Multi-Robot Mapping and Exploration of Environments with Hazards

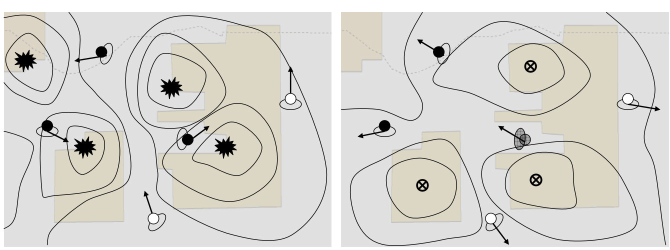
Mac Schwager
GASP Lab, University of Pennsylvania
Philadelphia, PA 19104, USA, and
CSAIL, MIT
Cambridge, MA 02139, USA
Email: schwager@mit.edu

Philip Dames and Vijay Kumar
GRASP Lab
University of Pennsylvania
Philadelphia, PA 19104, USA
Email: {pdames, kumar}@seas.upenn.edu

Daniela Rus
CSAIL, MIT
Cambridge, MA 02139, USA
Email: rus@csail.mit.edu

I. Introduction

Networks of robotic sensors have the potential to safely collect data over large-scale, unknown environments. They can be especially useful in situations where the environment is unsafe for humans to explore. In many such situations, robots are also susceptible to hazards. It is important to design exploration and mapping algorithms that are hazard-aware, so that the robotic sensor network can effectively carry out its task while minimizing the impact of individual robot failures. In this work we describe an algorithm, based on an analytic expression for the gradient of mutual information, that enables a robotic sensor network to estimate a map of events in the environment while avoiding failures due to unknown hazards.



(a) Event map (b) Hazard map

Fig. 1. The tragic accident at the Fukushima nuclear power plant in Japan is a fitting scenario for our algorithm. Hypothetical maps of the events and hazards are shown over an image of the Fukushima plant from `http://maps.google.com/`. On the left, the events of interest are structural defects represented by the explosion symbols, and the contour lines represent the probability of detecting these defects. Sensors move to determine where the structural defects are by increasing the informativeness of their sensor readings. Black circles represent sensors that see a defect while white circles do not see a defect. On the right, the hazards are radiation sources represented by the $\otimes$ symbol, and the contours represent the probability of failure due to radiation damage. By moving to increase informativeness, the robots implicitly avoid hazards that may cause failure thereby preventing information from being collected. The grayed-out robot in the center has failed due to the high radiation levels.

Consider, for example, the recent tragic accident at the Fukushima nuclear power plant in Japan, which sustained critical damage from a large earthquake and tsunami in March, 2011. The algorithm we describe here could be used by a team of flying quadrotor robots with cameras to inspect the plant for structural damage, keeping human workers at a safe distance. With our algorithm the robots could build a map of the areas that are likely to have structural damage, while simultaneously building a map of the radiation hazards, to avoid failure due to radiation exposure. This scenario is illustrated in Fig. 1. Both the event map and the hazard map are estimated online using a recursive Bayesian filter, where the event map is estimated from evidence of structural damage seen by the cameras, and the hazard map is estimated by the previous failures of other robots. The robots move along the gradient of mutual information, which gives the direction of expected maximum information gain given the current maps, thereby driving the exploration of the environment. Our algorithm could also be used, for example, to monitor forest fires while avoiding fire damage, taking ocean measurements while avoiding damage from adverse weather, or mapping a chemical spill site while avoiding failure from caustic chemicals.

In all of these examples, the robots must move to both avoid hazards and provide useful sensor information. Although these two objectives may seem to be in conflict with one another, they are in fact complementary. If we want to map the events as precisely as possible, we implicitly want the robots to avoid hazardous areas, since the failure of a robot makes it unable to contribute to estimating the event map in the future. We use the gradient of mutual information to move the sensors so that their next measurements are as informative as possible. The gradient strategy blends the avoidance of hazards and the seeking of information into one probabilistically consistent objective. The mutual information gradient is then used to control the robots. We do not consider decentralization of the algorithm, though that will be a central concern of future work, and several existing methods can be adapted for decentralization.

A. Related Work

Mutual information is one of the fundamental quantities in information theory [1, 2] and has been used extensively as a metric for robotic sensor network control and static sensor placement. For example, in [3] and [4] mutual information is used as a metric for driving robotic sensor networks in gridded environments for target tracking and exploration tasks. Also, [5] focused on decentralization and scalability using particle filters to approximate mutual information for target tracking. Recently in [6] the gradient of mutual information was used to drive a network of robots, and a sampling method was employed to improve computational efficiency. The property of submodularity of mutual information was used in [7] for placing static sensors at provably near-optimal positions for information gain. Approximations on information gain for static sensor placement were derived in [8] and an informative trajectory planning algorithm was presented in [9]. In a different but related application, [10] uses mutual information to place static sensors to provide localization information to a mobile robot.

In contrast to the literature described above, our work is specifically concerned with estimating and avoiding environmental hazards as well as estimating the environment state. Also, we use an analytically derived expression for the gradient of mutual information for control, whereas previous works have used grid-based finite difference methods to increase mutual information (with the exception of [6]). A more complete presentation of the work described in this abstract will appear in [11].

