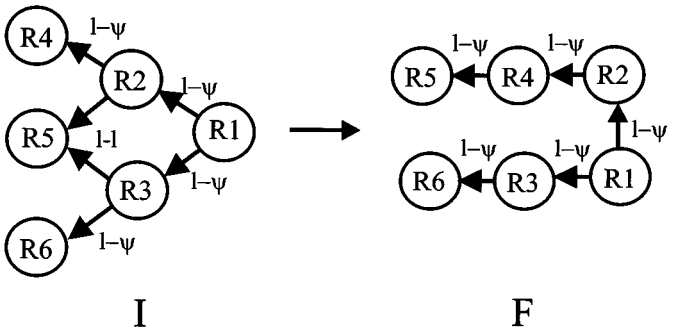


Fig. 7. Initial and final control graph for Example 2.

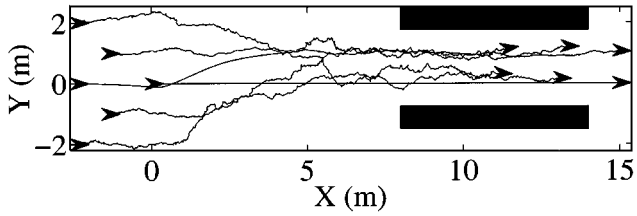


Fig. 8. Formation change for six robots in the presence of sensory noise.

V. Conclusion

In this paper, we have studied strategies for controlling formations of mobile robots using methods from nonlinear control theory and graph theory. We have focused on decomposing the problem of controlling a formation of nonholonomic mobile robots into: 1) controlling a single lead robot and 2) controlling other follower robots in the team. We used the terms $l-\psi$ and $l-l$ control to reflect whether the control laws are based on tracking the position and orientation of the robot relative to a leader, or the position relative to two leaders, respectively. We also defined the concept of a transition matrix, which governs the addition and deletion of edges in the control graph and hence the change in the communication protocol. Based on this, we presented an exhaustive list of all possible transitions that can occur within the robots in the formation and the corresponding transition matrix column.

There are several important issues that need to be addressed in future research in this area, including: 1) how to choose a control graph and the desired shape based on the constraints in the environment; 2) how to plan changes in $(g, r, \mathcal{H})$ depending on sensor constraints; 3) how to allow formations to be split into sub-formations, leading to multiple lead robots; and 4d) though the transition matrix gives us the information needed to change formations, it is not clear if there is an optimal way for carrying out these changes, rather than the sequential algorithm presented here. Some of these topics are the focus of our present research.

References

[1] J. C. Latombe, *Robot Motion Planning*. Boston, MA: Kluwer, 1991.
[2] R. W. Beard, J. Lawton, and F. Y. Hadaegh, "A feedback architecture for formation control," in *Proc. American Control Conf.*, vol. 6, 2000, pp. 4087–4091.
[3] J.-P. Laumond, P. E. Jacobs, M. Taix, and R. M. Murray, "A motion planner for nonholonomic mobile robots," *IEEE Trans. Robot. Automat.*, vol. 10, pp. 577–593, Oct. 1994.
[4] M. Egerstedt and H. Xiaoming, "Formation constrained multi-agent control," in *IEEE Int. Conf. Robotics and Automation*, vol. 4, 2001, pp. 3961–3966.
[5] J. P. Desai and V. Kumar, "Nonholonomic motion planning for multiple mobile manipulators," in *Proc. 1997 IEEE Int. Conf. on Robotics and Automation*, vol. 4, Albuquerque, NM, Apr. 1997, pp. 3409–3414.
[6] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *Int. J. Robot. Res.*, vol. 5, no. 1, pp. 90–99, 1986.
[7] T. Balch and R. C. Arkin, "Behavior-based formation control for multi-robot teams," *IEEE Trans. Robot. Automat.*, vol. 14, pp. 926–939, Dec. 1999.
[8] D. Swaroop and J. K. Hedrick, "String stability of interconnected systems," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 349–357, Mar. 1996.
[9] D. Yanakiev and I. Kanellakopoulos, "A simplified framework for string stability in AHS," in *Preprints of the 13th IFAC World Congress*, vol. Q, 1996, pp. 177–182.
[10] J. P. Desai, J. P. Ostrowski, and V. Kumar, "Controlling formations of multiple mobile robots," in *Proc. 1998 IEEE Int. Conf. Robotics and Automation*, vol. 4, Leuven, Belgium, May 1998, pp. 2864–2869.
[11] J. P. Desai, "Motion planning and control of cooperative robotic systems," Ph.D. dissertation, Univ. of Pennsylvania, Philadelphia, Oct. 1998.
[12] J. P. Desai, V. Kumar, and J. P. Ostrowski, "Control of changes in formation for a team of mobile robots," in *Proc. 1999 IEEE Int. Conf. on Robotics and Automation*, vol. 2, Detroit, MA, May 1999, pp. 1556–1561.
[13] N. Deo, *Graph Theory With Applications to Engineering and Computer Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

Multisensor Fusion for Simultaneous Localization and Map Building

J. A. Castellanos, J. Neira, and J. D. Tardós

***Abstract*—This paper describes how multisensor fusion increases both reliability and precision of the environmental observations used for the simultaneous localization and map-building problem for mobile robots. Multisensor fusion is performed at the level of landmarks, which represent sets of related and possibly correlated sensor observations. The work emphasizes the idea of partial redundancy due to the different nature of the information provided by different sensors. Experimentation with a mobile robot equipped with a multisensor system composed of a 2-D laser rangefinder and a charge coupled device camera is reported.**

***Index Terms*—Correlation, landmark, mobile robot, multisensor fusion, simultaneous localization and map building.**

I. Introduction

Reliable and accurate sensing of the environment of a mobile robot is an important task both in localizing the robot and in building a detailed map of such an environment. One of the fundamental ideas to achieve this reliability is the use of redundancy, that is, to combine environmental information obtained by several sensors [1]–[3]. Dealing with redundancy requires both the availability of a systematic description of uncertain geometric information and a consistent multisensor fusion mechanism [4].

Different approaches to the simultaneous localization and map-building (SLAM) problem for mobile robots have been reported in the literature after the seminal paper of Smith *et al.* [5] and the early

Manuscript received October 12, 2000; revised June 6, 2001. This paper was recommended for publication by Associate Editor N. Xi and Editor S. Hutchinson upon evaluation of the reviewers' comments. This work was supported by the Spanish Dirección General de Investigación Projects DPI2000-1265 and DPI2000-1272.

