Simultaneous Map Building and Localization for Mobile Robots: A Multisensor Fusion Approach*

J. A. Castellanos  J. M. Martínez  J. Neira  J. D. Tardós

Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza
c/María de Luna 3, E-50015 Zaragoza, Spain
email : {jcaste, josemari, jneira, tardos}@posta.unizar.es

Abstract

During mobile robot navigation, position estimates obtained by odometry drift with time, therefore becoming unrealistic and useless. This work enhances the use of external mechanisms by considering a multisensor system, composed of a 2D laser rangefinder and an off-the-self CCD camera, which provides redundancy and assures reliability and precision of the observed features. We simultaneously consider both the map building and the localization problems using a state vector approach, which is related to the location estimations of both the robot and the map features, whilst its covariance matrix reflects the relationships between them. Relevance and importance of its off-diagonal elements is demonstrated by their contributions to "backwards estimations" whenever the vehicle returns to places in the navigation area which have been already visited and learned. Real experiments are presented, considering a LabMate mobile robot navigating in an static indoor environment.

1 Introduction

Navigation of a mobile robot requires both a sufficiently reliable estimation of the current vehicle location, and a sufficiently precise map of the navigation area. A priori model maps are rarely available and when they are, they usually introduce inaccuracies in the planning tasks. Therefore an automatic construction of the map of the environment in which the robot navigates is desirable, and it has become one important research direction in today's robotics community.

\*This work was supported by spanish CICYT projects TAP94-0390 and TAP97-0892-C02-01.

Lately, the problem of map building has been considered simultaneously to the relocalization of the mobile robot, approximation based on the works of Smith et al. [10] who introduced the stochastic map. Ongoing research on this field has been recently described in the works [1, 6, 7, 8]. In this work, we take a step further from our previous works [5] by taking advantage of a multisensor system formed by two different sensors whose complementary nature allows to detect highly reliable observations in the navigation area. Namely, we use a 2D laser rangefinder and an off-the-self CCD camera, which have been properly calibrated for our experimentation. Monocular vision provides redundant information about the location of the geometric entities detected by the laser rangefinder. Thus, fusion is performed at the level of observations (figure 1), therefore increasing their robustness and
credibility from an early stage of the processing.

Simultaneous map building and mobile robot localization are considered with a state vector approach, which is related to the locations of the entities involved in the problem. Suboptimal estimation is subsequently applied using a recursive formulation, the Extended Kalman Filter, EKF. Dependencies between the location estimations of features are appropriately represented by a system covariance matrix [8]. We consider a probabilistic representation of uncertain geometric information, the 2D version of the Symmetries and Perturbations Model (SPmodel) [11]. We focus on the importance of the covariance matrix which reflects the relationships between entities. Relevance and importance of its off-diagonal elements is demonstrated by their contributions to “backwards estimations” whenever the vehicle returns to places in the navigation area which have been already visited and learned, that is, the improvement in the location estimation of features previously learned, which are not visible from the current robot location but are related to current observations through those off-diagonal elements.

The rest of the paper is structured as follows. Section 2 summarizes the multisensor fusion process to detect highly reliable features in the navigation area. Section 3 uses these features both to incrementally build a map of the environment and to relocalize the mobile robot. Finally, section 4 presents experimental results obtained with the mobile robot Labmate and section 5 discusses the work and further research directions.

2 Robust Processing of Sensor Information: Local Maps

Redundant sensing of the environment of the mobile robot increases robustness by combining information gathered by a multisensor system.

2.1 Range Readings Processing

A segmentation technique is considered to obtain meaningful information from a set of range readings gathered by the laser rangefinder. A reference $E$ is attached to each detected straight-line segment, such that its X-axis is aligned with the segment (figure 2), and its location (i.e. translation and rotation) with respect to a reference frame associated to the mobile robot $R$, is given by the composition of location vec-

\begin{equation}
\mathbf{x}_{RE} = \mathbf{x}_{RE} \oplus \mathbf{d}_E = \mathbf{x}_{RE} \oplus B_E^T \mathbf{p}_E
\end{equation}

where $\mathbf{x}_{RE}$ is the estimated location vector of the straight-line segment; $\mathbf{d}_E$ is the differential location vector; $\mathbf{p}_E \sim N(\hat{\mathbf{p}}_E, \mathbf{C}_E)$ is the normally distributed perturbation vector of the segment, which takes into account sensor imprecision; and $B_E$ was the self-binding matrix of the segment, which takes into account the symmetry of translation along the supporting line of the segment, thus $\mathbf{d}_E = (0, d_{y_E}, d_{\phi_E})^T$, and:

\[
B_E = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}; \quad \mathbf{p}_E = (d_{y_E}, d_{\phi_E})^T
\]

