Efficient Computation of Dynamic Probabilistic Maps*

João P. Hespanha* hespanha@ece.ucsb.edu

Hüseyin H. Kızılocak† kiziloca@usc.edu

* *Dept. of Electrical and Computer Engineering, University of California, Santa Barbara*
† *Dept. of Electrical Engineering, University of Southern California, Los Angeles*

Keywords: Probabilistic maps, State estimation, Kalman filter, Markov chains.

Abstract

We develop efficient algorithms to compute probabilistic maps for a large number of mobile objects. By a probabilistic map we mean the probability density of the objects' positions, conditioned to noisy sensor measurements.

The main result of the paper is that, under suitable assumptions, the joint probability distribution of n objects that lie in a region partitioned into N cells can be approximately determined from an *aggregate measurement function* that can be represented with memory complexity $o(N)$, *regardless of the number of objects* n. This is far more compact than an extensive representation of the joint distribution, whose memory complexity is $o(N^n)$.

In general, the approximation error introduced by this compact representation can be made arbitrarily small. Three main assumptions make this valid: the objects are indistinguishable from the sensor viewpoint, there is a minimum distance between any two of them, and their motion is essentially independent when they are far from each other.

1 Introduction

This paper addresses the navigation of autonomous vehicles that attempt to execute a mission under partial information. A typical example is that of a vehicle that is asked to move to a given destination while avoiding objects (or other vehicles) whose positions are only partially known. In another possible example, a vehicle is asked to search for objects whose positions are not accurately known. We consider situations in which noisy sensors are used to gather information about the localization of the objects to avoid or search for. This naturally leads to a probabilistic framework, in which sensors produce erroneous measurements with given probabilities.

A key concept used here is that of a "probabilistic map" [8, 5]. In usual maps, one finds markings that represent locations of objects within a region. By placing one of these markings in a location on the map, one is expressing the fact that the object represented can be found at that particular location with probability one. However, when the information used to build a map (or more precisely to locate interesting objects in it) is not perfect, it is not possible to mark a specific position on the map as the location of the object. In this case, the best one can do is to superimpose the probability density of the location of the object in question on the whole map. It is this probability density that we call the probabilistic map for the object. From a formal point of view, probabilistic maps are actually probability densities of object positions, conditioned to the available sensor measurements.

The computation of probabilistic maps can be viewed as a Kalman filtering problem, in which one propagates in time a probability density, with occasional Bayesian sensor fusion of noisy measurements. It is well known that if the *a priori* probability density of the objects is jointly Gaussian, their motion can be expressed by a linear model with Gaussian disturbances, and the measurements are obtained by adding Gaussian noise to the true object position, then the probabilistic map is Gaussian and can be computed using a finite dimensional Kalman filter.

*This material is based upon work supported by the Space and Naval Warfare Systems Center, San Diego under Contract Number N66001-01-C-8076 and by DARPA. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

However, here we are faced with probabilistic maps that are not Gaussian. Several factors contribute to this:

1. The *a priori* probability densities for the objects positions are not Gaussian. These densities may be multi-modal and typically have finite support because the objects are known to be in a prespecified region.

2. The models for the object's motion are not necessarily linear with Gaussian disturbances.

3. We want to consider non-Gaussian sensors, e.g., with limited range, which occasionally produce measurements that do not correspond to any object ("false positives"), which occasionally fail to produce measurements regarding objects that are within their range ("false negatives"); etc. In practice, these situations occur with any real sensor.

In general, without the Linear/Gaussian assumptions mentioned above, there is no finite-dimensional representation of the *a posteriori* probability densities. A finite discretization of the state-space overcomes this difficulty but one is faced with an exponential explosion on the number of states as the number of objects being mapped grows. For example, if one partitions the region where n objects lie into N cells, the memory complexity associated with representing the joint probability distribution of the objects is $o(N^n)$.

The main contribution of this paper is an algorithm to compute probabilistic maps that avoids the exponential explosion on the number of objects being mapped. We show that under suitable assumptions, the joint probability distribution of n objects that lie in a region partitioned into N cells can be approximately computed from a function defined on a set with only N elements. The memory complexity associated with representing this function, which we call the *aggregate measurement function*, is simply $o(N)$. In general, the approximation error introduced by this compact representation can be made arbitrarily small. Three main assumptions make this valid: the objects are indistinguishable from the sensor viewpoint, there is a minimum distance between any two of them, and their motion is essentially independent when they are far from each other.

Aggregate measurement functions were introduced in [5] for the probabilistic mapping of static aircraft-tracking radars using radiation detection sensors. Here we extend their use to more general types of sensors and objects being mapped. We also consider probabilistic mapping of mobile objects.

Probabilistic maps are becoming an important topic in the robotics literature. Especially for indoors robotics, map building is often combined with the problem of localizing the robot within a global coordinate system [3, 6, 7, 8, 2, 1]. This is required since the measurements used to build the map are a function of the position of the robot, which is hard to determine accurately due to accumulated odometry errors. In outdoors robotics, the localization problem can be mitigated through the use of GPS. Occupancy grid mapping is currently the most widely used class of mapping algorithms for mobile robots (cf. , e.g., [7, 9]). Occupancy maps provide the probability of each individual cell in a grid being occupied by an object. However, this information is not sufficient to compute the probability of a particular configuration of objects.

2 Probabilistic maps

Suppose that there exist n objects in a region $\mathcal{R}$ consisting of a finite collection of cells. We model the position of each object as a stochastic process[1] taking values in $\mathcal{R}$. We denote by $\mathbf{r}(t) := \{\mathbf{r}_1(t), \mathbf{r}_2(t), \ldots, \mathbf{r}_n(t)\} \in \mathcal{R}^n,\ t \geq t_0$, the positions of the objects at time t, where $\mathbf{r}_i(t) \in \mathcal{R}$ denotes the position of the ith object. We assume given an *a priori* joint probability distribution for the objects at time t_0

$$p_0(r, t_0) := \mathrm{P}(\mathbf{r}(t_0) = r), \qquad r \in \mathcal{R}^n$$

and that the motion for $t \geq t_0$ is Markovian. The motion model is specified by a known *n-object transition probability function* of the form

$$\Phi(r', r; t + dt, t) := \mathrm{P}(\mathbf{r}(t + dt) = r' \mid \mathbf{r}(t) = r)$$

$r, r' \in \mathcal{R}^n,\ dt \geq 0,\ t \geq t_0$, with the standing assumption that $\mathbf{r}(t+dt)$ is conditionally independent of any random variable at time t or before, given $\mathbf{r}(t)$.

