Robotica (2009) volume 27, pp. 653–662. © 2008 Cambridge University Press
doi:10.1017/S026357470800502X Printed in the United Kingdom

Video data validation by sonar measures for robot localization and environment feature estimation

A. Bonci, G. Ippoliti, A. La Manna and S. Longhi*

Dipartimento di Ingegneria Informatica, Gestionale e dell'Automazione, Università Politecnica delle Marche Via Brecce Bianche, 60131 Ancona, Italy.

(Received in Final Form: July 18, 2008. First published online: August 29, 2008)

SUMMARY

In this paper, the robust robot localization problem with respect to uncertainties on environment features is formulated in a stochastic setting, and an extended Kalman filtering approach is proposed for the integration of odometric, video camera, and sonar measures. The environment is supposed to be only partially known, and a probabilistic method for sensor data fusion aimed at increasing the environment knowledge is considered.

KEYWORDS: Robot localization; Mobile robots; SLAM; Mechatronic systems; Pose estimation and registration.

1. Introduction

Many practical applications, where a real robot autonomy is necessary, require an accurate localization system based on an efficient integration of multiple sensor information.[1–3] A mobile robot must be able to localize itself, i.e. to determine its own configuration (position and orientation) with respect to the surrounding environment.[4–6] The localization problem is generally solved by relative or absolute approaches. The first ones are based on the data provided by sensors measuring the dynamics of variables internal to the vehicle; absolute localization requires sensors measuring some parameters of the environment in which the robot is operating. In this case the definition of appropriate environment maps is also required. The current trend is to exploit the complementary nature of these two kinds of sensory information to improve the precision of the localization procedure (see e.g. [7, 8]), at expense of an increased cost and computational complexity. The aim is to improve the mobile robot autonomy by enhancing its capability of localization with respect to the surrounding environment.

In this framework the research interests have been focused on multisensor systems because of the limitations inherent in any single sensory device that can only supply partial information on the environment, thus limiting the ability of the robot to localize itself. The methods and algorithms proposed in the literature for an efficient integration of multisensor information differ according to the *a priori* information on the environment, which may be almost known and static, or almost unknown and dynamic. Therefore, in this context different solutions to the problem of map building have been developed in order to allow a precise and robust localization in real-world environments.[9,10] An extensive survey of various probabilistic techniques for robotic mapping has been issued.[11]

The problem of simultaneous localization and map building has been deeply investigated in the literature, and different solutions have been proposed (see e.g. refs [12–16]). A comprehensive overview of the commonest approaches is available.[17,18] Topological, metric, and hybrid schemes have been proposed in order to improve the efficiency and accuracy of the simultaneous localization and map building, where algorithms that can build a map of the environment, use the same map to localize the mobile robot. Other solutions are based on the particle filtering approach.[19–21] The key idea of these algorithms is that, if one could exactly know the robot trajectory, observations of different features would be statistically independent, then each feature in the map is updated separately.

At present, the most widespread approach to the simultaneous localization and map building is the use of stochastic estimation techniques to build and maintain current estimates of the vehicle position and of the environment map comprehending specific feature locations. The use of stochastic estimation techniques such as a suitably defined extended Kalman filter (EKF) appears to be the right tool to cope with this problem (see e.g. refs. [22–27]). The filter estimates the current position and orientation of the mobile robot, which is subsequently fed to an on-line map building algorithm. To obtain an efficient integration of map building and localization, the acquired knowledge on the environment must be represented by parametric features with the associated uncertainty. In this approach the computational complexity can be reduced if a hybrid solution is implemented[28]; the technique presented in ref. [29] is based on a global topological map which connects local metric maps with a reduced number of features.

In this paper, a localization technique based on the data fusion of internal and external sensors with improved robustness against possible uncertainties on the environment is proposed. A metric representation of the environment is considered with the advantage of offering a lower computational complexity in a hybrid solution to the simultaneous localization and map building problem. A procedure is developed for the consistent and simultaneous

* Corresponding author. E-mail: sauro.longhi@univpm.it

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

on-line estimation of the vehicle position and the environment features both, based on an EKF approach. This is obtained by defining a state–space model whose state vector contains both the state variables of the vehicle model and the state variables of the environment features.

A video camera and a proper set of sonars are used as exteroceptive sensors. In the developed strategy, a probabilistic method for the fusion of sonar and video data is considered for improving the reliability of the vision system to detect environment landmarks. This procedure results in a computationally efficient solution to the localization problem.

The main technical novelties of this paper are: (i) the introduction of a set of landmark measures acquired by a video camera and validated by a probabilistic map of the environment which is an occupancy grid map built making use of the sonar readings[1]; (ii) the integration in the same framework of the robot pose estimation and of the environment feature estimation problems.

The experimental tests, performed on the LabMate mobile robot in an indoor environment, show a significant improvement in the performance of the vehicle localization and environment estimation. In the performed tests the uncertainties on the mobile base localization and on the estimate of environment features are reduced without a significant increment of computational efforts.

The paper is organized in the following way. The robot model and the sensor device equipment are described in Section 2. The procedure for ultrasonic and video data fusion is given in Section 3. Details on the robust localization are reported in Section 4. The sonar reading selection is discussed in Section 5. Section 6 contains a detailed description of the experimental tests performed on the mobile base.

2. The Sensor Equipment

2.1. Odometric measures

The considered model of the vehicle is that of a unicycle-like mobile robot with two driving wheels, mounted on the left and right sides of the robot, with their common axis passing through the center of the robot. Let (O, X, Y) be the inertial coordinate system, and (O', X', Y') the coordinate system fixed to the mobile robot. The localization of this robot in a two-dimensional space requires the knowledge of the coordinates x and y [referred to (O, X, Y)] of the midpoint between the two driving wheels (the origin O' of the coordinate system fixed to the mobile robot) and of the angle θ between the main axis of the robot (coincident with the X'-direction) and the X-direction. The kinematic model of the unicycle robot is described by the following equations:

$$\dot{x}(t) = v(t)\cos\theta(t), \tag{1}$$

$$\dot{y}(t) = v(t)\sin\theta(t), \tag{2}$$

$$\dot{\theta}(t) = \omega(t), \tag{3}$$

where $v(t)$ and $\omega(t)$ are the displacement and angular velocities of the robot, respectively, and are expressed as

$$v(t) = \frac{\omega_{\mathrm{r}}(t) + \omega_{\mathrm{l}}(t)}{2} r, \tag{4}$$

$$\omega(t) = \frac{\omega_{\mathrm{r}}(t) - \omega_{\mathrm{l}}(t)}{d} r, \tag{5}$$

where $\omega_{\mathrm{r}}(t)$ and $\omega_{\mathrm{l}}(t)$ are the angular velocities of the right and left wheels respectively, r is the wheel radius, and d is the distance between the wheels.

Assuming $\omega_{\mathrm{r}}(t)$ and $\omega_l(t)$ constant over a sufficiently small sampling period $\Delta t_k \triangleq t_{k+1} - t_k$, the increments in the position and orientation of the robot during Δt_k can be expressed as[30]

$$\Delta x(t_{k+1}) = x(t_{k+1}) - x(t_k)$$

$$= \bar{v}(t_k)\Delta t_k \frac{\sin((\bar{\omega}(t_k)\Delta t_k)/2)}{(\bar{\omega}(t_k)\Delta t_k)/2} \cos\left(\theta(t_k) + \frac{\bar{\omega}(t_k)\Delta t_k}{2}\right), \tag{6}$$

$$\Delta y(t_{k+1}) = y(t_{k+1}) - y(t_k)$$

$$\bar{v}(t_k)\Delta t_k \frac{\sin((\bar{\omega}(t_k)\Delta t_k)/2)}{(\bar{\omega}(t_k)\Delta t_k)/2} \sin\left(\theta(t_k) + \frac{\bar{\omega}(t_k)\Delta t_k}{2}\right), \tag{7}$$

$$\Delta\theta(t_{k+1}) = \theta(t_{k+1}) - \theta(t_k) = \bar{\omega}(t_k)\Delta t_k, \tag{8}$$

where $\bar{v}(t_k)\Delta t_k$ and $\bar{\omega}(t_k)\Delta t_k$ are as follows:

$$\bar{v}(t_k)\Delta t_k = \frac{\Delta q_{\mathrm{r}}(t_k) + \Delta q_{\mathrm{l}}(t_k)}{2} r, \tag{9}$$

$$\bar{\omega}(t_k)\Delta t_k = \frac{\Delta q_{\mathrm{r}}(t_k) - \Delta q_{\mathrm{l}}(t_k)}{d} r. \tag{10}$$

The terms $\Delta q_{\mathrm{r}}(t_k)$ and $\Delta q_{\mathrm{l}}(t_k)$ are the incremental distances covered on the interval Δt_k by the right and left wheels of the robot, respectively.