From straight-line segments we easily identify higher level features [4] in the environment of the mobile robot (figure 2):

- First, corners are found at the intersection of two consecutive segments. We attach a reference $C$ to each corner such that its X-axis is aligned with its bisection. A corner $C$ is considered as an uncertain geometric entity without symmetries; therefore its self-binding matrix is the identity matrix $I_3$.

- Second, semiplanes derive from the free endpoints of segments, which might correspond to frames of open doors, convex corners, obstacles, etc.. Robustness in their detection is enhanced by avoiding false semiplanes arising from occlusions. We attach a reference $S$ to each semiplane such that its X-axis is aligned with the associated segment. A semiplane $S$ is also an uncertain geometric feature without
symmetries with a self-binding matrix equal to the identity matrix $I_3$. However, the positions of semiplanes are not as precise as the position of corners because their estimations are greatly influenced by the angular resolution of the laser rangefinder.

Consecutive segments produce sets of uncertain geometric features whose estimated locations are correlated because they are computed from a common set of segments (e.g. S2-E2-C-E3-S3 in figure 2). Therefore, estimation must be performed considering the state vector composed of the perturbation vectors of the correlated features:

$$x = (p_S, p_E_1, p_C_1, \ldots, p_{E_{n-1}}, p_{C_{n-1}}, p_E_n, p_{S_j})^T$$  \hspace{1cm} (2)

An estimation of this vector and its covariance are computed using the Extended Information Filter [3], where measurements are given by the estimated locations of each consecutive segments (e.g. segments E2 and E3).

2.2 Monocular Vision Processing

Grey-level images, obtained by an off-the-shelf CCD camera, are processed using the Burns’s segment extractor [2]. We are mainly interested in long vertical edges, corresponding to corners and door frames in the robot’s environment. Vision edges are represented by an attached reference V with its origin on the optical center of the camera and its X-axis pointing to the projection on a horizontal plane, containing the camera’s optical center, of the middle point of the segment detected on the image (figure 2). An angular error, represented by $\sigma^2$, is associated to the uncertain location of each vision edge, which is related to the detection error of the vertical edge in the image [9].

2.3 Range and Vision Fusion

Both laser rangefinder and monocular vision information have been expressed in the reference frame of the mobile robot R after calibrating the relative transformation between the sensors. Hence, matching between corners and vision edges on the one hand, and semiplanes and vision edges on the other hand can now be obtained. Compatibility between geometric features derives from a $\chi^2$ test based on the squared Mahalanobis distance $D^2$ [4]. Multiple validations of the $\chi^2$ test are resolved by minimizing $D^2$. Redundant information coming from the observation of vision edges is used to both increase the credibility in the detection, and to reduce the uncertainty in the estimation of features. Precision is especially increased for semiplanes due to the higher angular resolution of vision compared to that of the laser rangefinder.

3 Simultaneous Map Building and Localization using Robust Features

In this section, we consider the simultaneous map building and localization problem by using the local maps previously obtained.

3.1 Global Map State Vector

The set of geometric entities involved in the problem is composed of a mobile robot $R$, a set of segments $E_i$, $i \in \{1 \ldots n_E\}$, a set of semiplanes $S_j$, $j \in \{1 \ldots n_S\}$, and a set of corners $C_k$, $k \in \{1 \ldots n_C\}$ all of them expressed with respect to a global reference W. The system state vector is composed of the differential location vector of the robot $d_R$, and the perturbation vector of each map feature. Defining the subvectors:

$$x_E = \{p_{E_i}\} \quad i \in \{1 \ldots n_E\}$$

$$x_S = \{p_{S_j}\} \quad j \in \{1 \ldots n_S\}$$

$$x_C = \{p_{C_k}\} \quad k \in \{1 \ldots n_C\}$$

we obtain an $n$-dimensional state vector, with $n = 3 + 2n_E + 3n_S + 3n_C$, given by:

$$x = (d_R, x_E, x_S, x_C)^T$$

whose covariance matrix can be written as:

$$P = \begin{bmatrix}
C_{d_R} & C_{d_R,x_E} & C_{d_R,x_S} & C_{d_R,x_C} \\
C_{d_R,x_E}^T & C_{x_E} & C_{x_E,x_S} & C_{x_E,x_C} \\
C_{d_R,x_S}^T & C_{x_S,x_E} & C_{x_S} & C_{x_S,x_C} \\
C_{d_R,x_C}^T & C_{x_C,x_S} & C_{x_C} & C_{x_C}
\end{bmatrix}$$

where the elements in the diagonal represent the covariance of the subvectors and the off-diagonal elements represent the cross-covariance matrices between subvectors.

3.2 Robot’s Displacement

Dead-reckoning estimates the displacement of the robot between two consecutive points along the trajectory. Let $R_{k-1}$ and $R_k$ be the references attached to these points respectively, we may write:

$$x_{R_k-1} = x_{R_{k-1}} + d_{R_{k-1}R_k}$$

where $d_{R_{k-1}R_k} \sim \mathcal{N}(d_{R_{k-1}R_k}, C_{R_{k-1}R_k})$ represents dead-reckoning errors. The plant model of the system

\[^1\text{Observe that } d_R = p_R \text{ because } B_R = I_3.\]
<table>
<thead>
<tr>
<th>Local $k$</th>
<th>Global $k - 1$</th>
<th>Relation</th>
<th>Global $k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_k$</td>
<td>$E_G$</td>
<td>Collinearity</td>
<td>$E_G$</td>
</tr>
<tr>
<td>$S_k$</td>
<td>$S_G$</td>
<td>Coincidence</td>
<td>$S_G$</td>
</tr>
<tr>
<td>$S_k$</td>
<td>$S_G$</td>
<td>Position</td>
<td>$S_G$</td>
</tr>
<tr>
<td>$C_k$</td>
<td>$C_G$</td>
<td>Position</td>
<td>$C_G$</td>
</tr>
<tr>
<td>$C_k$</td>
<td>$C_G$</td>
<td>Coincidence</td>
<td>$C_G$</td>
</tr>
</tbody>
</table>

Table 1: Possible pairings between local and global features: segments $E$, semiplanes $S$, and corners $C$.

is obtained as (see [5] for details):

\[
x_k = \begin{bmatrix}
J_{R_k R_{k-1}} & 0_{3x3} & 0_{3x1} \\
0_{3x3} & I_{3x3} & 0_{3x1}
\end{bmatrix}
\begin{bmatrix}
x_{k-1} \\
I_{3x1}
\end{bmatrix}
\begin{bmatrix}
d_{R_k R_{k-1}}
\end{bmatrix}
\]

where $J_{R_k R_{k-1}}$ is the Jacobian of the transformation $x_{R_k R_{k-1}}$ between the position of the robot at time $k$ and that at time $k-1$. Thus, during the prediction phase only the estimated location of the robot changes. On the contrary, all the off-diagonal elements of the system covariance matrix concerning the correlation between the robot location and the map features locations are updated.

3.3 Integrating Local into Global Maps

Measurement equations are obtained by pairing the features of the local map at time $k$ with those of the global map known up to time $k - 1$. Different pairings have been considered (table 1) between two given features. Compatibility between paired features is decided by calculating the Mahalanobis distance $D$. Under the gaussianity hypothesis, the local feature $L$ is compatible with the global feature $G$ only if $D^2 \leq \chi^2_m$, where $\chi^2_m$ is a threshold value, obtained from the $\chi^2$ distribution with $m$ degrees of freedom, such that the probability of rejecting a good matching is $\alpha$. The most restrictive type of pairing is selected in first place (e.g., considering two semiplanes, the coincidence relation is selected). When a given type of pairing cannot be satisfied, another less restrictive pairing is selected from the table, provided that it exists. Correlated features, stacked together in an state vector, must be simultaneously integrated, because they represent a single, although complex, measurement.