The authors are with the Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza, E-50015 Zaragoza, Spain (e-mail: jacaste@posta.unizar.es; jneira@posta.unizar.es; tardos@posta.unizar.es).

Publisher Item Identifier S 1042-296X(01)10904-3.

experiments of Moutarlier *et al.* [6] and Leonard *et al.* [7]. Active research work is being developed in different groups worldwide, typically considering point features gathered by ultrasonic sensors [8] or laser rangefinders [9] which are directly input to the SLAM algorithm. Those approaches present high performance in sparse environments, however, they present low reliability and low robustness of data association in dense environments due to the high ambiguity of the observations used. Alternative approaches are reported in the literature to avoid those limitations: the use of raw sensor data and correlation techniques for data association [10], [11] or the reduction of sensor data ambiguity by detecting more meaningful features, such as 2-D segments from laser data [12], [13].

The work reported in this paper represents, to our knowledge, the first application of multisensor fusion to SLAM for mobile robots and extends the work of [12]. Our main contributions are as follows:

- Perceptual grouping, based on topologic and/or geometric relations between sensor observations, which provides a semantically upgraded landmark-based representation of the navigation area.
- Multisensor fusion at the level of landmarks provides redundancy, usually only partial due to the different nature of the sensors, about the features used for both localizing the robot and mapping the navigation area. Therefore, reliability and precision are increased from early stages of the processing.
- Data association for SLAM benefits from the landmark-based representation due to the low ambiguity of the features involved. Thus, even simple strategies such as nearest-neighbor demonstrate very reliable and robust.

The rest of the paper is structured as follows. We present in Section II the landmark-based description of the environment of a mobile robot based on the probabilistic representation of uncertain geometric information. In Section III, we present the multisensor fusion scheme, where the idea of *partial redundancy* is emphasized. The problem of data association is considered by taking into account both individual compatibility between features and also joint compatibility due to the presence of correlations between the estimated locations of some features. A case study is described in Section IV where experimentation with a mobile robot equipped with a 2-D laser rangefinder and a charge coupled device (CCD) camera is reported. To conclude, Section V contains a discussion of the proposed approach and further research directions.

II. Landmark-Based Representation of the Environment

A. Representing Uncertain Geometric Information

The location of a 2-D uncertain geometric element F, with respect to a base reference W, can be represented [14] by a location vector $\mathbf{x}_{WF} = (x,\ y,\ \phi)^T$ computed as the composition "$\oplus$" of an estimated location vector $\hat{\mathbf{x}}_{WF}$ taken as the base for perturbations, and a differential location vector $\mathbf{d}_F$

$$\mathbf{x}_{WF} = \hat{\mathbf{x}}_{WF} \oplus \mathbf{d}_F. \tag{1}$$

In general, due to the symmetries of the geometric element, some of the components of $\mathbf{d}_F$ do not represent an effective location error, thus a perturbation vector $\mathbf{p}_F$ is formed by the meaningful components of $\mathbf{d}_F$

$$\mathbf{d}_F = \mathbf{B}_F^T \mathbf{p}_F; \qquad \mathbf{p}_F = \mathbf{B}_F \mathbf{d}_F \tag{2}$$

where the row-selection matrix $\mathbf{B}_F$ is named the *self-binding matrix* of the geometric element. The perturbation vector $\mathbf{p}_F$ is normally distributed $\mathbf{p}_F \sim \mathcal{N}(\hat{\mathbf{p}}_F,\ \mathbf{C}_F)$ with mean $\hat{\mathbf{p}}_F$ and covariance matrix $\mathbf{C}_F$.

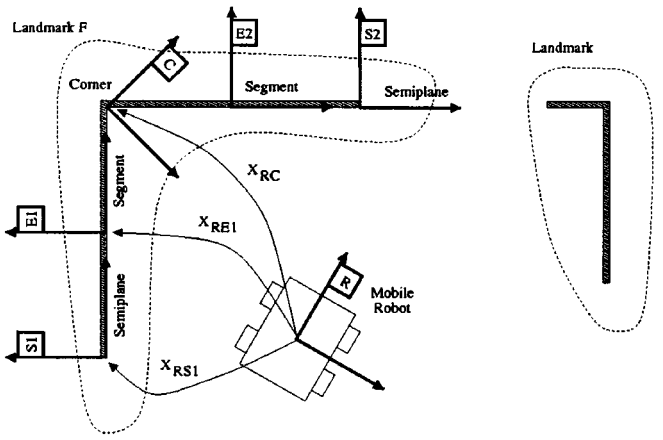


Fig. 1. Landmark $\mathcal{F} = \{S_1,\ E_1,\ C,\ E_2,\ S_2\}$ formed by a set of correlated features of the local environment of the mobile robot.

When $\hat{\mathbf{p}}_F = \mathbf{0}$, we say that the estimation is *centered.* Our approach to the SLAM problem for mobile robots [12] considers a state vector $\hat{\mathbf{x}}^W$ which represents the estimated location of the vehicle and the map features with respect to a base reference W, and a Gaussian perturbation vector $\mathbf{p}^W \sim \mathcal{N}(\hat{\mathbf{p}}^W,\ \mathbf{C}^W)$ which takes into account estimation of effective errors.

B. Features and Landmarks

Similar to other work in this area, we follow a feature-based approach to map building. Then, raw sensor data are first processed to obtain a set of low-level features such as segments obtained from laser scans or vertical edges obtained from gray-level images. In this work, we go a step further and explore the use of more distinct and meaningful observations composed of several related features.

We define a *landmark* $\mathcal{F}$ as a set of nearby features $F_i, i \in \{1 \ldots n\}$ derived from sensor data that verify some topologic and/or geometric properties. It should be noted that perceptual grouping is in general a complex problem, and the reliable detection of landmarks usually require the use of sensor-specific properties. For example, Fig. 1 illustrates a landmark $\mathcal{F}$ defined by the nearby segments E_1 and E_2 and other derived features like the corner C found at the intersection of the segments and the semiplanes S_1 and S_2 found at their free endpoints. Other examples of more complex landmarks, not used in this paper, can be doors detected in images or laser scans, corridor intersections, etc.