The question we address is: How do we choose the next positions $x_1, \ldots, x_n$ to make the next Bayesian estimate of the event state as precise as possible? As already described, implicit in this problem is the tendency to avoid hazards because a failed robot is an uninformative robot. However, in our scenario, as in real life, all the robots will eventually fail. To counteract the depletion of robots, we let there be a base station located in the environment that deploys new robots to replace ones that have failed. We let the rate of releasing new robots balance the rate of failed ones, so that the total number of robots is constant at all times. However many other interesting possibilities exist. One important metric of performance is the asymptotic rate at which robots are lost due to the hazards, which we would expect to be lower than if we were using an information based approach without regard to the hazards.

II. Control for Active Sensing with Hazard Avoidance

Consider a situation in which n robots move in a 3D environment $Q \subset \mathbb{R}^3$. The robots have positions $x_i(t) \in Q$ and we want to use them to sense the state of the environment while avoiding hazardous areas that may cause the robots to fail. Let the positions of all the robots be given by the vector $x = [x_1^T \cdots x_n^T]^T$. The robots give simple sensor measurements y_i indicating wether or not they have sensed an event of interest near by. They also give a signal to indicate their failure status f_i. Denote the random vector of all sensor outputs by $y = [y_1 \cdots y_n]^T$ and the vector of all failure statuses as $f = [f_1 \cdots f_n]^T$.

The task of the robot network is to estimate the state of the environment state, $s = [s_1 \quad \cdots \quad s_{m_s}]^T$, with as little uncertainty as possible, while avoiding the unknown hazards $h = [h_1 \quad \cdots \quad h_{m_h}]^T$. Let $\phi_0(s)$ and $\psi_0(h)$ denote the robots' initial guess at the distribution of the state and the hazards, respectively, which can be uniform if we have no prior information about the events or hazards.

Let the robots sensors give readings at time t according to a likelihood distribution $\mathbb{P}(y_i^t \mid f_i^t, s)$, and suppose that the sensors conditioned on the environment state and hazard state are independent, so that

$$\mathbb{P}(y^t \mid f^t, s) = \Pi_{i=1}^n \mathbb{P}(y_i^t \mid f_i^t, s).$$

Similarly, let the probability of failure for a robot at time t be given the environment and hazard state be $\mathbb{P}(f_i^t \mid h)$, and assume that the failures are independent conditioned on the environment and hazard state, so that

$$\mathbb{P}(f^t \mid h) = \Pi_{i=1}^n \mathbb{P}(f_i^t \mid h)$$

This gives the model for our environment and robots that can be used to derive an information seeking, hazard avoiding controller. Using the standard formulation of Bayes rule we can find recursive filters to update the event distribution as

$$\phi_t(s) = \frac{\mathbb{P}(y^t \mid f^t, s)\phi_{t-1}(s)}{\sum_{s \in \mathcal{S}} \mathbb{P}(y^t \mid f^t, s)\phi_{t-1}(s)},$$

and the hazard distribution is updated by

$$\psi_t(h) = \frac{\mathbb{P}(f^t \mid h)\psi_{t-1}(h)}{\sum_{h \in \mathcal{H}} \mathbb{P}(f^t \mid h)\psi_{t-1}(h)}.$$

We use these updated estimates to move in the direction of highest expected information gain.

In Shannon information theory [2, 1], the mutual information between two random variables indicates how much information one random variable gives about the other. In our scenario, we want to move our robots so that their next measurement gives the maximum amount of information about the event distribution. Therefore the positions of the robots are parameters of the join distribution of the measurements and the events. We write $\mathbb{P}_{s,y}(x)$ to emphasize this dependence. Then, letting $\mathbb{P}_s(x) := \int_{y \in \mathcal{Y}} \mathbb{P}_{s,y}(x)\, dy$ and $\mathbb{P}_y(x) := \int_{s \in \mathcal{S}} \mathbb{P}_{s,y}(x)\, ds$, the mutual information between measurements and events is given by

$$I_{s,y}(x) := \int_{y \in \mathcal{Y}} \int_{s \in \mathcal{S}} \mathbb{P}_{s,y}(x) \log \frac{\mathbb{P}_{s,y}(x)}{\mathbb{P}_s(x)\mathbb{P}_y(x)}\, ds\, dy.$$

We will show that the gradient of the mutual information with respect to the robot positions, $\partial I_{s,y}(x)/\partial x$, can be analytically calculated and has a simple form. Despite its simplicity, this analytic expression for the gradient of mutual information is not yet widely used in the robotics community. Furthermore, we can compute the mutual information gradient at time t using $\phi_t(s)$ and $\psi_t(h)$ from the Bayesian filter equations above.