Throughout this paper we assume that the distance between any two objects is never smaller than a minimum distance $d_{\min} > 0$, where "distance" refers to a given metric $d : \mathcal{R} \times \mathcal{R} \to [0, \infty)$ on $\mathcal{R}$. Denoting by $\delta(r, d_{\min})$ a function that is equal to zero whenever $r := \{r_1, \ldots, r_n\} \in \mathcal{R}^n$ has two $r_i,\ r_j,\ i \neq j$

[1] We denote random processes by **bold** symbols and their realizations by non-bold ones.

with $d(r_i, r_j) < d_{\min}$ and equal to one otherwise, this assumption can be formalized as follows:

Assumption 1. The *a priori* joint probability distribution p_0 and the n-object transition probability function Φ are such that $\delta(\mathbf{r}(t), d_{\min}) > 0$ with probability one for every $t \geq t_0$.

Our knowledge about the position of the objects improves with time through the measurements collected by the sensors. These measurements are denoted by $\mathbf{y}_k$, $k \in \{1, 2, \dots\}$, and each $\mathbf{y}_k$ consists of a triple of the form

$$\{\mathbf{t}_k, \mathbf{a}_k, \mathbf{s}_k\},$$

where $\mathbf{t}_k \geq t_0$ denotes the time at which the measurement was made and $\mathbf{a}_k \in \mathcal{R}$ the sensor position. The variable $\mathbf{s}_k$ has the value $\mathbf{s}_k = \emptyset$ when nothing is detected or some value $\mathbf{s}_k \in \mathcal{R}$ when an object is detected at position $\mathbf{s}_k$. However, we assume that the sensors are not perfect and therefore $\mathbf{s}_k = \emptyset$ simply means that they were not able to detect an object but one must take into account that this may be a "false negative." Similarly, $\mathbf{s}_k \in \mathcal{R}$ may be caused by a "false positive," or there may be a object close to $\mathbf{s}_k$ but not exactly at that location. The set of all possible measurement triples is denoted by $\mathcal{Y}$.

Sensors of this type can be modeled by a likelihood function of the form

$$\ell_n(y, r) := \mathrm{P}(\mathbf{y}_k = y \mid \mathbf{r}(t) = r), \tag{1}$$

$y := \{t, a, s\} \in \mathcal{Y}$, $r \in \mathcal{R}^n$, $k \geq 1$. We refer to ℓ_n as the *n-object sensor likelihood function*. We assume that each $\mathbf{y}_k$ is conditionally independent of any other measurements, given a particular configuration $\mathbf{r}(\mathbf{t}_k)$ for the objects. Given a list of k measurements

$$Y_k := \{y_1, y_2, \dots, y_k\},$$

$y_i := \{t_i, a_i, s_i\}$, $i \in \{1, \dots, k\}$, we define the *(a posteriori) probabilistic map at time $t \geq t_0$ given Y_k* by

$$m_r(r; t|Y_k) := \mathrm{P}(\mathbf{r}(t) = r \mid \mathbf{Y}_k = Y_k), \quad r \in \mathcal{R}^n.$$

For simplicity, we assume that the measurements in Y_k are ordered by time and that $t \geq t_k$. The goal of this paper is to determine efficient algorithms to compute probabilistic maps. Inspired by the recursive formulas for a Kalman filter, we will construct a *prediction operator* $T_{\mathrm{pred}}[t, \tau, \cdot\,]$ that takes a probabilistic map at time τ and propagates it forward in time to obtain a map at time $t > \tau$; and a *fusion operator* $T_{\mathrm{fuse}}[y, \cdot\,]$ that takes a new measurement y taken at time t and an existing map also at time t to produce a new map at t, given y together with all the measurements already considered in the old map. With these operators we can compute the probabilistic map $m_r(\cdot\,; t|Y_k)$, $t \geq t_k$ recursively as depicted in Figure 1. Formally, the recursion can be written as follows:

$$\begin{aligned}
m_r(\cdot\,; t_0|Y_0) &= p_0(\cdot\,, t_0), \\
m_r(\cdot\,; t_i|Y_{i-1}) &= T_{\mathrm{pred}}[t_i,\ t_{i-1},\ m_r(\cdot\,; t_{i-1}|Y_{i-1})], \\
m_r(\cdot\,; t_i|Y_i) &= T_{\mathrm{fuse}}[y_i,\ m_r(\cdot\,; t_i|Y_{i-1})], \\
m_r(\cdot\,; t|Y_k) &= T_{\mathrm{pred}}[t,\ t_k,\ m_r(\cdot\,; t_k|Y_k)],
\end{aligned}$$

$i \in \{1, 2, \dots, k\}$, where Y_0 denotes an empty list of measurements.

When the objects being mapped are static, $\mathbf{r}(t)$ is constant over t and therefore

$$m_r(r; t|Y_k) = m_r(r; \tau|Y_k), \qquad \forall t, \tau.$$

This means that the prediction operator does not alter the probabilistic map and is therefore not needed. In the following sections we address the prediction and fusion problems separately. We start with the fusion operator, which is sufficient for static objects. However, before that we need to specify the class of sensor likelihood functions under consideration.

3 Sensor likelihood functions

In practice, the precise shape of the sensor likelihood function should be estimated experimentally. We derive next a simple parametric model for it that can be "best-fitted" to experimental data. For simplicity of notation, in this section we consider a fixed time $t \geq t_0$ and drop the time-dependence of $\mathbf{r}$.

3.1 Single object

We start by considering the case of a single object, i.e., $n = 1$ and $\mathbf{r} = \mathbf{r}_1$. To this effect consider the parameters in Table 1. In general, these parameters may depend on the time at which the measurement was taken, the sensor position, and the object position but for simplicity we take them as constant. The generalization for non-constant parameters is straightforward. Also for simplicity, we assume that the distractors can appear anywhere in the sensor range with equal probability.