From the above example, it can be deduced that the features composing a landmark are not required to be independent. Therefore, generalizing (1) and (2), an estimated location vector and a perturbation vector can be associated to a landmark $\mathcal{F}$

$$\hat{\mathbf{x}}_{R\mathcal{F}} = \begin{bmatrix} \hat{\mathbf{x}}_{RF_1} \\ \vdots \\ \hat{\mathbf{x}}_{RF_n} \end{bmatrix}, \qquad \mathbf{p}_{\mathcal{F}} = \begin{bmatrix} \mathbf{p}_{F_1} \\ \vdots \\ \mathbf{p}_{F_n} \end{bmatrix} \sim \mathcal{N}(\hat{\mathbf{p}}_{\mathcal{F}},\ \mathbf{C}_{\mathcal{F}}). \tag{3}$$

The diagonal elements of the covariance matrix $\mathbf{C}_{\mathcal{F}}$ consider the location uncertainty for each feature of the landmark $\mathcal{F}$ whilst the off-diagonal elements represent the cross covariances between the estimations of the different features of the landmark $\mathcal{F}$. When independence between the estimation of the features holds, $\mathbf{C}_{\mathcal{F}}$ is reduced to a block-diagonal matrix.

The set of landmarks obtained from the information provided by a sensor F when the vehicle is at a particular location k along its trajectory is subsequently referred to as *local map* $\mathbf{LM}_k^F$.

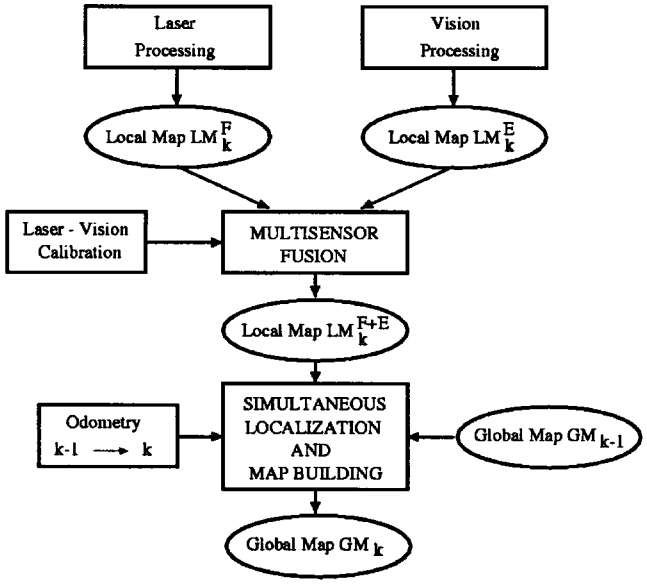


Fig. 2. A case study: fusion of laser and monocular vision for simultaneous localization and map building.

III. Multisensor Fusion

For ease of explanation, suppose that the mobile robot is equipped with a sensor F, considered as the fundamental sensing device, which provides the landmark-based local map $\mathbf{LM}_k^F$, and a sensor E, considered as a secondary sensor, which obtains partially redundant information $\mathbf{LM}_k^E$ from the environment of the robot at location k. Fig. 2 depicts the particular case of a laser rangefinder as the primary sensor and monocular vision as the secondary sensor of a multisensor system. Then, the multisensor fusion algorithm processes, sequentially, the set of landmarks of $\mathbf{LM}_k^F$ and searches for redundancy using the features of $\mathbf{LM}_k^E$. Thus, for each landmark of $\mathbf{LM}_k^F$, we have:

1) Data association, using a nearest neighbor approach and an innovation test based on the Mahalanobis distance [15] decides which features of the local map $\mathbf{LM}_k^E$ are compatible with those of the local map $\mathbf{LM}_k^F$. Due to the existence of correlated features within a given landmark, not only individual compatibility but also joint compatibility of the different matchings must be validated. This step benefits from the availability of an accurate sensor–sensor calibration.
2) Suboptimal estimation based on the EKF integrates the set of matchings found by data association to improve the estimation of $\mathbf{LM}_k^F$ by the partially redundant information provided by $\mathbf{LM}_k^E$. As result, the multisensor-based local map $\mathbf{LM}_k^{F+E}$ is obtained, which is subsequently used as input to the SLAM algorithm [12].

A. Validation of Individual Compatibility

Let $(E_i,\ F_{j_i})$ be a candidate pairing between a feature E_i of landmark $\mathcal{E}$ of $\mathbf{LM}_k^E$ and a feature F_{j_i} of a landmark $\mathcal{F}$ of $\mathbf{LM}_k^F$ which must be validated. The matching imposes a constraint (Fig. 3) on their relative transformation [14] that can be expressed by means of an implicit measurement equation

$$\begin{aligned}\mathbf{f}_i(\mathbf{p}_{\mathcal{F}},\ \mathbf{p}_{E_i}) &= \mathbf{B}_{E_iF_{j_i}}\mathbf{x}_{E_iF_{j_i}} \\ &= \mathbf{B}_{E_iF_{j_i}}(\ominus\mathbf{B}_{E_i}^T\mathbf{p}_{E_i} \oplus \hat{\mathbf{x}}_{E_iF_{j_i}} \oplus \mathbf{B}_{F_{j_i}}^T\mathbf{p}_{F_{j_i}}) \\ &= \mathbf{0}. \end{aligned} \tag{4}$$

The nature of the features F_{j_i} and E_i depends on the characteristics of the sensors used. In general, F_{j_i} and E_i are of different nature, thus only *partial redundancy* can be obtained. This fact is represented by the row-selection matrix $\mathbf{B}_{E_iF_{j_i}}$, which selects the meaningful components of the relative transformation which must be equal to zero. Due to orientation terms, (4) is generally nonlinear, therefore, we consider its first-order Taylor approximation around the best available centered estimation ($\hat{\mathbf{p}}_{\mathcal{F}} = \mathbf{0},\ \hat{\mathbf{p}}_{E_i} = \mathbf{0}$) with[1]