2.2. Sonar measures

The distance readings by sonar sensors are related to the indoor environment model and to the configuration of the mobile robot. Considering a planar distribution of n_{s} sonar sensors, denote by x'_h, y'_h, θ'_h the position of the hth sonar, $1 \le h \le n_{\mathrm{s}}$, in the coordinate system (O', X', Y') fixed to the mobile robot. At the sampling time t_k, the position x_h, y_h, and the orientation θ_h of the hth sonar, in the inertial coordinate system (O, X, Y), have the following form:

$$x_h(t_k) = x(t_k) + x'_h\cos\theta(t_k) - y'_h\sin\theta(t_k), \tag{11}$$

$$y_h(t_k) = y(t_k) + x'_h\sin\theta(t_k) + y'_h\cos\theta(t_k), \tag{12}$$

$$\theta_h(t_k) = \theta(t_k) + \theta'_h. \tag{13}$$

The walls and the obstacles in an indoor environment are represented by a proper set of planes orthogonal to the XY plane of the inertial coordinate system. Each plane P^j, $j = 1, 2, \ldots, n_{\mathrm{p}}$ (where n_{p} is the number of planes which describe the indoor environment), is represented by the triplet P_{r}^j, P_{n}^j, P_{v}^j, where P_{r}^j is the normal distance of the plane from the origin O, P_{n}^j is the angle between the normal line to the plane and the X-direction, and P_{v}^j is a binary variable, $P_{\mathrm{v}}^j \in \{-1, 1\}$, which defines the face of the plane reflecting the sonar beam. This parametric representation is suitable for environments with a simple geometry; however, the presence

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

of possible unexpected obstacles can be properly filtered by the method of sensor reading selection presented in Section 5.

In such a notation, the expectation $d_h^j(t_k)$ for the present distance of the hth sonar from the plane P^j has the following expression[31]:

$$d_h^j(t_k) = P_{\mathrm{v}}^j \cdot \left(P_{\mathrm{r}}^j - x_h(t_k)\cos P_{\mathrm{n}}^j - y_h(t_k)\sin P_{\mathrm{n}}^j\right), \quad (14)$$

if $P_{\mathrm{n}}^j \in [\theta_h(t_k) - \delta/2, \theta_h(t_k) + \delta/2]$, where δ is the beamwidth of the sonar sensor. To simplify the position estimation algorithm without an appreciable reduction of accuracy, the echoes traveling along the edges of the sonar cone, that is the pattern where most of the power of the ultrasound wave is concentrated, have been neglected.

2.3. Video camera measures

A video camera and the associated image processing procedures have been used for extracting environment features that are related to the indoor environment model and to the robot configuration. The structured environment, as stated in the previous subsection, is characterized by walls and obstacles that are represented by a proper set of planes orthogonal to the XY plane of the inertial coordinate system. The video camera is used for extracting the walls and obstacle borders that are straight lines on the XY plane. Therefore, in the XY plane a feature-based modeling (FBM) of the environment is based on a set of straight line segments.

A CCD video camera installed on the mobile robot is used for the image formation; reference is made to the pinhole model.[32] To reduce the computational efforts, the "visible space" is considered. In the pinhole model, the viewing frustum of the video camera is the projection of the image plane corners from the pinhole, which is located one focal length behind the image plane.[33] Pointing the camera down in the left side with respect to the forward direction of the robot, the "visible space" on the ground plane is defined by the projection of the frustum vertices on the floor plane[1] (see Fig. 1). For improving the detection of straight lines representing environment features, the ultrasonic and video data fusion procedure presented in the following section is used. This procedure is able to match the lines extracted from the video images and from the occupancy grid map; the considered line detection tool is the Hough transform (HT).[34] In the HT, the straight-line equation is expressed by

$$\rho = u\cos\phi + v\sin\phi, \quad (15)$$

where ρ is the distance between the straight line and the origin of the coordinate system in the image plane, ϕ is the orientation of the line, and u, v are the coordinates of the edge point belonging to the line. The HT requires an accumulator array $H(\rho, \phi)$, called Hough space, to represent the possible values of (ρ, ϕ); it is generally approximated by a discrete array. The edge points (u, v) are detected by means of an orthogonal differential operator (e.g., the Sobel operator[35]); for each detected edge point the parameters (ρ, ϕ) are estimated and quantized, and the accumulator array is increased accordingly. After this preliminary edge-point processing, the accumulator array is searched for peaks. The peaks identify the parameters of the highest probability lines. In the standard HT the accumulator is increased by the same quantity for each edge point, assuming that each of them contributes equally to the line features. Here, a new algorithm for the accumulator updating[36] is considered which is able to produce an estimate of the line probability.

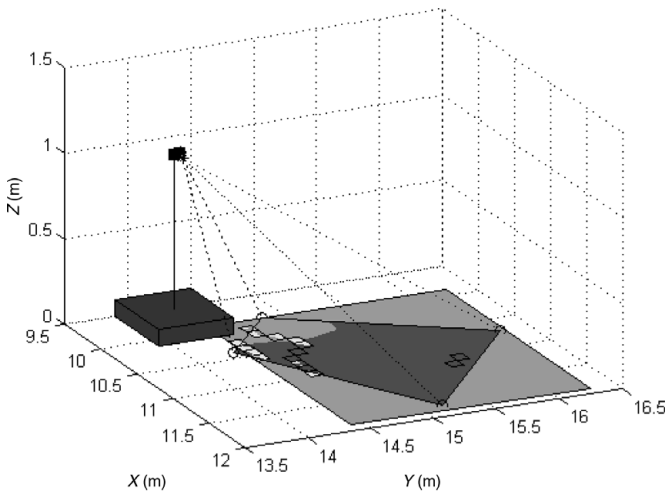


Fig. 1. Visible space of the CCD camera installed on the robot.

3. Ultrasonic and Video Data Fusion

Ultrasonic map building and HT are two standard and well-known techniques that are traditionally used separately for processing different information related to environment features. A new method for integrating these two different kinds of information is proposed here. A procedure of video data validation has been developed for improving the reliability of the vision system to detect environment line features. This procedure makes use of past sonar readings stored in the occupancy grid map. For each environment border detected by the vision system and described by a line of parameters (ρ_j, ϕ_j), a preliminary selection is performed by the use of the occupancy grid map. Here, some technical aspects of the developed validation procedure are introduced. The assumption of straight lines for the environment borders on the floor is considered. Moreover, assume that images of the environment are acquired every M sonar measures. Each sonar measure is composed of n sonar reading, where $n \leq n_{\mathrm{s}}$. The integer M depends on the computational efforts of implemented image processing procedures.

