Comparative Study of Inverse and Forward Sensor Models in Occupancy Grid Mapping Using Sonars [1]

João Carvalho
joaopscarvalho@ist.utl.pt
Rodrigo Ventura
rodrigo.ventura@isr.ist.utl.pt

Institute for Systems and Robotics
Instituto Superior Técnico - Technical University of Lisbon
Lisbon, Portugal

Abstract

Building occupancy grid maps with sonar sensors is a challenging task due to angular uncertainty, specular reflections and crosstalk. This paper presents a qualitative comparison of two probabilistic approaches to the robotic mapping — inverse and forward sensor models — and proposes a different formulation to the latter. The inverse one assumes independence of the cells while the forward is formulated as a maximum likelihood problem over a binary grid.

1 Introduction

One of the most used map representation in robotics is the *occupancy grid map* (OccGrid map), which aims to geometrically represent the environment through a grid discretization of the space. To build this maps, one commonly used sensor is the sonar. Sonars are cheap and allow the construction of maps even with a low number of sensors. Despite these advantages, sonars suffer from angular uncertainty, specular reflections and crosstalk between each other, causing erroneous and conflicting measurements [5].

The typically used method to build OccGrid maps is the one proposed by Elfes, making use of *inverse sensor models* [2]. In this approach, the occupancy of each cell is computed disregarding the rest of the map. A different approach was proposed by Thrun, using *forward sensor models* [4]. This method approaches the mapping problem in the high-dimensional space of all maps, trying to solve erroneous and conflicting sonar measurements which affect Elfe's method results.This paper aims to briefly compare this two approaches and propose a slightly different formulation to the latter.

2 OccGrid Maps with Inverse Sensor Models

In this approach, the mapping problem is treated inversely to how sonar data is generated, being formulated as

$$p(M|z_{1:T}, x_{1:T}) \tag{1}$$

where M represents the complete map, $z_{1:T}$ represents the complete set of measurements and $x_{1:T}$ are the corresponding poses. This is the denominated *inverse sensor model*.

To simplify the mapping problem, it is assumed that the occupancy of a given cell is not important to the computation of the occupancy of its neighbour cells, i.e., cells are conditionally independent given measurements and the robot trajectory, transforming the mapping problem into a binary estimation problem,

$$p(m_i|z_{1:T}, x_{1:T}), \tag{2}$$

where m_i is an individual cell of the complete map. A second assumption made is the *static world* assumption, considering a measurement t conditionally independent from the previous measurements given the map knowledge. This is a common assumption in mapping but given the decomposition into a binary problem this becomes a much stronger and also incorrect assumption, since it considers conditional independence given only a map cell and not the complete map. Additionally, the pose in the instant t is independent from the poses in previous instant times. So, for time t:

$$p(z_t|z_{1:t-1}, x_{1:t}, m_i) = p(z_t|m_i, x_t) \tag{3}$$

Given these assumptions and applying the Bayes rule to (2), we have:

$$p(m_i|z_{1:t}, x_{1:t}) = \frac{p(m_i|z_t, x_t)p(z_t|x_t)p(m_i|z_{1:t-1}, x_{1:t-1})}{p(m_i)p(z_t|z_{1:t-1}, x_{1:t})}. \tag{4}$$

As is common practice, we will compute the log odds of this probability instead of the probability itself:

$$l_i^t = \log\frac{p(m_i|z_t, x_t)}{1-p(m_i|z_t, x_t)} - \log\frac{p(m_i)}{1-p(m_i)} + l_i^{t-1}, \tag{5}$$

where l_i^t represents $\log\frac{p(m_i|z_{1:t}, x_{1:t})}{1-p(m_i|z_{1:t}, x_{1:t})}$. The term l_i^{t-1} equals $\log\frac{p(m_i)}{1-p(m_i)}$ when $t=1$. The probability $p(m_i)$ is the prior of occupancy of the cell i of the map. A typical and simple approach is to model the posterior not as a fixed functional form but by a finite number of values which roughly approximate the posterior [5]. For the cells at distances between 0 and the neighbourhood of the measurement the occupancy probability has a low value, in the neighbourhood it has a high value and 0.5 beyond.

Making use of (5) the log-odds occupancy representation can be easily computed for each cell that falls into the coverage cone of the sonar measurements. So, finally, the desired occupancy probability of the cells can be recovered through:

$$p(m_i|z_{1:t}, x_{1:t}) = 1 - \frac{1}{1+e^{l_i^t}}. \tag{6}$$

In the implementation, and in order to make it more robust to specular reflections, it was given less weight to larger measurements. So in (5), the term $\log\frac{p(m_i|z_t, x_t)}{1-p(m_i|z_t, x_t)}$ comes multiplied by a variable, restricted between 0 and 1, that is inversely proportional to the sonar measurement z_t. The final map is obtain after submitting each cell to a threshold. If the probability of occupancy is inferior to the threshold, the cell is considered unoccupied, being considered occupied otherwise.

