no goal to guide the selection of the best branch from the tree, we simply select the branch with the minimum cost endpoint in the entire tree. This enables the planner to compute paths that aim to maximize the predicted information gain.

Define $\mathcal{X}_{\text{free}}$ to be the space spanned by the unoccupied cells in the occupancy grid.

Algorithm 1 CL-RRT: Tree Expansion

```
Sample point x_s from the environment
Select min-cost node from n nearest in tree
k ← 0
x̂(t+k) ← last state at n
while x̂(t+k) ∈ X_free(t+k) and x̂(t+k) ≠ x_s do
    Compute control input to drive system to x_s
    Forward simulate system dynamics
    Compute next state x̂(t+k+1) from propagation model
    k ← k + 1
end whileN ← r_final
for each feasible node N produced do
    Update cost estimates for N
    Add N to tree
end for
```

VI. State Estimation

State estimation addresses the problem of determining the robot's pose in the environment given noisy sensor observations. This state estimation pipeline's pose output is a superior alternative to directly feeding in exteroceptive sensor observations to the controller and planner. In this section, we present an Unscented Kalman Filter for fusing inertial measurement unit (IMU) and localization observations for a 2D robot.

A. Process Model

The vehicle state $\mathbf{x}$ consists of the global position $\mathbf{p} = [p_x \; p_y]^T$, global heading angle θ, global velocity $\mathbf{v} = [v_x \; v_y]^T$, IMU angular velocity bias in the z-direction b_ω and IMU linear acceleration biases in the x- and y-directions $\mathbf{b}_a = [b_{ax} \; b_{ay}]^T$.

$$\dot{\mathbf{p}} = \mathbf{v} \tag{10}$$
$$\dot{\theta} = \omega - b_\omega - n_\omega \tag{11}$$
$$\dot{\mathbf{v}} = \mathbf{C}(\theta)\left(\mathbf{a} - \mathbf{b}_a - \mathbf{n}_a\right) \tag{12}$$
$$\dot{b}_\omega = n_{b\omega} \tag{13}$$
$$\dot{\mathbf{b}}_a = \mathbf{n}_{ba} \tag{14}$$

The IMU measurements are comprised of the z-direction angular velocity ω as well as the body frame x- and y-direction linear accelerations $\mathbf{a} = [a_x \; a_y]^T$. Both measurements are modelled as being corrupted by additive Gaussian white noise and a random walk bias driven by Gaussian white noise (13), (14).

$$\mathbf{n} = \begin{bmatrix} \mathbf{n}_a \\ n_\omega \\ \mathbf{n}_{ba} \\ n_{b\omega} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}) \tag{15}$$
$$\mathbf{Q} = \text{diag}\{\sigma_a^2, \sigma_a^2, \sigma_\omega^2, \sigma_{ba}^2, \sigma_{ba}^2, \sigma_{b\omega}^2\} \tag{16}$$

Since the IMU measurements are in the body frame, we use a 2×2 rotation matrix $\mathbf{C}(\theta)$ to rotate them into the global reference frame (12). Each component of the noise vector in (15) is assumed to be independent, and their corresponding covariance sigma values are chosen by offline sensor characterization tests (16).

B. Correction Model

The laser scans are used to construct a map of the environment, and a grid-based localization algorithm is used to compute the global pose (position and heading) of the vehicle with respect to the map. The sensor model for the localization algorithm is given by

$$\mathbf{z} = \begin{bmatrix} \mathbf{p} \\ \theta \end{bmatrix} + \begin{bmatrix} \mathbf{C}(\theta) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{w} \tag{17}$$
$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}) \tag{18}$$

The observation noise covariance matrix $\mathbf{R}$ (18) is computed by fitting a multivariate Gaussian to the posterior probability grid of the localization algorithm. Because this covariance is expressed in the scanner frame (assumed to be coincident with the body frame), we rotate its associated noise vector $\mathbf{w}$ into the world frame in (17).

