Information Based Adaptive Robotic Exploration

Frédéric Bourgault, Alexei A. Makarenko, Stefan B. Williams,
Ben Grocholsky, Hugh F. Durrant-Whyte

Australian Centre for Field Robotics
The University of Sydney, Sydney NSW 2006, Australia
{f.bourgault, a.makarenko, stefanw, ben, hugh}@acfr.usyd.edu.au

Abstract

Exploration involving mapping and concurrent localization in an unknown environment is a pervasive task in mobile robotics. In general, the accuracy of the mapping process depends directly on the accuracy of the localization process. This paper address the problem of maximizing the accuracy of the map building process during exploration by adaptively selecting control actions that maximize localisation accuracy. The map building and exploration task is modeled using an Occupancy Grid (OG) with concurrent localisation performed using a feature-based Simultaneous Localisation And Mapping (SLAM) algorithm . Adaptive sensing aims at maximizing the map information by simultaneously maximizing the expected Shannon information gain (Mutual Information) on the OG map and minimizing the uncertainty of the vehicle pose and map feature uncertainty in the SLAM process. The resulting map building system is demonstrated in an indoor environment using data from a laser scanner mounted on a mobile platform.

1 Introduction

The main objective of exploration is to create an accurate map of an unknown area. To perform this mapping task a robot must localize itself accurately within the environment as the fidelity of the resulting map depends directly on the accuracy of the localisation process. Exploration must therefore satisfy the dual, and sometimes conflicting objectives, of localisation accuracy, for high fidelity mapping, and exploring unknown territory.

The problem of adaptive exploration and mapping of an unknown environment can be cast as an information, or an information gain, maximization problem. Shannon and Fisher information measures provide a natural way of quantifying localisation accuracy and map fidelity. Formal information measures provide a natural mean for describing uncertainty and for evaluating the effectiveness of future control actions in reducing this uncertainty [9]. Practically, information methods also provide a natural way of mixing measures of continuous information gain (in localisation for example) with discrete information gain (through area exploration for example). The adaptive exploration policies presented in this paper are based on the use of information as a measure of utility for taking exploration control actions. Such techniques are also referred to as "Information-Theoretic" in the literature (see [9, 7]).

The exploration framework proposed in this paper uses an Occupancy Grid (OG) environment model [4]. This provides a simple straight-forward spatial representation of the environment within which sensor readings are easily fused. The representation is particularly helpful in indoor exploration problems. However, the OG method requires an accurate estimate of the location of the robot from which data is being acquired. Failure to have an accurate location estimate results in new sensor data, taken relative to the robot, being used to update the wrong OG cells. In turn, this results in a low fidelity map.

The localisation framework employed in this paper uses a feature-based Simultaneous Localisation And Mapping (SLAM) algorithm [2, 8]. The SLAM method is most appropriate for exploring unknown environments in the absence of an absolute position sensor such as GPS. In its simplest form, the SLAM algorithm is an Extended Kalman Filter (EKF) that generates estimates of vehicle location and estimates of the location of landmark features in the environment.

While SLAM can provide accurate localisation information, the maps built usually consist of only a sparse set of easily seen features. This contrasts with the spatially dense and rich maps composed in the OG method and argues that *both* SLAM features and OG representations should be used as part of an exploration strategy.

Information metrics have previously been proposed for OG techniques [5], for SLAM methods [6], and for hybrid sensor control tasks [9]. The contribution of this paper is to fuse these ideas in the competing requirements of adaptive exploration. Section 2 provides some background to both the SLAM and OG methods. Section 3 develops information metrics for each representation and shows how these are combined in a single measure of exploration performance. Section 5 shows how these metrics are then used to determine the future motion or trajectory of the sensor platform. Section 6 describes the implementation and evaluation of these techniques on an indoor vehicle equipped with a laser ranging system.

2 Background

This section briefly describes the two main algorithms employed in this work: the SLAM algorithm for localisation and the OG method for map building.

2.1 The SLAM Algorithm

The SLAM algorithm employs an EKF to generate and maintain a combined estimate of the vehicle and map feature locations together with a measure of uncertainty in these estimates [2, 10, 8]. The SLAM algorithm requires a vehicle model, $\mathbf{f}$, and an observation (sensor) model, $\mathbf{h}$.