At each point of the robot’s trajectory it is desirable to obtain as much pairings as possible because they represent the links between new observations and old stored knowledge of the navigation area. They are the toeholds on which the map building task is fundamental. As yet, we assume that non-paired local observations represent knowledge about the environment of the robot which has not yet been learned. Usually, whenever the robot changes from one well-defined navigation area to another, there are many non-paired observations which are directly added to the map. Further work will focus on the possibility of non-paired observations being spurious data or measurements derived from moving obstacles.

3.4 Estimating Features Location

Finally, an EKF estimates both the state vector $x_{1:k}$ and its covariance matrix $P_{1:k}$ using the classical equations (see [3] for details). Integration of new information produces a reduction in the location uncertainty corresponding not only to the paired map features, but also to the whole set of map features, learned in previous time instants, because they are all included in the same state vector. The system covariance matrix is also updated to reestablish the relationships between map features. When the whole local map is composed of non-paired observations, the location estimation of the robot obtained by odometry cannot be improved, and the robot might even get lost.

4 Experimental Results

For experimentation, a Labmate mobile robot was programmed to follow a predefined trajectory along our laboratory (53 m aprox.). The robot stopped at regular intervals to obtain environmental information through its multisensor system. There, range readings and grey-level images were taken with a 2D laser rangefinder and an off-the-self CCD camera respectively. Complementary information were taken, by hand, with a pair of theodolites which provided real locations of the robot with respect to a base reference.

At each point along the trajectory, a local map of the environment of the robot was obtained. Figure 3 presents the processing of the first local map, with the segmentation of both range readings and the grey-level image, and the final solution obtained by fusing laser rangefinder observations with monocular vision information. Corners and semiplanes have been represented by their 95% error ellipses.

From figure 4 it can be observed the role of the system covariance matrix as a memory of the relationships between entries created in previous time instants. When the robot navigates in a previously
unknown area, uncertainty increases (figure 4, top), whilst, when the robot visits places in the environment already learned, uncertainty decreases for all features (figure 4, bottom), even for those not visible from the current robot location but related to current observations (i.e. "backwards estimation") through the off-diagonal elements of the covariance matrix (dashed area). This is a well-known effect historically used by the topographers to build maps by distributing errors among measurements after reobserving already measured points.

Complete solutions from the available data have been obtained considering first, a laser-only-based solution where the map is formed just by segments, and second a multisensor-based solution, where the map is not only formed by segments but also corners and semiplanes which derive from multisensor fusion, thus, achieving a semantically upgraded representation of the navigation area. Both approaches have successfully solved the problem of recognizing places of the navigation area already visited and learned, loosely speaking, the robot has been able to say: 'I have been here before!'. Figure 5 shows a comparison between the robot's location estimation errors in which each technique has incurred compared to the real location of the robot obtained by a pair of theodolites. The figure represents the distance error and the angular error between real and estimated locations for both odometry estimates, laser-only-based estimates and multisensor-based estimates. Errors obtained by either a laser-only-based approach or a multisensor-based approach are comparable, and clearly lower than odometric errors. Nevertheless, the multisensor-based solution obtained, in almost 80% of the cases, a better estimate for the robot location than the laser-only-based solution. Also, we have performed a $\chi^2$ test based on the squared Mahalanobis distance to verify the appropriateness of the solutions from a probabilistic point-of-view, observing that both solutions approached the 95% theoretical threshold.

5 Conclusions

Our work has concentrated on the simultaneous map building and localization problem for a mobile robot navigating indoors. Environmental information has been obtained by a multisensor system (laser and vision). Information fusion has been performed at the level of observations, resulting in highly reliable and robust observations from an early stage of the processing. Suboptimal estimation has been combined with the SPmodel. The work has highlighted the idea of "backwards estimations" due to the dependencies between entities stored in the covariance matrix. Satisfactory results have been obtained concerning the problem of revisiting previously learned places of the navigation area, both using a laser-only-based approach and a multisensor-based approach. The multisensor-based approach obtained better robot's location estimations in almost 80% of the trajectory points, as well as it provided a semantically upgraded representation of the environment.

Further work is planned to increase the semantical meaning of the representation of the map of the environment by labelling doors, corridors, rooms, etc.
Figure 4: Integration of local maps into the global map with multisensor-based features. Trajectory points: 29 (top) and 54 (bottom). A model map has been drawn for reference purposes.

References


Figure 5: Comparison of errors along the trajectory.