3 OccGrid Maps with Forward Sensor Models and 1D Clustering

Static world assumption is also made in this approach but, in order to address the listed sonar problems, it does not make a map decomposition, dealing with mapping problem in its complete state space. Additionally it uses *forward sensor models*, being able to make use more complex sensor models. The forward mapping approach is modelled as a likelihood

$$p(z_{1:T}|M, x_{1:T}). \tag{7}$$

This is a generative model, being formulated as the phenomenon happens; given the world (represented by the map M) and a given set of poses, a particular set of sonar reading is generated. The goal is to maximize (7), iteratively adjusting M till no better model is found. This problem can then be formulated as a maximum likelihood estimation problem.

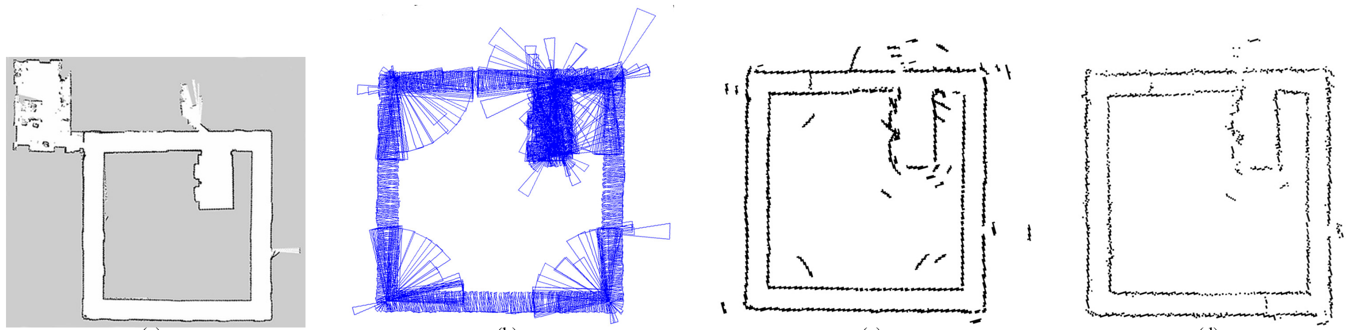
Rather then assuming that all measurements are caused by an obstacle, three possible cases of beam reflection are considered, *maximum reading*, *random* and *non-random*. A non-random measurement is caused by an obstacle in the sonar beam. A maximum value reading happens with the failure in detecting all the obstacles, when present, and returning the maximum range value, z_{max}. The random case models the remaining causes. Since in practice the true cause of the sonar reading is not known, a classifier has to be used to identify it.

For the measurement with index t, consider K_t to be the number of obstacles present in the sonar cone, $d_{t,k}$ the distance from the k'th obstacle in the cone and D_t the set of obstacle distances in ascending order. The model (7) is defined as the combination of the models of each possible cause. Consider the binary variables $c_{t,*}$, $c_{t,k}$, $c_{t,0}$, restricted to:

$$c_{t,*} + \sum_{k=0}^{K_t} c_{t,k} = 1. \tag{8}$$

These variables are equal to 1 when the measurement is random, caused by obstacle k or equal to the maximum range, respectively. The random case is modeled as a uniform distribution in the entire sonar range, since the reading could have been caused in any part of the sonar cone. When the beam is reflected by an obstacle, it is considered that it is affected by

[1] This work was supported by the FCT project [PEst-OE/EEI/LA0009/2011]



(a) (b) (c) (d)

Figure 1: Experimental Results: (a) ground truth map obtained with GMapping and a Sick laser rangefinder; (b) measurements taken; (c) OccGrid map using inverse sensor model; (d) OccGrid map using forward sensor model.

additive white gaussian noise. In the case where $z_t = z_{max}$, since it is a discrete event, a Dirac delta function is considered.

$$p\left(z_t|M,x_t,c_{t,*}=1\right) = \begin{cases} \frac{1}{z_{max}} & \text{if } 0 \leq z_t < z_{max}, \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

$$p\left(z_t|M,x_t,c_{t,k}=1\right) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z_t-d_{t,k})^2}{2\sigma^2}} & \text{if } 0 \leq z_t < z_{max}, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

$$p\left(z_t|M,x_t,c_{t,0}=1\right) = \delta\left(z_t - z_{max}\right) \quad (11)$$

In a single expression it can be written as

$$p\left(z_t|M,x_t,c_t\right) = p\left(z_t|M,x_t,c_{t,*}=1\right)^{c_{t,*}} \prod_{k=0}^{K_t} p\left(z_t|M,x_t,c_{t,k}=1\right)^{c_{t,k}} \quad (12)$$

