

Robust Mapping and Localization in Indoor Environments using Sonar Data *

Juan D. Tardós, José Neira
Dept. Informática e Ingeniería de Sistemas,
Universidad de Zaragoza
María de Luna 3, E-50018 Zaragoza, Spain
tardos@posta.unizar.es, jneira@posta.unizar.es

Paul M. Newman, John J. Leonard
MIT Dept. of Ocean Engineering
77 Massachusetts Av., Cambridge, MA 02139-4307 USA
pnewman@mit.edu, jleonard@mit.edu

June 24, 2002

Abstract

In this paper we describe a new technique for the creation of feature-based stochastic maps using standard Polaroid sonar sensors. The fundamental contributions of our proposal are: (1) a perceptual grouping process that permits the robust identification and localization of environmental features, such as straight segments and corners, from the sparse and noisy sonar data; (2) a map joining technique that allows the system to build a sequence of independent limited-size stochastic maps and join them in a globally consistent way; (3) a robust mechanism to determine which features in a stochastic map correspond to the same environment feature, allowing the system to update the stochastic map accordingly, and perform tasks such as revisiting and loop closing. We demonstrate the practicality of this approach by building a geometric map of a medium size, real indoor environment, with several people moving around the robot. Maps built from laser data for the same experiment are provided for comparison.

Key words

Map Building, Local Maps, Data Association, Sonar Sensors, Hough Transform

*To appear in The International Journal of Robotics Research, 2002

1 Introduction

The problem of concurrent mapping and localization (CML) for an autonomous mobile robot is stated as follows: starting from an initial position, a mobile robot travels through a sequence of positions and obtains a set of sensor measurements at each position. The goal is for the mobile robot to process the sensor data to produce an estimate of its position while concurrently building a map of the environment. The difficulty of this problem lies in the fact that, to obtain a good map, an accurate estimation of the vehicle trajectory is required, but reducing the unbounded growing odometry errors requires to associate sensor measurements with a precise map. The problem of CML, also referred to as SLAM (simultaneous localization and map building), presents a number of difficult issues, including (1) efficient mapping of large-scale environments, (2) correct association of measurements, and (3) robust estimation of map and vehicle trajectory information. This paper presents contributions to each of these three areas.

1.1 Choice of Representation

As with many problems in robotics and artificial intelligence, the issue of choosing a representation is perhaps the key step in developing an effective CML solution. A central requirement is the ability to represent uncertainty (Brooks 1984, Lozano-Pérez 1989). Popular choices for the map representation include grid-based (Elfes 1987, Schultz and Adams 1998), topological (Kuipers 2000, Choset and Nagatani 2001), feature-based models (Moutarlier and Chatila 1989, Ayache and Faugeras 1989), and sequential Monte Carlo methods (Thrun 2001, Doucet, de Freitas and Gordan 2001).

This paper adopts a feature-based approach to CML, in which the locations of geometric features in the environment and the position of the vehicle are jointly estimated in a stochastic framework (Smith, Self and Cheeseman 1988, Moutarlier and Chatila 1989). CML is cast as a variable-dimension state estimation problem in which the size of the state space is increased or decreased as features are added or removed from the map. As the robot moves through its environment, it uses new sensor measurements to perform two basic operations: (1) adding new features to its state vector, and (2) updating concurrently its estimate of its own state and the locations of previously observed features in the environment.

Grid-based representations seek to determine *where* particular measurements come from. Feature-based approaches additionally try to estimate *what* measurements come from, or, what processes give rise to particular measurements. Part of the motivation here is based on how the environment is modelled. In many cases, the physical processes underlying sensor measurement can be broken into discrete mathematical models. For example, sonar returns are characterized as specular or diffuse. Feature-based representation of the environment allows the use of multiple models to describe the measurement processes for different parts of the environment.

One reason for avoiding a cell-based approach is the effect of data smearing. Cell-based approaches often smear measurements onto a region of the map to account for measurement and navigational uncertainty. However, these two types of uncertainty are fundamentally different. Navigational uncertainty is an *a posteriori* amalgam of measurement and process noises. Measurement noise is stipulated *a priori*. By combining these uncertainties for data smearing, information is lost. If a feature-based approach is taken, a distinction is made between modelling features themselves and mapping the features globally. Characterization of a feature and relative positioning with nearby features can be obtained with low uncertainty (near the level of measurement noise) even when the vehicle navigational uncertainty is high. In a cell-based approach, the local information would be lost by reducing the information content to global levels of certainty.

Related recent research by ourselves and others that adopts a feature-based representation can be found in Feder, Leonard and Smith (1999), Castellanos and Tardós (1999), Dissanayake, Newman, Durrant-Whyte, Clark and Csorba (2001) and Guivant and Nebot (2001). Alternative approaches include the work of Lu and Milios (1997), Gutmann and Konolige (1999) and Thrun (2001). These methods do not need to explicitly associate individual measurements with features in the environment, but seem to rely on the high quality of laser scanner data. Thrun (2001) writes “It is unclear how the performance of our approach degrades with inaccuracy of the sensors. For example, it is unclear if sonar sensors are sufficiently accurate to yield good results”.

The scientific motivation behind our research is to develop new methods for sensor data interpretation that can cope with sparse data sets that contain a large amount of outliers and that are contaminated by significant levels of navigation uncertainty. Attempting CML with range-only measurements, instead of using advanced sensor arrays (Chong and Kleeman 1999b), forces one to confront challenging issues such as partial observability. We believe that this line of investigation can lead to algorithms that will apply in more general situations, such as multi-robot mapping of complex natural environments.

1.2 The Sonar Mapping Problem

Most would agree that mobile robot navigation and mapping in indoor environments is far more difficult to perform with sonar than with laser data. Figure 1 provides a comparison of data from a SICK laser scanner and a ring of 24 sonar sensors taken at a single position in a typical environment. The lack of information in the sonar data in comparison to the laser data is evident: only half of the Polaroid sensors obtain a return, with a high proportion of outliers. As a result, the underlying structure of the scene is less visually apparent to a human observer. Despite the increased difficulty of sonar interpretation, we feel that it is interesting to perform research with sonar for a variety of reasons. From the perspective of cost, laser scanners are much more expensive than sonar sensors. From the perspective of basic science, questions such as the basic mechanisms of bat and dolphin echolocation are highly important (Au 1993). Finally, the

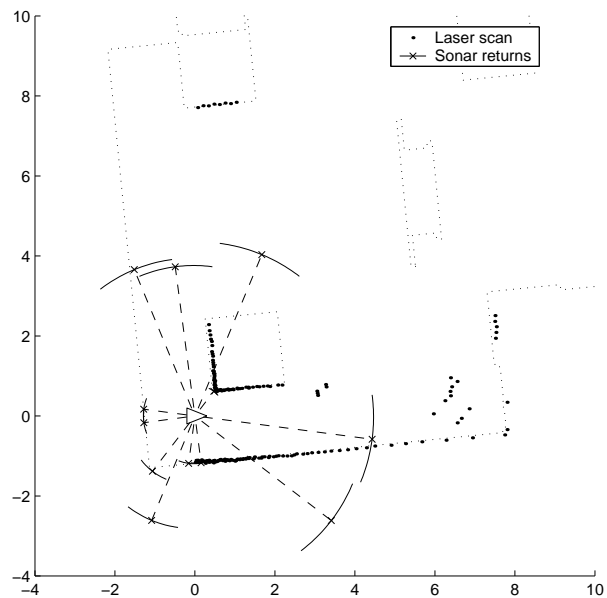


Figure 1: Sensor information obtained from a single robot position in a typical environment using a 180 degree SICK laser scanner and a ring of 24 Polaroid sonar sensors. The true environment map, obtained by hand, is shown in dotted lines.

fundamental characteristics of high range accuracy and wide beam-width are shared with many types of sonars of which underwater mapping sonars are a prime example. Our long-term goal is to develop interpretation methods that are applicable to a wide variety of different types of sonars and environments.

The information given by laser sensors in a single scan is quite dense and has good angular precision, usually better than one degree (see figure 1). Feature-based approaches to CML using laser data typically perform the data association task in two steps:

1. A segmentation step, where the laser returns in each scan are grouped and processed to obtain simple geometric features such as lines corresponding to walls (Castellanos, Montiel, Neira and Tardós 1999) or circles corresponding to trees (Guivant and Nebot 2001). Also more complex features as corners or door frames can be easily obtained from laser data (Castellanos, Neira and Tardós 2001).
2. During the map building process, a second data association step looks for matches between the features obtained from different scans, based on a probabilistic model of the sensor and of the vehicle motion.

Other approaches do not rely on the assumption that a certain type of geometric feature will be present in the environment, and use raw laser data in the

map building process. The data association is solved implicitly by computing the robot locations that maximize scan-to-scan (Lu and Milios 1997, Gutmann and Konolige 1999) or scan-to-map (Thrun 2001) correlations. In these approaches, the density and precision of the laser data is central in achieving robustness in the scan matching process.

In contrast, information given by a ring of sonar sensors is sparse and its angular precision is limited by the large sonar beam-width, typically in the range of 20 to 30 degrees. In man-made environments, most surfaces are specular reflectors for sonar (Kuc and Siegel 1987). This has two undesirable effects: only surfaces whose perpendicular is inside the sonar beam are usually detected, and many ghost returns are obtained as a result of multiple specular reflections. A rotating sonar scanner has been used by Leonard and Durrant-Whyte (1992) to obtain regions of constant depth (RCDs), allowing to determine with good precision the location of simple features as planes, corners, edges and cylinders. However, from a single scan, the RCDs produced by edges and corners are indistinguishable from the RCDs produced by planes.

One avenue for research to overcome difficulties with sonar data is to develop advanced sensor arrays that allow feature discrimination and precise localization from a single vehicle location. For example, Barshan and Kuc (1990) developed an intelligent sonar sensor that was able to disambiguate plane and corner reflectors using amplitude and travel time information. Peremans, Audenaert and Van (1993) developed a trinaural sensor that could measure object curvature. One of the most notable land robot sonars in recent years was developed by Kleeman and Kuc (1995), who were able to achieve remarkably precise localization and classification of targets from a rotating sensor array. This sensor was used in a CML implementation to perform large-scale CML using multiple submaps (Chong and Kleeman 1999a, Chong and Kleeman 1999b). All of these advanced sensors are based on the use of several sonar transducers in a careful configuration with known baselines. The different echoes received by several transducers from a single target allow the determination of the feature type and its precise location. With accurate knowledge of the relative positions of transducers, a variety of powerful array processing algorithms already developed for radar and underwater sonar become possible (Johnson and Dudgeon 1993).

In some ways, the use of multiple transducers in a sonar array is similar to the use of several cameras in classical stereo vision: the use of different viewpoints with precisely calibrated baselines allows the computation of depth information to permit estimation of the 3D structure of the environment (Faugeras 1993). However, a great deal of research in the last decade has shown that, using a single moving camera, it is possible to determine both the camera motion and the environment structure (Hartley and Zisserman 2000, Faugeras and Luong 2001). Ortín and Montiel (2001) present an interesting example of the use of these techniques to determine the 2D robot motion in an indoor environment using robust techniques to find matchings in the sequence of images. This parallelism with vision techniques raises a fundamental question: is it possible to determine the vehicle motion and the environment structure with a sonar scanner or a sonar ring? In this paper, we aim to pose the problem of interpretation of range-

only data from multiple uncertain vantage points, and to present an effective solution, using the Hough transform to identify points and line segments, for the common scenario of a land robot equipped with a ring of Polaroid sensors.

A somewhat related technique called triangulation-based fusion (TBF) has been developed by Wijk and Christensen (2000) for point objects only. In TBF, the basic operations are the computation of circle intersection points and application of angle constraints for multiple sonar returns obtained from adjacent positions as the robot moves. This approach has been documented with extensive experimental results, including grid-based mapping, continuous localization using Kalman filtering, and absolute relocation using particle filters in a large indoor environment (Wijk 2001). In addition, the method has been used for CML with point features (Zunino and Christensen 2001). We feel that the Hough transform approach presented in this paper offers advantages over triangulation-based fusion because it can directly identify specular planar reflectors from sonar data, which is vitally important in typical man-made environments with many smooth walls.