Vehicle and Landmark Models. The augmented state vector at a time t_k contains both the state of the vehicle, $\mathbf{x}_v(k)$, and the state of all the landmarks, $\mathbf{x}_m$,

$$\mathbf{x}(k) = [\mathbf{x}_v^T(k), \mathbf{x}_1^T, \cdots, \mathbf{x}_{n_f}^T]^T = [\mathbf{x}_v^T(k), \mathbf{x}_m^T]^T.$$

The state transition model for the vehicle and landmarks is described by

$$\mathbf{x}_v(k+1) = \mathbf{f}(\mathbf{x}_v(k), \mathbf{u}(k)) + \mathbf{v}(k) \quad (1)$$

$$\mathbf{x}_m(k+1) = \mathbf{x}_m(k) \quad (2)$$

where $\mathbf{v}(k)$ is taken to be a zero mean uncorrelated random variable with covariance $\mathbf{Q}(k)$ describing platform motion uncertainty.

Sensor Model. The sensor observation is described by a model in the form

$$\mathbf{z}(k) = \mathbf{h}(\mathbf{x}_v(k), \mathbf{x}_m) + \mathbf{w}(k) \quad (3)$$

where $\mathbf{w}(k)$ is taken to be a zero mean uncorrelated random variable with covariance $\mathbf{R}(k)$ describing observation uncertainty.

Estimation Process. The estimation process is a recursive three step algorithm consisting of prediction, observation and update steps.

Prediction: An estimate $\hat{\mathbf{x}}(k \mid k)$ of the state $\mathbf{x}(k)$ at time t_k is assumed. A prediction $\hat{\mathbf{x}}(k+1 \mid k)$ is formed by taking expectations of the state transition model conditioned on the observation sequence up to time t_k as follows

$$\begin{bmatrix}\hat{\mathbf{x}}_v(k+1 \mid k)\\ \hat{\mathbf{x}}_m(k+1 \mid k)\end{bmatrix} = \begin{bmatrix}\mathbf{f}(\hat{\mathbf{x}}_v(k \mid k), \mathbf{u}(k))\\ \hat{\mathbf{x}}_m(k \mid k)\end{bmatrix} \quad (4)$$

A covariance $\mathbf{P}(k \mid k)$ in the estimate $\hat{\mathbf{x}}(k \mid k)$ is assumed and the prediction covariance is then computed as

$$\mathbf{P}(k+1 \mid k) = \mathbf{F}\mathbf{P}(k \mid k)\mathbf{F}^T + \mathbf{Q}(k) \quad (5)$$

where $\mathbf{F} = \nabla_v \mathbf{f}(k)$ is the Jacobian of $\mathbf{f}$ with respect to the vehicle states evaluated at $\hat{\mathbf{x}}(k \mid k)$.

Observation: An observation is made according to Equation 3. A predicted observation is computed by taking expected values Equation 3 as

$$\hat{\mathbf{z}}(k+1 \mid k) = \mathbf{h}(\hat{\mathbf{x}}_v(k+1 \mid k), \hat{\mathbf{x}}_m(k+1 \mid k)). \quad (6)$$

The difference between the actual and the predicted observation, (the innovation) is then computed from

$$\nu(k+1) = \mathbf{z}(k+1) - \hat{\mathbf{z}}(k+1 \mid k) \quad (7)$$

along with the innovation covariance

$$\mathbf{S}(k+1) = \mathbf{H}\mathbf{P}(k+1 \mid k)\mathbf{H}^T + \mathbf{R}(k+1) \quad (8)$$

where $\mathbf{H} = \nabla_x \mathbf{h}(k+1)$ is the Jacobian of the observation function with respect to *all the states*.