Since there is no prior knowledge of the measurement's cause we define the posterior probability

$$p\left(c_t|M,x_t\right) = p\left(c_{t,*}=1|M,x_t\right)^{c_{t,*}} \prod_{k=0}^{K_t} p\left(c_{t,k}=1|M,x_t\right)^{c_{t,k}} \quad (13)$$

$$p\left(c_t|M,x_t\right) = \begin{cases} p_{rand} & \text{if } c_{t,*}=1, \\ p_{max} & \text{if } c_{t,0}=1, K_t \geq 1. \\ \left(1-p_{rand}-p_{max}\right)\prod_{i=1}^{k-1}\left[\left(1-p_{hit}^{(i)}\right)\right] p_{hit}^{(k)} & \text{if } c_{t,k}=1, k \geq 1. \end{cases} \quad (14)$$

where p_{rand} is the prior probability of a measurement being random, p_{max} is the prior probability of a measurement being maximum and $p_{hit}^{(i)}$ is the prior of the obstacle i to reflect the sonar beam. The p_{hit} probability is function of the obstacle's width coverage in the sonar cone, varying linearly between a minimum and a maximum value and being equal to the maximum value when the obstacle covers 100% of the cone width. Therefore, an obstacle might be formed by one or more occupied cells, forming a *cluster*. Cells are clustered having as criterium its distance to the sonar cone origin. A cluster is initially formed by a single cell in which further cells are added if the difference between its distance, $d_{t,k}$, and the cluster center of mass is smaller then a given threshold. When a cell does not meet this criterium, a new cluster is created with it.

Summarizing, we have $p\left(z_t|M,x_t,c_t\right)$ and $p\left(c_t|M,x_t\right)$ but we want $p\left(z_t|M,x_t\right)$. However, we can write

$$p\left(z_t,c_t|M,x_t\right) = p\left(z_t|M,x_t,c_t\right) p\left(c_t|M,x_t\right), \quad (15)$$

which regarding all sensor data and using the logarithm becomes

$$\log p\left(z_{1:T},c_{1:T}|M,x_{1:T}\right) = \sum_t^T \log p\left(z_t,c_t|M,x_t\right). \quad (16)$$

Given the unobservability of c_t, we now compute the expected value of (16) in order of c_t, arriving to the *expected log-likelihood* to maximize:

$$E\left[\log p\left(z_{1:T},c_{1:T}|M,x_{1:T}\right)|z_{1:T},M\right] = E\left\{\sum_t^T \log p\left(z_t,c_t|M,x_t\right)|z_{1:T},M\right\}. \quad (17)$$

In the computation of this likelihood (12) and (13) are used. Being c_t a Bernoulli random variable, its expected value equals its probability, (14).

Not considering the maximum reading event as a particular case of the non-random case and defining $p(z_t|M,x_t,c_{t,0}=1)$ as a Delta dirac function makes those readings have no influence in the likelihood and in the process of maximization, contrary to what happens in the original formulation. Making p_{hit} function of the coverage and the introduction of clustering allows the representation of the angular uncertainty, which is a process not clear in [4].

To maximize (17), a variation of Dempster's Expectation-Maximization algorithm is used, where c_t is a vector of hidden variables [1]. No terms are discarded in (17), since any change in M might produce significant value variations in those terms. To find the map M that maximizes the likelihood, the occupancy of the cells that fall into the measurements cone is flipped and maintained if its new value increases the likelihood value. Given the discretization made, this results in a very greedy algorithm, in which the final result highly depends on the cell flipping order. Since it gave empirically good results, in this implementation we chose to first flip the cells closest to the measurement and progressively moving away. The Dirac delta function in (11) is implemented as a gaussian distribution with a very low variance.

4 Results

The robot used was the Pioneer P3-AT, equipped with SensComp 600 Series sonar sensors and a Sick LMS200 laser rangefinder. A ground truth map was built using the laser and the *GMapping* algorithm [3], Figure 1(a). Since it is assumed that the robot's pose is known, the GMapping's pose estimative was assumed as the true pose.

To build the maps two sonars were used, placed orthogonally to the robot's movement and one on each side. The measurements were taken with the robot moving approximately at $0.6m/s$ and measurements being taken with a $4Hz$ frequency on a single lap to the corridor.

Both approaches present an overall good representation of the environment. The forward method presents less artifact obstacles while fails to represent some obstacles in the corridor atrium, moreover it is prone to local maxima and is very computationally intensive.

5 Conclusion

This paper presented a comparison between OccGrid mapping using inverse and forward sensor models. The preliminary results showed that the latter solves some problems that affect the inverse but having as drawback the computational effort. Future work consists in presenting a more quantitative and systematic comparison and in studying better approaches using forward sensor models, to improve computational efficiency and to deal with outliers.

References

[1] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

[2] A. Elfes. Sonar-based real-world mapping and navigation. *Robotics and Automation, IEEE Journal of*, 3(3):249–265, 1987.

[3] G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grid mapping with rao-blackwellized particle filters. *Robotics, IEEE Transactions on*, 23(1):34–46, 2007.

[4] S. Thrun. Learning occupancy grid maps with forward sensor models. *Autonomous robots*, 15(2):111–127, 2003.

[5] S. Thrun, W. Burgard, and D. Fox. Probabilistic robotics. 2005. isbn 0262201623, 2005.