1.3 Mapping Large Scale Environments

Stochastic mapping techniques to solve the continuous mapping and localization problem were first introduced by Smith and Cheeseman (1986). The classical technique for updating the state vector of a stochastic map and its covariance uses EKF. Given that the full covariance matrix must be maintained to assure consistency, updating a full stochastic map of n features is well known to be $O(n^2)$ (Castellanos et al. 1999, Guivant and Nebot 2001). Many recent efforts have concentrated on reducing this computational complexity. Decoupled Stochastic Mapping (Leonard and Feder 2000, Jensfelt 2001), the Local Mapping Algorithm (Chong and Kleeman 1999b), Suboptimal SLAM (Guivant and Nebot 2001), and the Sparse Weight Filter (Julier 2001), reduce the computational complexity of map updating to $O(1)$ the first two and $O(n)$ the others, by obtaining a pessimistic solution. In contrast, Postponement (Davison 1998, Knight, Davison and Reid 2001) and the Compressed Filter (Guivant and Nebot 2001) delay the global map update, obtaining a suboptimal solution (subject to linearization approximations made by EKF) and reducing the amount of computations by a constant factor. With these methods, the order of growth of complexity is still $O(n^2)$.

In this work we propose a new solution, denominated *Map Joining*, a general procedure to consistently convey the information of several independent stochastic maps. Map joining is a very useful tool in many situations, including:

- *Robot relocation*. In this problem, also known as the First Location problem, or the Kidnapped Robot problem, a global stochastic map of the environment where the vehicle is navigating is available, but the location of the vehicle is completely unknown. In this case, the vehicle can build a local map of the environment, which will be independent from the global

map. If we identify a feature in the local map whose location is available in the global map, we can join the two maps. The result will be a global map which includes the vehicle location.

- *Multirobot map building.* We can deploy a team of vehicles to independently build maps of different areas of an environment. Whenever we can determine that two vehicles have included the same environment feature in their maps, we can join their maps into a full map that also includes the location of one vehicle with respect to the other.
- *Local map sequencing.* In this paper we concentrate on showing that map joining can be used to obtain a full, consistent global map of any size from a sequence of independent local stochastic maps of limited size. First, we use classical stochastic mapping techniques to build local maps along limited portions of the robot trajectory. Then, we join each local map with the global map, find matchings between global and local features, and fuse them to update the global map. The computational complexity of Local Map Sequencing is equivalent to that of the Compressed Filter and Postponement, with the additional advantage of updating local maps where errors remain small, therefore reducing the harmful effects of linearization errors. During the last steps of writing this paper we have become aware of the independent development by Williams (2001) of the Constrained Local Submap Filter, a map building technique equivalent to Local Map Sequencing.

The structure of this paper is as follows: Section 2 explores the minimal necessary conditions allowing the determination of robot motion and environment structure from range-only sonar data, and presents a new technique for classification of points and line segments from sonar data acquired from multiple vantage points, using the Hough transform. Section 3 presents a new technique for joining and updating two submaps when performing CML, providing increased computational efficiency and robustness. Section 4 illustrates the application of these techniques to data acquired by a mobile robot in an environment containing several loops, with spurious data from people walking near the robot. A side-by-side comparison is provided between the map built from sonar and a map built from laser scanner data. Finally, Section 5 draws conclusions and discusses future research topics.

2 Robust map building with sonar

2.1 The structure and motion problem

In this section we will analyze the conditions under which a sonar scanner or a sonar ring mounted on a vehicle gives enough information to determine both the vehicle motion and the environment structure. Finding a general solution to this problem is an interesting research topic, beyond the scope of this work.

Our goal is to determine the potential of using the sonar data to solve the CML problem, where an estimation of robot motion is available through the use of odometry sensors.

For the purpose of this analysis we will first consider the case where data association is solved, i.e. we know which environment feature has given each sonar return. Assume a robot moves along k locations, and at each location observes n geometric features with some of the sensors in a sonar ring. Let us call r the degrees of freedom (d.o.f) of robot motion and f the d.o.f. that determine each feature location. Without loss of generality, we will use the first robot location as a reference system. The number of unknowns is $r(k-1)$ for the robot motion, plus nf for the feature locations. Given the high angular uncertainty of the sonar returns, we will only consider as data the nk range values given by the sonar sensor from each robot location to each feature. Provided that no degenerate situations arise, the system is solvable to find robot motion and feature locations iff:

$$nk \geq r(k-1) + nf \tag{1}$$

i.e., iff:

$$n \geq r \frac{(k-1)}{(k-f)}$$

or

$$k \geq \frac{(fn-r)}{(n-r)}$$

In particular, we can also state two simple necessary conditions:

$$k > f$$

$$n > r$$

For example, consider the particular case of a robot moving in 2D ($r = 3$), observing the distance to lines or points ($f = 2$). If data association is known, it is possible to compute robot motion and feature location if the robot observes 4 features for at least 5 steps.

Of course, with standard sonar sensors, the difficult part is to solve the data association problem: finding groups of returns coming from the same feature and determining the feature type. When there is enough sensor information (the $>$ condition in eq. (1) holds), the excess data can be used in a hypothesis-verification scheme to obtain data association, and solve the structure and motion problem.

2.2 The CML problem

In this paper, we will concentrate on solving the CML problem: given a set of robot motions, measured by odometry with some uncertainty, and a set of sonar

measurements, solve the data association problem and obtain the structure of the environment together with a better estimation of the robot trajectory.

If the robot motion and data association were perfectly known, from eq. (1), the condition to solve the structure problem would be:

$$k \geq f \tag{2}$$

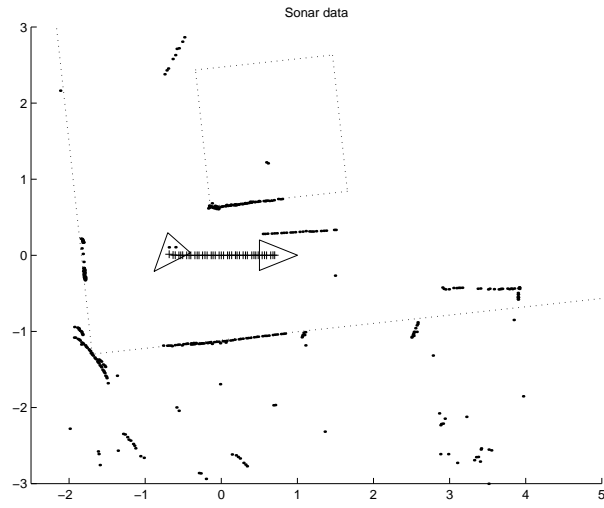
However, in order to robustly solve the data association problem, we need to obtain sensor data from a number of robot positions greater than the limit defined by eq. (2). When a feature is successfully matched for more than f robot locations, the excess data will be used to reduce the uncertainty in robot and feature location.

When the amount of sensor data successfully associated is large enough to also satisfy eq. (1), the robot motion can be completely determined from sonar data. In this case robot uncertainty will approach sensor precision. Otherwise, the robot uncertainty will only be marginally better than the bare odometric uncertainty.

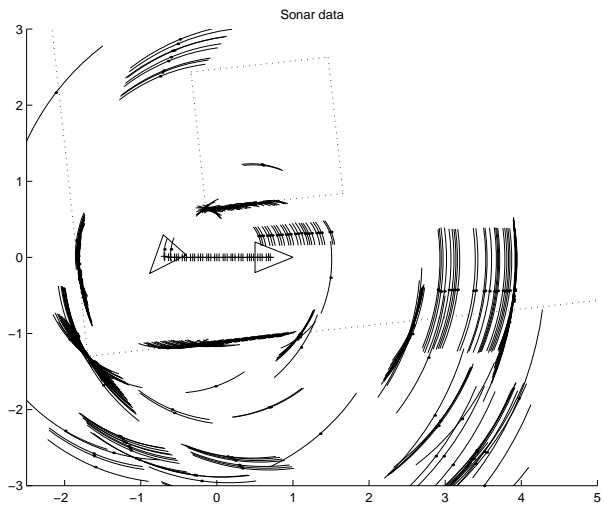
2.3 Sonar modelling

The above analysis suggests that by moving the robot several steps and using robot odometry, it should be possible to group sonar returns and determine target type. An example data set obtained with a 24 sonar ring along a simple motion (first a turn and then a straight motion) is shown in figure 2a. Each point represents the nominal position of a sonar return, computed along the central axis of the transducer, using the robot odometry. The dotted lines represent the ground truth environment map. The walls at both sides of the robot trajectory clearly show up as lines of sonar returns. However, some other false lines appear in front of the robot. Usually these false lines arise as a result of people moving in the robot’s environment, but also from some sonar artifacts such as sensor cross-talk. The lower left corner appears as an ambiguous arc of sonar returns. Other corners or edges appear as small lines of returns. Also, many phantom returns arise from multiple specular reflections. Simple segmentation techniques using only the position of the nominal sonar returns (Grossmann and Poli 2001) can successfully find groups of returns coming from the same target (real or phantom), but they will surely fail to distinguish between real and false targets and to identify the feature type.

On the contrary, if we use a more detailed sensor model, ambiguity can be significantly reduced. Figure 2b shows the same data set where each return is depicted with an arc showing the sonar angular uncertainty, i.e. showing all the possible locations of the object surface actually producing the sonar return. In this more meaningful representation, walls appear as sets of sonar arcs tangent to the wall, while point features, such as corners or edges, appear as sets of arcs intersecting at the feature location. In contrast, false features produce sets of incoherent sonar arcs and thus can be easily spotted. The conclusion from this



(a)



(b)

Figure 2: Sonar returns along 40 robot positions, with a reference map superimposed. (a) Corners and walls appear as arcs and lines of sonar returns, respectively. But also false lines can be seen as a result of people walking around the robot, specular reflections and other sonar artifacts. (b) When sonar returns are depicted with arcs showing its angular uncertainty, walls clearly show up as sets of tangent arcs, and corners as sets of intersecting arcs.

example is compelling and well known: using a careful sensor model is crucial to adequately interpret sonar returns.

This work is restricted to man-made indoor environments, and we will use two types of geometric features, extracted from the work of Leonard and Durrant-Whyte (1992): 2D lines to represent walls, and 2D points to represent corners and edges (figure 3). A wall, provided that its surface is smooth, will only produce a sonar return for sensor S_j when the perpendicular to the wall is contained in the sonar emission cone:

$$\theta_{S_j} - \frac{\beta}{2} \leq \theta_k^B \leq \theta_{S_j} + \frac{\beta}{2}$$

where β is the visibility angle for walls. The distance actually given by the sensor ρ^{S_j} will correspond to the perpendicular distance from the sensor to the line. Conversely, given the sensor location and the measured distances, all possible lines giving such a return are tangent to the arc depicted in figure 3a.

In a similar way, a point feature like a corner or an edge will produce a sonar return for sensor S_j when it is located inside the sonar emission cone. The distance measured will correspond to the distance between the point and the sensor. All possible points giving the same return are located along the arc depicted in figure 3b, where in this case β is the visibility angle for point features. Although the visibility angles for corners or for edges may be different (Leonard and Durrant-Whyte 1992), in this work we do not attempt to distinguish them, and we use the same value of β for all point features.

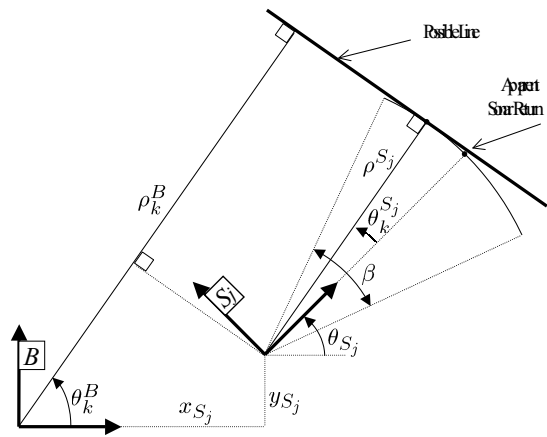
2.4 Sonar data association with Hough Transform

With the sonar model presented above, associating sonar returns to line and point features may be stated as finding groups of sonar arcs all tangent to the same line, and groups of sonar arcs intersecting on the same point, respectively. Given the large amount of spurious data coming from moving people, specular reflections and sonar artifacts, classical robust techniques such as RANSAC (Fischler and Bolles 1981, Hartley and Zisserman 2000) or the Hough transform (Ballard and Brown 1982, Illingworth and Kittler 1988) seem very appropriate.