$$\begin{aligned}\mathbf{h}_i &= \mathbf{f}_i(\hat{\mathbf{p}}_{\mathcal{F}},\ \hat{\mathbf{p}}_{E_i}) = \mathbf{B}_{E_iF_{j_i}}\hat{\mathbf{x}}_{E_iF_{j_i}} \\ \mathbf{H}_i &= \left.\frac{\partial \mathbf{f}_i}{\partial \mathbf{p}_{\mathcal{F}}}\right|_{(\hat{\mathbf{p}}_{\mathcal{F}},\hat{\mathbf{p}}_{E_i})} = \left(\mathbf{0}\ \ldots\ \mathbf{0}\ \mathbf{H}_i^{F_{j_i}}\ \mathbf{0}\ \mathbf{0}\ \ldots\ \mathbf{0}\right) \\ \mathbf{H}_i^{F_{j_i}} &= \left.\frac{\partial \mathbf{f}_j}{\partial \mathbf{p}_{F_{j_i}}}\right|_{(\hat{\mathbf{p}}_{\mathcal{F}},\hat{\mathbf{p}}_{E_i})} = \mathbf{B}_{E_iF_{j_i}}\mathbf{J}_{2\oplus}\{\hat{\mathbf{x}}_{E_iF_{j_i}},\ \mathbf{0}\}\mathbf{B}_{E_i}^T \\ \mathbf{G}_i &= \left.\frac{\partial \mathbf{f}_i}{\partial \mathbf{p}_{E_i}}\right|_{(\hat{\mathbf{p}}_{\mathcal{F}},\hat{\mathbf{p}}_{E_i})} = -\mathbf{B}_{E_iF_{j_i}}\mathbf{J}_{1\oplus}\{\mathbf{0},\ \hat{\mathbf{x}}_{E_iF_{j_i}}\}\mathbf{B}_{F_{j_i}}^T \end{aligned} \tag{5}$$

where $\mathbf{J}_{1\oplus}$ and $\mathbf{J}_{2\oplus}$ are the Jacobians of the composition of location vectors [16]. Due to uncertainty, (4) only holds approximately, thus, for a given significance level α, features F_{j_i} and E_i are considered to be compatible if

$$D^2 = \mathbf{h}_i^T\left(\mathbf{H}_i\mathbf{C}_{\mathcal{F}}\mathbf{H}_i^T + \mathbf{G}_i\mathbf{C}_{E_i}\mathbf{G}_i^T\right)^{-1}\mathbf{h}_i \le \chi^2_{r,\,\alpha} \tag{6}$$

with $r = rank(\mathbf{h}_i)$ degrees of freedom and $\chi^2_{r,\,\alpha}$ a threshold obtained from the χ^2 probability distribution. Otherwise, the matching is discarded. Whenever multiple candidates pairings exist for a given feature E_i, a nearest-neighbor strategy is applied.

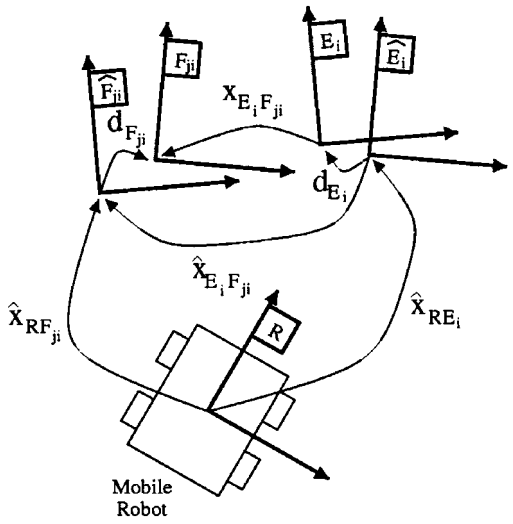


Fig. 3. Improving the estimation of F_{j_i} by the redundant information provided by E_i.

B. Validation of Joint Compatibility

Mutual compatibility of the matchings satisfying (6) must be verified due to the possible existence of correlations between the estimation of the features of a given landmark. The complete set of constraints, of the form described by (4), can be written as

$$\mathbf{f}(\mathbf{p}_{\mathcal{F}},\ \mathbf{p}_{\mathcal{E}}) = \begin{bmatrix} \mathbf{f}_1(\mathbf{p}_{\mathcal{F}},\ \mathbf{p}_{E_1}) \\ \vdots \\ \mathbf{f}_m(\mathbf{p}_{\mathcal{F}},\ \mathbf{p}_{E_m}) \end{bmatrix} = \mathbf{0} \tag{7}$$

[1]The column occupied by $\mathbf{H}_i^{F_{j_i}}$ in matrix $\mathbf{H}_i$ corresponds to the position of the feature F_{j_i} in the state vector of landmark $\mathcal{F}$.

which can be linearized around the best available centered estimation ($\hat{\mathbf{p}}_{\mathcal{F}} = \mathbf{0}$, $\hat{\mathbf{p}}_{\mathcal{E}} = \mathbf{0}$) with coefficients given by

$$\mathbf{h} = \mathbf{f}(\hat{\mathbf{p}}_{\mathcal{F}}, \hat{\mathbf{p}}_{\mathcal{E}}) = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_m \end{bmatrix}$$

$$\mathbf{H} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{p}_{\mathcal{F}}}\right|_{(\hat{\mathbf{p}}_{\mathcal{F}}, \hat{\mathbf{p}}_{\mathcal{E}})} = \begin{pmatrix} \mathbf{0} & \mathbf{H}_1^{F_{j_1}} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_m^{F_{j_m}} & \dots & \mathbf{0} \end{pmatrix}$$

$$\mathbf{G} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{p}_{\mathcal{E}}}\right|_{(\hat{\mathbf{p}}_{\mathcal{F}}, \hat{\mathbf{p}}_{\mathcal{E}})} = diag\,(\mathbf{G}_1, \dots, \mathbf{G}_m)\,. \tag{8}$$

Finally, joint compatibility of the whole set of candidate matchings is satisfied, for a significance level α, if

$$D^2 = \mathbf{h}^T \left(\mathbf{H}\mathbf{C}_{\mathcal{F}}\mathbf{H}^T + \mathbf{G}\mathbf{C}_{\mathcal{E}}\mathbf{G}^T\right)^{-1} \mathbf{h} \leq \chi^2_{r,\,\alpha} \tag{9}$$

with $r = rank(\mathbf{h})$ degrees of freedom. Otherwise we search, among the matches which satisfy individual compatibility, for the largest subset (i.e., the one with greatest value of r) that satisfies joint compatibility. In the case of a tie, the one with a lower value of D^2 is chosen.