Pick an arbitrary object position $r \in \mathcal{R}$ and a nonzero probability realization $y := \{t, a, s\}$ for the

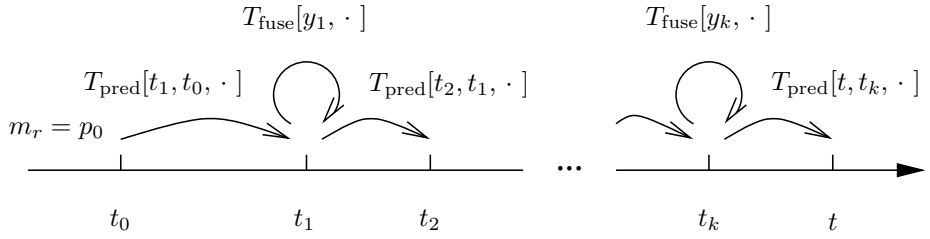

Figure 1: Recursive computation of $m_r(\,\cdot\,; t|Y_k)$, $t \geq t_k$ using the prediction and fusion operators.

$\rho_{\text{sensor}} > 0$	Range over which the sensor can detect objects.
p_1	Probability of not recognizing anything, given that there is an object in the range of the sensor.
p_2	Probability of recognizing a distractor instead of a real object, given that there is an object in the range of the sensor.
p_3	Probability of recognizing a distractor instead of a real object, given that there is no object in the range of the sensor.
e_{loc}	Maximum localization error for the position of the object when the sensor detects the presence of object. The actual error is assumed uniform within the range specified by e_{loc}. Without loss of generality, we assume that $e_{\text{loc}} \leq \rho_{\text{sensor}}$.

Table 1: Sensor parameters for the sensor likelihood function.

kth measurement $\mathbf{y}_k$. We consider first the case in which an object was detected, i.e., $s \in \mathcal{R}$. From the definition of ℓ_1,

$$\ell_1(y, r) = \mathrm{P}(\mathbf{s}_k = s \mid \mathbf{r}_1 = r, \mathbf{t}_k = t, \mathbf{a}_k = a) \\ \mathrm{P}(\mathbf{t}_k = t, \mathbf{a}_k = a \mid \mathbf{r}_1 = r). \quad (2)$$

We assume here that the time measurement $\mathbf{t}_k = t$ and the sensor position measurement $\mathbf{a}_k = a$ are error free and therefore $\mathrm{P}(\mathbf{t}_k = t, \mathbf{a}_k = a \mid \mathbf{r}_1 = r) = 1$. Using this, we can expand the right-hand side of (2) to conclude that

$$\ell_1(y, r) = \mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_1)\,\mathrm{P}(\mathfrak{E}_1 \mid \times) \\ + \mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_2)\,\mathrm{P}(\mathfrak{E}_2 \mid \times),$$

where $\times$ stands for $\{\mathbf{r}_1 = r, \mathbf{t}_k = t, \mathbf{a}_k = a\}$, and $\mathfrak{E}_1$ denotes the event that the algorithm recognized the real object and $\mathfrak{E}_2$ the event that it recognized a distractor instead of the real object. According to our sensor model:

$$\mathrm{P}(\mathfrak{E}_1 \mid \times) = \begin{cases} 1 - p_1 - p_2 & d(r, a) \leq \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{P}(\mathfrak{E}_2 \mid \times) = \begin{cases} p_2 & d(r, a) \leq \rho_{\text{sensor}} \\ p_3 & \text{otherwise} \end{cases}$$

$$\mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_1) = \begin{cases} \frac{1}{B} & d(s, r) \leq e_{\text{loc}} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_2) = \begin{cases} \frac{1}{A} & d(s, a) \leq \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases}$$

where B is the area of the region $\{\bar{s} \in \mathcal{R} : d(\bar{s}, r) \leq e_{\text{loc}}\}$ and A the area of the region $\{\bar{s} \in \mathcal{R} : d(\bar{s}, a) \leq \rho_{\text{sensor}}\}$. Therefore

$$\ell_1(y, r) = \begin{cases} \frac{1 - p_1 - p_2}{B} & d(r, a) \leq \rho_{\text{sensor}},\ d(s, r) \leq e_{\text{loc}} \\ 0 & \text{otherwise} \end{cases} \\ + \begin{cases} \frac{p_2}{A} & d(r, a) \leq \rho_{\text{sensor}},\ d(s, a) \leq \rho_{\text{sensor}} \\ \frac{p_3}{A} & d(r, a) > \rho_{\text{sensor}},\ d(s, a) \leq \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Once the likelihood function is known for each $s \in \mathcal{R}$, its values for $s = \emptyset$ can be obtained from

$$\ell_1\big(y, r\big) = 1 - \sum_{\bar{s} \in \mathcal{R}} \ell_1\big(\{t, a, \bar{s}\}, r\big) \\ = \begin{cases} p_1 & d(r, a) \leq \rho_{\text{sensor}} \\ 1 - p_3 & \text{otherwise} \end{cases} \quad (4)$$

3.2 Multiple objects

The computation of the likelihood function for multiple objects presented below is based on the following basic assumptions:

Assumption 2. From the sensor point of view, the objects are indistinguishable from each other.

Assumption 3. The distance between any two objects is no smaller than twice the sensor range ρ_{sensor}, i.e., $d_{\min} \geq 2\,\rho_{\text{sensor}}$.

For sensors with a wide range, Assumption 3 may be too restrictive. The following alternative, assumption covers the extreme case of sensors with a global range:

Assumption 3'. The sensor range covers the entire region and the distance between any two objects is no smaller than twice the maximum localization error e_{loc}, i.e., $d(x_1, x_2) \leq \rho_{\text{sensor}}$, $\forall x_1, x_2 \in \mathcal{R}$ and $d_{\min} \geq 2\,e_{\text{loc}}$.

