Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Sonar and Video Data Fusion for Robot Localization and Environment Feature Estimation

A. Bonci, G. Ippoliti, A. La Manna, S. Longhi and L. Sartini

***Abstract*— In this paper the localization and environment feature estimation problems are formulated in a stochastic setting, and an Extended Kalman Filtering (EKF) 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 sensory data fusion aimed at increasing the environment knowledge is considered.**

I. 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]–[4]. If the environment is only partially known, the localization algorithm needs a preliminary definition of a suitable environment map; some methods which enable a mobile robot to build by itself a map of its environment are presented in [5], [6].

Recently, the Simultaneous Localization and Map Building problem (SLAM problem) has been deeply investigated for increasing the autonomy of navigation of mobile robots. The idea is to develop a localization algorithm that can build a map of the environment while simultaneously using the same map to localize the mobile robot. This approach promises to allow these vehicles to really operate autonomously for long periods of time in unknown environments. In the literature many contributions aimed at solving the SLAM problem by means of different techniques have been issued; it is worth mentioning the probabilistic method shown in [7], which uses particle filtering and is based on the factorization of the SLAM posterior into a localization problem and K independent landmark estimation problems (where K is the number of landmarks) conditioned on the robot pose estimate. The variety of the available contributions arises from the requirement to cope with different practical situations; for example, in [8] the connection of SLAM with the servoing problem is examined.

At present, the most widespread approach to SLAM 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 (frequently landmark 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. [9]–[11]). 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 (frequently landmark locations) with the associated uncertainty.

The contribution of this paper is to develop a localization technique with improved robustness against possible uncertainties on the environment. To this purpose a procedure is defined for the consistent and simultaneous on-line estimation of both the vehicle position and the environment features, 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 landmarks. 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; 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.

II. THE SENSOR EQUIPMENT

A. Odometric Measures

Consider 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. Consider the inertial coordinate system (O, X, Y) and the coordinate system (O', X', Y') fixed to the mobile robot. The localization of this mobile robot in a two-dimensional space requires the knowledge of

The authors are with the Dipartimento di Ingegneria Informatica, Gestionale e dell'Automazione, Università Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy {a.bonci, g.ippoliti, a.lamanna}@diiga.univpm.it sauro.longhi@univpm.it

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.

Assuming the right and left wheel velocities constant over a sufficiently small sampling period $\Delta t_k := t_{k+1} - t_k$, the increments Δx, Δy and $\Delta\theta$ in the position and orientation of the robot during Δt_k specify the odometric measures [4], [12].

B. Sonar Measures

The distance readings by sonar sensors are related to the indoor environment model and to the configuration of the mobile robot.

Consider a planar distribution of n_s sonar sensors. Denote by x'_h, y'_h, θ'_h the position of the h-th sonar, $1 \leq h \leq n_s$, referred to 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 h-th sonar, referred to the inertial coordinate system (O, X, Y), have the following form:

$$x_h(t_k) = x(t_k) + x'_h \cos\theta(t_k) - y'_h \sin\theta(t_k), \quad (1)$$

$$y_h(t_k) = y(t_k) + x'_h \sin\theta(t_k) + y'_h \cos\theta(t_k), \quad (2)$$

$$\theta_h(t_k) = \theta(t_k) + \theta'_h. \quad (3)$$

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, $j = 1, 2, \ldots, n_p$ (where n_p is the number of planes which describe the indoor environment), is represented by the triplet P^j_r, P^j_n, P^j_ν, where P^j_r is the normal distance of the plane from the origin O, P^j_n is the angle between the normal line to the plane and the X-direction, and P^j_ν is a binary variable, $P^j_\nu \in \{-1, 1\}$, which defines the face of the plane reflecting the sonar beam. In such a notation, the expectation $d^j_h(t_k)$ for the present distance of the h-th sonar from the plane P^j has the following expression [13]:

$$d^j_h(t_k) = P^j_\nu \cdot \left(P^j_r - x_h(t_k)\cos P^j_n - y_h(t_k)\sin P^j_n\right), \quad (4)$$

if $P^j_n \in [\theta_h(t_k) - \delta/2, \theta_h(t_k) + \delta/2]$, where δ is the beamwidth of the sonar sensor. To simplify the position estimation algorithm without an appreciable reduction of accuracy, the sonar echoes traveling along the cone edges have been omitted. In fact, the measures along the cone edges require an *a priori* model of the environment including the different roughness of the walls and are less accurate than the distance measures given by (4).