C. Formulation of the Multisensor Fusion Equations

An improved perturbation vector $\mathbf{p}_{\mathcal{F}+\mathcal{E}} \sim \mathcal{N}(\hat{\mathbf{p}}_{\mathcal{F}+\mathcal{E}}, \mathbf{C}_{\mathcal{F}+\mathcal{E}})$ for landmark $\mathcal{F}$ is obtained by fusing its previously available information $\mathbf{p}_{\mathcal{F}} \sim \mathcal{N}(\hat{\mathbf{p}}_{\mathcal{F}}, \mathbf{C}_{\mathcal{F}})$ and the partially redundant information $\mathbf{p}_{\mathcal{E}} \sim \mathcal{N}(\hat{\mathbf{p}}_{\mathcal{E}}, \mathbf{C}_{\mathcal{E}})$ provided by landmark $\mathcal{E}$. Using the classical EKF equations, we have

$$\begin{aligned} \mathbf{K} &= \mathbf{C}_{\mathcal{F}}\mathbf{H}^T(\mathbf{H}\mathbf{C}_{\mathcal{F}}\mathbf{H}^T + \mathbf{G}\mathbf{C}_{\mathcal{E}}\mathbf{G}^T)^{-1} \\ \hat{\mathbf{p}}_{\mathcal{F}+\mathcal{E}} &= \hat{\mathbf{p}}_{\mathcal{F}} + \mathbf{K}(-\mathbf{h}) \\ \mathbf{C}_{\mathcal{F}+\mathcal{E}} &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{C}_{\mathcal{F}}. \end{aligned} \tag{10}$$

Finally, after sequentially processing the set of landmarks of $\mathbf{LM}_k^F$, the algorithm comes up with a landmark-based local map $\mathbf{LM}_k^{F+E}$ composed of more accurate features. Reliability and robustness of the features are increased due to the partial redundancy provided by the secondary sensor.

IV. A Case Study: 2-D Laser + Monocular Vision

A supervised exploration of a human-made indoor environment was conducted by using a mobile robot equipped with a 2-D laser rangefinder and a CCD camera. Environmental information was regularly obtained from static robot locations. Ground-truth, obtained by a pair of theodolites, was available to validate the results. An overview of the case study was described in Fig. 2 which is subsequently detailed.

A. Laser-Based Local Map, $\mathbf{LM}_k^F$

First, the 2-D laser data [Fig. 4(b)] are processed by a segmentation algorithm [14]. Different uncertain geometric features are obtained: 1) segments, which are considered as low-level features, and 2) corners and semiplanes which semantically upgrade the representation of the environment. Corners are found at the intersection of two consecutive segments whilst semiplanes are found at the free endpoints of segments, and might correspond to door frames or convex obstacles. In the detection of semiplanes, robustness has been enhanced by avoiding false

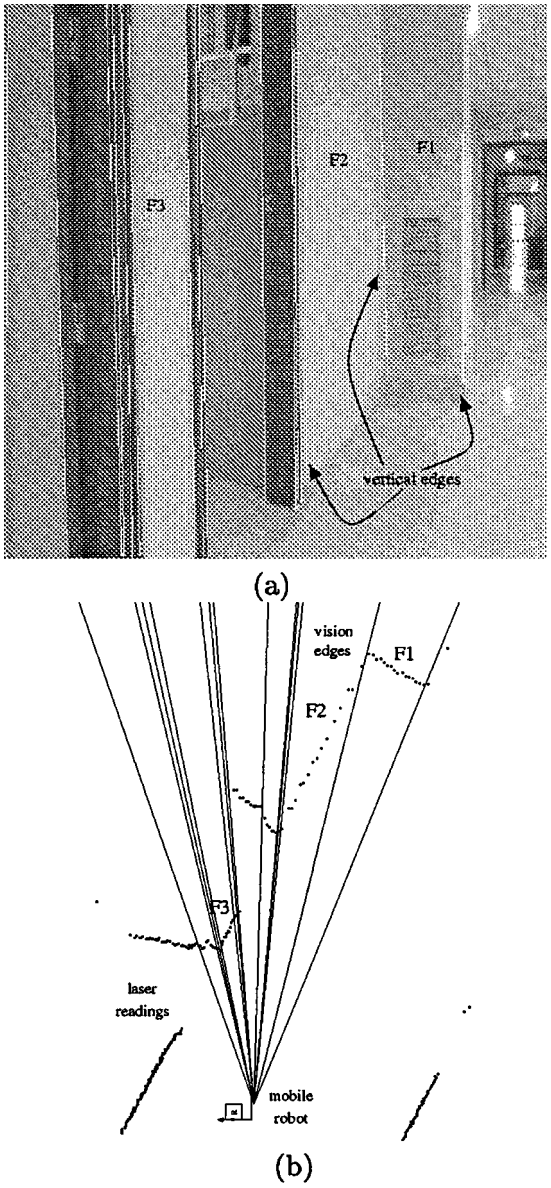


Fig. 4. Sensor data. (a) Vertical edges (white arrow-headed lines) detected in a gray-level image. (b) Two-dimensional laser readings and vision edges.

semiplanes derived from occlusions. Finally, a landmark is formed by each set of consecutive segments and their derived corners and semiplanes [Fig. 5(a)].

B. Vision-Based Local Map, $\mathbf{LM}_k^E$

Redundant information about the location of laser corners and semiplanes can be obtained by processing the gray-level images taken by a CCD camera. The camera has been calibrated using the Tsai method [17] which includes compensation of the lens distorsion. Vertical edges, longer than 150 pixels, are extracted from the image, which might correspond to corners and door frames [Fig. 4(a)]. Each vertical edge detected on the gray-level image is represented by a *vision edge* [Fig. 4(b)], that is, a 2-D line defined by the optical center $\mathcal{O}$ of the camera and the projection of the middle point of the undistorted vertical edge on an horizontal plane containing $\mathcal{O}$. A standard deviation of 0.1 deg is assigned to the angular uncertainty of each vision edge, which corresponds to a detection error of the vertical segment on the image of around 4 pixels. Vision edges are considered as statistically independent features.