For sensors that do not produce negative readings $\mathbf{s}_k = \emptyset$, the following assumption will also suffice:

Assumption 3". The sensor does not produce negative readings $\mathbf{s}_k = \emptyset$, the probability of recognizing a distractor does not depend on the existence of an object within the sensor range, and the distance between any two objects is no smaller than twice the maximum localization error e_{loc}, i.e., $p_2 = p_3$ and $d_{\min} \geq 2\,e_{\text{loc}}$.

The following is the main result of this section:

Theorem 1. *Under Assumption 2 and either Assumptions 3, 3', or 3", for every $k \geq 1$ and every $r \in \mathcal{R}^n$ such that $\delta(r, d_{\min}) > 0$,*

$$\ell_n(\mathbf{y}_k, r) = c \prod_{i=1}^{n} \ell_1(\mathbf{y}_k, r_i), \tag{5}$$

with probability one, where c is a normalizing constant[2] that does not depend on r and ℓ_1 is the single-object likelihood function defined by (3)–(4).

Proof of Theorem 1. Consider an arbitrary $r := \{r_1, \ldots, r_n\} \in \mathcal{R}^n$ such that $\delta(r, d_{\min}) > 0$ and a nonzero probability realization $y := \{t, a, s\} \in \mathcal{Y}$ for the kth measurement triple $\mathbf{y}_k$.

We consider first the case of a measurement y for which an object was not detected, i.e., $s = \emptyset$. Because of the object indistinguishability Assumption 2,

$$\begin{aligned} \ell_n(y, r) &= \begin{cases} p_1 & \exists i : d(r_i, a) \leq \rho_{\text{sensor}} \\ 1 - p_3 & \text{otherwise} \end{cases} \\ &= \ell_1(y, r_{i^+}), \end{aligned} \tag{6}$$

where i^+ is the index of the object that is closest to the sensor position a. Moreover, for every $i \neq i^+$, under Assumption 3 we have that

$$\ell_1(y, r_i) = \begin{cases} p_1 & d(r_i, a) \leq \rho_{\text{sensor}} \\ 1 - p_3 & d(r_i, a) > \rho_{\text{sensor}} \end{cases} = 1 - p_3$$

because if the distance between any two objects is no smaller than $2\rho_{\text{sensor}}$, then at most i^+ can be in the range of the sensor. Alternatively, under Assumption 3', we have that

$$\ell_1(y, r_i) = p_1, \quad \forall i$$

In either case, $\ell_1(y, r_i)$ is independent of r_i for every $i \neq i^+$ and therefore, because of (6), equation (5) holds with $c := \frac{1}{(1-p_3)^{n-1}}$ or $c := \frac{1}{p_1^{n-1}}$, depending on which Assumption 3 or 3' holds. Under Assumption 3" we do not need to consider this case.

Let us consider now the case of a measurement y for which an object was detected, i.e., $s \in \mathcal{R}$. Similarly to what was done for the single object case, we expand

$$\begin{aligned} \ell_n(y, r) = {} & \mathrm{P}(\mathbf{s}_k = s \mid \mathbf{r}_{i^*} = r_{i^*}, \times, \mathfrak{E}_1)\,\mathrm{P}(\mathfrak{E}_1 \mid \times) \\ & + \mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_2)\,\mathrm{P}(\mathfrak{E}_2 \mid \times), \end{aligned}$$

where $\times$ stands for $\{\mathbf{r} = r, \mathbf{t}_k = t, \mathbf{a}_k = a\}$, and $\mathfrak{E}_1$ denotes the event that the algorithm recognized a real object and $\mathfrak{E}_2$ the event that it recognized a distractor instead of a real object. Given $\mathfrak{E}_1$, i^* denotes the index of the object that was recognized and therefore $d(r_{i^*}, a) \leq \rho_{\text{sensor}}$. Since there can only be one object within e_{loc} of the sensor position s, the index i^* is uniquely determined. According to our sensor model:

$$\mathrm{P}(\mathfrak{E}_1 \mid \times) = \begin{cases} 1 - p_1 - p_2 & \exists i : d(r_i, a) \leq \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathrm{P}(\mathfrak{E}_2 \mid \times) = \begin{cases} p_2 & \exists i : d(r_i, a) \leq \rho_{\text{sensor}} \\ p_3 & \text{otherwise} \end{cases}$$

$$\mathrm{P}(\mathbf{s}_k = s \mid \mathbf{r}_{i^*} = r_{i^*}, \times, \mathfrak{E}_1) = \begin{cases} \frac{1}{B} & d(s, r_{i^*}) \leq e_{\text{loc}}, \\ 0 & \text{otherwise} \end{cases}$$

[2] With some abuse of notation, we use the same symbol c to denote different normalizing constants when this does not lead to ambiguity.

$$\mathrm{P}(\mathbf{s}_k = s \mid \times, \mathfrak{E}_2) = \begin{cases} \frac{1}{A} & d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases}$$

where B is the area of the region $\{\bar{s} \in \mathcal{R} : d(\bar{s}, r_{i^*}) \le e_{\text{loc}}\}$ and A the area of the region $\{\bar{s} \in \mathcal{R} : d(\bar{s}, a) \le \rho_{\text{sensor}}\}$. The object indistinguishability Assumption 2 is implicit in the first two equations.

Suppose first there is one object with index i^+ within e_{loc} of the sensor reading s and in the range of the sensor, i.e., $d(s, r_{i^+}) \le e_{\text{loc}},\ d(r_{i^+}, a) \le \rho_{\text{sensor}}$. Since the minimum distance between two objects is $2\, e_{\text{loc}}$, the object i^+ is the only one than could have produced the sensor reading and therefore $i^* = i^+$, given $\mathfrak{E}_1$. In this case,

$$\begin{aligned} \ell_n(y, r) &= \frac{1 - p_1 - p_2}{B} + \begin{cases} \frac{p_2}{A} & d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases} \\ &= \ell_1(y, r_{i^+}). \end{aligned} \tag{7}$$

Moreover, for every $i \neq i^+$, we have that

$$\ell_1(y, r_i) = \begin{cases} \frac{p_2}{A} & d(r_i, a) \le \rho_{\text{sensor}},\ d(s,a) \le \rho_{\text{sensor}} \\ \frac{p_3}{A} & d(r_i, a) > \rho_{\text{sensor}},\ d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases}$$

because no other object can be within e_{loc} of the sensor reading. Then, under Assumption 3,

$$\ell_1(y, r_i) = \begin{cases} \frac{p_3}{A} & d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

(because in this case i^+ is the only object within the range of the sensor) and, under Assumption 3',

$$\ell_1(y, r_i) = \begin{cases} \frac{p_2}{A} & d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

(because in this case all objects are within the range of the sensor). Under Assumption 3" both (8) and (9) hold because $p_2 = p_3$. From (7) and (8) or (9) we conclude that (5) holds with $c := \frac{A^{n-1}}{p_3^{n-1}}$ or $c := \frac{A^{n-1}}{p_2^{n-1}}$. Note that for a nonzero probability measurement $d(s,a) \le \rho_{\text{sensor}}$.