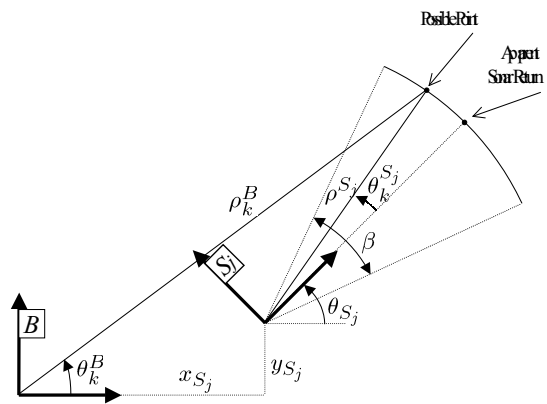
The Hough transform is a voting scheme where each piece of sensor information accumulates evidence about the presence of certain features compatible with the actual measurement. Voting is performed in a discretized parametric space, known as the Hough space, representing all possible feature locations. The most voted cells in the Hough space should correspond to the features actually present in the environment. By keeping track of the votes, it is very easy to obtain the groups of sensor data coming from each feature detected.

We have found that this technique is particularly well suited to efficiently solve the 2D sonar data association problem, for the following reasons:

- The location of point and line features can be easily described with two parameters, giving a 2D Hough space in which the voting process and the search for maxima can be done quite efficiently.



(a)



(b)

Figure 3: Model of sonar sensor for (a) line features and (b) point features

```

for i := 1 to num_positions do
  for j := 1 to num_sensors do
    compute sensor location  $\mathbf{x}_{S_j}^B = (x_{S_j}, y_{S_j}, \theta_{S_j})^T$ 
    for  $\theta_k^{S_j} := -\beta/2$  to  $\beta/2$  step  $\delta_\theta$  do
      ;compute line parameters and vote
       $\theta_k^B := \theta_{S_j} + \theta_k^{S_j}$ ;
       $\rho_k^B := \rho_{S_j} + x_{S_j} \cos(\theta_k^{S_j}) + y_{S_j} \sin(\theta_k^{S_j})$ ;
      vote( $\theta_k^B, \rho_k^B, i, j$ );
    od
  od
od

```

Figure 4: Basic Hough voting algorithm for sonar lines

- The sonar model presented above can be used to restrict the votes generated by each sonar return to be located along the corresponding transformed sonar arc.
- Since each sonar return emits a constant number of votes, the whole Hough process is linear with the number of returns processed.
- Being a voting scheme, it is intrinsically very robust against the presence of many spurious sonar returns.

One of the key issues of its practical implementation is choosing the parameters defining the Hough space and their quantization (Illingworth and Kittler 1988). In our implementation, the Hough transform is applied to the sonar returns obtained along short trajectories (around 1-2m of total travel), in order to keep the odometry errors small. Lines are represented in a base reference B , located in the central robot position, using parameters θ^B and ρ^B defining the line orientation and its distance to the origin (see figure 3a). In the case of points, we have chosen to use a polar representation relative to the same reference B (see figure 3b), instead of a cartesian representation, because the discretization of this Hough space represents more accurately the precision obtained in point location from range sensors. This also has the advantage that the line and point Hough spaces are very similar. The quantization of the Hough voting table was tuned to approximately match the typical accumulated odometry errors during the trajectory (around 3-5cm in position and 2-5deg in orientation).

With this definition of the Hough space, the basic voting algorithm for lines is shown in figure 4, where δ_θ is the angular quantization of the Hough space. The voting algorithm for the case of points can be easily derived from that of lines by changing the geometric definition of ρ^B and θ^B as illustrated in Figure 3b. Figure 5 (left) shows the resulting voting tables for the example trajectory of figure 2.

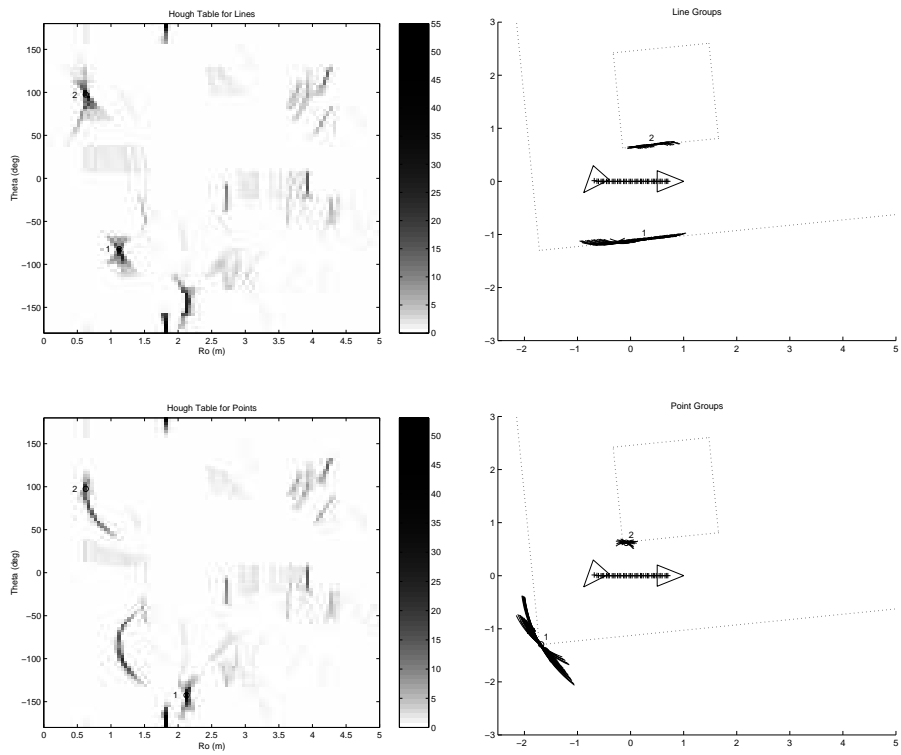


Figure 5: Hough voting tables for lines and points, and groups of sonar returns corresponding to the most voted line and point features

Next, the two voting tables are searched for local maxima having a number of votes above a certain threshold. However, in real indoor environments, an ambiguity usually arises: some groups of sonar returns give maxima in both tables, meaning that they could be interpreted as a line or as a point feature. An typical example can be seen in the figure around $\rho = 2.1$ and $\theta = -150deg$. This problem can be successfully solved by sorting all maxima by the number of votes received, and using a winner-takes-all strategy. In our example, the point hypothesis wins and gets all votes, discarding the line hypothesis. A different case can be seen at $\rho = 0.6$ and $\theta = 90deg$, where the line hypotheses wins and takes some votes, but there are still enough sonar returns left to also obtain a point hypothesis. However, there are some singular cases where the ambiguity between point and line cannot be decided: a sonar sensor moving perpendicular to a wall, or in a straight line towards a point would get exactly the same returns. To avoid this problem a further verification is made: to accept a point or line feature we require to have sonar returns taken from sensor positions spread around the point or along the line, respectively.

In figure 5 we can see the point and line features identified, with their corresponding sonar data groups. Despite the large amount of spurious data, our method was able to successfully detect two line features corresponding to walls and two point features corresponding to a corner and an edge. This perceptual data grouping technique yields a *robust* interpretation of the sonar data in terms of point and line features, in the sense of it being insensitive to the presence of spurious measurements.

3 Consistent Stochastic Mapping using Sequences of Local Maps

In classical stochastic mapping, the environment information related to a set of elements $\mathcal{F} = \{B, F_0, F_1, \dots, F_m\}$ is represented by a global map $\mathcal{M}_{\mathcal{F}}^B = (\hat{\mathbf{x}}_{\mathcal{F}}^B, \mathbf{P}_{\mathcal{F}}^B)$, where:

$$\hat{\mathbf{x}}_{\mathcal{F}}^B = \begin{bmatrix} \hat{\mathbf{x}}_{F_0}^B \\ \vdots \\ \hat{\mathbf{x}}_{F_m}^B \end{bmatrix}; \mathbf{P}_{\mathcal{F}}^B = \begin{bmatrix} \mathbf{P}_{F_0 F_0}^B & \cdots & \mathbf{P}_{F_0 F_m}^B \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{F_m F_0}^B & \cdots & \mathbf{P}_{F_m F_m}^B \end{bmatrix} \quad (3)$$

The state vector $\hat{\mathbf{x}}_{\mathcal{F}}^B$ contains the estimated location of the vehicle F_0 and the environment features $F_1 \dots F_m$, all with respect to a base reference B . In the case of the vehicle, its location vector $\hat{\mathbf{x}}_{F_0}^B$ describes the transformation from B to F_0 . In the case of an environment feature f , the parameters that compose its location vector $\hat{\mathbf{x}}_{F_f}^B$ depend on the feature type (see appendix A). Matrix $\mathbf{P}_{\mathcal{F}}^B$ is the estimated error covariance of $\hat{\mathbf{x}}_{\mathcal{F}}^B$.

Instead of building one global map from the beginning, we propose to build a sequence of local maps of limited size, and later join them together, to obtain the global map. This process is explained next.

3.1 Local map building

Each local map can be built as follows: at a given instant t_0 , a new map is initialized using the current vehicle location as base reference B . Then, the vehicle performs a limited motion (say p steps) acquiring sensor information about m environment features. At each step k , the local map is updated with standard stochastic mapping techniques (Castellanos and Tardós 1999), using the vehicle motion estimated by odometry, the sensor observations, and a data association hypothesis, relating measurements with map features.

For sonar data, classification and assignment for measurements become much easier if it is possible to consider sets of data obtained from multiple vantage points. Better performance will be achieved if one can delay difficult decisions until enough redundant information has been obtained to make the decisions easier. In this work, we perform a Hough transform along the vehicle locations of each local map to associate observations to map features. The EKF prediction and updating algorithms are run after the vehicle has completed the motion, to estimate its current location and the location of environment features. The detailed formulation of our approach includes three steps:

1. Map Initialization: as the base B of the local stochastic map is the vehicle location before the first motion, the map for time t_0 is initialized with perfect knowledge of the vehicle location: $\hat{\mathbf{x}}_{R_0}^B = (0, 0, 0)^T$, $P_{R_0 R_0}^B = \mathbf{0}$. The Hough transform provides initial estimates for the m feature locations $\hat{\mathbf{x}}_{F_{j,0}}^B$, which are incorporated into the local map:

$$\hat{\mathbf{x}}_{\mathcal{F}_0}^B = \begin{bmatrix} \hat{\mathbf{x}}_{R_0}^B \\ \hat{\mathbf{x}}_{F_{1,0}}^B \\ \vdots \\ \hat{\mathbf{x}}_{F_{m,0}}^B \end{bmatrix}; \mathbf{P}_{\mathcal{F}_0}^B = \begin{bmatrix} P_{R_0 R_0}^B & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & P_{F_{m,0} F_{m,0}}^B \end{bmatrix} \quad (4)$$

As it is standard EKF practice, in order to avoid optimistic estimations, we use a non-informative prior, associating a large initial covariance $P_{F_{j,0} F_{j,0}}^B$ to each feature. This means that the prior estimations given by the Hough transform are only used as initial values for the linearization of the EKF equations. As their precision is quite good, linearization errors will be small.

2. Robot motion: the vehicle motion from position $k - 1$ to position k is estimated by odometry:

$$\begin{aligned} \mathbf{x}_{R_k}^{R_{k-1}} &= \hat{\mathbf{x}}_{R_k}^{R_{k-1}} + \mathbf{v}_k \\ E[\mathbf{v}_k] &= \mathbf{0} \\ E[\mathbf{v}_k \mathbf{v}_j^T] &= \delta_{kj} \mathbf{Q}_k \end{aligned} \quad (5)$$

where \mathbf{v}_k (process noise) is assumed to be additive, zero-mean and white. That is, all the \mathbf{v}_k are mutually independent (Jazwinski 1970). After this motion, the robot location will be:

$$\mathbf{x}_{R_k}^B = \mathbf{x}_{R_{k-1}}^B \oplus \mathbf{x}_{R_k}^{R_{k-1}} \quad (6)$$

where \oplus represents the composition of transformations (see appendix A). Thus, given the local map $\mathcal{M}_{\mathcal{F}_{k-1}}^B = (\hat{\mathbf{x}}_{\mathcal{F}_{k-1}}^B, \mathbf{P}_{\mathcal{F}_{k-1}}^B)$ at step $k-1$, the predicted local map $\mathcal{M}_{\mathcal{F}_{k|k-1}}^B$ at step k is obtained as follows:

$$\begin{aligned} \hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B &= \begin{bmatrix} \hat{\mathbf{x}}_{R_{k-1}}^B \oplus \hat{\mathbf{x}}_{R_k}^{R_{k-1}} \\ \hat{\mathbf{x}}_{F_{1,k-1}}^B \\ \vdots \\ \hat{\mathbf{x}}_{F_{m,k-1}}^B \end{bmatrix} \\ \mathbf{P}_{\mathcal{F}_{k|k-1}}^B &= \mathbf{F}_k \mathbf{P}_{\mathcal{F}_{k-1}}^B \mathbf{F}_k^T + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^T \end{aligned} \quad (7)$$

where:

$$\begin{aligned} \mathbf{F}_k &= \begin{bmatrix} \mathbf{J}_{1\oplus} \left\{ \hat{\mathbf{x}}_{R_{k-1}}^B, \hat{\mathbf{x}}_{R_k}^{R_{k-1}} \right\} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & & \vdots \\ \vdots & & \ddots & \\ \mathbf{0} & \cdots & & \mathbf{I} \end{bmatrix} \\ \mathbf{G}_k &= \begin{bmatrix} \mathbf{J}_{2\oplus} \left\{ \hat{\mathbf{x}}_{R_{k-1}}^B, \hat{\mathbf{x}}_{R_k}^{R_{k-1}} \right\} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \end{aligned}$$

where $\mathbf{J}_{1\oplus}$ and $\mathbf{J}_{2\oplus}$ are the Jacobians of transformation composition (appendix A).