C. Video Camera Measures

A video camera and related 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 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 [14]. 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 [15]. 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 [2], [16].

For improving the detection of straight lines representing environment features, the Hough Transform (HT) is used [17]. In the HT the straight line equation is expressed by $\rho = u\cos\phi + v\sin\phi$, 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 whatever 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 [18]); 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. Recently, a new algorithm for the accumulator updating has been introduced which is able to produce an estimate of the line probability [19].

III. ULTRASONIC AND VIDEO DATA FUSION

A procedure of video data validation has been developed for improving the reliability of the vision system to detect environment line features. This procedure carries out a preliminary video reading selection on the basis of past sonar readings stored in a probabilistic map of the environment, which is an 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. In particular, a video straight line of parameters (ρ_j, ϕ_j) with a high probability also in the occupancy grid map is selected as a reliable video reading of an environment feature. In the following some technical aspects of the developed video data validation procedure are recalled.

The sonar readings are used as distance readings for the pose estimation of the mobile robot and for building an occupancy grid map which is a probabilistic map of the environment [2], [16].

During the robot exploration, the values of the cells in the occupancy grid are updated using the probabilistic signal level fusion of sonar readings proposed in [20]. In this way, the reliability and accuracy of sonar measures are improved. In fact, these measures are corrupted by specular reflections and modest angular resolution. The on-line updating of this map requires a low computational cost at each sampling instant. The solution proposed in [20], based on the probabilistic approach, is able to find a trade-off between accuracy and computational cost. The main features of the map building procedure that has been implemented are the following ones [20]: i) the sonar map consists of square cells (10 cm each side), each one containing a value between 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 dispersion caused by the mobile base); iv) the updating of the obstacle probability of each cell is done by Bayes's theorem. A numerical approximation is generally introduced to reduce the computational efforts. The map is updated without any *a priori* environment knowledge.

Straight line segments can be found in the occupancy grid as aligned cells with high probability of occupancy. By interpreting a grid and its probabilities as an image with different levels of intensity (grey level), it is possible to apply the HT to detect straight lines and to associate a probability to each detected line. Straight line segments represent the environment borders estimated by the sonar sensors.

The initial state of the occupancy grid is completely unknown because an *a priori* model of the environment is not provided. During the robot navigation the sonar readings are integrated into the occupancy grid and, at fixed time intervals, images of a part of the environment floor are acquired.

The environment borders on the floor have a geometry which is described by straight lines. The proposed updating process makes use of a probabilistic approach to the Hough Transform for extracting line features and the associated certainty values both from video data and from the occupancy grid. Matching between the lines extracted from video data and from the occupancy grid is performed for environment features belonging to the part of the floor visible from the CCD camera (here called "visible space" as in Subsection II-C). The proposed matching algorithm is based on the combination of the line probabilities encoded in both Hough accumulators as stated in the following.

All the straight lines detected from the video image are matched with the straight lines detected from the portion of the occupancy grid 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: $H_C(\rho, \phi)$ for the lines of the video image projected on the floor, and $H_G(\rho, \phi)$ for the lines detected from the occupancy grid. The matching algorithm is based on the combination of the line probabilities encoded in both Hough accumulators. A Bayesian estimator is considered [19]. For each j-th line feature, described by its vector of parameters $\Theta_j = (\rho_j, \phi_j)$, two probabilistic estimates are provided: $P_C = P(\Theta_{Cj}|\Theta_j)$ stored in H_C and $P_G = P(\Theta_{Gj}|\Theta_j)$ stored in H_G, where Θ_{Cj} is the estimate of the line parameters Θ_j obtained by using the video image and Θ_{Gj} is the estimate of the line parameters Θ_j obtained by using the occupancy grid. Using Bayes's theorem, the combined estimate $P(\Theta_j|\Theta_{Cj} \cup \Theta_{Gj})$ is given by