Update: The state estimate and the corresponding state estimate covariance matrix are now updated in the normal manner:

$$\hat{\mathbf{x}}(k+1 \mid k+1) = \hat{\mathbf{x}}(k+1 \mid k) + \mathbf{W}\nu(k+1) \quad (9)$$

$$\mathbf{P}(k+1 \mid k+1) = \mathbf{P}(k+1 \mid k) - \mathbf{W}\mathbf{S}(k+1)\mathbf{W}^T \quad (10)$$

where the optimal Kalman gain is

$$\mathbf{W} = \mathbf{P}(k+1 \mid k)\mathbf{H}^T\mathbf{S}^{-1}(k+1) \quad (11)$$

2.2 Occupancy Grid

An OG is a discretised random field in which the probability of occupancy of each independent cell in a spatial lattice is maintain [4]. The state x_i of the i^{th} cell in the grid may be occupied or unoccupied

$$x_i = \{OCC, EMP\}, \qquad \forall i = 1, \cdots, N.$$

The OG method aims to compute a probability of occupancy $P_i(x_i|Z^k)$, for each cell, given sequence of observations $Z^k = \{z_1, ..., z_k\}$. Given a sensor model $p(z_k|x_i)$, this probability may be computed recursively from [4, 5]

$$P_i(x_i|Z^k) = \frac{p(z_k|x_i)P_i(x_i|Z^{k-1})}{\sum_{x_i \in X} p(z_k|x_i)P_i(x_i|Z^{k-1})} \quad (12)$$

The sensor model for this problem $p(z_k|x_i)$ can be quite complex as it must be defined over all all possible grid configurations (see [3, 4, 5] for derivations). The grid is initialised be setting all cells to having a probability of 0.5 of being occupied and being unoccupied. This initial uniform distribution for all states of each cell corresponds to the least informative, or maximum entropy distribution.

3 Information Metrics

3.1 Info Metrics for Localizing (SLAM)

A multivariate Gaussian distribution has entropy proportional to the logarithm of the determinant of its covariance matrix $\mathbf{P}$ ([1]):

$$\frac{1}{2}ln[(2\pi e)^n \det \mathbf{P}]. \tag{13}$$

Since the determinant of a matrix is a measure of its volume (product of its eigenvalues), the entropy measures the compactness of the corresponding Gaussian distribution. Maximizing information about a state estimate is equivalent to minimizing the product of the eigenvalues of the corresponding covariance matrix.

A suitable information-based cost function $C(\cdot)$ for evaluating the utility of a candidate localisation control action proposed in ([6]) is given by

$$\begin{aligned} C(\mathbf{P}) &= \pi \prod_j \sqrt{\lambda_j(\mathbf{P}_{vv})} + \pi \sum_{i=1}^{n_f} \prod_j \sqrt{\lambda_j(\mathbf{P}ii)} \\ &= \pi \sqrt{\det(\mathbf{P}_{vv})} + \pi \sum_{i=1}^{n_f} \sqrt{\det(\mathbf{P}_{ii})} \end{aligned} \tag{14}$$

where $\lambda_j(.)$ is the j^{th} eigenvalue of its argument. The cost $C(\mathbf{P})$ represents the sum of the areas of the features and vehicle uncertainty ellipses for the predicted covariance matrix $\mathbf{P}$ after the expected observation from the predicted state.

3.2 Info Metrics for Mapping (OG)

For each cell c_i of the map, it is possible to compute the entropy (Shannon information) [1] as a metric of the knowledge acquired so far. The *a priori* entropy at time t_k for the i^{th} grid cell is defined as the expectation of the logarithm of the probability distribution $P_i(x_i)$

$$H_{k,i} \equiv -E[\ln P_i(x_i)] = -\sum_{x_i \in X_i} P_i(x_i) \ln P_i(x_i) \tag{15}$$

Since there are two possible states (occupied and empty), with $p = P_i(OCC)$ the probability of occupancy, and $q = P_i(EMP) = 1 - p$ its complement, we can write

$$H_{k,i}[X_i] = -p \ln p - q \ln q \tag{16}$$

Given an observation z_k at time t_k, the conditional entropy [1] for cell i is defined as

$$\begin{aligned} H_i(z_k) &\equiv -E[\ln P_i(x_i|z_k)] \\ &= -\sum_{x_i \in X_i} P_i(x_i|z_k) \ln P_i(x_i|z_k). \end{aligned}$$

The posterior probability $P_i(x_i|z_k)$ can be obtained from Bayes rule as

$$P_i(x_i|z_k) = \frac{P_i(z_k|x_i)P_i(x_i)}{P_i(z_k)}. \tag{17}$$

Note that the conditional entropy is a function of the observation. The *mean* conditional entropy (over all possible observations) is an *a priori* measure of the value of an observation. It is the expectation of entropy left after observation and is given by [1]

$$\overline{H}_i \equiv E[H_i(z_k)] = \int H_i(z_k)P_i(z_k)dz_k \tag{18}$$

In practice, this integral is evaluated numerically over the sensor range. Also, since the observations to be made z_k depend directly on the vehicle pose from which the observations are made, the mean conditional entropy becomes a function of the estimated vehicle pose prior to observation.