3. Feature observations: at instant k the sonar sensors give distance measurements $\mathbf{z}_{k,i}$ with $i = 1 \dots s$. The Hough transform gives a hypothesis $\mathcal{H}_k = [j_1, j_2, \dots, j_s]$ associating each sonar return i with its corresponding feature F_{j_i} ($j_i = 0$ indicates that return $\mathbf{z}_{k,i}$ has been considered spurious). For each associated measurement, the theoretic distance from the sensor i to feature F_{j_i} is a nonlinear function \mathbf{h}_{ij_i} of the vehicle and feature location, that are contained in the map state vector $\mathbf{x}_{\mathcal{F}_k}^B$. The full set of associated features is given by:

$$\begin{aligned}
\mathbf{z}_k &= \mathbf{h}_k(\mathbf{x}_{\mathcal{F}_k}^B) + \mathbf{w}_k \\
\mathbf{h}_k &= \begin{bmatrix} \mathbf{h}_{1j_1} \\ \mathbf{h}_{2j_2} \\ \vdots \\ \mathbf{h}_{sj_s} \end{bmatrix}
\end{aligned} \tag{8}$$

where \mathbf{w}_k (measurement noise) is assumed to be additive, zero-mean, white and independent of the process noise:

$$\begin{aligned}
E[\mathbf{w}_k] &= \mathbf{0} \\
E[\mathbf{w}_k \mathbf{w}_j^T] &= \delta_{kj} \mathbf{R}_k \\
E[\mathbf{w}_k \mathbf{v}_j^T] &= \mathbf{0}
\end{aligned} \tag{9}$$

Linearization around the current map estimate yields:

$$\begin{aligned}
\mathbf{z}_k &\simeq \mathbf{h}_k(\hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B) + \mathbf{H}_k(\mathbf{x}_{\mathcal{F}_k}^B - \hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B) \\
\mathbf{H}_k &= \left. \frac{\partial \mathbf{h}_k}{\partial \mathbf{x}_{\mathcal{F}_k}^B} \right|_{(\hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B)}
\end{aligned} \tag{10}$$

Measurement \mathbf{z}_k is furthermore verified using Joint Compatibility (Neira and Tardós 2001), and used to obtain a new estimation of the state using the standard EKF update equations:

$$\begin{aligned}
\hat{\mathbf{x}}_{\mathcal{F}_k}^B &= \hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B + \mathbf{K}_k(\mathbf{z}_k - \mathbf{h}_k(\hat{\mathbf{x}}_{\mathcal{F}_{k|k-1}}^B)) \\
\mathbf{P}_{\mathcal{F}_k}^B &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{\mathcal{F}_{k|k-1}}^B \\
\mathbf{K}_k &= \mathbf{P}_{\mathcal{F}_{k|k-1}}^B \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{\mathcal{F}_{k|k-1}}^B \mathbf{H}_k^T + \mathbf{R}_k)^{-1}
\end{aligned} \tag{11}$$

In this way, after processing the p steps, the resulting stochastic map $\mathcal{M}_{\mathcal{F}_p}^B$ includes an estimation of the feature locations, as well as their correlations. It also includes an estimation of the current vehicle location, where odometry errors have been corrected by the use of sonar returns.

Two important properties of this local map building process should be pointed out:

1. The local map $\mathcal{M}_{\mathcal{F}_p}^B$ is independent of any prior estimation of the vehicle location because it is built relative to the initial vehicle location $B = R_0$.

2. Map $\mathcal{M}_{\mathcal{F}_p}^B$ depends only on the sequence of odometry readings \mathbf{u}_k and sensor data \mathbf{z}_k obtained during the p steps:

$$\begin{aligned} D^{1\dots p} &= \{\mathbf{u}_1 \mathbf{z}_1 \dots \mathbf{u}_p \mathbf{z}_p\} \\ \mathbf{u}_k &= \hat{\mathbf{x}}_{R_k}^{R_{k-1}} \end{aligned} \quad (12)$$

and the data association hypotheses:

$$\mathcal{H}^{1\dots p} = \{\mathcal{H}_1 \dots \mathcal{H}_p\} \quad (13)$$

Strictly speaking, given that the motion and measurement equations (6) (8) are nonlinear, the EKF is the best linear minimum mean-square error (MMSE) estimator (Bar-Shalom and Fortmann 1988), that gives an approximation of the conditional mean:

$$\hat{\mathbf{x}}_{\mathcal{F}_p}^B \simeq E \left[\mathbf{x}_{\mathcal{F}_p}^B \mid D^{1\dots p}, \mathcal{H}^{1\dots p} \right] \quad (14)$$

It is fairly clear that two independent robots mapping the same environment will produce two statistically independent maps. Under the common assumption that process and measurement noise are *white* random sequences, the above two properties assure that two local maps built with the same robot from *disjoint* sequences of steps are functions of independent stochastic variables. Therefore, the two maps will be *statistically independent* and *uncorrelated* (Papoulis 1991).

This fact has a very important consequence: during local map building, we do not need to compute the correlations between features in the current local map and features in any other local map, because they are known to be zero by construction. Thus, the cost of local map building is independent from the size of the global map.

3.2 Changing the base reference of a stochastic map

Suppose that we choose to change the base reference of a map $\mathcal{M}_{\mathcal{F}}^B = (\hat{\mathbf{x}}_{\mathcal{F}}^B, \mathbf{P}_{\mathcal{F}}^B)$ from B to F_j , where reference F_j is associated to the vehicle, or to any map feature. Using the composition and inversion operations (see appendix A), the resulting state vector would be:

$$\hat{\mathbf{x}}_{\mathcal{F}}^{F_j} = \begin{bmatrix} \hat{\mathbf{x}}_{F_0}^{F_j} \\ \vdots \\ \hat{\mathbf{x}}_B^{F_j} \\ \vdots \\ \hat{\mathbf{x}}_{F_m}^{F_j} \end{bmatrix} = \begin{bmatrix} \ominus \hat{\mathbf{x}}_{F_j}^B \oplus \hat{\mathbf{x}}_{F_0}^B \\ \vdots \\ \ominus \hat{\mathbf{x}}_B^B \\ \vdots \\ \ominus \hat{\mathbf{x}}_{F_j}^B \oplus \hat{\mathbf{x}}_{F_m}^B \end{bmatrix} \quad (15)$$

To make this operation reversible, we have incorporated the location of the former base reference B with respect to F_j in the new state vector, replacing the location of feature F_j . The covariance of $\mathbf{x}_{\mathcal{F}}^{F_j}$ can be obtained as follows:

$$\mathbf{P}_{\mathcal{F}}^{F_j} = \mathbf{J}_B^{F_j} \mathbf{P}_{\mathcal{F}}^B \mathbf{J}_B^{F_j T} \quad (16)$$

where:

$$\mathbf{J}_B^{F_j} = \left. \frac{\partial \mathbf{x}_{\mathcal{F}}^{F_j}}{\partial \mathbf{x}_{\mathcal{F}}^B} \right|_{(\hat{\mathbf{x}}_{\mathcal{F}}^B)} = \begin{bmatrix} \mathbf{J}_{00} & \cdots & \mathbf{J}_{0j} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{J}_{jj} & \cdots & \mathbf{0} \\ \vdots & & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{J}_{mj} & \cdots & \mathbf{J}_{mm} \end{bmatrix}$$

$$\begin{aligned} \mathbf{J}_{jj} &= \mathbf{J}_{\ominus} \left\{ \hat{\mathbf{x}}_{F_j}^B \right\} \\ \mathbf{J}_{ij} &= \mathbf{J}_{1\oplus} \left\{ \ominus \hat{\mathbf{x}}_{F_j}^B, \hat{\mathbf{x}}_{F_i}^B \right\} \mathbf{J}_{\ominus} \left\{ \hat{\mathbf{x}}_{F_j}^B \right\} \quad i = 0..m, i \neq j \\ \mathbf{J}_{ii} &= \mathbf{J}_{2\oplus} \left\{ \ominus \hat{\mathbf{x}}_{F_j}^B, \hat{\mathbf{x}}_{F_i}^B \right\} \quad i = 0..m, i \neq j \end{aligned}$$

A particular case of this transformation, which produces a stochastic map relative to the vehicle reference ($j = 0$), was proposed by Castellanos and Tardós (1999).

3.3 Local Map Joining

Given two uncorrelated local maps:

$$\begin{aligned} \mathcal{M}_{\mathcal{F}}^B &= (\hat{\mathbf{x}}_{\mathcal{F}}^B, \mathbf{P}_{\mathcal{F}}^B) \quad ; \quad \mathcal{F} = \{B, F_0, F_1, \dots, F_n\} \\ \mathcal{M}_{\mathcal{E}}^{B'} &= (\hat{\mathbf{x}}_{\mathcal{E}}^{B'}, \mathbf{P}_{\mathcal{E}}^{B'}) \quad ; \quad \mathcal{E} = \{B', E_0, E_1, \dots, E_m\} \end{aligned}$$

where a common reference has been identified:

$$F_i = E_j,$$

the goal of map joining is to obtain one full stochastic map:

$$\mathcal{M}_{\mathcal{F}+\mathcal{E}}^B = (\hat{\mathbf{x}}_{\mathcal{F}+\mathcal{E}}^B, \mathbf{P}_{\mathcal{F}+\mathcal{E}}^B)$$

containing the estimations of the features from both maps, relative to a common base reference B , and to compute the correlations appearing in the process. This can be done in two steps:

1. Change the base reference of map $\mathcal{M}_{\mathcal{E}}^{B'}$ from B' to E_j , using eqs. (15) and (16), obtaining:

$$\mathcal{M}_{\mathcal{E}}^{E_j} = (\hat{\mathbf{x}}_{\mathcal{E}}^{E_j}, \mathbf{P}_{\mathcal{E}}^{E_j})$$

2. Given that the features from the first map are expressed relative to reference B , to form the joint state vector $\mathbf{x}_{\mathcal{F}+\mathcal{E}}^B$ we only need to transform the features of the second map to reference B using the fact that $F_i = E_j$:

$$\mathbf{x}_{E_j}^B = \mathbf{x}_{F_i}^B \oplus \mathbf{x}_{E_j}^{E_j} \quad (17)$$

The estimation of the joined map is thus computed as follows:

$$\hat{\mathbf{x}}_{\mathcal{F}+\mathcal{E}}^B = \begin{bmatrix} \hat{\mathbf{x}}_{\mathcal{F}}^B \\ \hat{\mathbf{x}}_{\mathcal{E}}^B \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}_{\mathcal{F}}^B \\ \hat{\mathbf{x}}_{F_i}^B \oplus \hat{\mathbf{x}}_{E_0}^{E_j} \\ \vdots \\ \hat{\mathbf{x}}_{F_i}^B \oplus \hat{\mathbf{x}}_{E_m}^{E_j} \end{bmatrix} \quad (18)$$