During the robot exploration, the sonar measures are corrupted by specular reflections and poor angular resolution, and a simple method for improving the reliability and accuracy of these measures is to integrate them in an occupancy grid map where the values of the cells are updated using the probabilistic signal level of each sonar reading. In general, the on-line updating of this map requires a low computational cost at each sampling instant. In the proposed ultrasonic and video data fusion the computation of the occupancy grid map is based on the probabilistic approach,[37] which is able to find a trade-off between accuracy and computational cost. The main features of the grid map building procedure that has been implemented are the following: (i) the occupancy grid map consists of square cells (10 cm each side), each one containing a value between

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

0 and 1 representing the probability to find an obstacle in the considered cell; (ii) each sonar measure is used to update the obstacle probability of the cells in the measured zone using the probability profiles estimated by experimental tests; (iii) in each measure, only the central part of the sonar beam is considered (this choice reduces the computational efforts and the measure of dispersion caused by the mobile base); and (iv) the updating of the obstacle probability (probability of occupancy) of each cell is done by Bayes' theorem. A numerical approximation is generally introduced to reduce the computational efforts. The occupancy grid map is updated without any *a priori* environment knowledge.

Straight line segments can be found in the occupancy grid map as aligned cells with high probability of occupancy. By interpreting a grid and its probabilities as an image with different levels of intensity (gray level), the application of the HT to this image allows to detect straight lines and to associate a probability to each detected line in the occupancy grid map. These straight line segments represent the environment borders estimated by the sonar sensors. The initial state of the occupancy grid map is completely unknown because an *a priori* model of the environment is not provided.

The proposed updating process makes use of a probabilistic approach to the HT for extracting line features and the associated certainty values both from video data and from the occupancy grid map. Matching between the lines extracted from video data and from the occupancy grid map is performed for environment features belonging to the part of the floor visible from the video camera (here called "visible space" as in Subsection 2.3). The proposed matching algorithm is based on the combination of the line probabilities encoded in both Hough accumulators.

All the straight lines detected from the video image are matched with the straight lines detected from the portion of the occupancy grid map corresponding to the portion of floor falling in the "visible space" by projecting the straight lines of the video image on the floor plane. The existence probability of the straight lines is stored in both Hough accumulators: $H_{\mathrm{C}}(\rho, \phi)$ for the lines of the video image projected on the floor, and $H_{\mathrm{G}}(\rho, \phi)$ for the lines detected from the occupancy grid map. The matching algorithm is based on the combination of the line probabilities encoded in both Hough accumulators. A Bayesian estimator is considered.[36] For each jth line feature, described by its parameters (ρ_j, ϕ_j), two probabilistic estimates are provided: $P_{\mathrm{C}} = P((\rho_{\mathrm{C}j}, \phi_{\mathrm{C}j})|(\rho_j, \phi_j))$ stored in H_{C}, and $P_{\mathrm{G}} = P((\rho_{\mathrm{G}j}, \phi_{\mathrm{G}j})|(\rho_j, \phi_j))$ stored in H_{G}, where $(\rho_{\mathrm{C}j}, \phi_{\mathrm{C}j})$ is the estimate of the line parameters (ρ_j, ϕ_j) obtained by using the video image, and $(\rho_{\mathrm{G}j}, \phi_{\mathrm{G}j})$ is the estimate of the line parameters (ρ_j, ϕ_j) obtained by using the occupancy grid map. Using Bayes' theorem, the combined estimate $P((\rho_j, \phi_j)|(\rho_{\mathrm{C}j}, \phi_{\mathrm{C}j}) \cup (\rho_{\mathrm{G}j}, \phi_{\mathrm{G}j}))$ is given by

$$P((\rho_j, \phi_j)|(\rho_{\mathrm{C}j}, \phi_{\mathrm{C}j}) \cup (\rho_{\mathrm{G}j}, \phi_{\mathrm{G}j})) = \frac{(P_{\mathrm{C}} P_{\mathrm{G}})/P(\rho_j, \phi_j)}{(P_{\mathrm{C}} P_{\mathrm{G}})/P(\rho_j, \phi_j) + ((1 - P_{\mathrm{C}})(1 - P_{\mathrm{G}}))/(1 - P(\rho_j, \phi_j))}, \tag{16}$$

that is also known as the independent opinion pool.[38]

In this way, the straight lines of parameters (ρ_j, ϕ_j) detected by the vision system and characterized by a high probability are considered as reliable environment features to be introduced as video readings in the simultaneous robot pose and environment estimation, as stated in the next section.

Each video image has a greater resolution than the grid map, and therefore it is likely to extract more than one video straight line corresponding to each sonar line. If more than one line from the video image corresponds to a single line extracted from the sonar grid map, the algorithm selects the line which is closest to the vehicle; this is the most conservative choice from the perspective of safety. Note that the sonar readings are taken with the same sampling period T of the odometric measures, while the video data are acquired at every M sampling periods. The sonar readings acquired at each sampling time, whose number is $n \leq n_{\mathrm{s}}$, are used as distance readings for the pose estimation of the mobile robot, while the validation of the line features extracted from the video image is performed by the occupancy grid map built using all past sonar measures. This means that the line features in the video image, acquired at every M sampling periods, are validated by the line features in the grid map extracted only until the $M - 1$ sampling period. Therefore it is possible to assume that video and sonar measurements are uncorrelated during the fusion process.

4. Robust Localization with Respect to Uncertainties on Environment Features

The algorithm operates in a stochastic framework. In ref. [30], it has been shown that the incremental errors on the encoder readings especially affect the estimate of the orientation, and reduce their applicability to short trajectories. A state–space approach is adopted with the purpose of defining a more general method merging the information carried by the vehicle model with that provided by the sensor equipment. The estimation algorithm is an EKF defined on the basis of a state equation derived from (1) to (3) and from the environment geometric features, and of a measurement equation containing the odometric measures, the measures obtained through the video system, and the distance measures provided by the set of sonar sensors.

4.1. Robot state

Denote by $X(t) \triangleq [x(t)\ y(t)\ \theta(t)]^T$ the true robot state, and by $U(t) \triangleq [v(t)\ \omega(t)]^T$ the robot control input. The kinematic model of the robot can be written in the compact form of the following stochastic differential equation:

$$\dot{X}(t) = F(X(t), U(t)) + W(t), \tag{17}$$

where $F(X(t), U(t))$ represents the set of Eqs. (1)–(3), and $W(t)$ is a White noise process with mean zero and covariance matrix Q representing the model inaccuracies (parameter uncertainties, slippage, dragging). It is assumed that $Q = \mathrm{diag}(\sigma_1^2, \sigma_1^2, \sigma_2^2)$. The diagonal form of Q is based on the hypothesis that model (17) describes the true dynamics of the three state variables with independent errors.

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

Let $\Delta t_k = T$ be the constant sampling period, and denote t_k by kT, assume $U(t) = U(kT) \triangleq U(k)$ for $t \in [kT, (k+1)T)$, and denote by $X(k)$ and by $\hat{X}(k|k)$ the current state and its filtered estimate, respectively, at time instant $t = kT$. For obtaining an EKF with an effective state prediction equation in a simple form, and for estimating both the gain matrix $K(k+1)$ and the covariance matrix $P(k|k)$ of the filter, (17) has been linearized about $\hat{X}(k|k)$ and $U(k-1)$, and subsequently discretized with period T. This gives rise to the following equation:

$$X(k+1) = A_d(k)X(k) + L(k)U(k) + D(k) + W(k). \quad (18)$$

The expressions of vectors and matrices on the right-hand side of Eq. (18) are the following:

$$A_d(k) = \exp(A(k)T) = \begin{bmatrix} 1 & 0 & -T\nu(k-1)\sin\hat{\theta}(k|k) \\ 0 & 1 & T\nu(k-1)\cos\hat{\theta}(k|k) \\ 0 & 0 & 1 \end{bmatrix}, \quad (19)$$

where

$$A(k) \triangleq \left[\frac{\partial F(X(t),U(t))}{\partial X(t)}\right]_{\substack{X(t)=\hat{X}(k|k) \\ U(t)=U(k-1)}} = \begin{bmatrix} 0 & 0 & -\nu(k-1)\sin\hat{\theta}(k|k) \\ 0 & 0 & \nu(k-1)\cos\hat{\theta}(k|k) \\ 0 & 0 & 0 \end{bmatrix}, \quad (20)$$

$$L(k) = \begin{bmatrix} T\cos\hat{\theta}(k|k) & -\frac{T^2}{2}\nu(k-1)\sin\hat{\theta}(k|k) \\ T\sin\hat{\theta}(k|k) & \frac{T^2}{2}\nu(k-1)\cos\hat{\theta}(k|k) \\ 0 & T \end{bmatrix}, \quad (21)$$

$$D(k) = \begin{bmatrix} T\nu(k-1)\hat{\theta}(k|k)\sin\hat{\theta}(k|k) \\ -T\nu(k-1)\hat{\theta}(k|k)\cos\hat{\theta}(k|k) \\ 0 \end{bmatrix}. \quad (22)$$