Using Equations 15) and (18), it is possible to compute the *expected mutual information gain* following an observation ([1]):

$$\hat{I}_i(x_i) \equiv -E\left[\ln \frac{P_i(x_i|z_k)}{P_i(x_i)}\right] = H_i - \overline{H}_i(x_i|z_k) \tag{19}$$

The actual information gain estimate can only be evaluated after the control action and the observation have been made.

$$I_i = H_{k-1,i}(x_i) - H_{k,i}(x_i) \tag{20}$$

The total mutual information (total expected information gain) for a scan from a scanning sensor can be defined as

$$\hat{I}_{S_j}(x_i|z_k) = \sum_{i \in S_j} I_i(x_i|z_k) \tag{21}$$

where S_j is the region covered by the scan, and N_{S_j} is the number of grid cells included in S_j. If S_j is taken as the scanning region associated with predicted vehicle pose $\hat{\mathbf{x}}_v(k+1 \mid k, j) = \mathbf{f}(\hat{\mathbf{x}}_v(k \mid k), \mathbf{u}_j(k)$ for a given control action $\mathbf{u}_j(k) \in \mathbf{U}(k)$, where $\mathbf{U}(k) = \{\mathbf{u}_1(k), \cdots, \mathbf{u}_n(k)\}$ is the discretized set of possible control action at time t_k, then $\hat{I}_{S_j}$ can be used directly as a utility or value measure.

$$U(\mathbf{x}, \mathbf{u}_j(k)) = \hat{I}_{S_j} \tag{22}$$

3.3 Combined Information Utilities: Integration

As described in [9], for any multi-objective optimization problem, for which the individual utilities are based on information metrics, a composite utility can be constructed simply by linear combination. Hence, for the dual objective of mapping (OG) and localizing (SLAM) the composite utility can be defined as

$$\begin{aligned} U_k &= I_{composite}(\mathbf{x}, \mathbf{x}_c, \mathbf{u}_j(k)) \\ &= w_1 I_{SLAM}(\mathbf{x}, \mathbf{u}_j(k)) + w_2 I_{OG}(\mathbf{x}_c, \mathbf{u}_j(k)) \end{aligned} \quad (23)$$

where the weights $w_1(k) = \alpha / I_{SLAM_{MAX}}(k)$ and $w_2 = (1-\alpha)/I_{OG_{MAX}}$ are the respective normalization factors multiplied by a constant. $I_{SLAM_{MAX}}(k)$ is obtained from the upper bound on the SLAM covariance matrix for a given number of landmarks (see [2]), and $I_{OG_{MAX}}$ is simply the total information of a perfectly known OG map (probability of 0 or 1 for every cell). The default value of the tuning parameter α is 0.5, hence making the upper bound of the combined utility to sum to 1. Increasing α would create a more accurate OG map by putting the emphasis on the localization, and reducing it would make the robot explore more deliberately to the detriment of accuracy.

4 Optimal Trajectory

4.1 Local Optimization

The optimal control action $\mathbf{u}^*(k)$ at time t_k, for a one step look-ahead is easily selected as the action which maximizes the total expected utility at time t_k

$$\mathbf{u}^*(k) = \arg\max_{\mathbf{u}_j} U_k(\mathbf{x}, \mathbf{x}_c, \mathbf{u}_j(k)) \quad (24)$$

Viewing the information space as a potential field, the optimal control action at any given time t_k is the action that constantly directs the vehicle towards the maximum increase in global information. The control strategy is to "surf" towards the highest rate of mutual information. Notice though that unlike control actions driven by potential fields, information surfing strategies do not get trapped in local minima as the mutual information gain decreases exponentially as observations are made. Thus the local optimum in the field reduces in value as observations are made.