The covariance $\mathbf{P}_{\mathcal{F}+\mathcal{E}}^B$ of the joined map is obtained from the linearization of eq. (17), and is given by:

$$\begin{aligned} \mathbf{P}_{\mathcal{F}+\mathcal{E}}^B &= \mathbf{J}_{\mathcal{F}} \mathbf{P}_{\mathcal{F}}^B \mathbf{J}_{\mathcal{F}}^T + \mathbf{J}_{\mathcal{E}} \mathbf{P}_{\mathcal{E}}^{E_j} \mathbf{J}_{\mathcal{E}}^T \\ &= \begin{bmatrix} \mathbf{P}_{\mathcal{F}}^B & \mathbf{P}_{\mathcal{F}}^B \mathbf{J}_1^T \\ \mathbf{J}_1 \mathbf{P}_{\mathcal{F}}^B & \mathbf{J}_1 \mathbf{P}_{\mathcal{F}}^B \mathbf{J}_1^T \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \mathbf{P}_{\mathcal{E}}^{E_j} \mathbf{J}_2^T \end{bmatrix} \end{aligned} \quad (19)$$

where:

$$\begin{aligned} \mathbf{J}_{\mathcal{F}} &= \left. \frac{\partial \mathbf{x}_{\mathcal{F}+\mathcal{E}}^B}{\partial \mathbf{x}_{\mathcal{F}}^B} \right|_{(\hat{\mathbf{x}}_{\mathcal{F}}^B, \hat{\mathbf{x}}_{\mathcal{E}}^{E_j})} = \begin{bmatrix} \mathbf{I} \\ \mathbf{J}_1 \end{bmatrix} \\ \mathbf{J}_{\mathcal{E}} &= \left. \frac{\partial \mathbf{x}_{\mathcal{F}+\mathcal{E}}^B}{\partial \mathbf{x}_{\mathcal{E}}^{E_j}} \right|_{(\hat{\mathbf{x}}_{\mathcal{F}}^B, \hat{\mathbf{x}}_{\mathcal{E}}^{E_j})} = \begin{bmatrix} \mathbf{0} \\ \mathbf{J}_2 \end{bmatrix} \\ \mathbf{J}_1 &= \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{J}_{1\oplus} \left\{ \hat{\mathbf{x}}_{F_i}^B, \hat{\mathbf{x}}_{E_0}^{E_j} \right\} & \cdots & \mathbf{0} \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{J}_{1\oplus} \left\{ \hat{\mathbf{x}}_{F_i}^B, \hat{\mathbf{x}}_{E_m}^{E_j} \right\} & \cdots & \mathbf{0} \end{bmatrix} \\ \mathbf{J}_2 &= \begin{bmatrix} \mathbf{J}_{2\oplus} \left\{ \hat{\mathbf{x}}_{F_i}^B, \hat{\mathbf{x}}_{E_0}^{E_j} \right\} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{J}_{2\oplus} \left\{ \hat{\mathbf{x}}_{F_i}^B, \hat{\mathbf{x}}_{E_m}^{E_j} \right\} \end{bmatrix} \end{aligned}$$

Obtaining vector $\hat{\mathbf{x}}_{\mathcal{F}+\mathcal{E}}^B$ with eq. (18) is an $O(m)$ operation. Given that the number of non-zero elements in \mathbf{J}_1 and \mathbf{J}_2 is $O(m)$, obtaining matrix $\mathbf{P}_{\mathcal{F}+\mathcal{E}}^B$ with eq. (19) is an $O(nm + m^2)$ operation. Thus when $n \gg m$, map joining is linear with n .

3.4 Matching and Fusion after Map Joining

The map resulting from map joining is statistically consistent, but may contain features that, coming from different local maps, correspond to the same environment feature. To eliminate such duplications and obtain a more precise map we need a data association algorithm to determine correspondences, and a feature fusion mechanism to update the global map. Determining correspondences between features in two successive local maps is a fairly simple problem, because the relative location uncertainty is small. Classical techniques, such as the Nearest Neighbor (Castellanos et al. 1999, Feder et al. 1999), will successfully solve this problem. However, to find matchings in loop closing situations, a more robust method is required.

We use the Joint Compatibility Branch and Bound (JCBB) algorithm (Neira and Tardós 2001). JCBB combines a branch-and-bound search technique with a powerful test to determine the joint compatibility of a set of matchings, taking all feature correlations into account. This method is guaranteed to provide the largest set of correspondences that are jointly consistent. By favoring the hypothesis with the largest consensus, robustness of data association is significantly improved. Experiments will show the superiority of this algorithm in loop closing situations.

Next, we explain the details of feature matching and fusion for global map updating. For simplicity, let $\mathbf{x} = \mathbf{x}_{\mathcal{F}+\mathcal{E}}^B$ and $\mathbf{P} = \mathbf{P}_{\mathcal{F}+\mathcal{E}}^B$ represent the state and covariance of the joined map. Let us consider also that references F_0 and E_0 represent the vehicle location in each corresponding map. Let $\mathcal{H} = \{j_1, \dots, j_m\}$ be a hypothesis that pairs each feature E_i coming from the local map with a feature F_{j_i} coming from the global map. When $j_i = 0$, feature E_i is considered new. The condition that features E_i and F_{j_i} coincide can be expressed by an ideal measurement equation without noise:

$$\mathbf{z}_i = \mathbf{h}_{i j_i}(\mathbf{x}) = \mathbf{0} \quad (20)$$

Thus, the state estimation can be updated using a modified version of the EKF update equations, with $\mathbf{z} = \mathbf{0}$ and noise covariance $\mathbf{R} = \mathbf{0}$. Since $\mathbf{h}_{i j_i}$ is usually non-linear, linearization around the current estimation is necessary:

$$\mathbf{h}_{i j_i}(\mathbf{x}) \simeq \mathbf{h}_{i j_i}(\hat{\mathbf{x}}) + \mathbf{H}_{i j_i}(\mathbf{x} - \hat{\mathbf{x}})$$

where:

$$\mathbf{H}_{i j_i} = \left. \frac{\partial \mathbf{h}_{i j_i}}{\partial \mathbf{x}} \right|_{(\hat{\mathbf{x}})} = [0 \cdots \mathbf{H}_{F_{j_i}} \cdots \mathbf{H}_{E_i} \cdots 0]$$

$$\mathbf{H}_{F_{j_i}} = \left. \frac{\partial \mathbf{h}_{i j_i}}{\partial \mathbf{x}_{F_{j_i}}} \right|_{(\hat{\mathbf{x}})}$$

$$\mathbf{H}_{E_i} = \left. \frac{\partial \mathbf{h}_{i j_i}}{\partial \mathbf{x}_{E_i}} \right|_{(\hat{\mathbf{x}})}$$

Vector $\mathbf{h}_{i_j}(\hat{\mathbf{x}})$ represents the innovation of the pairing between E_i and F_{j_i} . The joint implicit function of hypothesis \mathcal{H} is $\mathbf{h}_{\mathcal{H}}(\mathbf{x}) = \mathbf{0}$, where:

$$\mathbf{h}_{\mathcal{H}}(\mathbf{x}) = \begin{bmatrix} \mathbf{h}_{1j_1}(\mathbf{x}) \\ \vdots \\ \mathbf{h}_{mj_m}(\mathbf{x}) \end{bmatrix} \simeq \mathbf{h}_{\mathcal{H}}(\hat{\mathbf{x}}) + \mathbf{H}_{\mathcal{H}}(\mathbf{x} - \hat{\mathbf{x}})$$

$$\mathbf{H}_{\mathcal{H}} = \left. \frac{\partial \mathbf{h}_{\mathcal{H}}}{\partial \mathbf{x}} \right|_{(\hat{\mathbf{x}})} = \begin{bmatrix} \mathbf{H}_{1j_1} \\ \vdots \\ \mathbf{H}_{mj_m} \end{bmatrix}$$

The validity of \mathcal{H} can be determined using an innovation test on the joint innovation $\mathbf{h}_{\mathcal{H}}(\hat{\mathbf{x}})$ as follows:

$$D_{\mathcal{H}}^2 = \mathbf{h}_{\mathcal{H}}(\hat{\mathbf{x}})^T (\mathbf{H}_{\mathcal{H}} \mathbf{P} \mathbf{H}_{\mathcal{H}}^T)^{-1} \mathbf{h}_{\mathcal{H}}(\hat{\mathbf{x}}) < \chi_{d,\alpha}^2$$

The value of α is the desired confidence level, and $d = \dim(\mathbf{h}_{\mathcal{H}})$. This innovation test is used inside the Joint Compatibility Branch and Bound algorithm to search for the hypothesis with the largest set of compatible matchings. Once this hypothesis has been found, it can be used to obtain a new estimate $\hat{\mathbf{x}}'$ of the state vector and its covariance \mathbf{P}' , by applying the modified EKF update equations:

$$\begin{aligned} \hat{\mathbf{x}}' &= \hat{\mathbf{x}} - \mathbf{K} \mathbf{h}_{\mathcal{H}}(\hat{\mathbf{x}}) \\ \mathbf{P}' &= (\mathbf{I} - \mathbf{K} \mathbf{H}_{\mathcal{H}}) \mathbf{P} \\ \mathbf{K} &= \mathbf{P} \mathbf{H}_{\mathcal{H}}^T (\mathbf{H}_{\mathcal{H}} \mathbf{P} \mathbf{H}_{\mathcal{H}}^T)^{-1} \end{aligned}$$

Once the matching constraints have been applied, the corresponding matching features become fully correlated, with the same estimation and covariance. Thus, one of them can be eliminated.

3.5 Local Map Sequencing

Map joining is used in Local Map Sequencing to obtain a full consistent stochastic map of any size by joining a sequence of local maps of limited size. In order to satisfy the conditions under which map joining can be used, we proceed as follows:

1. Using the initial robot location as base reference, say B_1 , we perform a limited motion sequence of k_1 steps acquiring odometry and sensorial information:

$$D^{1\dots k_1} = \{\mathbf{u}_1 \mathbf{z}_1 \dots \mathbf{u}_{k_1} \mathbf{z}_{k_1}\} \quad (21)$$

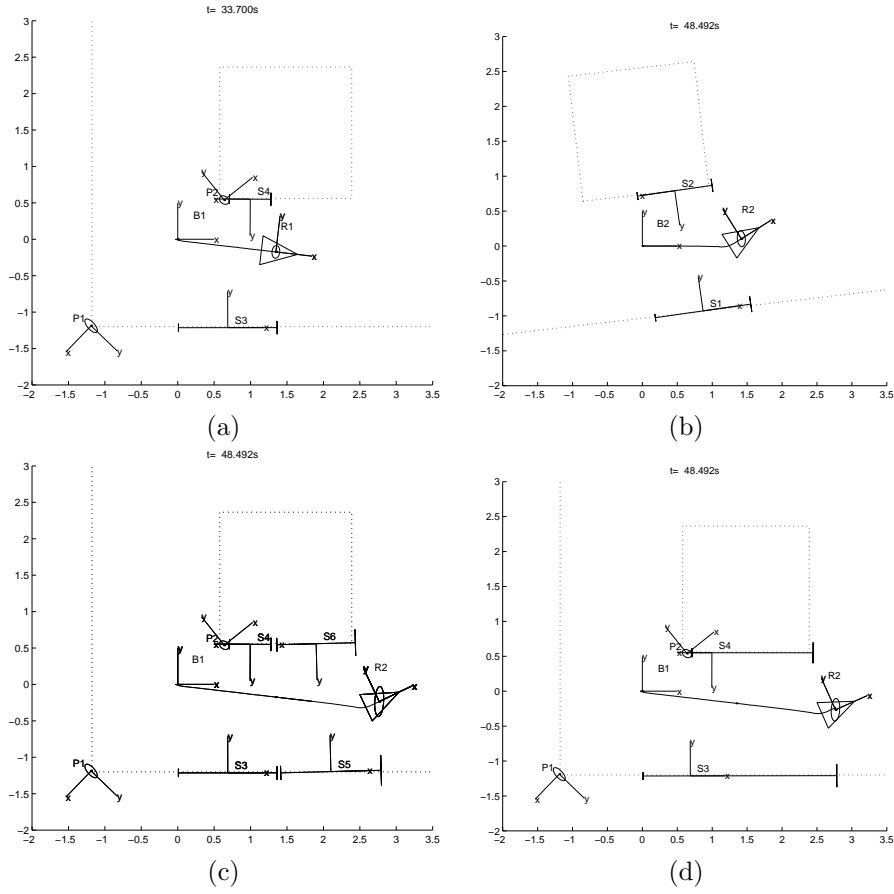


Figure 6: (a) Local map $\mathcal{M}_{\mathcal{F}_1}^{B_1}$ with four features, P1, P2, S3 and S4, with respect to reference B_1 ; (b) local map $\mathcal{M}_{\mathcal{F}_2}^{B_2}$ with two features, S1 and S2, with respect to reference B_2 ; (c) both maps are joined to obtain $\mathcal{M}_{\mathcal{F}_1+\mathcal{F}_2}^{B_1}$; (d) map $\mathcal{M}_{\mathcal{F}_{1:2}}^{B_1}$ after updating by fusing S3 with S5, and S4 with S6. Uncertainty is depicted by the 2σ bounds for point and segment tip locations.