Suppose now that there is no object both within e_{loc} of the sensor reading a and in the range of the sensor, i.e., there is no i^* for which $d(s, r_{i^*}) \le e_{\text{loc}},\ d(r_{i^*}, a) \le \rho_{\text{sensor}}$. In this case

$$\begin{aligned} \ell_n(y, r) &= \begin{cases} \frac{p_2}{A} & d(r_{i^+}, a) \le \rho_{\text{sensor}},\ d(s,a) \le \rho_{\text{sensor}} \\ \frac{p_3}{A} & d(r_{i^+}, a) > \rho_{\text{sensor}},\ d(s,a) \le \rho_{\text{sensor}} \\ 0 & \text{otherwise} \end{cases} \\ &= \ell_1(y, r_{i^+}), \end{aligned}$$

where i^+ is the index of the object that is closest to the sensor position a. Also now, for every $i \neq i^+$, we have that (8) or (9) hold, from which (5) follows. ■

4 Recursive probabilistic map building

The goal of this section is to determine the *prediction operator* $T_{\text{pred}}[t, \tau, \cdot\,]$ that takes a probabilistic map at time τ and propagates it forward in time to obtain a map at time $t > \tau$; and the *fusion operator* $T_{\text{fuse}}[y, \cdot\,]$ that takes a new measurement y taken at time t and an existing map also at time t to produce a new map at t, given y together with all the measurements already considered in the old map. With these operators we can compute the probabilistic map $m_r(\cdot\,; t|Y_k)$, $t \ge t_k$ recursively as depicted in Figure 1.

We say that a probabilistic map m_r is *separable* if it can be written as

$$m_r(r; t|Y) := c\Big(\prod_{i=1}^{n} f(r_i; t|Y)\Big)\delta(r; d_{\min}), \tag{10}$$

$\forall r := \{r_1, \ldots, r_n\} \in \mathcal{R}^n$, where c is a normalizing constant, f is an appropriately defined function, and $\delta(r, d_{\min})$ is zero whenever two r_i, r_j, $i \neq j$ are closer than some minimum distance $d_{\min} > 0$ and one otherwise. The function $f(\cdot\,; t|Y)$ is called the *aggregate measurement function* and encapsulates the information contained in all the measurement. Together with $d_{\min}$, it unambiguously represents the probabilistic map. This representation is very useful because for any region of reasonable size, the posterior probabilistic map $m_r(\cdot\,; t|Y_k)$ cannot be represented exhaustively as such representation has memory complexity $o(N^n)$, where N is the number of cells in $\mathcal{R}$. However, using (10) one can represent $m_r(\cdot\,; t|Y_k)$ using a single function $f(\cdot\,; t|Y_k)$, whose representation has memory complexity $o(N)$.

We shall see shortly that both the fusion and the prediction operators preserve separability. This means that if the *a priori* map is separable then any *a posteriori* map will also be separable. Moreover, it is possible to apply these operators directly to the aggregate measurement functions, avoiding the explicit computation of the intermediate *a posteriori* maps. This makes recursive probabilistic map building extremely efficient and scalable. It should be noted

however, that the prediction operator only "approximately" preserves separability. This issue is discussed below.

4.1 Fusion Operator

Consider a list $Y_k := \{y_1, y_2, \ldots, y_k\}$ of k measurements with $y_k := \{t_k, a_k, s_k\}$, corresponding to a nonzero probability realization of $\mathbf{Y}_k$. Using the Bayes rule, we conclude that

$$m_r(r; t_k|Y_k) = c\,\mathrm{P}(\mathbf{y}_k = y_k \mid \mathbf{r}(t_k) = r, \mathbf{Y}_{k-1} = Y_{k-1})\, m_r(r; t_k|Y_{k-1}),$$

where $Y_{k-1} := \{y_1, y_2, \ldots, y_{k-1}\}$, $r \in \mathcal{R}^n$, and $c > 0$ is a normalizing constant independent of r chosen so that $\sum_{r\in\mathcal{R}} m_r(r, Y_k) = 1$. From this and the definition of the sensor likelihood function in (1) we obtain

$$m_r(r; t_k|Y_k) = c\,\ell_n(y_k, r)\, m_r(r; t_k|Y_{k-1}), \qquad (11)$$

which defines the fusion operator $T_{\text{fuse}}[y_k, \cdot\,]$. Note that this formula takes a map at time t_k, given a list of old measurements Y_{k-1} and a new measurement y_k taken at time t_k, to produce a new map given both the old and new measurements.