While this control strategy is not necessarily globally optimal, it exhibits an interesting reactive property by constantly adapting the control decisions to the current information state. In an exploration scenario, the utility for mapping I_{OG} will attract the robot towards unknown areas (areas of high entropy, see Section 5), while the utility for localization I_{SLAM} competes to ensures that the robot stays well localized relative to known features in the map. The relative importance of localization will change with localization uncertainty. The robot will adapt by taking control action that will increase or decrease sensitivity to localization uncertainty. This is a useful result as there is little value in a robot exploring and mapping new areas when it has no idea of how accurately it knows it's own location.

4.2 Global Optimization

The generalization of the integrated utility function over a finite horizon of length $\Delta t = Nk$ can be defined in the form

$$U_N(\pi_j) = \sum_{k=1}^{N} U_k(\mathbf{x}, \mathbf{x}_c, \mathbf{u}_j(k)) \quad (25)$$

where the $\pi_j = \{\mathbf{u}_j(1), \cdots, \mathbf{u}_j(N)\}$ is a policy vector of N control decisions. The optimal control strategy is given by

$$\pi^* = \{\mathbf{u}_*(1), \cdots, \mathbf{u}_*(N)\} = \arg\max_{\pi_j} U_N(\pi_j) \quad (26)$$

Multi-step solutions are appropriate in problems such as path planning. However, the computational cost will grow very rapidly with the depth of look-ahead unless the information metrics are very simple.

5 Implementation

Figure 1 shows the information-based exploration algorithm implemented, while Figure 2 shows an illustration of what the algorithm does.

The control action space is first discretised, $\{\mathbf{u}_1(k), \mathbf{u}_2(k), \mathbf{u}_3(k)\}$. Then, following Figure 1, the vehicle model (Equations 4 and 5) is used to predict the states and covariance matrix for each of these control inputs (a). From each projected state, the OG Mutual Information in (b), (Equations 19 and 21), and the expected covariance matrix in (c), (Equation 14) are evaluated, taking into account the uncertainty in the predicted state. Then, in (d), the composite utility is computed (Equation 23), and in (e) the "optimal" control action (Equation 24) is performed. In (f), the observation is made (Equation 3), and the Kalman filter updates the state estimate and covariance matrix (Equations 9 and 10). Finally in (g), the Occupancy Grid is updated (Equation 12).

6 Experiment and Results

The mobile sensing platform (see Figure 3) consists of a Pioneer2 robot equipped with a laser ranging device. Figure 4a shows the Occupancy Grid and Figure 4b its corresponding entropy map. The robot position is marked by the small circle. On Figure 5 is shown a more intuitive 3-D image of the same entropy map. The low parts corresponding to the very well known region (maximum information). The mapping utility will attract the robot towards region of high entropy