We use this information to build a standard stochastic map of limited size $\mathcal{M}_{\mathcal{F}_1}^{B_1} = (\hat{\mathbf{x}}_{\mathcal{F}_1}^{B_1}, \mathbf{P}_{\mathcal{F}_1}^{B_1})$, which includes the final robot location R_1 and the set of perceived features \mathcal{F}_1 (fig. 6a). Formally, the map estimation obtained is:

$$\hat{\mathbf{x}}_{\mathcal{F}_1}^{B_1} \simeq E \left[\mathbf{x}_{\mathcal{F}_1}^{B_1} \mid D^{1\dots k_1}, \mathcal{H}^{1\dots k_1} \right] \quad (22)$$

- Using the current vehicle location as base reference B_2 (fig. 6b), we perform some motion to obtain measurements:

$$D^{k_1+1\dots k_2} = \{ \mathbf{u}_{k_1+1} \mathbf{z}_{k_1+1} \dots \mathbf{u}_{k_2} \mathbf{z}_{k_2} \} \quad (23)$$

With this information we obtain a new map $\mathcal{M}_{\mathcal{F}_2}^{B_2} = (\hat{\mathbf{x}}_{\mathcal{F}_2}^{B_2}, \mathbf{P}_{\mathcal{F}_2}^{B_2})$, where:

$$\hat{\mathbf{x}}_{\mathcal{F}_2}^{B_2} \simeq E \left[\mathbf{x}_{\mathcal{F}_2}^{B_2} \mid D^{k_1+1\dots k_2}, \mathcal{H}^{k_1+1\dots k_2} \right] \quad (24)$$

- Given that no sensorial information has been shared to build both maps:

$$D^{1\dots k_1} \cap D^{k_1+1\dots k_2} = \emptyset \quad (25)$$

Thus, vectors $\hat{\mathbf{x}}_{\mathcal{F}_1}^{B_1}$ and $\hat{\mathbf{x}}_{\mathcal{F}_2}^{B_2}$ are uncorrelated. Additionally, we know that the last vehicle position in map 1 is the base reference of map 2: i.e. $B_2 = R_1$. This correspondence gives us the second condition that allows us to join both maps into map $\mathcal{M}_{\mathcal{F}_1+\mathcal{F}_2}^{B_1}$ (fig. 6c), where:

$$\hat{\mathbf{x}}_{\mathcal{F}_1+\mathcal{F}_2}^{B_1} \simeq E \left[\mathbf{x}_{\mathcal{F}_1+\mathcal{F}_2}^{B_1} \mid D^{1\dots k_2}, \mathcal{H}^{1\dots k_1}, \mathcal{H}^{k_1+1\dots k_2} \right] \quad (26)$$

It might be the case that the same environment feature has been perceived by the vehicle along the two trajectories. In figure 6c, features $S3$ and $S5$ correspond to the same environment wall (the same applies to features $S4$ and $S6$). This means that both local maps provide two statistically independent estimations of the feature location, relative to each local base reference. When both maps are joined, the estimations are no longer uncorrelated, since they are expressed in the same base reference B_1 . However, their correlation has been correctly calculated in equation (19).

An interesting detail that can be observed in figure 6c is that there is no overlapping between the segments coming from each map. This is because of the circular disposition of the sonar sensors in the B21 robot used in the experiments and the fact that sonars receive returns in directions orthogonal to the wall (figure 3a): the gap between two segments approximately corresponds to the perpendicular projection on the wall of the robot motion between step k_1 and step $k_1 + 1$. During the matching process, this gap is considered acceptable to match and fuse both segments.

4. To avoid this feature duplication, matching takes place after map joining, producing a matching hypothesis $\mathcal{H}^{1:2}$, which is used to obtain the updated map $\mathcal{M}_{\mathcal{F}_{1:2}}^{B_1}$ (figure 6d), where:

$$\hat{\mathbf{x}}_{\mathcal{F}_{1:2}}^{B_1} \simeq E \left[\mathbf{x}_{\mathcal{F}_{1:2}}^{B_1} \mid D^{1\dots k_2}, \mathcal{H}^{1\dots k_1}, \mathcal{H}^{k_1+1\dots k_2}, \mathcal{H}^{1:2} \right] \quad (27)$$

After s steps of repeating this process, we have a full stochastic map $\mathcal{M}_{\mathcal{F}_{1:s}}^{B_1}$, with the initial vehicle position as base reference.

3.6 Computational efficiency of Local Map Sequencing

Assume that a vehicle navigates in an environment of n features, and that each feature is perceived by the vehicle from k different locations. In full stochastic mapping, the cost related to this feature is $O(n)$ when it is included in the map, plus $O(n^2)$ for map updating each time it is re-observed (Castellanos and Tardós 1999), giving a total cost of $O((k-1)n^2)$.

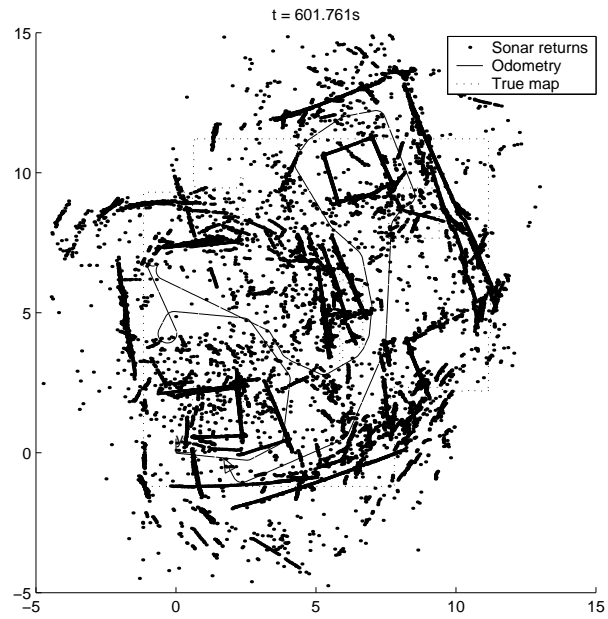
In contrast, using Local Map Sequencing, assume that each feature appears in q different local maps. The cost of including and updating the feature in each local map does not increase with n , and therefore it is negligible. Its inclusion in the full stochastic map will cost $O(n)$ when the first local map where the feature appears is joined, plus $O(n^2)$ for each full map update where the feature reappears, giving a total cost of $O((q-1)n^2)$.

Therefore, Local Map Sequencing cuts processing time by a factor of $(k-1)/(q-1)$. In our experiments, each feature has been observed from a mean of $k = 77$ locations, in a mean of $q = 2.3$ local maps, giving an asymptotic speedup factor of around 58.5.

4 Experimental results

Experiments were carried out using a B21 mobile robot equipped with a SICK laser scanner and a ring of 24 Polaroid sensors (the enclosure sensors were used). The robot carried out a guided trajectory in the ‘Compton Gallery’, a $12m \times 12m$ exhibition gallery at MIT. Several people were visiting the gallery during the experiment. The total trajectory of the robot, which included several loops, was around 101m, lasting around 18 minutes, with an average speed of 9.2cm/s. During the robot motion, the laser sensor and the sonar ring acquired range scans at average frequencies of 2.28Hz and 3.72Hz, respectively. Figure 7a shows the raw sonar returns obtained during the first 10 minutes of the trajectory. Two important facts should be noticed in the figure:

- Although the environment structure is still perceptible, there are huge amounts of spurious data coming from moving people, specular reflections and other sonar artifacts.
- Robot odometry suffered a severe drift, probably due to the carpeted floor of the Compton Gallery. In the upper right corner, after a travel of some



(a)



(b)

Figure 7: (a) Sonar returns obtained during the first 10 minutes of the experiment. (b) Lines and points detected using the Hough transform.

35m, the accumulated error is about 3.9m and 21deg; when the robot is closing the loop, in the lower left corner, the error is about 2.4m and 22deg.

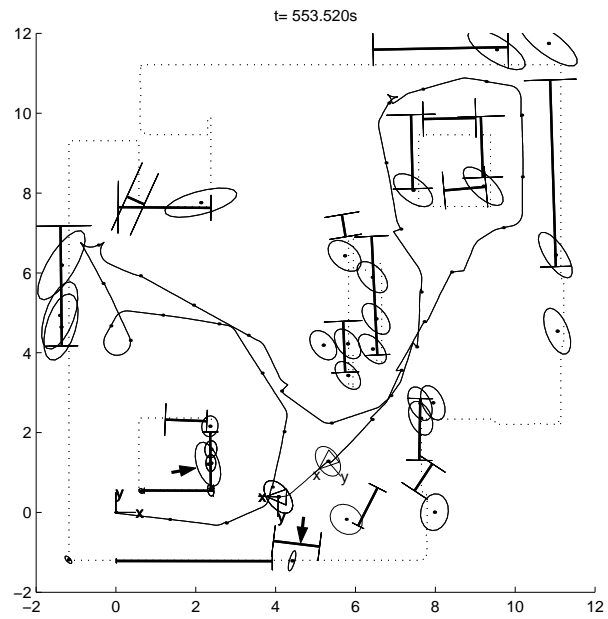
The grouping scheme described in section 2 was applied to the sonar returns obtained along short sequences of 40 robot steps, with a separation of 5cm between them. The lines and points detected are shown in figure 7b. The small dots along the robot trajectory represent the locations where the Hough transform was applied. Most lines detected correspond to the walls, while point features correspond to corners, edges, picture frames and other objects placed on the walls. The majority of spurious sonar returns were successfully rejected by our grouping scheme.

For every sequence of 40 steps, a local stochastic map with points and lines was built, using the technique described in section 3.1. The data associations given by the Hough transform were also stochastically verified by testing the joint compatibility between the returns and their corresponding map features (Neira and Tardós 2001). Each local map was built relative to the robot location at the beginning of the sequence. Using the techniques described in section 3, a global stochastic map was built by joining every new local map as it was available. Then, the largest set of jointly compatible pairings between the new features coming from the local map and those of the global map was obtained. This set of pairings was used to update the global map. The sonar grouping, local map building, map joining, matching and updating processes were repeated along the robot trajectory, obtaining a growing global map.

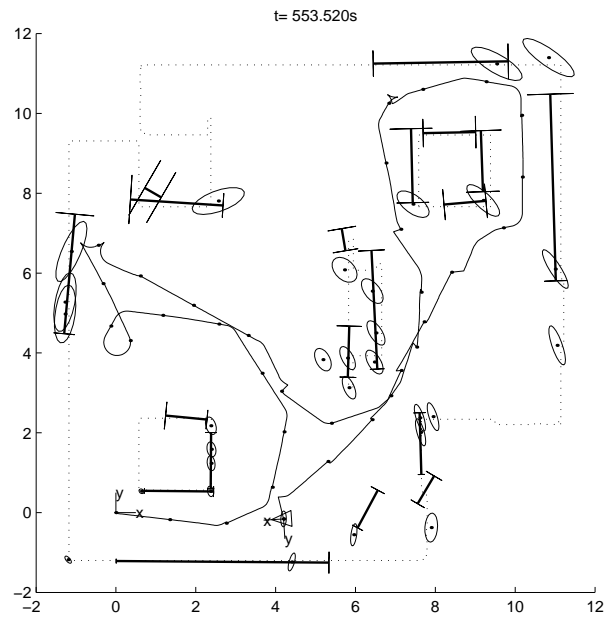
The robustness of our approach is exemplified by the loop closing operation shown in figure 8. Due to the big odometry errors accumulated, simple data association algorithms, such as the Nearest Neighbor, would incorrectly match the signaled point with a point feature previously observed in the pillar. Accepting an incorrect matching will cause the EKF to diverge, obtaining an inconsistent map. On the other hand, our joint compatibility algorithm takes into account the relative location between the point and the segment (Neira and Tardós 2001) and has no problem in finding the right associations. The result is a consistent and more precise global map.

The final global map obtained at the end of the experiment is shown in figure 9a. For comparison purposes, the laser scans obtained every 5cm along the same trajectory were processed to extract straight segments, and the same map joining approach was used to build a segment map of the environment (figure 9b). Comparing both figures, it is clear that the laser sensor is able to obtain a more detailed map. The laser map suffers from a small magnification error, probably due to a systematic error in range measurements. The compensation of this error with a suitable calibration technique would produce a map more precise than the sonar map. Nevertheless, the sonar map obtained is consistent with ground truth and is complete and precise enough to be reliably used for navigation purposes.