Suppose now that (5) holds and that the probabilistic map given Y_{k-1} is separable, i.e., that (10) holds with $t := t_k$ and $Y := Y_{k-1}$. Then, from (11) one concludes that

$$m_r(r; t_k|Y_k) = c\Big(\prod_{i=1}^{n} f(r_i; t_k|Y_k)\Big)\delta(r; d_{\min}),$$

where c is a different normalizing constant, and

$$f(x; t_k|Y_k) := \ell_1(y_k, x)\, f(x; t_k|Y_{k-1}), \quad \forall x \in \mathcal{R}. \qquad (12)$$

This means that the map given Y_k is still separable and therefore that $T_{\text{fuse}}[y_k, \cdot\,]$ indeed preserves separability. The following was proved:

Theorem 2. *The fusion operator that produces the probabilistic map $m_r(\cdot\,; t_k|Y_k)$ from the measurement $y_k := \{t_k, a_k, s_k\}$ and the map $m_r(\cdot\,; t_k|Y_{k-1})$ is given by* (11)*. Moreover, in case* (5) *holds and $m_r(\cdot\,; t_k|Y_{k-1})$ is separable, then $m_r(\cdot\,; t_k|Y_k)$ is also separable and the aggregate measurement function for the latter can be computed from the one for the former using* (12)*.*

4.2 Prediction Operator

Consider again a list $Y_k := \{y_1, y_2, \ldots, y_k\}$ of k measurements with $y_k := \{t_k, a_k, s_k\}$, corresponding to a nonzero probability realization of $\mathbf{Y}_k$ and pick an arbitrary time $t \geq t_k$. By the Total Probability Theorem, we have

$$m_r(r'; t + dt|Y_k) = \sum_{r\in\mathcal{R}} \mathrm{P}(\mathbf{r}(t + dt) = r' \mid \mathbf{r}(t) = r, \mathbf{Y}_k = Y_k)\, m_r(r; t|Y_k).$$

Moreover, since $t \geq t_k$, $r(t + dt)$ is conditionally independent of $\mathbf{Y}_k$, given $\mathbf{r}(t)$. Therefore

$$m_r(r'; t + dt|Y_k) = \sum_{r\in\mathcal{R}} \Phi(r', r; t + dt, t) m_r(r; t|Y_k), \qquad (13)$$

which defines the prediction operator $T_{\text{pred}}[t+dt, t, \cdot\,]$ that takes a probabilistic map at time t and propagates it forward in time to obtain a map at time $t+dt$.

Suppose now that the probabilistic map at time t is separable, i.e., that (10) holds with $Y := Y_k$. Our goal for the remaining of this section is to determine if the probabilistic map $m_r(r'; t + dt|Y_k)$ is still separable.

4.2.1 Single object

We start by considering the case $n = 1$ and therefore the *single-object transition probability function*

$$\varphi_1(r', r; t + dt, t) := \Phi(r', r; t + dt, t),$$

$r, r' \in \mathcal{R},\ t \geq t_0, dt \geq 0$, for which

$$\sum_{r'\in\mathcal{R}} \varphi_1(r', r; t + dt, t) = 1, \qquad (14)$$

$\forall r \in \mathcal{R},\ t \geq t_0, dt \geq 0$. To compute the object's probabilistic map at some time $t + dt$, $dt > 0$, we use (13) and the maps' separability at time t to obtain

$$m_r(r'; t + dt|Y_k) = c \sum_{r\in\mathcal{R}} \varphi_1(r', r; t + dt, t) f(r; t|Y_k).$$

Note that for a single object $\delta(r; d_{\min})$ is always equal to one. Therefore, the map at time $t + dt$ is still separable and can be written as in (10) with $Y := Y_k$ and

$$f(r'; t + dt|Y_k) = \sum_{r\in\mathcal{R}} \varphi_1(r', r; t + dt, t) f(r; t|Y_k), \qquad (15)$$

for every $r' \in \mathcal{R}$.

4.2.2 Multiple objects

We consider now the general case $n \geq 1$. We assume that the multiple objects essentially move independently of each other according to a model similar to that in the previous section, except that we enforce the constraint that two objects will never get closer than $d_{\min}$ from each. This is needed for consistency with Assumption 1. The following *n-object transition probability function* captures this:

$$\Phi(r', r; t+dt, t) = c(r)\Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t)\Big)\delta(r'; d_{\min}) \quad (16)$$

$r, r' \in \mathcal{R}^n$, where $c(r)$ is the following normalizing constant

$$c(r)^{-1} = \sum_{\bar{r} \in \mathcal{R}^n} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r_k; t+dt, t)\delta(\bar{r}; d_{\min}). \quad (17)$$

We assume here that the objects' velocity is bounded and therefore, for small values of $dt > 0$, $\varphi_1(r, r'; t, t+dt)$ is nonzero only when r and r' correspond to the same or adjacent cells. Moreover, in the latter case $\varphi_1(r, r'; t, t+dt)$ is of order dt, i.e., $\varphi_1(r, r'; t, t+dt) \leq c\,dt$ for some finite constant c. This corresponds to the following assumption:

Assumption 4. The n-object transition probability function is given by (16), with

$$\varphi_1(r', r; t+dt, t) = \begin{cases} 1 - o(dt) & r = r' \\ o(dt) & r,\ r' \text{ adjacent} \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

$\forall r, r' \in \mathcal{R},\ t \geq t_0, dt \geq 0.$

We show next that under this assumption the probabilistic map is approximately separable at time $t+dt$, for small dt. Moreover, the error is $o(dt^2)$ for most object configurations and $o(dt)$ for a small set $\mathcal{A}_{\text{bad}} \subset \mathcal{R}^n$ of "bad" configurations r for which

1. $\delta(r; d_{\min}) \neq 0$, i.e., no two objects are closer than $d_{\min}$ from each other; and
2. there exists some configuration $\bar{r} \in \mathcal{R}^n$ for which $\delta(\bar{r}; d_{\min}) = 0$ and it is possible to go from configuration $\bar{r}$ to configuration r by a single object moving to an adjacent cell.

In practice, $\mathcal{A}_{\text{bad}}$ denotes the set of configurations for which the objects still satisfy the minimum distance constrain but are very close to violating it.

Theorem 3. *The prediction operator that produces the probabilistic map $m_r(\,\cdot\,; t+dt|Y_k)$, $dt \geq 0$ from the map $m_r(\,\cdot\,; t|Y_k)$ is given by (13). Moreover, in case $m_r(\,\cdot\,; t|Y_k)$ is separable and Assumption 4 holds, then*

$$m_r(r'; t+dt|Y_k) = c\Big(\prod_{k=1}^{n} f(r'_k; t+dt|Y_k)\Big)\delta(r'; d_{\min}) + \epsilon(r'; dt), \qquad \forall r' \in \mathcal{R}^n,$$

where $f(\,\cdot\,; t+dt|Y_k)$ is defined as in (15), and $\epsilon(r'; dt) = o(dt^2)$ for $r' \notin \mathcal{A}_{\text{bad}}$ and $o(dt)$ otherwise.