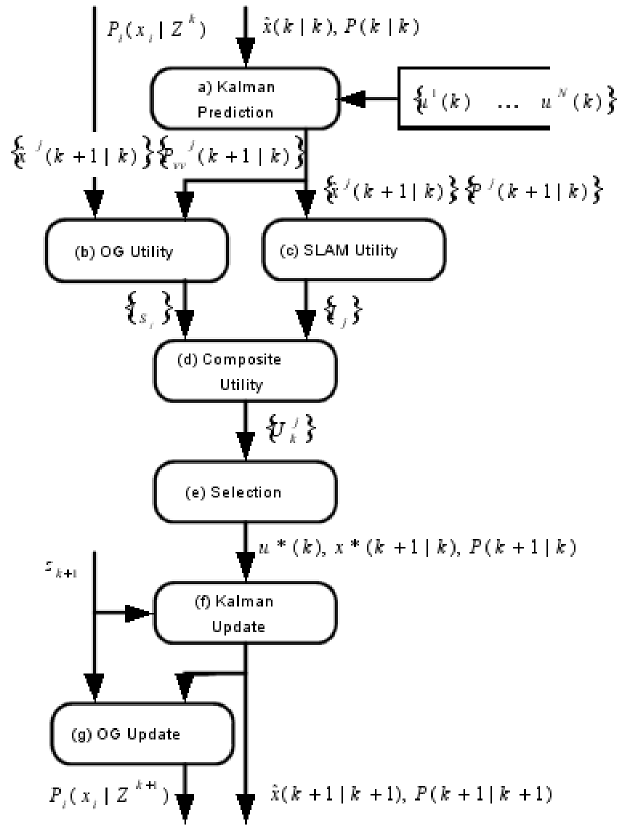


Figure 1: *Integrated Adaptive Information-based Exploration Algorithm.*

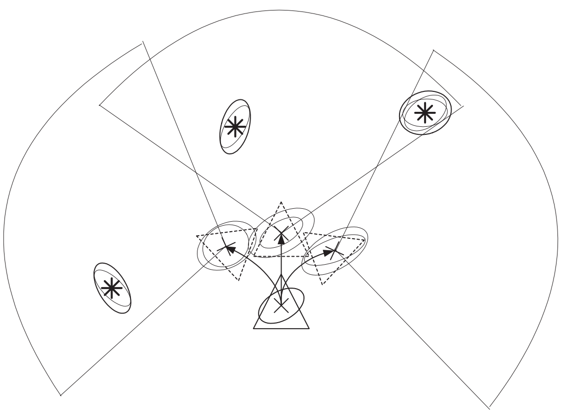
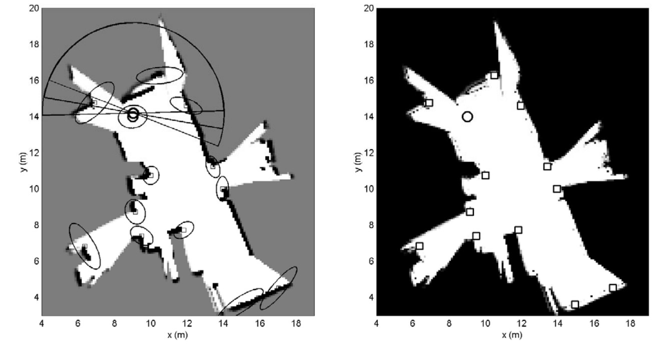


Figure 2: *Projected vehicle state estimates, $\hat{\mathbf{x}}_{v_j}(k \mid k)$, for three discrete control inputs, $\{\mathbf{u}_1(k), \mathbf{u}_2(k), \mathbf{u}_3(k)\}$ and their corresponding sensor coverage. The * are the SLAM feature estimates with their actual and predicted uncertainty ellipses.*

(unexplored) that are actually in the line of sight of the sensor. Figure 6 shows the Mutual Information gain for three potential control actions. In this case, the third action (c) is the most appealing from the mapping point of view (highest mutual information) while action (a) is the most appealing from the localization point of view as it is the one that is expected to shrink the most the uncertainty of the landmark and vehicle estimates (ellipses).

Figure 3: *The Pioneer2 mobile robot equipped with a laser.*

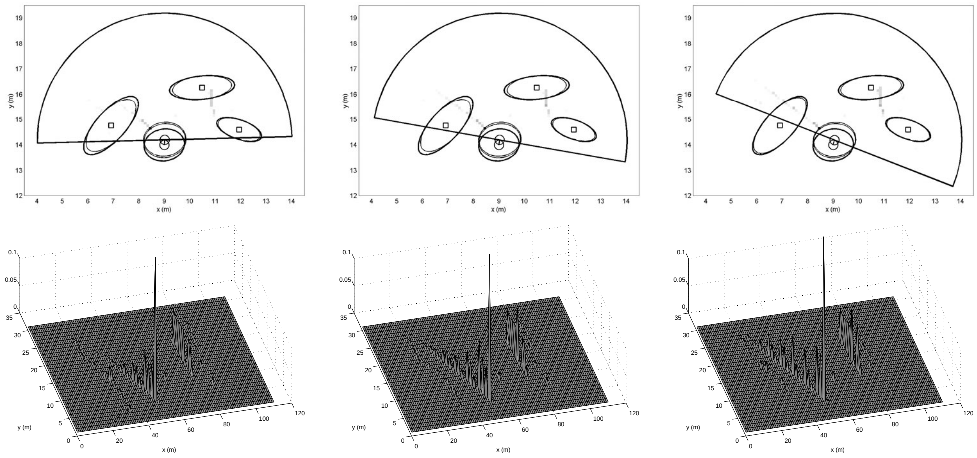


(a) (b)