The covariance matrix of the discretized noise $W(k) \in \mathbb{R}^3$, $k = 0, 1, 2\cdots$ is expressed by $Q_d(k) \triangleq E(W(k)W(k)^T)$, where

$$Q_{d_{1,1}}(k) = T\sigma_1^2 + \frac{T^3}{3}\sigma_2^2\nu^2(k-1)\sin^2\hat{\theta}(k|k), \quad (23)$$

$$Q_{d_{2,2}}(k) = T\sigma_1^2 + \frac{T^3}{3}\sigma_2^2\nu^2(k-1)\cos^2\hat{\theta}(k|k), \quad (24)$$

$$Q_{d_{1,2}}(k) = Q_{d_{2,1}}(k) = -\frac{T^3}{3}\sigma_2^2\nu^2(k-1) \times \sin\hat{\theta}(k|k)\cos\hat{\theta}(k|k), \quad (25)$$

$$Q_{d_{1,3}}(k) = Q_{d_{3,1}}(k) = -\frac{T^2}{2}\sigma_2^2\nu(k-1)\sin\hat{\theta}(k|k), \quad (26)$$

$$Q_{d_{2,3}}(k) = Q_{d_{3,2}}(k) = \frac{T^2}{2}\sigma_2^2\nu(k-1)\cos\hat{\theta}(k|k), \quad (27)$$

$$Q_{d_{3,3}}(k) = T\sigma_2^2. \quad (28)$$

4.2. Robot and environment state

A poor *a priori* knowledge of the environment features would negatively affect the reliability and autonomy of a localization algorithm not taking into account the environment parameters in the formulation of the EKF. To deal with this problem, a stochastic state–space model has been defined with the state vector composed of the robot state and of the environment geometric features, which are assumed to be stationary: $\tilde{X}(k) \triangleq [X(k)^T \; X'(k)^T]^T$. $X'(k) \triangleq [P_{\mathrm{r}}^1 \; P_{\mathrm{n}}^1 \; P_{\mathrm{r}}^2 \; P_{\mathrm{n}}^2 \; \dots \; P_{\mathrm{r}}^{n_{\mathrm{p}}} \; P_{\mathrm{n}}^{n_{\mathrm{p}}} T]^T$ is the environment state.

The dynamic state–space equation for $\tilde{X}(k)$ is directly derived from (18)

$$\tilde{X}(k+1) = \tilde{A}(k)\tilde{X}(k) + \tilde{L}(k)U(k) + \tilde{D}(k) + \tilde{W}(k), \quad (29)$$

where

$$\tilde{A}(k) = \begin{bmatrix} A_d(k) & 0_{3,2n_{\mathrm{p}}} \\ 0_{2n_{\mathrm{p}},3} & I_{2n_{\mathrm{p}},2n_{\mathrm{p}}} \end{bmatrix}, \quad (30)$$

$$\tilde{L}(k) = \begin{bmatrix} L(k) \\ 0_{2n_{\mathrm{p}},2} \end{bmatrix}, \quad (31)$$

$$\tilde{D}(k) = \begin{bmatrix} D(k) \\ 0_{2n_{\mathrm{p}},1} \end{bmatrix}, \quad (32)$$

$$\tilde{W}(k) = \begin{bmatrix} W(k) \\ 0_{2n_{\mathrm{p}},1} \end{bmatrix}, \quad (33)$$

$0_{i,j}$ being the $i \times j$ null matrix.

4.3. Measurement equations

Denote by $Z(k)$ the measurement vector at time instant kT. Vector $Z(k)$ is composed of the following three subvectors:

- $Z_1(k) = [z_{\mathrm{o},1}(k) \; z_{\mathrm{o},2}(k) \; z_{\mathrm{o},3}(k)]^T$ (odometric measures),
- $Z_2(k) = [z_{\mathrm{v},1}(k) \; z_{\mathrm{v},2}(k) \; \dots \; z_{\mathrm{v},2n_{\mathrm{vm}}}(k)]^T$ (video measures), and
- $Z_3(k) = [z_{\mathrm{s},1}(k) \; z_{\mathrm{s},2}(k) \; \dots \; z_{\mathrm{s},n_{\mathrm{sm}}}(k)]^T$ (sonar measures).

The dimensions $2n_{\mathrm{vm}}$ of $Z_2(k)$ and n_{sm} of $Z_3(k)$ ($n_{\mathrm{sm}} \leq n_{\mathrm{s}}$) are not constant, but depend on the number of measures that are actually used at each time, after the selection procedure described in Section 5.

The elements of $Z_1(k)$ are modeled by

$$z_{\mathrm{o},1}(k) = x(k) + v_{\mathrm{o},1}(k), \quad (34)$$

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

$$z_{o,2}(k) = y(k) + v_{o,2}(k), \tag{35}$$

$$z_{o,3}(k) = \theta(k) + v_{o,3}(k), \tag{36}$$

where $z_{o,1}(k) = \hat{x}(k-1|k-1) + \Delta x(k)$, $z_{o,2}(k) = \hat{y}(k-1|k-1) + \Delta y(k)$, and $z_{o,3}(k) = \hat{\theta}(k-1|k-1) + \Delta\theta(k)$, with the odometric increments expressed by Eqs. (6)–(8).

The elements of $Z_2(k)$ are modeled by

$$z_{v,2i-1}(k) = P_r^i(k) + v_{v,2i-1}(k), \tag{37}$$

$$z_{v,2i}(k) = P_n^i(k) + v_{v,2i}(k), \tag{38}$$

where $i = 1, 2, \ldots, n_{vm}$, and $z_{v,2i-1}(k) = \rho_i(k)$ and $z_{v,2i}(k) = \phi_i(k)$ are the environment feature measures extracted by the vision system whose reliability has been verified on the occupancy grid map, as discussed in Section 3.

The elements of $Z_3(k)$ are modeled by $z_{s,i}(k) = d_{h,i}^j(k) + v_{s,i}(k)$, $i = 1, 2, \ldots, n_{sm}$, $1 \le j \le n_p$, $1 \le h \le n_s$, with $d_{h,i}^j(k)$ given by (14); $z_{s,i}(k)$ is the ith sonar measure, representing the distance of the hth sonar sensor from the plane P^j. The environment map provides the information needed to detect which is the plane P^j in front of the hth sonar.