In practice, for small dt we can approximate

$$m_r(r'; t+dt|Y_k) \approx c\Big(\prod_{k=1}^{n} f(r'_k; t+dt|Y_k)\Big)\delta(r'; d_{\min}). \quad (19)$$

Moreover, if one wants to compute $m_r(r'; t'|Y_k)$ for some $t' > t$ for which $t' - t$ is not small, one can divide the interval $[t', t]$ into N subintervals of length $dt := (t'-t)/N$ and apply (19) N times to propagate the map from t to t'. Outside $\mathcal{A}_{\text{bad}}$, each of the N steps will introduce an error $o\big((t'-t)^2/N^2\big)$, which adds up to a total worst-case error $o\big((t'-t)^2/N\big)$. Therefore the error can be made small by making N large, which means that the map at time t' is still approximately separable. For the configurations in $\mathcal{A}_{\text{bad}}$, in general it will not be possible to make the error arbitrarily small. However, in practice $\mathcal{A}_{\text{bad}}$ is a small subset of $\mathcal{R}^n$.

Proof of Theorem 3. To compute the object's probabilistic map at some time $t+dt$, $dt > 0$, we use (13), (16), and the maps' separability at time t to obtain

$$\begin{aligned} &m_r(r'; t+dt|Y_k) \\ &\quad = c\sum_{r \in \mathcal{R}} c(r)\Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t) f(r_k; t|Y_k)\Big) \\ &\qquad\qquad \delta(r; d_{\min})\delta(r'; d_{\min}) \\ &\quad = c\sum_{r \in \mathcal{R}} \Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t) f(r_k; t|Y_k)\Big) \\ &\qquad\qquad \delta(r'; d_{\min}) - \epsilon_1 + \epsilon_2 \end{aligned}$$

$$= c\Big(\prod_{k=1}^{n} f(r'_k; t+dt|Y_k)\Big)\delta(r'; d_{\min}) - \epsilon_1 + \epsilon_2,$$

where

$$\epsilon_1 := c \sum_{\substack{r\in\mathcal{R}:\\ \delta(r;d_{\min})=0\\ \delta(r';d_{\min})\neq 0}} \Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t) f(r_k; t|Y_k)\Big)$$

$$\epsilon_2 := c \sum_{\substack{r\in\mathcal{R}:\\ \delta(r;d_{\min})\neq 0\\ \delta(r';d_{\min})\neq 0}} (c(r)-1)\Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t) f(r_k; t|Y_k)\Big).$$

We show next that both ϵ_1 and ϵ_2 are $o(dt)$ for $r' \in \mathcal{A}_{\text{bad}}$ and $o(dt^2)$ otherwise. To this effect, we consider three cases separately:

Case 1: For any $r' \in \mathcal{R}^n$ (including $r' \in \mathcal{A}_{\text{bad}}$), we have $r \neq r'$ in all terms of the summation that defines ϵ_1. Therefore, there must always exist at least one $\varphi_1(r'_k, r_k; t+dt, t)$ for which $r_k \neq r'_k$. Because of (18) that term will be $o(dt)$ and therefore $\epsilon_1 = o(dt)$.

As for ϵ_2, note that because of (17) and (14), we have that

$$\begin{aligned} c(r)^{-1} &= \sum_{\bar{r}\in\mathcal{R}^n} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r_k; t+dt, t) \\ &\quad - \sum_{\substack{\bar{r}\in\mathcal{R}^n:\\ \delta(\bar{r};d_{\min})=0}} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r_k; t+dt, t) \\ &= \prod_{k=1}^{n} \sum_{r_k\in\mathcal{R}} \varphi_1(\bar{r}_k, r_k; t+dt, t) \\ &\quad - \sum_{\substack{\bar{r}\in\mathcal{R}^n:\\ \delta(\bar{r};d_{\min})=0}} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r_k; t+dt, t) \\ &= 1 - \sum_{\substack{\bar{r}\in\mathcal{R}^n:\\ \delta(\bar{r};d_{\min})=0}} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r_k; t+dt, t). \quad (20) \end{aligned}$$

Moreover, for any r such that $\delta(r; d_{\min}) \neq 0$, all the $\bar{r}$ in the summation must be different than r and therefore there will always be one $\bar{r}_k \neq r_k$. Because of (18) we conclude that $c(r)^{-1} = 1 - o(dt)$ and therefore $c(r) - 1 = o(dt)$. From this and the definition of ϵ_2 it follows that $\epsilon_2 = o(dt)$.

Case 2: In case $\delta(r'; d_{\min}) = 0$, then $\epsilon_1 = \epsilon_2 = 0$ by definition. Therefore ϵ_1 and ϵ_2 are $o(dt^2)$.

Case 3: In case $r' \notin \mathcal{A}_{\text{bad}}$ and $\delta(r'; d_{\min}) \neq 0$, at least two objects must change their positions to reach r' from any configuration r for which $\delta(r; d_{\min}) = 0$. Therefore, in the definition of ϵ_1, each term of the summation must always have at least two $\varphi_1(r'_k, r_k; t+dt, t)$ for which $r_k \neq r'_k$. Because of (18) that term will be $o(dt^2)$ and therefore $\epsilon_1 = o(dt^2)$. Regarding ϵ_2, we can re-write its definition as

$$\begin{aligned} \epsilon_2 &:= c(c(r')-1)\Big(\prod_{k=1}^{n} \varphi_1(r'_k, r'_k; t+dt, t) f(r'_k; t|Y_k)\Big) \\ &\quad + c \sum_{\substack{r\neq r':\\ \delta(r;d_{\min})\neq 0\\ \delta(r';d_{\min})\neq 0}} (c(r)-1)\Big(\prod_{k=1}^{n} \varphi_1(r'_k, r_k; t+dt, t) f(r_k; t|Y_k)\Big) \end{aligned}$$

The second term is $o(dt^2)$ because $c(r) - 1 = o(dt)$ and all terms in the summation are also $o(dt)$ because $r \neq r'$. As for the first term, because of (20), $c(r')-1$ is of the same order as

$$\sum_{\substack{\bar{r}\in\mathcal{R}^n:\\ \delta(\bar{r};d_{\min})=0}} \prod_{k=1}^{n} \varphi_1(\bar{r}_k, r'_k; t+dt, t).$$