To further analyze the consistency of our approach, we obtained a ground truth solution for the robot trajectory by matching the segments obtained from

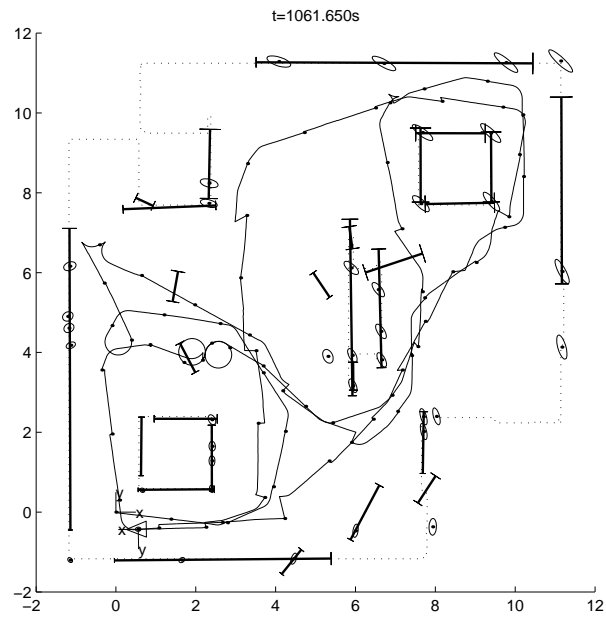


(a)

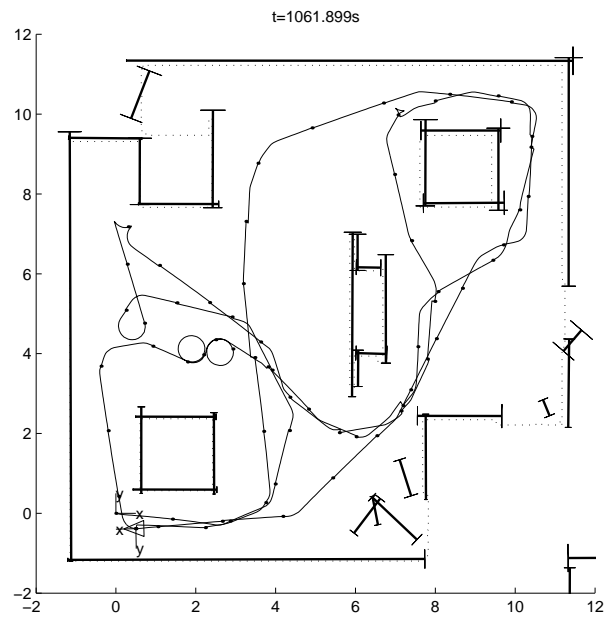


(b)

Figure 8: Global stochastic maps in a loop closing situation. (a) Before loop closing, a local map with features signaled with an arrow has been joined to the global map. (b) The two features have been correctly matched with the corner and the lower wall, and the global map has been updated.



(a)



(b)

Figure 9: Global maps obtained using (a) sonar and (b) laser. Small dots along the robot trajectory represent the locations where map joining and updating was performed.

laser and the true environment map, measured by hand. Figure 10 shows the errors in robot location, relative to this ground truth, during the sonar map building process. We can clearly see that most of the time, errors remain inside their 2σ uncertainty bounds.

Regarding computational efficiency, all techniques presented run faster than real-time for this environment. In our actual implementation, for each sequence of 40 steps (about 2m travel), typical computing times on a Pentium III at 600MHz are: 300ms for sonar grouping, 1.5s for local map building, and up to 1s for fusing the local map with the global map (joining, matching and updating). The first two processes use local information and their computing times are independent of environment size. The computing time for map fusion grows with the number of features in the global map. Although its asymptotic complexity is quadratic, the moderate size of the final map in our experiment (63 features) makes the increase to be almost linear, dominated by the cost of the matching operation. For bigger environments, the quadratic cost of the global map update would become dominant.

5 Conclusion

This paper has presented several new techniques for mobile robot navigation and mapping with sonar. Our results include a new method for detection of points and line segments from sonar data and new techniques for joining and combining several stochastic maps. Results have been presented for the implementation of these methods using data from a standard ring of Polaroid sonars mounted on a B21 mobile robot. The results are compared to a map produced from laser scanner data for exactly the same environment. To our knowledge, this constitutes the first published side-by-side comparison of laser and sonar maps produced via CML.

Many early implementations of navigation and mapping algorithms that used a first-order representation of uncertainty suffered the criticism of “brittleness” (Lozano-Pérez 1989). When incorrect decisions are made concerning the origins of measurements, dramatically erroneous results can be produced. These issues motivated the development of methods for sensor data interpretation that can make delayed decisions, such as multiple hypothesis tracking (Cox and Leonard 1994), to achieve better classification of measurements. The methods presented in this paper can effectively make delayed decisions about the assignment of individual measurements to achieve robustness despite the difficulties inherent in sonar data. The voting scheme used to aggregate data from multiple vantage points, acts analogously to methods in computer vision such as RANSAC that use the principle of consensus to resolve data association ambiguity. We believe that in the future it should be possible to generalize the concept behind the voting scheme used in this paper to provide a more generic capability for computing structure-from-motion from acoustic sensors in more complex environments (i.e., three-dimensional and underwater).

We anticipate that the techniques presented here for joining and combining

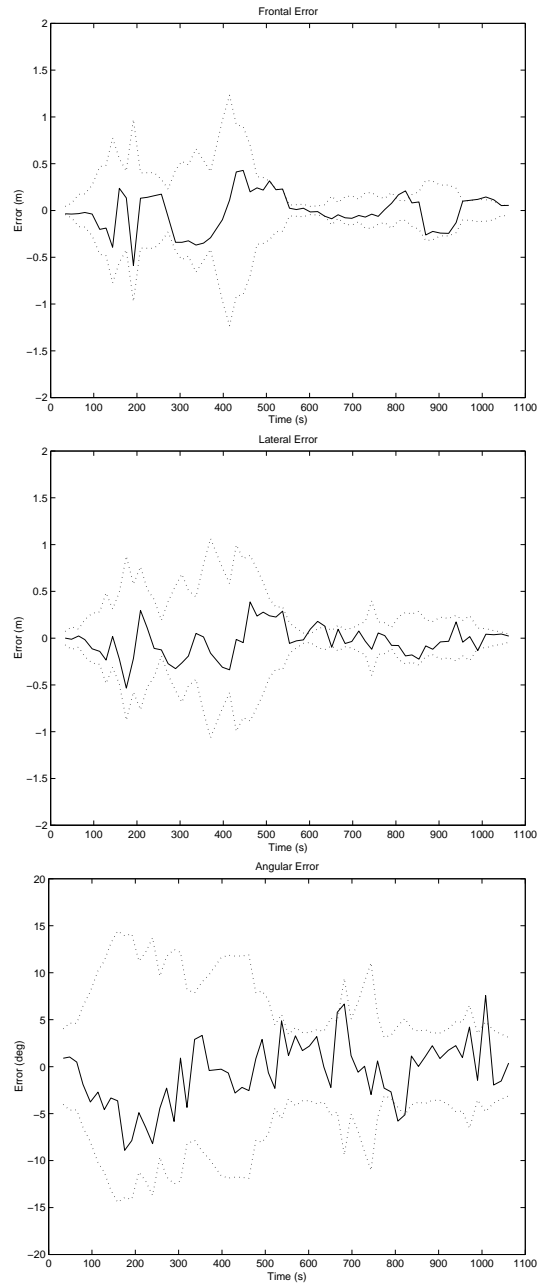


Figure 10: Errors in robot location during the sonar map building process. Dotted lines represent 2σ uncertainty bounds.

multiple submaps in CML will be useful in the development of new computationally efficient methods for large-scale CML. By performing the operations of data association and feature initialization in local submaps, sensitivity to linearization errors in the state estimation process is reduced. Indeed, it is possible to pose the problem such that all operations for creation of local submaps could be performed without Kalman filtering, for example using nonlinear least-squares minimization techniques that have effectively been employed for the structure from motion problem in computer vision (Hartley and Zisserman 2000).

In the example in this paper, all maps were combined relative to a single, globally-referenced coordinate frame. In situations with sufficiently large angular errors, it will be impossible to consistently reference all submaps to a common reference frame using a first-order representation of uncertainty. For this situation, we envision that effective solutions can be developed using a representation consisting of a network of different coordinate frames, each related through a sequence of approximate transformations. The techniques of joining and combining submaps are expected to be invaluable as a means of building and maintaining the components of the map. A variation of such an approach might be to adopt the semantic spatial hierarchy proposed by Kuipers (2000), with local maps serving as vertices in a very large-scale topological graph of the environment.

Acknowledgements

This research has been funded in part by the National Science Foundation under Career Award BES-9733040, the National Oceanic and Atmospheric Administration via the MIT Sea Grant College Program under grant NA86RG0074 (project RCM-3), the Dirección General de Investigación of Spain under project DPI2000-1265, the Ministerio de Educación y Cultura of Spain under grant PR2000-0104, and the Spain-US Commission for Cultural, Educational and Scientific Exchange (Fulbright) under grant 20079.

Appendix A: Transformations, points and lines in 2D

Two basic operations used in stochastic mapping are transformation inversion and composition, which were represented by Smith et al. (1988) using operators \ominus and \oplus :

$$\begin{aligned}\hat{\mathbf{x}}_A^B &= \ominus \hat{\mathbf{x}}_B^A \\ \hat{\mathbf{x}}_C^A &= \hat{\mathbf{x}}_B^A \oplus \hat{\mathbf{x}}_C^B\end{aligned}$$

In this work, we generalize the \oplus operator to also represent the composition of transformations with feature location vectors, which results in the change of base reference of the feature. The Jacobians of these operations are defined as:

$$\begin{aligned}
\mathbf{J}_\ominus \{ \hat{\mathbf{x}}_B^A \} &= \left. \frac{\partial (\ominus \mathbf{x}_B^A)}{\partial \mathbf{x}_B^A} \right|_{(\hat{\mathbf{x}}_B^A)} \\
\mathbf{J}_{1\oplus} \{ \hat{\mathbf{x}}_B^A, \hat{\mathbf{x}}_C^B \} &= \left. \frac{\partial (\mathbf{x}_B^A \oplus \mathbf{x}_C^B)}{\partial \mathbf{x}_B^A} \right|_{(\hat{\mathbf{x}}_B^A, \hat{\mathbf{x}}_C^B)} \\
\mathbf{J}_{2\oplus} \{ \hat{\mathbf{x}}_B^A, \hat{\mathbf{x}}_C^B \} &= \left. \frac{\partial (\mathbf{x}_B^A \oplus \mathbf{x}_C^B)}{\partial \mathbf{x}_C^B} \right|_{(\hat{\mathbf{x}}_B^A, \hat{\mathbf{x}}_C^B)}
\end{aligned}$$

Transformations in 2D

In 2D, the location of a reference B relative to a reference A (or transformation from A to B) can be expressed using a vector with three d.o.f.:

$$\mathbf{x}_B^A = \begin{bmatrix} x_1 \\ y_1 \\ \phi_1 \end{bmatrix}$$

The location of A relative to B is computed using the inversion operation:

$$\mathbf{x}_A^B = \ominus \mathbf{x}_B^A = \begin{bmatrix} -x_1 \cos \phi_1 - y_1 \sin \phi_1 \\ x_1 \sin \phi_1 - y_1 \cos \phi_1 \\ -\phi_1 \end{bmatrix}$$

The Jacobian of transformation inversion is:

$$J_\ominus \{ \mathbf{x}_B^A \} = \begin{bmatrix} -\cos \phi_1 & -\sin \phi_1 & -x_1 \sin \phi_1 - y_1 \cos \phi_1 \\ \sin \phi_1 & -\cos \phi_1 & x_1 \cos \phi_1 + y_1 \sin \phi_1 \\ 0 & 0 & -1 \end{bmatrix}$$

Let $\mathbf{x}_C^B = [x_2, y_2, \phi_2]^T$ be a second transformation. The location of reference C relative to A is obtained by the composition of transformations \mathbf{x}_B^A and \mathbf{x}_C^B :

$$\mathbf{x}_C^A = \mathbf{x}_B^A \oplus \mathbf{x}_C^B = \begin{bmatrix} x_1 + x_2 \cos \phi_1 - y_2 \sin \phi_1 \\ y_1 + x_2 \sin \phi_1 + y_2 \cos \phi_1 \\ \phi_1 + \phi_2 \end{bmatrix}$$