Figure 4: *Occupancy Grid (a), and corresponding entropy map (b). The square symbols represents the landmarks, and the circle represent the vehicle with a pose of about 80 degrees. The ellipses are the actual 3σ uncertainty bounds of the landmarks and the vehicle amplified five times for better visualization.*

7 Conclusion

This paper has shown how to combined OG and SLAM information metrics in exploration of an unknown environment. Good preliminary results indicate the validity of the approach. While this paper is concerned with the control and trajectory optimiza-



(a) (b) (c)

Figure 6: *Mutual Information for three discrete control inputs* $\{\mathbf{u}^1(k), \mathbf{u}^2(k), \mathbf{u}^3(k)\}$. *The small circles on the top figures represent the vehicle initial position and projected state estimate* $\hat{\mathbf{x}}_{vj}[k|k]$ *with their respective uncertainty. The semicircles are the corresponding sensor coverage (range = 5 m), and the squares are the SLAM features. Again the uncertainty bounds have been amplified five times for visualization. The lower figure are the three-dimensional representation of the same mutual information as the grey traces in the upper figures.*

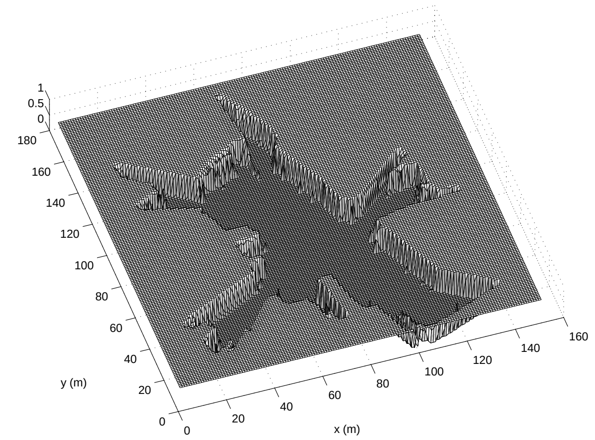


Figure 5: *3-D representation of the entropy map.*

tion problem of a single mobile platform for exploration and mapping tasks, it constitutes a building block of the much broader problem of management in Decentralized systems. The various solutions to information maximization problems presented here, while being suboptimal, give practical and quantifiable algorithms with the potential to scale to multi-robot systems.

References

[1] T.M. Cover and J.A. Thomas. *Elements of information theory*. Wiley series in telecommunications. Wiley, New York, 1991.

[2] M.W.M.G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte, and M. Csobra. A solution to the simultaneous localization and map building (slam) problem. *Robotics and Automation*, 17(3):229–241, 2001.

[3] A. Elfes. *Occupancy Grids : A Probabilistic Framework for Robot Perception and Navigation*. Phd, Carnegie Mellon University, 1989.

[4] A. Elfes. Occupancy grids: A stochastic spatial representation for active robot perception. In *the Sixth Conference on Uncertainty in AI*, pages 60–70, 1990.

[5] A. Elfes. Robot navigation: Integrating perception, environmental constraints and task execution within a probabilistic framework. In *Reasoning with Uncertainty in Robotics. International Workshop (RUR '95)*, pages 93–129, Amsterdam, Netherlands, 1995.

[6] H.J.S. Feder, J.J. Leonard, and C.M. Smith. Adaptive mobile robot navigation and mapping. *Int. Journal of Robotics Research*, 18(7):650–668, 1999.

[7] B. Grocholsky. *Information-Theoretic Control of Multiple Sensor Platforms*. Phd, The University of Sydney, 2002.

[8] J.J. Leonard and H.F. Durrant-Whyte. Simultaneous map building and localisation for an autonomous robot. In *IEEE/RSJ International Workshop on Intelligent Robots and Systems (IROS '91)*, pages 1442–1447, Osaka, Japan, 1991.

[9] J. Manyika and H.F. Durrant-Whyte. *Data Fusion and Sensor Management: a decentralized information-theoretic approach*. Ellis Horwood Series in Electrical and Electronic Engineering. Ellis Horwood, 1994.

[10] S.B. Williams. *Efficient Solutions to Autonomous Mapping and Navigation Problems*. Phd, The University of Sydney, 2001.