The observation noise

$$V(k) = [v_{o,1}(k)\ v_{o,2}(k)\ v_{o,3}(k)\ v_{v,1}(k)\ \ldots\ v_{v,2n_{vm}}(k) \times v_{s,1}(k)\ \ldots\ v_{s,n_{sm}}(k)]^T$$

is a white Gaussian sequence with mean zero and covariance matrix R where R has the structure $R \triangleq$ block diag (R_1, R_2, R_3), with $R_1 \triangleq \text{diag}(\sigma_{o,1}^2, \sigma_{o,2}^2, \sigma_{o,3}^2)$, $R_2 \triangleq \text{diag}(\sigma_{v,1}^2, \sigma_{v,2}^2, \ldots, \sigma_{v,2n_{vm}}^2)$, and $R_3 \triangleq \text{diag}(\sigma_{s,1}^2, \sigma_{s,2}^2, \ldots, \sigma_{s,n_{sm}}^2)$; R_2 and R_3 represent the covariance matrices of the independent errors affecting the video and sonar measures, respectively. The diagonal form of R follows from the independence of the encoder, video, and sonar measures. In particular, the independence between video and sonar measures holds because the latter are not used for the validation of video straight lines acquired at the same sampling instant.

Vectors $Z(k)$ and $\tilde{X}(k)$ are related by a nonlinear measurement equation. Linearization of the measurement equation about the optimal prediction $\hat{X}(k|k-1)$ results in

$$\bar{Z}(k) = \tilde{C}(k)\tilde{X}(k) + V(k), \tag{39}$$

where $\tilde{C}(k) = [\tilde{C}_1(k)^T\ \tilde{C}_2(k)^T\ \tilde{C}_3(k)^T]^T$. We have $\tilde{C}_1(k) = [I_3\ 0_{3,2n_p}]$; moreover, $\tilde{C}_2(k)$ is a matrix whose ith row is zero, except for the entry corresponding to the parameter of the state vector which is updated by the ith video measure; this entry is equal to 1. Finally, $\tilde{C}_3(k) = [\tilde{c}_1(k)^T\ \tilde{c}_2(k)^T\ \ldots\ \tilde{c}_{n_{sm}}(k)^T]^T$, where

$$\tilde{c}_i(k) = P_v^j \cdot \begin{bmatrix} -\cos P_n^j \\ -\sin P_n^j \\ \tilde{c}'_{i,3}(k) \\ 0_{2(j-1),1} \\ 1 \\ \tilde{c}'_{i,2j+3}(k) \\ 0_{2(n_p-j),1} \end{bmatrix}^T, \tag{40}$$

$$\tilde{c}'_{i,3}(k) = x'_h \sin\left(\hat{\theta}(k|k-1) - P_n^j\right) + y'_h \cos\left(\hat{\theta}(k|k-1) - P_n^j\right), \tag{41}$$

$$\tilde{c}'_{i,2j+3}(k) = x_h(k) \sin P_n^j - y_h(k) \cos P_n^j, \tag{42}$$

and $i = 1, 2, \ldots, n_{sm}$, $1 \le j \le n_p$, $1 \le h \le n_s$.

Equations (29) and (39) represent the linearized and discretized state–space form to which the classical EKF algorithm has been applied.

5. Sensor Reading Selection

It is known that more sensor data do not always mean better estimates. The reason may be an inadequate interpretation of erroneous data, due to sensor failures or to the inconsistency of the underlying filter equations. The second issue is discussed in ref. [39]. In the present case, sonar measures may not only be affected by White Gaussian noise, but also by crosstalk interferences and multiple echoes.[40] Techniques for the treatment of these systematic errors based on the construction of a navigation map and/or on a time pulse modulation have been proposed in refs. [40, 41], respectively, and have been implemented here. Also prefiltering procedure for improving the quality of video data have been used. The prefiltering procedure makes use of a threshold applied to the probability of the lines in the Hough space. Lines that have not reached a defined probability level will be discarded. The threshold level is selected experimentally based also on the texture of the floor.

To reduce the probability of an inadequate interpretation of erroneous sonar sensor data and video data, a method is proposed to deal with the undesired interferences produced by the presence of unknown obstacles in the environment or by uncertainty on the sensor readings. Notice that for the problem handled here both the events are equally distributed. A simple and efficient way to perform this preliminary measure selection is to compare the actual video and sonar readings with their expected values. Measures are discarded if the difference exceeds a time-varying threshold. The value of the threshold has to be chosen on the basis of two opposite considerations: too large values would accept undesired interferences as reliable sensor readings, while too low values would produce an undesired loss of useful information. For this reason, it is preferable not to choose a unique fixed threshold, but to adapt the preliminary choice on-line, on the basis of the reliability of the current predicted sensor measure.

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

This is done here in the following way: at each step, for each couple of video measures $z_{\mathrm{v},2i-1}$ and $z_{\mathrm{v},2i}$, representing the distance and orientation of a detected straight line i in the Hough transform notation, the residuals $\delta_{2i-1} = z_{\mathrm{v},2i-1} - \hat{P}^i_{\mathrm{r}}$ and $\delta_{2i} = z_{\mathrm{v},2i} - \hat{P}^i_{\mathrm{n}}$ represent the differences between the distance $z_{\mathrm{v},2i-1}$ and orientation $z_{\mathrm{v},2i}$ of the detected straight line and their expected values $\hat{P}^i_{\mathrm{r}}$ and $\hat{P}^i_{\mathrm{n}}$, respectively, $i = 1, 2, \ldots, n_{\mathrm{vm}}$. In the same way, at each step, for each measure $z_{\mathrm{s},\ell}(k)$ of a sonar sensor, the residual $\gamma_\ell(k) = z_{\mathrm{s},\ell}(k) - \hat{d}^j_{h,\ell}$ represents the difference between the actual sonar measure $z_{\mathrm{s},\ell}(k)$ and its expected value $\hat{d}^j_{h,\ell}$, $\ell = 1, 2, \ldots, n_{\mathrm{sm}}$, $1 \le j \le n_{\mathrm{p}}$, $1 \le h \le n_{\mathrm{s}}$, which is computed by (14) on the basis of the current estimate of $\tilde{X}(k)$ and of the *a priori* knowledge of the environment.

As $\delta_{2i-1}(k) \sim N(0, \bar{\sigma}^2_{\mathrm{v},2i-1})$ and $\delta_{2i}(k) \sim N(0, \bar{\sigma}^2_{\mathrm{v},2i})$, the current values of video measures $z_{\mathrm{v},2i-1}$ and $z_{\mathrm{v},2i}$ are accepted if both $|\delta_{2i-1}(k)| \le 2\bar{\sigma}_{\mathrm{v},2i-1}$ and $|\delta_{2i}(k)| \le 2\bar{\sigma}_{\mathrm{v},2i}$.[31] Namely, variable thresholds of video measures are considered that are equal to twice the standard deviations of the innovation processes. In the same way, as $\gamma_i(k) \sim N(0, \bar{\sigma}^2_i(k))$, the current value of sonar measures $z_{\mathrm{s},i}(k)$ is accepted if $|\gamma_i(k)| \le 2\bar{\sigma}_{\mathrm{s},i}(k)$. Namely, the variable threshold of sonar measures is chosen as twice the standard deviation of the innovation process.

Simple calculations show that these techniques are equivalent to the sensor data prefiltering based on the Mahalanobis distance,[22, 42] where the cutoff value is chosen as $(\chi^2_{1,1-\alpha})^{1/2}$ with $\alpha \in (0.95, 0.96)$.

6. Experimental Results

The experimental tests have been performed on the LabMate mobile base in indoor environments with different geometries. This vehicle is equipped with two driving wheels. The odometric system is composed of two optical encoders connected to the axes of the driving wheels. A sampling period T of 0.4 s has been used; this value is small enough if we consider that the velocity of the robot during the experiments is relatively low (less than 1m/s).

The odometric data are the incremental measures that at each sampling interval are provided by the encoders attached to the right and left robot wheels. These measures are directly acquired by the low-level controller of the mobile base. The sonar measures have been carried out by the standard proximity system of the vehicle, composed of a half ring of 15 Polaroid sonar sensors. The video measures have been collected by a low-cost CCD webcam Philips PCVC 680K installed on top of the vehicle and pointed down in the left side.