$$P(\Theta_j|\Theta_{Cj} \cup \Theta_{Gj}) = \frac{\frac{P_C P_G}{P(\Theta_j)}}{\frac{P_C P_G}{P(\Theta_j)} + \frac{(1-P_C)(1-P_G)}{1-P(\Theta_j)}}, \tag{5}$$

that is also known as the Independent Opinion Pool [21].

In this way, in the developed validation procedure, 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 Section IV.

IV. SIMULTANEOUS ESTIMATION OF ROBOT LOCATION AND ENVIRONMENT FEATURES

The algorithm operates in a stochastic framework. In [12], 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 Extended Kalman Filter (EKF) defined on the basis of a state equation derived from the kinematic model of the unicycle vehicle and from the environment geometric features, and of a measure equation containing the odometric measures, the measures obtained through the video system, and the distance readings provided by the set of sonar sensors.

A. Robot State

Denote by $X(t) := [x(t), y(t), \theta(t)]^T$ the true robot state and by $U(t) := [\nu(t), \omega(t)]^T$ the robot control input. Let $\Delta t_k = T$ be the constant sampling period and denote t_k by kT, assume $U(t) = U(kT) := U(k)$ 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 $t = kT$.

B. 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 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: $\tilde{X}(k) := [X(k)^T, X'(k)^T]^T$, where $X'(k) := [P_r^1, P_n^1, P_r^2, P_n^2, \ldots, P_r^{n_p}, P_n^{n_p}]^T$ is the environment state.

C. Measurement Equations

Denote by $Z(k)$ the measurement vector at time instant kT. Vector $Z(k)$ is composed of three subvectors, $Z_1(k) = [z_{o,1}(k), z_{o,2}(k), z_{o,3}(k)]^T$ (odometric measures), $Z_2(k) = [z_{f,1}(k), z_{f,2}(k), \ldots, z_{f,2n_{fm}}(k)]^T$ (video measures), and $Z_3(k) = [z_{s,1}(k), z_{s,2}(k), \ldots, z_{s,n_{sm}}(k)]^T$ (sonar measures). The dimensions $2n_{fm}$ of $Z_2(k)$ and n_{sm} of $Z_3(k)$ ($n_{sm} \leq n_s$) are not constant, but depend on the number of measures that are actually used at each time.

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

$$
\begin{aligned}
z_{o,1}(k) &= x(k) + v_{o,1}(k), && (6)\\
z_{o,2}(k) &= y(k) + v_{o,2}(k), && (7)\\
z_{o,3}(k) &= \theta(k) + v_{o,3}(k), && (8)
\end{aligned}
$$

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 specified in [4], [12].

The elements of $Z_2(k)$ are modeled by $z_{f,2i-1}(k) = P_r^i(k) + v_{f,2i-1}(k)$ and $z_{f,2i}(k) = P_n^i(k) + v_{f,2i}(k)$, where $i = 1, 2, \ldots, n_{fm}$, and $z_{f,2i-1}(k) = \rho_i(k)$ and $z_{f,2i}(k) = \phi_i(k)$ are the environment feature measures extracted by the vision system whose reliability has been verified on the occupancy grid map, as recalled in Section III.

The elements of $Z_3(k)$ are modeled by $z_{s,i}(k) = d_{h,i}^j(k) + v_{s,i}(k)$, $i = 1, 2, \ldots, n_{sm}$, $1 \leq j \leq n_p$, $1 \leq h \leq n_s$, with $d_{h,i}^j(k)$ given by (4); $z_{s,i}(k)$ is the i-th sonar measure, representing the distance of the h-th 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 h-th sonar.