C. Unscented Kalman Filter

The nonlinearities introduced by the rotation in the process and correction models motivated the choice of the Unscented Kalman Filter in this project. For brevity, we omit the complete presentation of all the steps involved in the UKF (we refer the reader to [4]). We apply the process model and correction model equations when the relevant measurement is received by the state estimator. Outlier exteroceptive observations are rejected by a chi-squared innovation gate.

VII. Results

To test the exploration stategy experimentally, we implemented the formulations and algorithms discussed in Sections III-VI in a C++ simulation environment. The simulator assumes a ground robot constrained to a 2D plane, which is equipped with a noisy laser scanner (range $= 30$ m, $\sigma = 0.1$ cm), and IMU. Our simulator generates horizontal laser scans from a meshed point cloud input, and generates IMU observations according to the true state of the robot. Prior to evaluating CSQMI or building an RRT, we build a map and localize with respect to it using a custom SLAM implementation that was developed prior to this project. Our SLAM implementation leverages ICP for laser odometry [6], a histogram filter for localization [8], and a custom 3D mapping framework.

VIII. Conclusion and Future Work

References

[1] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte. Information based adaptive robotic exploration. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, volume 1, pages 540–545. IEEE, 2002.

[2] B. Charrow, S. Liu, V. Kumar, and N. Michael. Information-theoretic mapping using cauchy-schwarz quadratic mutual information. technical report. *IEEE International Conference on Robotics and Automation (ICRA)*, 2015.

[3] B. Julian, S. Karaman, and D. Rus. On mutual information-based control of range sensing robots for mapping applications. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5156–5163. IEEE, 2013.

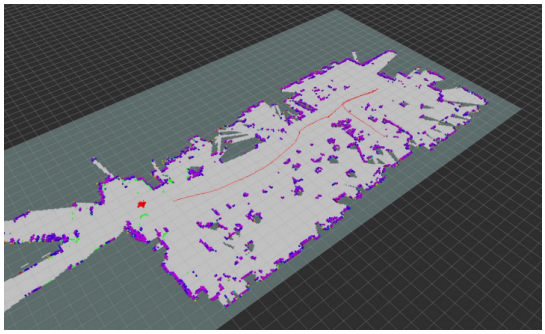
[4] S. J. Julier and J. K. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Proc. SPIE*, volume 3068, pages 182–193, July 1997.

[5] Y. Kuwata, S. Karaman, J. Teo, E. Frazzoli, J. P. How, and G. Fiore. Real-time motion planning with applications to autonomous urban driving. *IEEE Transactions on Control Systems Technology*, 17(5):1105–1118, Sepetember 2009.

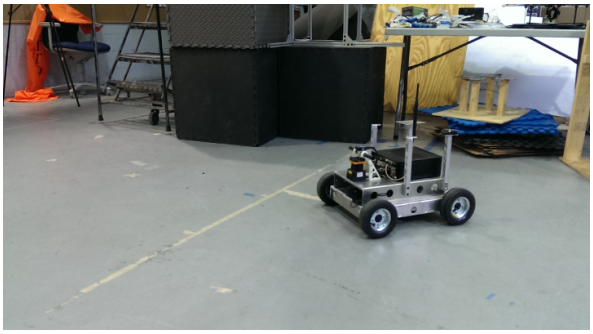
[6] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat. Comparing icp variants on real-world data sets. *Autonomous Robots*, 34(3):133–148, 2013.

[7] C. Stachniss, G. Grisetti, and W. Burgard. Information gain-based exploration using rao-blackwellized particle filters. In *Robotics: Science and Systems*, volume 2, 2005.

[8] S. Thrun, W. Burgard, and D. Fox. *Probabilistic robotics*. MIT press, 2005.



(a) Map and trajectory



(b) Ground robot

Figure 2: A ground robot mapping while being driven through a cluttered environment.