Using the same argument used to show that $\epsilon_1 = o(dt^2)$, we conclude that this term is also $o(dt^2)$ and therefore $\epsilon_2 = o(dt^2)$. ■

5 Occupancy maps

It is often convenient to work with occupancy maps instead of probabilistic maps. Given a list of k measurements Y_k, we define the *(a posteriori) occupancy map at time $t \geq t_0$ given Y_k* by

$$m_o(x; t|Y_k) := \mathrm{P}(\exists i : \mathbf{r}_i(t) = x \mid \mathbf{Y}_k = Y), \quad x \in \mathcal{R}.$$

It turns out that under suitable assumptions one can compute the occupancy map $m_o(\,\cdot\,; t|Y)$ directly from the aggregate measurement function $f(\,\cdot\,; t|Y)$, without explicitly computing the probabilistic map $m_r(\,\cdot\,; t|Y)$. Indeed, suppose that Assumption 1 holds; that the probabilistic map at time t is separable, i.e., that (10) holds with $Y := Y_k$; and that there exists a partition $\mathcal{R}_0$, $\mathcal{R}_1$, $\mathcal{R}_2$, …, $\mathcal{R}_m$ of $\mathcal{R}$ such that

1. all points in each $\mathcal{R}_i$, $i \in \{1, 2, \ldots, m\}$, are away from each other by no more than $d_{\min}$;
2. all the $\mathcal{R}_i$, $i \in \{1, 2, \ldots, m\}$, are away from each other by more than $d_{\min}$;
3. the $f(\,\cdot\,; t|Y_k)$ are sufficiently uniform over $\mathcal{R}_0$ so that, for any $\bar{\mathcal{R}} \subset \mathcal{R}_0$ with small area compared to $\mathcal{R}_0$,

$$\sum_{x\in\mathcal{R}_0\setminus\bar{\mathcal{R}}} f(x; t|Y_k) \approx \sum_{x\in\mathcal{R}_0} f(x; t|Y_k).$$

For each $x \in \mathcal{R}$, we denote by $\sigma := \{\sigma_1, \sigma_2, \ldots, \sigma_k\}$ a combination of k elements taken from the set $\{i : x \notin \mathcal{R}_i\} \subset \{1, 2, \ldots, m\}$ and by $\mathcal{S}^{(k)}[x]$ the set of all such combinations. The following Lemma can be adapted to the setting of this paper from the results in [4, 5]:

Lemma 1. *The occupancy map at time t given Y_k is given by*

$$m_o(x; t|Y_k) = \sum_{r \in \mathcal{R}^n : \exists i : r_i = x} m_r(r; t|Y_k).$$

Moreover, in case Assumption 1 holds and the map at time t is separable, then

$$m_o(x; t|Y_k) \approx c f(x; t|Y_k) \Bigg(n f_0^{n-1} + \sum_{k=1}^{n-1} \frac{n!\, f_0^{n-k-1}}{(n-1-k)!} \sum_{\sigma \in \mathcal{S}^{(k)}[x]} f_{\sigma_1} f_{\sigma_2} \ldots f_{\sigma_k} \Bigg),$$

where $f_i := \sum_{x \in \mathcal{R}_i} f(x; t|Y_k)$, $i \in \{0, 1, \ldots, m\}$ with the regions $\mathcal{R}_i$ defined as above, and c a normalizing constant chosen so that $\sum_{x \in \mathcal{R}} m_o(x; t|Y_k) = n$.

6 Conclusions

We proposed an efficient algorithm to compute probabilistic maps. This algorithm scales well with the number of objects being mapped and is characterized by a memory complexity that is linear on the area of the region where the objects lie. To achieve this, we represent the joint probability distribution of the objects' positions approximately using an appropriately defined aggregate measurement function. In general, the approximation error can be made arbitrarily small. Three main assumptions make this approximation valid: the objects are indistinguishable from the sensor view-point, there is a minimum distance between any two objects, and their motion is essentially independent when the objects are far from each other.

We are currently investigating implementations of the algorithm proposed here that allow the measurements $\mathbf{y}_k$ to arrive out-of-order. We are also investigating distributed versions of these algorithms to address the efficient fusion of measurements coming from different agents.

References

[1] J. A. Castellanos, J. M. M. Montiel, J. Neira, , and J. D. Tardós. The SPmap: a probabilistic framework for simultaneous localization and map building. *IEEE Trans. Robot. Automat.*, 15(5): 948–52, Oct. 1999.

[2] G. Dedeoglu, M. J. Matarić, and G. S. Sukhatme. Incremental online topological map building with a mobile robot. In *Proc. of the SPIE*. International Society for Optical Engineering, Sept. 1999.

[3] S. P. Engelson and D. V. McDermott. Error correction in mobile robot map learning. In *Proc. of the IEEE Int. Conf. on Robot. and Automat.*, pages 2555–2560, May 1992.

[4] J. P. Hespanha, H. H. Kızılocak, and Y. S. Ateşkan. Probabilistic map building for aircraft-tracking radars. Technical report, University of Southern California, Los Angeles, CA, Dec. 2000. This report supersedes a previous report with the same title, dated September 2000.

[5] J. P. Hespanha, H. H. Kızılocak, and Y. S. Ateşkan. Probabilistic map building for aircraft-tracking radars. In *Proc. of the 2001 Amer. Contr. Conf.*, June 2001.

[6] F. Lu and E. Milios. Globally consistent range scan alignment for environment mapping. *Autonomous Robots*, 4:333–349, 1997.

[7] S. Thrun. Learning metric-topological maps for indoor mobile robot navigation. *Artificial Intelligence*, 99(1):21–71, Feb. 1998.

[8] S. Thrun, W. Burgard, and D. Fox. A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning* and *Autonomous Robots* (joint issue), 31(5):29–53, 1998.

[9] B. Yamauchi and P. Langley. Place recognition in dynamic environments. *J. of Robot. Systems*, 14:107–120, 1997.