The Jacobians of transformation composition are:

$$\begin{aligned}
\mathbf{J}_{1\oplus} \{ \mathbf{x}_B^A, \mathbf{x}_C^B \} &= \begin{bmatrix} 1 & 0 & -x_2 \sin \phi_1 - y_2 \cos \phi_1 \\ 0 & 1 & x_2 \cos \phi_1 - y_2 \sin \phi_1 \\ 0 & 0 & 1 \end{bmatrix} \\
\mathbf{J}_{2\oplus} \{ \mathbf{x}_B^A, \mathbf{x}_C^B \} &= \begin{bmatrix} \cos \phi_1 & -\sin \phi_1 & 0 \\ \sin \phi_1 & \cos \phi_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

Point features

If we consider a point feature P , we can represent its location with respect to reference B in cartesian coordinates:

$$\mathbf{x}_P^B = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

The composition operation to obtain the location of the point with respect to reference A is as follows:

$$\mathbf{x}_P^A = \mathbf{x}_B^A \oplus \mathbf{x}_P^B = \begin{bmatrix} x_1 + x_2 \cos \phi_1 - y_2 \sin \phi_1 \\ y_1 + x_2 \sin \phi_1 + y_2 \cos \phi_1 \end{bmatrix}$$

The Jacobians for transforming points are:

$$\begin{aligned} J_{1\oplus}\{\mathbf{x}_B^A, \mathbf{x}_P^B\} &= \begin{bmatrix} 1 & 0 & -x_2 \sin \phi_1 - y_2 \cos \phi_1 \\ 0 & 1 & x_2 \cos \phi_1 - y_2 \sin \phi_1 \end{bmatrix} \\ J_{2\oplus}\{\mathbf{x}_B^A, \mathbf{x}_P^B\} &= \begin{bmatrix} \cos \phi_1 & -\sin \phi_1 \\ \sin \phi_1 & \cos \phi_1 \end{bmatrix} \end{aligned}$$

Line features

The location of a line feature L relative to reference B can be represented using the perpendicular distance from the origin of reference B to the line, and the line orientation:

$$\mathbf{x}_L^B = \begin{bmatrix} \rho_2 \\ \theta_2 \end{bmatrix}$$

The composition operation to obtain the location of the line with respect to reference A is as follows:

$$\begin{aligned} \mathbf{x}_L^A &= \mathbf{x}_B^A \oplus \mathbf{x}_L^B \\ &= \begin{bmatrix} x_1 \cos(\phi_1 + \theta_2) + y_1 \sin(\phi_1 + \theta_2) + \rho_2 \\ \phi_1 + \theta_2 \end{bmatrix} \end{aligned}$$

The Jacobians for transforming lines are:

$$\begin{aligned} J_{1\oplus}\{\mathbf{x}_B^A, \mathbf{x}_L^B\} &= \begin{bmatrix} \cos(\phi_1 + \theta_2) & \sin(\phi_1 + \theta_2) \\ 0 & 0 \\ -x_1 \sin(\phi_1 + \theta_2) + y_1 \cos(\phi_1 + \theta_2) \\ 1 \end{bmatrix} \\ J_{2\oplus}\{\mathbf{x}_B^A, \mathbf{x}_L^B\} &= \begin{bmatrix} 1 & -x_1 \sin(\phi_1 + \theta_2) + y_1 \cos(\phi_1 + \theta_2) \\ 0 & 1 \end{bmatrix} \end{aligned}$$

References

- Au, W.: 1993, *The Sonar of Dolphins*, Springer-Verlag, New York.
- Ayache, N. and Faucher, O. D.: 1989, Maintaining representations of the environment of a mobile robot, *IEEE Trans. Robot. Automat.* **5**(6), 804–819.
- Ballard, D. H. and Brown, C. M.: 1982, *Computer Vision*, Prentice Hall, Englewood Cliffs, N. J.
- Bar-Shalom, Y. and Fortmann, T. E.: 1988, *Tracking and Data Association*, Academic Press, Boston, MA.
- Barshan, B. and Kuc, R.: 1990, Differentiating sonar reflections from corners and planes by employing an intelligent sensor, *IEEE Trans. Pattern Analysis and Machine Intelligence* **12**(6), 560–569.
- Brooks, R. A.: 1984, Aspects of mobile robot visual map making, *Second Int. Symp. Robotics Research*, MIT Press, Tokyo, Japan, pp. 287–293.
- Castellanos, J. A., Montiel, J. M. M., Neira, J. and Tardós, J. D.: 1999, The SPmap: A probabilistic framework for simultaneous localization and map building, *IEEE Trans. Robot. Automat.* **15**(5), 948–953.
- Castellanos, J. A., Neira, J. and Tardós, J. D.: 2001, Multisensor fusion for simultaneous localization and map building, *IEEE Trans. Robot. Automat.* **17**(6), 908–914.
- Castellanos, J. A. and Tardós, J. D.: 1999, *Mobile Robot Localization and Map Building: A Multisensor Fusion Approach*, Kluwer Academic Publishers, Boston, Mass.
- Chong, K. S. and Kleeman, L.: 1999a, Feature-based mapping in real, large scale environments using an ultrasonic array, *Int. J. Robotics Research* **18**(1), 3–19.
- Chong, K. S. and Kleeman, L.: 1999b, Mobile-robot map building from an advanced sonar array and accurate odometry, *Int. J. Robotics Research* **18**(1), 20–36.
- Choset, H. and Nagatani, K.: 2001, Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization, *IEEE Trans. Robot. Automat.* **17**(2), 125–137.
- Cox, I. J. and Leonard, J. J.: 1994, Modeling a dynamic environment using a Bayesian multiple hypothesis approach, *Artificial Intelligence* **66**(2), 311–344.
- Davison, A. J.: 1998, *Mobile Robot Navigation Using Active Vision*, PhD thesis, University of Oxford. Available from <http://www.robots.ox.ac.uk/~ajd/>.

- Dissanayake, M. W. M. G., Newman, P., Durrant-Whyte, H. F., Clark, S. and Csorba, M.: 2001, A solution to the simultaneous localization and map building (SLAM) problem, *IEEE Trans. Robot. Automat.* **17**(3), 229–241.
- Doucet, A., de Freitas, N. and Gordan, N. (eds): 2001, *Sequential Monte Carlo Methods in Practice*, Springer-Verlag.
- Elfes, A.: 1987, Sonar-based real-world mapping and navigation, *IEEE Journal of Robotics and Automation* **3**(3), 249–265.
- Faugeras, O.: 1993, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, Mass.
- Faugeras, O. and Luong, Q. T.: 2001, *The Geometry of Multiple Images: The laws that govern the formation of multiple images of a scene and some of their applications*, The MIT Press, Cambridge, Mass.
- Feder, H. J. S., Leonard, J. J. and Smith, C. M.: 1999, Adaptive mobile robot navigation and mapping, *Int. J. Robotics Research* **18**(7), 650–668.
- Fischler, M. A. and Bolles, R. C.: 1981, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, *Comm. Assoc. Comp. Mach.* **24**(6), 381–395.
- Grossmann, A. and Poli, R.: 2001, Robust mobile robot localisation from sparse and noisy proximity readings using Hough transform and probability grids, *Robotics and Autonomous Systems* **37**, 1–18.
- Guivant, J. E. and Nebot, E. M.: 2001, Optimization of the simultaneous localization and map building algorithm for real time implementation, *IEEE Trans. Robot. Automat.* **17**(3), 242–257.
- Gutmann, J. S. and Konolige, K.: 1999, Incremental mapping of large cyclic environments, *IEEE Int. Symp. on Computational Intelligence in Robotics and Automation CIRA*, pp. 318–325.
- Hartley, R. and Zisserman, A.: 2000, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U. K.
- Illingworth, J. and Kittler, J.: 1988, A survey of the Hough transform, *Computer Vision, Graphics and Image Processing* **44**, 87–116.
- Jazwinski, A. H.: 1970, *Stochastic Processes and Filtering Theory*, Academic Press, New York.
- Jensfelt, P.: 2001, *Approaches to Mobile Robot Localization in Indoor Environments*, PhD thesis, Royal Institute of Technology, Sweden. Available from <http://www.nada.kth.se/~patric/>.
- Johnson, D. and Dudgeon, D.: 1993, *Array Signal Processing*, Prentice Hall.

- Julier, S. J.: 2001, A sparse weight Kalman filter approach to simultaneous localisation and map building, *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Vol. 3, Maui, Hawaii, pp. 1251–1256.
- Kleeman, L. and Kuc, R.: 1995, Mobile robot sonar for target localization and classification, *Int. J. Robotics Research* **14**(4), 295–318.
- Knight, J., Davison, A. and Reid, I.: 2001, Towards constant time SLAM using postponement, *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems IROS*, Vol. 1, Maui, Hawaii, pp. 405–413.
- Kuc, R. and Siegel, M. W.: 1987, Physically based simulation model for acoustic sensor robot navigation, *IEEE Trans. Pattern Analysis and Machine Intelligence* **9**(6), 766–778.
- Kuipers, B. J.: 2000, The spatial semantic hierarchy, *Artificial Intelligence* **119**, 191–233.
- Leonard, J. J. and Durrant-Whyte, H. F.: 1992, *Directed Sonar Sensing for Mobile Robot Navigation*, Kluwer Academic Pub., London.
- Leonard, J. J. and Feder, H. J. S.: 2000, A computationally efficient method for large-scale concurrent mapping and localization, in D. Koditschek and J. Hollerbach (eds), *Robotics Research: The Ninth International Symposium*, Springer Verlag, Snowbird, Utah, pp. 169–176.
- Lozano-Pérez, T.: 1989, Foreword, in I. J. Cox and G. T. Wilfong (eds), *Autonomous Robot Vehicles*, Springer-Verlag.
- Lu, F. and Milios, E.: 1997, Globally consistent range scan alignment for environment mapping, *Autonomous Robots* **4**, 333–349.
- Moutarlier, P. and Chatila, R.: 1989, Stochastic multisensory data fusion for mobile robot location and environment modeling, *5th Int. Symp. on Robotics Research*, Tokyo, Japan.
- Neira, J. and Tardós, J. D.: 2001, Data association in stochastic mapping using the joint compatibility test, *IEEE Trans. Robot. Automat.* **17**(6), 890–897.
- Ortín, D. and Montiel, J. M. M.: 2001, Indoor robot motion from monocular images, *Robotica* **19**(3), 331–342.
- Papoulis, A.: 1991, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill.
- Peremans, H., Audenaert, K. and Van, C. J. M.: 1993, A high-resolution sensor based on tri-aural perception, *IEEE Trans. Robot. Automat.* **9**(1), 36–48.
- Schultz, A. C. and Adams, W.: 1998, Continuous localization using evidence grids, *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 2833–2839.

- Smith, R. C. and Cheeseman, P.: 1986, On the representation and estimation of spatial uncertainty, *Int. J. Robotics Research* **5**(4), 56–68.
- Smith, R., Self, M. and Cheeseman, P.: 1988, A stochastic map for uncertain spatial relationships, in O. Faugeras and G. Giralt (eds), *Robotics Research, The Fourth Int. Symp.*, The MIT Press, pp. 467–474.
- Thrun, S.: 2001, An online mapping algorithm for teams of mobile robots, *Int. J. Robotics Research* **20**(5), 335–363.
- Wijk, O.: 2001, *Triangulation Based Fusion of Sonar Data with Application in Mobile Robot Mapping and Localization*, PhD thesis, Royal Institute of Technology, Stockholm, Sweden. Available from <http://www.e.kth.se/~olle/>.
- Wijk, O. and Christensen, H. I.: 2000, Triangulation based fusion of sonar data with application in robot pose tracking, *IEEE Trans. Robot. Automat.* **16**(6), 740–752.
- Williams, S.: 2001, *Efficient Solutions to Autonomous Mapping and Navigation Problem*, PhD thesis, Dept. of Mechanical and Mechatronic Engineering, University of Sydney. Available from <http://www.acfr.usyd.edu.au/>.
- Zunino, G. and Christensen, H. I.: 2001, Navigation in realistic environments, in M. Devy (ed.), *9th Int. Symp. on Intelligent Robotic Systems, SIRS*, Toulouse, France.