The observation noise $V(k) = [v_{o,1}(k), v_{o,2}(k), v_{o,3}(k), v_{f,1}(k), \ldots, v_{f,2n_{fm}}(k), v_{s,1}(k), \ldots, v_{s,n_{sm}}(k)]^T$ is a white sequence $\sim N(0, R)$, where R has the structure $R :=$ block diag$[R_1, R_2, R_3]$, with $R_1 :=$ diag$[\sigma_{o,1}^2, \sigma_{o,2}^2, \sigma_{o,3}^2]$, $R_2 :=$ diag$[\sigma_{f,1}^2, \sigma_{f,2}^2, \ldots, \sigma_{f,2n_{fm}}^2]$, and $R_3 :=$ 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 measures and sonar measures, respectively. The diagonal form of R follows from the independence of the encoder, video and sonar measures.

Linearization of the non-linear measure equation about the optimal prediction $\hat{X}(k|k-1)$ results in $\bar{Z}(k) = \tilde{C}(k)\tilde{X}(k) + V(k)$, where $\tilde{C}(k) = [\tilde{C}_1(k)^T, \tilde{C}_2(k)^T, \tilde{C}_3(k)^T]^T$. One has that $\tilde{C}_1(k) = [I_3, 0_{3,2n_p}]$, $0_{i,j}$ being the $(i \times j)$ null matrix; moreover, $\tilde{C}_2(k)$ is a matrix whose i-th row is zero, except for the entry corresponding to the parameter of the state vector which is updated by the i-th 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

$$
\tilde{c}_i(k) = P_\nu^j \cdot \begin{bmatrix} -\cos P_n^j \\ -\sin P_n^j \\ \tilde{c}'_{i,3}(k) \\ 0_{2(j-1),1} \\ 1 \\ \tilde{c}'_{i,2j+3}(k) \\ 0_{2(n_p-j),1} \end{bmatrix}^T, \qquad (9)
$$

$$
\begin{aligned}
\tilde{c}'_{i,3}(k) &= x'_h \sin(\hat{\theta}(k|k-1) - P_n^j) \\
&\quad + y'_h \cos(\hat{\theta}(k|k-1) - P_n^j), && (10)\\
\tilde{c}'_{i,2j+3}(k) &= x_h(k) \sin P_n^j - y_h(k) \cos P_n^j && (11)
\end{aligned}
$$

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

On the basis of the considered assumption, the classical EKF algorithm has been applied. Also the necessary data association procedure has been implemented for the selection of sensor readings [2], [4].

V. 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 time T of 0.4 s has been used. 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.

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 [20]. The sonars have been mounted with a difference on the orientation of 15 degrees. 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 implemented data association procedure [2], [4].

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

Figures 1 and 2 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. In the dynamic state-space equation, noise inputs have been considered only on the component $X(\cdot)$ of the state vector, because the $X'(\cdot)$ component is known to be composed of really stationary environment geometric features. The diagonal elements of the covariance matrix R of the observation noise $V(\cdot)$ are determined by the sensor characteristics.

Figure 1 shows the trajectory of the vehicle estimated by the proposed algorithm. It can be seen that the robot 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 estimate errors corresponding to the final vehicle configuration (the distance between the actual and the estimated configuration) have been reported in Table I. Figure 2 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, while the dotted and dashed lines represent the initial and final estimates of the environment map landmarks, respectively. 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.

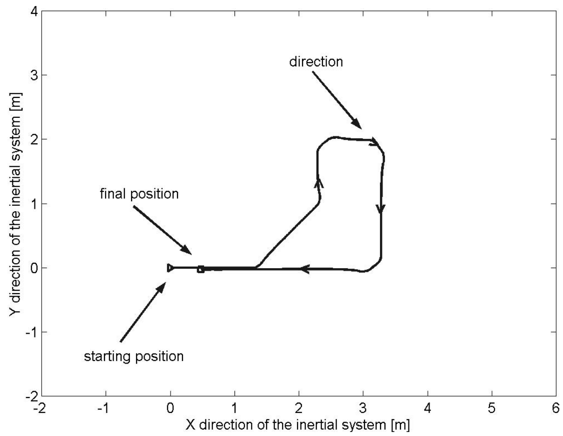

Fig. 1. Estimated trajectory.