The gradient of mutual information allows us to use an information seeking controller of the form

$$x_i(t+1) = x(t) + k \frac{\frac{\partial I_{s,y}(x)}{\partial x_i}}{\left\| \frac{\partial I_{s,y}(x)}{\partial x_i} \right\| + \epsilon},$$

where $k > 0$ is a maximum step size, and $\epsilon > 0$ is a small factor to prevent singularities when a local minimum of mutual information is reached. Note that although this looks like a gradient controller, the mutual information, $I_{s,y}$, changes with each new measurement, so it is a stochastic controller with nonlinear dependence on the robots positions. For this reason a formal stability and convergence analysis would be difficult. However, intuitively the controller moves the robots in the direction of the highest immediate expected information gain.

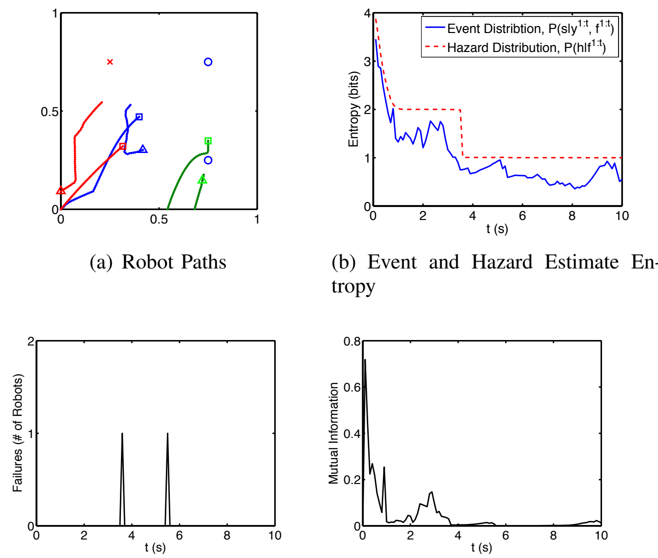
Empirically, the controller drives the robots to uncertain areas while veering away from suspected hazard sites, learned through the failures of previous robots. The robots eventually come to a stop when they estimate the events with high confidence (i.e. when the entropy of $\phi_t(s)$ approaches zero). While hazard avoidance does not explicitly appear in the control law, it is implicit as a robot failure will decrease the amount of information gained at the next time step.

III. Conclusions

In this work we propose a multi-robot control policy that utilizes measurements about events of interest to locally increase the mutual information, while also using the history of robot failures to avoid hazardous areas. We use an analytical expression for the gradient of mutual information, which provides a principled approach to active mapping by calculating robot trajectories that lead to the greatest immediate information gain. The event state and hazard fields over the environment are estimated using recursive Bayesian filters. The main drawback of the approach is high computational complexity. We are currently working to develop well-reasoned approximations to speed up computation. We are also investigating methods to decentralize the algorithm to run over a multi-robot network.

Acknowledgements

This work was funded in part by ONR MURI Grants N00014-07-1-0829, N00014-09-1-1051, and N00014-09-1-1031. We are grateful for this financial support.



(a) Robot Paths

(b) Event and Hazard Estimate Entropy

(c) Failure Rate of Robots

(d) Mutual Information

Fig. 2. This figure shows a representative simulation of the control algorithm for three robots in an environment with four possible event and hazard locations. Frame 2(a) shows the paths of the robots, where the triangles mark initial positions, the squares mark final positions, and a discontinuity in the path indicates a failure. The '$\times$' shows the location of a hazard and the 'o's give the locations of events. Frame 2(b) shows the entropy of the event (solid) and hazard (dotted) estimates. Both entropies decrease as the robots learn the event and hazard locations. The number of failures per time and the mutual information verses time are shown in frames 2(c) and 2(d), respectively.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, 1948.

[2] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley and Sons, 2006.

[3] B. Grocholsky, "Information-theoretic control of multiple sensor platforms," Ph.D. dissertation, University of Sydney, 2002.

[4] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, "Information based adaptive robotic exploration," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS 02)*, 2002, pp. 540–545.

[5] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 32–47, January 2010.

[6] B. J. Julian, M. Angermann, M. Schwager, and D. Rus, "A scalable information theoretic approach to distributed robot coordination," in *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2011, Sumbitted.

[7] A. Krause and C. Guestrin, "Near-optimal observation selection using submodular functions," in *Proceedings of 22nd Conference on Artificial Intelligence (AAAI)*, Vancouver, Canada, July 2007.

[8] H.-L. Choi, J. P. How, and P. I. Barton, "An outer-approximation algorithm for generalized maximum entropy sampling," in *Proceedings of the American Control Conference*, Seattle, WA, USA, June 2008, pp. 1818–1823.

[9] H.-L. Choi and J. P. How, "Continuous trajectory planning of mobile sensors for informative forecasting," *Automatica*, 2011, Accepted.

[10] M. P. Vitus and C. J. Tomlin, "Sensor placement for improved robotic navigation," in *In the Proceedings of Robotics: Science and Systems*, Zaragoza, Spain, June 2010.

[11] M. Schwager, P. Dames, D. Rus, and V. Kumar, "A multi-robot control policy for information gathering in the presence of unknown hazards," in *Proceedings of the International Symposium on Robotics Research (ISRR)*, Flagstaff, AZ, USA, August 2011, submitted.