The LabMate mobile robot with data acquisition system for video camera, incremental encoders, and sonar sensors is shown in Fig. 2. A preliminary reduction of sonar crosstalk interferences has been carried out by a proper distribution of the orientations of the sonar sensors, and by the implementation of the prefiltering techniques proposed in refs. [37, 40]. The sonars have been mounted with a difference on the orientation of 15°. A significant reduction of the wrong readings produced by the presence of unknown obstacles or by uncertainty on the sensor readings has been also carried out by the selection of sonar and video measures by the procedure described in Section 5.

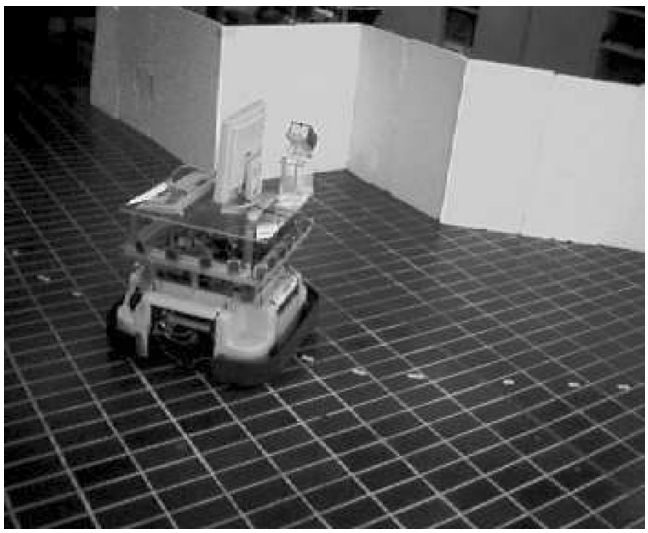


Fig. 2. Indoor environment with the LabMate mobile vehicle equipped with data acquisition system for video camera, incremental encoders, and sonar sensors.

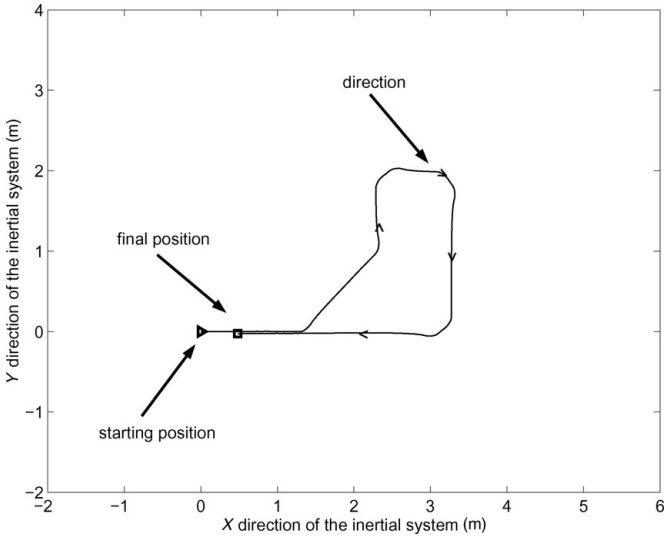


Fig. 3. First experiment: Estimated trajectory.

The EKF has been implemented on the navigation module of the mobile robot. The navigation system is connected directly to the low-level controller.

Figures 3, 4, and 5 illustrate a sample of the obtained results. The state vector $\tilde{X}(\cdot)$ of the EKF algorithm has been initialized by the actual robot state and by an inaccurate estimate of the environment geometric features, whose uncertainties are initially all independent from each other. This is reasonable because the robot starts from a docking position, and therefore the initial uncertainty on its pose is low; the errors affect the feature initial estimates only. The diagonal elements of the covariance matrix R of the observation noise $V(\cdot)$ are determined by an analysis of the sensor characteristics.

Figure 3 shows the trajectory of the vehicle estimated by the proposed algorithm. It can be seen that the robot

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

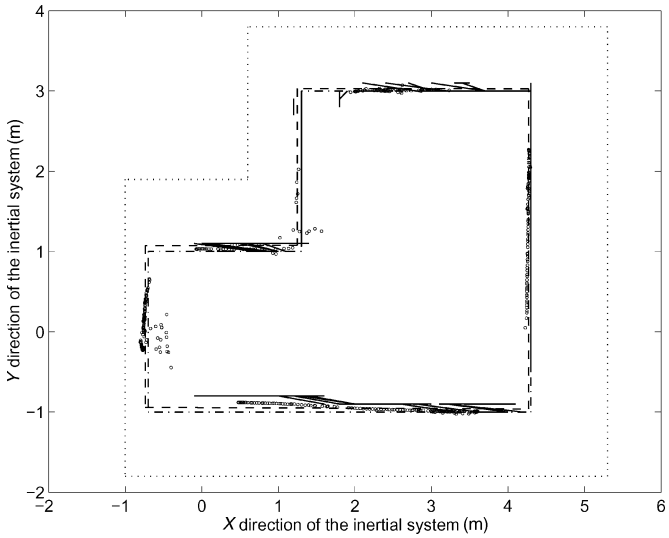


Fig. 4. First experiment: Environment map estimation using the proposed sonar and video data fusion procedure. The dash-dot lines are the actual locations of the environment map landmarks, while the dotted and the dashed lines represent the initial and the final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures, and the solid lines are the video measures.

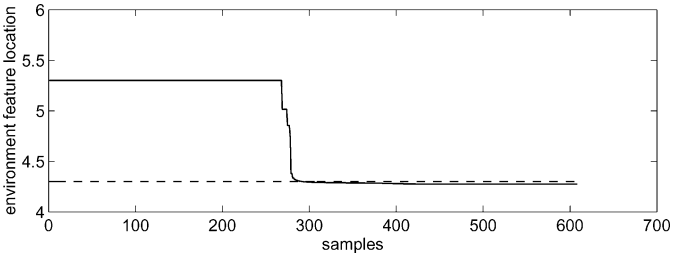


Fig. 5. First experiment: A typical estimate of an environment feature location. The continuous and the dashed lines represent the estimated and the actual environment feature location, respectively.

changes its direction more than once, which would negatively affect the performance of a localization procedure using odometric data only, while the proposed algorithm performs satisfactorily. The detected estimation errors corresponding to the final vehicle configuration (the distance between the actual and the estimated configuration) have been reported in Table I. The actual configuration of the robot (ground truth) has been obtained through direct measure. Figure 4 shows the considered indoor environment represented by a suitable set of planes orthogonal to the XY plane of the inertial system. The dash-dot lines are the actual locations of the environment map landmarks, the dotted lines represent

Table I. First experiment: Estimation errors (E) corresponding to the final vehicle configuration (distance between the actual (AC) and the estimated (EC) configuration).

	x (m)	y (m)	θ (rad)
AC	0.5000	0.0000	3.0543
EC	0.4823	−0.0217	3.0325
E	0.0177	0.0217	0.0218

Table II. First experiment: Differences in absolute value between the actual locations of the environment landmarks and their initial (IE) and final (FE) estimates. $L_i, i = 1, 2, \ldots, 6$, are the considered landmarks.