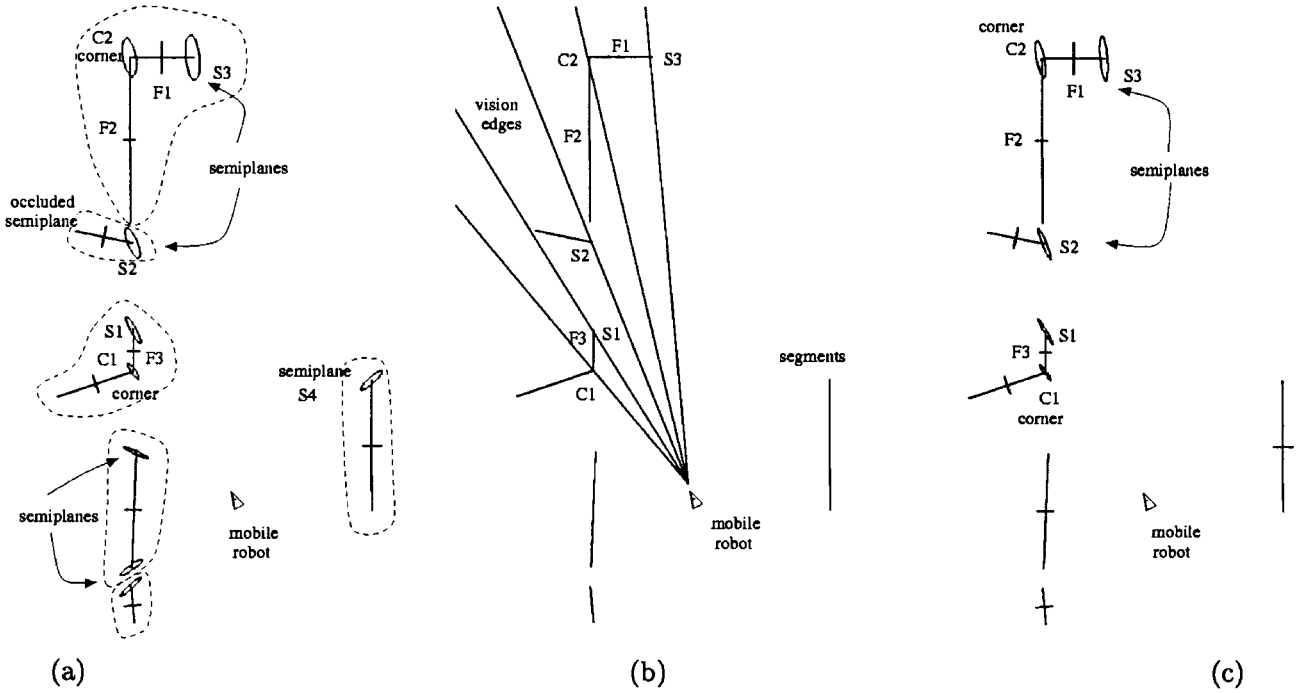


Fig. 5. (a) Laser-based local map with landmarks represented by dashed lines. (b) Vision edges matched with corners and semiplanes. (c) Local map after fusion. Location uncertainty is described by the 95% error bounds magnified $\times 3$.

C. Calibration of the Multisensor System

Multisensor fusion requires sensor–sensor calibration to represent features in a common reference frame. The calibration method consisted of matching observations of a pattern gathered by both sensors. The calibration pattern [14] was composed of two zones: a white-like zone and a black-like zone. The border line between those two zones was easily detected by the CCD camera. Additionally, the angle between the two pattern planes ensured the detection of two nonparallel segments by the laser rangefinder. A system of n nonlinear constraints obtained from the observations of the pattern from n different locations was solved by applying the Levenberg–Marquadt algorithm.

D. Fusion of Laser and Vision Local Maps, $\mathbf{LM}_k^{F+E}$

Data association between sensor observations benefits from an accurate sensor-sensor calibration and allows the use of simple algorithms, such as nearest neighbor. Fusion of laser and vision exploits the partially redundant information provided by the vision edges E_i about the location of corners and semiplanes F_{j_i} detected by the laser rangefinder. Individual compatibility between corners and vision edges on the one hand and semiplanes and vision edges on the other hand is validated using (5) and (6), with the self-binding matrices

$$\mathbf{B}_{F_{j_i}} = \mathbf{I}_3, \qquad \mathbf{B}_{E_i} = (\,0 \quad 0 \quad 1\,)$$

and with the binding matrix of the pairing

$$\mathbf{B}_{E_i F_{j_i}} = (\,0 \quad 1 \quad 0\,)$$

which indicates that only partial redundancy is achieved because monocular vision provides directional but not depth information. Due to correlations between the estimated locations of corners and semiplanes, joint compatibility must be also validated using (8) and (9). Fig. 5(b) describes the set of vision edges detected from the gray-level image which have been matched with features previously detected by the laser rangefinder. Fig. 5(c) presents the results of multisensor fusion for the considered local map. Due to the higher

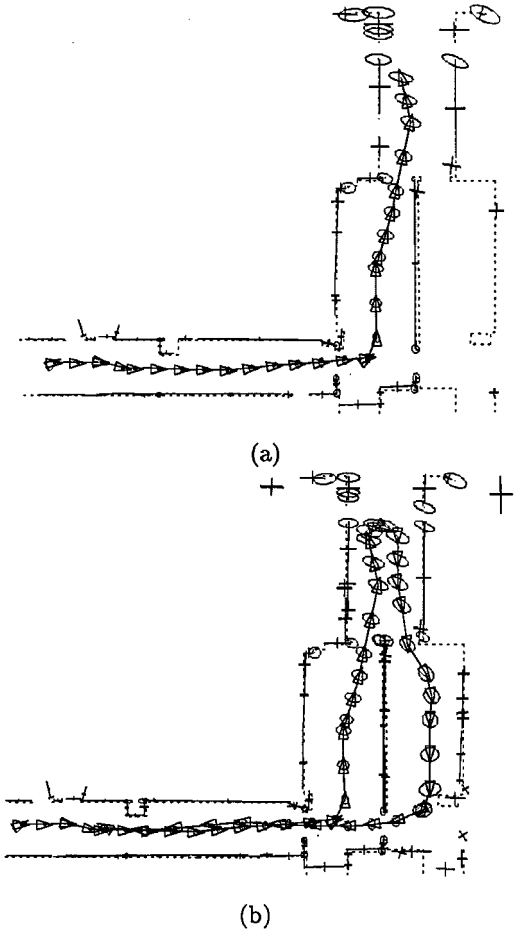
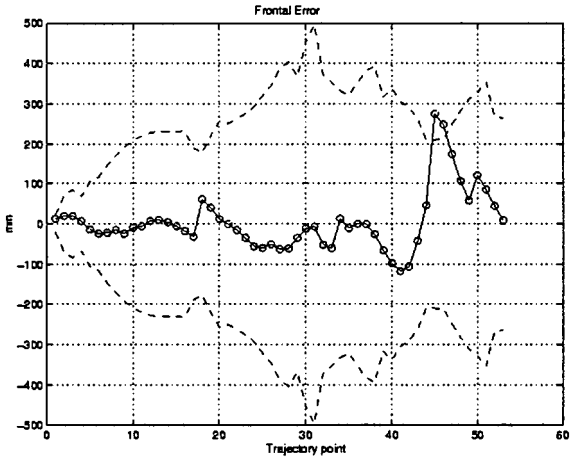
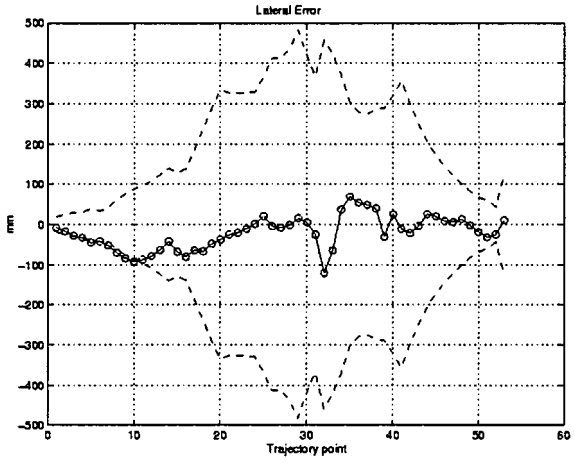