	L_1 (m)	L_2 (m)	L_3 (m)	L_4 (m)	L_5 (m)	L_6 (m)
IE	1.0000	0.8000	0.7000	0.9000	0.3000	0.8000
FE	0.0590	0.0532	0.0506	0.0721	0.0456	0.0788

the initial inaccurate estimates of the environment features, and the dashed lines represent their final estimates. The dots are the actually used sonar measures, and the solid lines are the video measures. Visually the overall error on the map is reduced at the end of the vehicle task, as confirmed by the numerical values of Table II showing the differences in absolute value between the actual locations of the considered environment landmarks and their initial and final estimates. A typical estimate (continuous line) of an environment geometric feature location is shown in Fig. 5 where, as the vehicle progresses through the environment, the error on the landmark location estimate decreases monotonically to a lower bound. The dashed line represents the actual environment geometric feature location. The fusion of ultrasonic and video data produces a significant improvement in the reliability of the environment feature estimation compared to the results obtained using sonar sensor readings only.[3] This is evident comparing Fig. 4 with Fig. 6, where the dash-dot lines are the actual locations of the environment map landmarks, while the dotted and the dashed lines represent the initial and the final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures. The proposed method for sensor data fusion is able to deal with uncertainties on environment feature locations on the order of 1 m and more, unlike a procedure using sonar readings only; this is

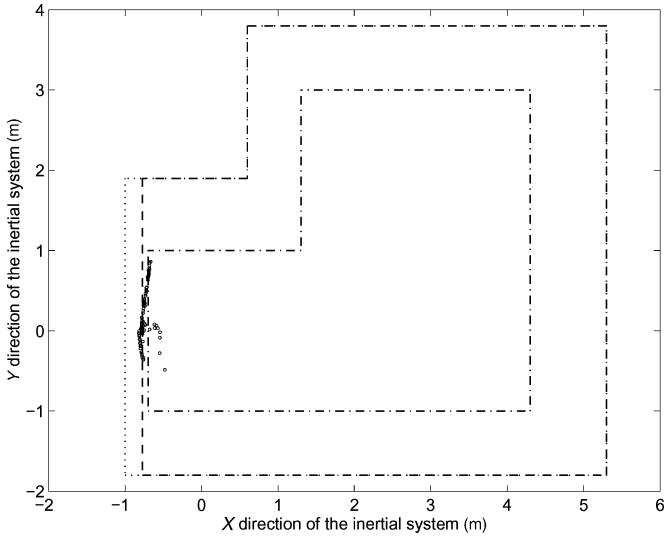


Fig. 6. First experiment: Environment map estimation using sonar sensor readings only. The dash-dot lines are the actual locations of the environment map landmarks, while the dotted and the dashed lines represent the initial and the final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures.

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

Table III. Second experiment: Differences in absolute value between the actual locations of the environment landmarks and their initial (IE) and final (FE) estimates. L_i, $i = 1, 2, \ldots, 7$, are the considered landmarks.

	L_1 (m)	L_2 (m)	L_3 (m)	L_4 (m)	L_5 (m)	L_6 (m)	L_7 (m)
IE	1.1000	0.2800	0.7800	0.4600	0.3900	0.4900	0.7300
FE	0.0646	0.0050	0.1341	0.1500	0.0681	0.0113	0.0360

Table IV. Second experiment: Estimation errors (E) corresponding to the final vehicle configuration (distance between the actual (AC) and the estimated (EC) configuration).

	x (m)	y (m)	θ (rad)
AC	0.0500	−0.8500	3.2289
EC	0.0309	−0.8215	3.1947
E	0.0191	0.0285	0.0342

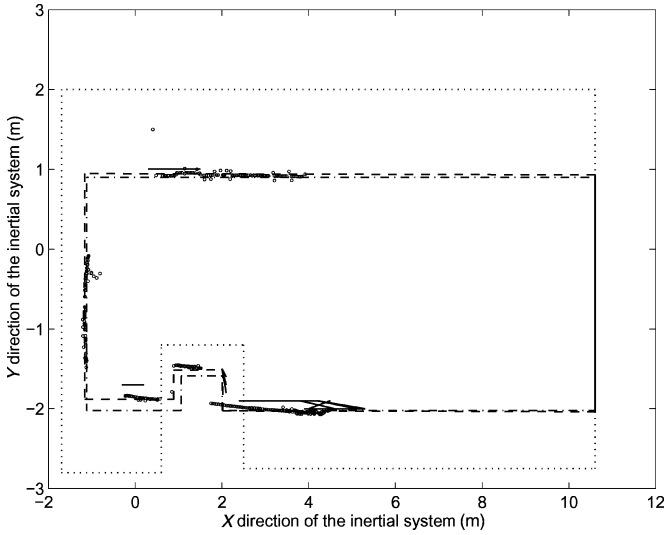


Fig. 7. Second experiment: Environment map estimation using the proposed sonar and video data fusion procedure. The dash-dot lines are the actual locations of the environment map landmarks, while the dotted and the dashed lines represent the initial and the final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures, and the solid lines are the video measures.

demonstrated by the small number of sonar readings actually used by the latter, and by the resulting inadequate correction of the initial uncertainty.

The detected estimation errors corresponding to the final vehicle configuration have been reported in Table IV.

Figure 7 shows the results obtained in an environment characterized by a different geometry. Also in this situation the overall error on the map is reduced at the end of the vehicle task, as confirmed by the numerical values of Table III.

7. Conclusions

In this paper, an EKF is proposed as a tool for the consistent fusion of information acquired by a set of internal and external sensors, characterized by different degree of accuracy, in an absolute approach to mobile robot localization with the double aim to reduce the uncertainties on the vehicle localization and on the estimate of environment features. The main characteristic of the proposed approach is the improved robustness of the localization procedure with respect to uncertainties on the environment map. The algorithm has been tested by a set of indoor experiments. Different kinds of sensors are considered: sonars, video camera, and encoders. Video and sonar data are fused together to increase the reliability of the environment landmark measures acquired by the vision system.

The algorithm operates in a state–space form where sensor and model uncertainties are intrinsically taken into account through the definition of a stochastic model whose state vector contains both the position and orientation of the robot, and the parameters of environment features. This introduces a very high degree of accuracy in the estimated vehicle pose, and makes the estimator more robust with respect to possible uncertain physical parameters and/or not exactly known initial conditions. The results in terms of errors in the simultaneous on-line estimation of both the vehicle position and the environment features are adequate for the safe navigation of the robot in the considered indoor environment. The work is in progress to validate the results in larger and more complex environments. Moreover, the problem of the high computational cost of the EKF for large environments can be solved by a hybrid approach where a global topological map connects local metric maps with a reduced number of features.[29] In this way the complexity can be reduced by a careful selection of the features that have to be treated by the EKF, based on the idea that processing all features at the same time can be uselessly demanding. The proposed procedure represents a first step toward simultaneous localization and map building; however, to become an actual implementation of SLAM, where no *a priori* information is available, the described procedure will have to perform the initialization of new features as they are detected for the first time in the environment.

References

1. A. Bonci, G. Ippoliti, L. Jetto, T. Leo and S. Longhi, "Methods and Algorithms for Sensor Data Fusion Aimed at Improving the Autonomy of a Mobile Robot," **In**: *Springer Tracts in Advanced Robotics, Vol. 10: Advances in Control of Articulated and Mobile Robots* (B. Siciliano, A. De Luca, C. Melchiorri and G. Casalino, eds.) (Springer, Berlin, Heidelberg, Germany, 2004) pp. 191–222.
2. J. Gu, M. Meng, A. Cook and P. X. Liu, "Sensor Fusion in Mobile Robot: Some Perspectives," *Proceedings of the 4th World Congress on Intelligent Control and Automation*, Vol. 2 (Shanghai, China, 2002) pp. 1194–1199.
3. G. Ippoliti, L. Jetto, A. La Manna and S. Longhi, "Improving the Robustness Properties of Robot Localization Procedures with Respect to Environment Features Uncertainties," *Proceedings of the IEEE International Conference on Robotics and Automation* (Barcelona, Spain, 2005) pp. 1451–1458.
4. J.-S. Gutmann, W. Burgard, D. Fox and K. Konolige, "An Experimental Comparison of Localization Methods," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2 (Victoria, Canada, 1998) pp. 736–743.

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press

5. J.-S. Gutmann and D. Fox, "An Experimental Comparison of Localization Methods Continued," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1 (Lausanne, Switzerland, 2002) pp. 454–459.
6. K. O. Arras, J. A. Castellanos, M. Schilt and R. Siegwart, "Feature-based multi-hypothesis localization and tracking using geometric constraints," *Rob. Autonom. Syst.* **44**(1), 41–53 (Jul. 2003).
7. J. Borenstein, H. R. Everett, L. Feng and D. Wehe, "Mobile robot positioning—Sensors and techniques," *J. Rob. Syst.* **14**(4), 231–249 (Apr. 1997).
8. P. Goel, S. I. Roumeliotis and G. S. Sukhatme, "Robust Localization Using Relative and Absolute Position Estimates," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2 (Kyonagju Korea, 1999) pp. 1134–1140.
9. J. D. Tardós, J. Neira, P. M. Newman and J. J. Leonard, "Robust mapping and localization in indoor environments using sonar data," *Int. J. Rob. Res.* **21**(4), 311–330 (Apr. 2002).
10. J. Borenstein and Y. Koren, "Histogramic in-motion mapping for mobile robot obstacle avoidance," *Int. J. Rob. Res.* **7**(4), 535–539 (Aug. 1991).
11. S. Thrun, "Robotic Mapping: A Survey," **In**: *Exploring Artificial Intelligence in the New Millenium* (G. Lakemeyer and B. Nebel, eds.) (Morgan Kaufmann, San Francisco, CA, 2002).
12. J. Folkesson and H. Christensen, "Graphical SLAM—A Self-Correcting Map," *Proceedings of the IEEE International Conference on Robotics and Automation*, Vol. 1 (New Orleans, LA, USA, 2004) pp. 383–390.
13. J. J. Leonard and H. J. S. Feder, "Decoupled stochastic mapping," *IEEE J. Oceanic Eng.* **26**(4), 561–571 (Oct. 2001).
14. S. Thrun, D. Fox and W. Burgard, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Mach. Learn.* **31**(1–3), 29–53 (1998). Also appeared in *Autonom. Rob.* **5**, 253–271 (joint issue).
15. F. Dellaert and M. Kaess, "Square root SAM: Simultaneous localization and mapping via square root information smoothing," *Int. J. Rob. Res.* **25**(12), 1181–1203 (Dec. 2006).
16. R. Eustice, H. Singh, J. Leonard, M. Walter and R. Ballard, "Visually Navigating the RMS Titanic with SLAM Information Filters," *Proceedings of Robotics: Science and Systems* (Cambridge, MA, USA, 2005) pp. 57–64.
17. H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Rob. Automat. Mag.* **13**(2), 99–110 (Jun. 2006).
18. T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Rob. Automat. Mag.* **13**(3), 108–117 (Sep. 2006).
19. C. Fulgenzi, G. Ippoliti and S. Longhi, "Experimental Validation of FastSLAM Algorithm Characterized by a Linear Features Based Map," *Proceedings of the 8th International IFAC Symposium on Robot Control* (Bologna, Italy, 2006).
20. M. Montemerlo, S. Thrun, D. Koller and B. Wegbreit, "FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem," *Proceedings of the AAAI National Conference on Artificial Intelligence* (Edmonton, Canada, 2002).
21. M. Montemerlo, S. Thrun, D. Koller and B. Wegbreit, "FastSLAM 2.0: An Improved Particle Filtering Algorithm for Simultaneous Localization and Mapping that Provably Converges," *Proceedings of the International Joint Conference on Artificial Intelligence* (Acapulco, Mexico, 2003) pp. 1151–1156.
22. M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. Rob. Automat.* **17**(3), 229–241 (Jun. 2001).
23. J. J. Leonard and H. F. Durrant-Whyte, "Simultaneous Map Building and Localization for an Autonomous Mobile Robot," *Proceedings of the IEEE/RSJ International Workshop on Intelligent Robots and Systems*, Vol. 3, (Osaka, Japan, 1991) pp. 1442–1447.
24. G. C. Anousaki and K. J. Kyriakopoulos. "Simultaneous localization and map building of skid-steered robots," *IEEE Rob. Automat. Mag.* **14**(1), 79–89 (Mar. 2007).
25. L. M. Paz, P. Jensfelt, J. D. Tardós and J. Neira, "EKF SLAM Updates in O(n) With Divide and Conquer SLAM," *Proceedings of the IEEE International Conference on Robotics and Automation* (Rome, Italy, 2007) pp. 1657–1663.
26. J. E. Guivant, F. R. Masson and E. M. Nebot, "Simultaneous localization and map building using natural features and absolute information," *Rob. Autonom. Syst.* **40**(2), 79–90 (Aug. 2002).
27. L. Armesto and J. Tornero, "SLAM Based on Kalman Filter for Multi-Rate Fusion of Laser and Encoder Measurements," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2 (Senda Japan, 2004) pp. 1860–1865.
28. M. C. Bosse, P. M. Newman, J. J. Leonard and S. Teller, "SLAM in large-scale cyclic environments using the Atlas framework," *Int. J. Rob. Res.* **23**(12), 1113–1139 (Dec. 2004).
29. N. Tomatis, I. Nourbakhsh and R. Siegwart, "Hybrid simultaneous localization and map building: A natural integration of topological and metric," *Rob. Autonom. Syst.* **44**(1), 3–14 (Jul. 2003).
30. C. M. Wang, "Location Estimation and Uncertainty Analysis for Mobile Robots," *Proceedings of the IEEE International Conference on Robotics and Automation* (Philadelphia, PA, USA, 1988) pp. 1230–1235.
31. L. Jetto, S. Longhi and G. Venturini, "Development and experimental validation of an adaptive extended Kalman filter for the localization of mobile robots," *IEEE Trans. Rob. Automat.* **15**(2), 219–229 (Apr. 1999).
32. N. Ayache and O. D. Faugeras, "Maintaining representations of the environment of a mobile robot," *IEEE Trans. Rob. Automat.* **5**(6), 804–819 (Dec. 1989).
33. A. C. Kak, "Depth Perception for Robots," **In**: *Handbook of Industrial Robotics* (S. Y. Nof, ed.) (John Wiley & Sons, New York, 1985) pp. 272–319.
34. P. V. C. Hough, "Method and Means for Recognizing Complex Patterns," U.S. Patent 3,069,654 (Dec. 1962).
35. I. E. Sobel, Camera Models and Machine Perception *Ph.D. Thesis* (Electrical Engineering Department, Stanford University, CA, 1970).
36. A. Bonci, G. Di Francesco and S. Longhi, "A Bayesian Approach to the Hough Transform for Video and Ultrasonic Data Fusion in Mobile Robot Navigation," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 3 (Hammamet, Tunisia, 2002).
37. A. Angeloni, T. Leo, S. Longhi and R. Zulli, "Real Time Collision Avoidance for Mobile Robots," *Proceedings of the 6th International Symposium on Measurement and Control in Robotics* (Brussels, Belgium, 1996) pp. 239–244.
38. E. A. Bender, *Mathematical Methods in Artificial Intelligence* (Wiley-IEEE Computer Society Press, Los Alamitos, CA, 1996).
39. W. D. Blair and Y. Bar-Shalom, "Tracking maneuvering targets with multiple sensors: Does more data always mean better estimates?," *IEEE Trans. Aerospace Electron. Syst.* **32**(1), 450–456 (Jan. 1996).
40. J. Borenstein and Y. Koren, "Error eliminating rapid ultrasonic firing for mobile robot obstacle avoidance," *IEEE Trans. Rob. Automat.* **11**(1), 132–138 (Feb. 1995).
41. M. Beckerman and E. M. Oblow, "Treatment of systematic errors in the processing of wide-angle sonar sensor data for robotic navigation," *IEEE Trans. Rob. Automat.* **6**(2), 137–145 (Apr. 1990).
42. Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association* (Academic Press, San Diego, CA, 1988).

https://doi.org/10.1017/S026357470800502X Published online by Cambridge University Press