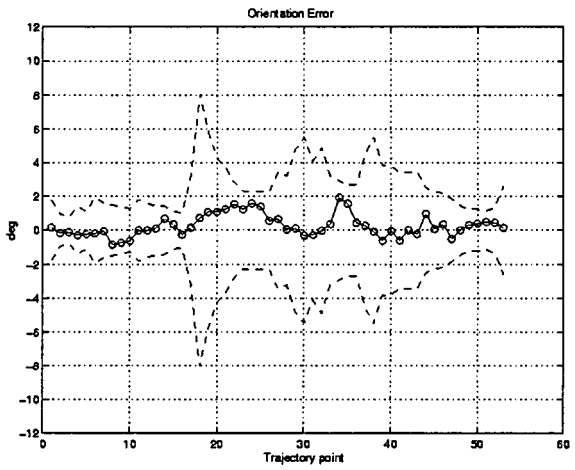
Fig. 6. Evolution of the solution of SLAM using multisensor features. (a) Navigation in unknown areas. (b) Returning to previously visited areas. Dashed-lines correspond to a reference model map.



(a)



(b)



(c)

Fig. 7. Estimation errors and 2σ error bounds for the location of the vehicle. (a) Frontal. (b) Lateral. (c) Orientation error.

angular resolution of the CCD camera as compared to the 2-D laser rangefinder, precision of the estimated location of geometric features increases. Only reliable features are kept after fusion. Thus, semiplanes observed by the laser rangefinder but not confirmed by the CCD camera are removed from the representation of the local map. In Fig. 5(a), the semiplane S_4 was detected and interpreted as a possible door-frame; however, redundancy about its location was not provided by the vision sensor [Fig. 5(b)], thus, the semiplane was removed from further processing [Fig. 5(c)]. Therefore, the multisensor fusion approach increases the reliability and robustness of the local map features.

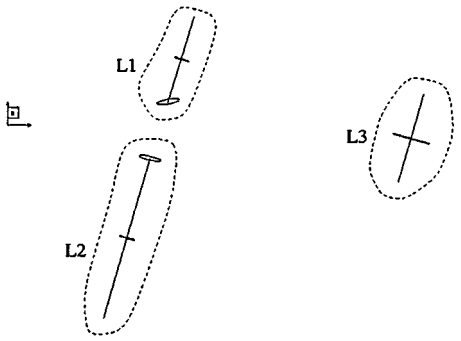


Fig. 8. Multisensor observations: landmarks L_1, L_2, and L_3. Observe that landmarks L_1 and L_2 are formed by a segment and a semiplane.

E. Simultaneous Localization and Map Building

As described in Fig. 2, at a given time instant k, the location of the mobile robot along its trajectory and the structure of the environment are estimated by matching the multisensor-based features $\mathbf{LM}_k^{F+E}$ and the available features of the environment up to time $k-1$, $\mathbf{GM}_{k-1}$. The robustness and efficiency of data association is improved for two reasons. First, the validation of joint compatibility reduces the number of spurious matches which satisfy the innovation test. Second, the use of more distinct and meaningful features reduces the computational complexity by reducing the number of possible candidate features for each feature of the local map.

Fig. 6 describes the solution to the SLAM problem by considering, at each point of the robot trajectory, the information provided by the multisensor system. In the figure, $\pm 2\sigma$ position uncertainty regions have been drawn for the mobile robot and the corners and semiplanes contained in the map. In the case of segments, only lateral uncertainty is represented. Initially, the mobile robot explores previously unknown regions of the environment [Fig. 6(a)], therefore its uncertainty keeps growing at a rate related to sensor imprecision. When the robot returns to previously visited regions [Fig. 6(b)], its uncertainty reduces to the level of uncertainty of the reobserved features. Along the trajectory, a maximum frontal error of 27 cm, a maximum lateral error of 12 cm, and a maximum orientation error of 2 deg was obtained.

Consistency of the estimated location of the mobile robot is a crucial aspect of SLAM. In our experimentation, a nondivergent solution was obtained, that is, the location uncertainty was not optimistically computed along the robot trajectory. Fig. 7 shows frontal, lateral, and orientation errors for the estimated location of the mobile robot along its trajectory, together with the $\pm 2\sigma$ bounds. Compatibility between the estimated robot trajectory and ground-truth reached 96.2% (i.e., only 3.8% of the estimated errors were outside the computed error bounds) where the significance level was set to 5%.

F. Robustness of Landmarks

The robustness of the landmark-based approach is demonstrated by an experiment that simulates the revisiting of a previously mapped area [upper part of Fig. 6(b)], which is one of the most critical data association issues in the SLAM problem. In the experiment, we used the observations obtained by the multisensor system (Fig. 8) to try to relocate the robot within the map. To analyze robustness, we generated a set of 100 simulated robot locations with a random Gaussian perturbation of ± 1 m in position and ± 10 deg in orientation, around the true robot location [Fig. 9(a)].

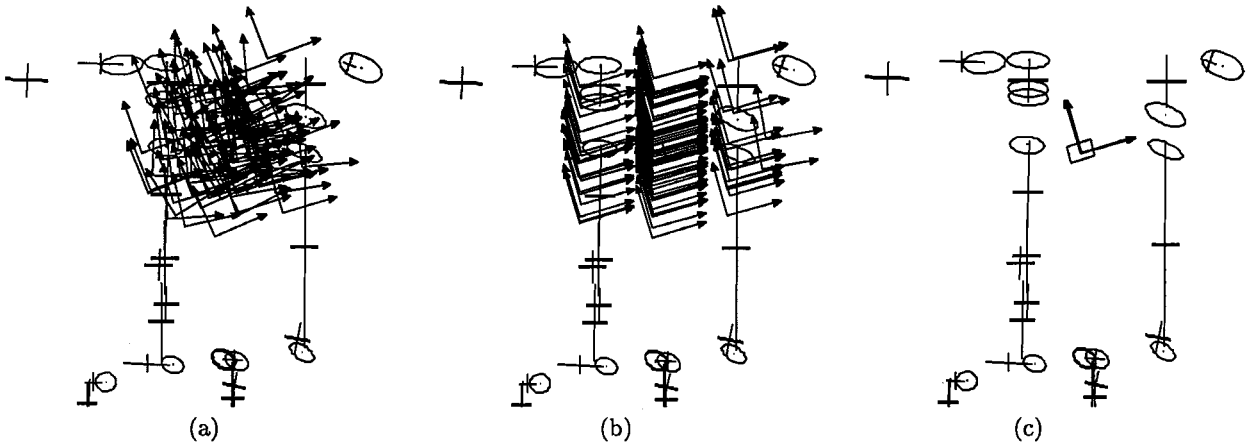


Fig. 9. Robustness of the landmark-based approach. (a) Simulated robot locations. (b) Segment-based robot locations. (c) Landmark-based robot locations. Observe that only the top-right part of the map is shown. The small rectangle represents the real location of the vehicle.

The nearest-neighbor approach discussed in Section III was used to match the observations of the multisensor system with the features of the map. The constraints imposed by the matchings were used, within the EKF framework, to improve the location of the vehicle. The results of the experiment demonstrate the low reliability of the solutions obtained by a segment-based approach [Fig. 9(b)] due to spurious matchings. However, the landmark-based approach always found the correct matching, given a consistent estimation of the robot location [Fig. 9(c)].

V. Conclusion

This paper describes the benefits of using multisensor fusion for the simultaneous localization and map building problem for mobile robots. Some of the key points described are: 1) the landmark-based representation of the environment based on topologic and/or geometric relations between sensor data; 2) multisensor fusion at the level of landmarks, which provides partial redundancy between sensor observations, and which increases both reliability and precision from early stages of the processing; and 3) data association for SLAM, which benefits from the landmark-based representation due to the low ambiguity of the features involved.

Further work concentrates on: 1) increasing the structuration and semantical meaning of the mapped area by identifying corridors, rooms, intersections, etc.; 2) reducing the computational complexity of the SLAM algorithm to achieve a real-time implementation; and 3) improving data association for mobile robot localization in previously mapped areas.

References

[1] P. Grandjean and A. Robert de Saint Vincent, "3D modeling of indoor scenes by fusion of noisy range and stereo data," in *IEEE Int. Conf. Robot. Automat.*, Scottsdale, AZ, 1989, pp. 681–687.

[2] G. Dudek, P. Freedman, and I. M. Rekleitis, "Just-in-time sensing: Efficiently combining sonar and laser range data for exploring unknown worlds," in *IEEE Int. Conf. Robot. Automat.*, Minneapolis, MN, 1996, pp. 667–672.

[3] K. T. Song and W. H. Tang, "Environment recognition for a mobile robot using double ultrasonic sensors and a CCD camera," in *Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, Las Vegas, NV, 1994, pp. 715–722.

[4] H. F. Durrant-Whyte, "Sensor models and multisensor integration," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 97–112, 1988.

[5] R. Smith, M. Self, and P. Cheeseman, "A stochastic map for uncertain spatial relationships," in *Robot. Res., the Fourth Int. Symp.*, R. Bolles and B. Roth, Eds. Cambridge, MA: MIT Press, 1988, pp. 467–474.

[6] P. Moutarlier and R. Chatila, "Stochastic multisensory data fusion for mobile robot location and environment modeling," in *Proc. 5th Int. Symp. on Robot. Res.*, Tokyo, Japan, 1989, pp. 207–216.

[7] J. J. Leonard, H. F. Durrant-Whyte, and I. J. Cox, "Dynamic map building for an autonomous mobile robot," *Int. J. Robot. Res.*, vol. 11, no. 4, pp. 286–298, 1992.

[8] H. J. S. Feder, J. J. Leonard, and C. M. Smith, "Adaptive mobile robot navigation and mapping," *Int. J. Robot. Res.*, vol. 18, no. 7, pp. 650–668, 1999.

[9] J. Guivant, E. Nebot, and H. F. Durrant-Whyte, "Simultaneous localization and map building using natural features in outdoor environments," in *Proc. Intelligent Autonomous Syst. 6*, Venice, Italy, July 2000, pp. 581–588.

[10] J. S. Guttmann and K. Konolige, "Incremental mapping of large cyclic environments," in *Int. Symp. Computat. Intell. Robot. Automat.*, Monterey, CA, Nov. 1999.

[11] S. Thrun, W. Burgard, and D. Fox, "A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping," in *IEEE Int. Conf. Robot. Automat.*, San Francisco, CA, Apr. 2000, pp. 321–328.

[12] J. A. Castellanos, J. M. M. Montiel, J. Neira, and J. D. Tardós, "The SPmap: A probabilistic framework for simultaneous localization and map building," *IEEE Trans. Robot. Automat.*, vol. 15, pp. 948–952, Oct. 1999.

[13] T. Bailey, E. M. Nebot, J. K. Rosenblatt, and H. F. Durrant-Whyte, "Data association for mobile robot navigation: A graph theoretic approach," in *IEEE Int. Conf. Robot. Automat.*, San Francisco, CA, Apr. 2000, pp. 2512–2517.

[14] J. A. Castellanos and J. D. Tardós, *Mobile Robot Localization and Map Building: A Multisensor Fusion Approach*. Boston, MA: Kluwer, 1999.

[15] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. New York: Academic, 1988.

[16] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Proc. 2nd Workshop on Uncertainty in Artificial Intelligence*, Philadelphia, PA, 1986, pp. 167–193.

[17] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Automat.*, vol. RA-4, pp. 323–344, June 1987.