TABLE I

ESTIMATION ERRORS (E) IN CORRESPONDENCE OF THE FINAL VEHICLE CONFIGURATION (DISTANCE BETWEEN THE ACTUAL (AC) AND THE CORRESPONDING ESTIMATED (EC) CONFIGURATION)

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

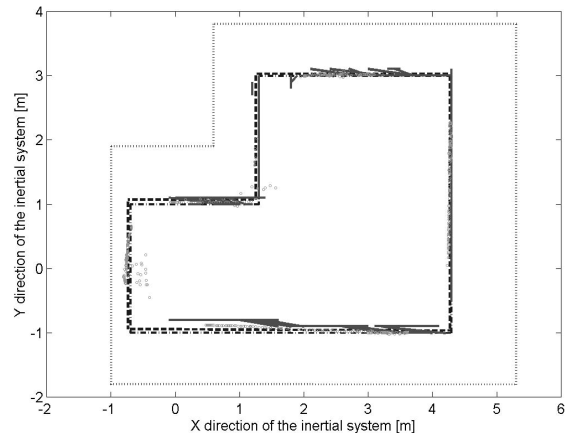

Fig. 2. The dash-dot lines are the actual locations of the environment map landmarks, while the dotted and dashed lines represent the initial and final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures, and the solid lines are the video measures.

TABLE II

DIFFERENCES IN ABSOLUTE VALUE BETWEEN THE ACTUAL LOCATIONS OF ENVIRONMENT LANDMARKS AND THEIR INITIAL (IE) AND FINAL (FE) ESTIMATES. L_i, $i = 1, 2, \ldots, 6$ ARE THE CONSIDERED LANDMARKS

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

The fusion of ultrasonic and video data produces a significant improvement in the reliability of the environment feature estimation with respect to the results obtained using sonar sensor readings only [4]. This is evident comparing Figure 2 with Figure 3, where the dash-dot lines are the actual locations of the environment map landmarks, while the dotted and dashed lines represent the initial and final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures.

VI. CONCLUSIONS

In this paper, an EKF is proposed as a tool for the consistent fusion of information acquired by the robot to reduce the uncertainties on the vehicle localization and on the estimate of environment features. 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

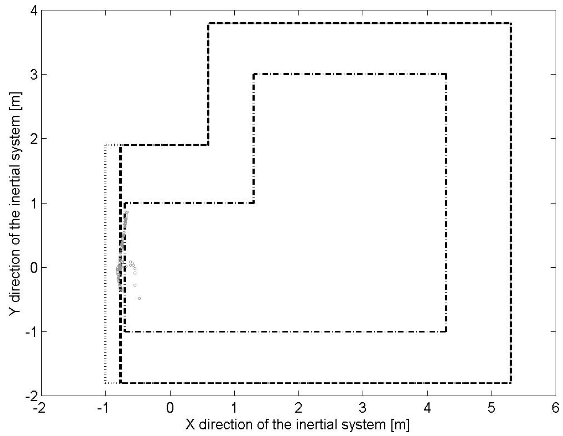

Fig. 3. The dash-dot lines are the actual locations of the environment map landmarks, while the dotted and dashed lines represent the initial and final estimates of the environment map landmarks, respectively. The dots are the actually used sonar measures.

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 state-space model whose state vector contains both the state variables of the robot model and the state variables of landmarks. This introduces a very high degree of accuracy in the estimated vehicle trajectory 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 LabMate in the considered indoor environment. The work is in progress to validate the results in larger and more complex environments. The appealing features of this approach are: i) the possibility of collecting all the available information and uncertainties of a different kind into a meaningful state-space representation; ii) the recursive structure of the solution; iii) the modest computational effort.

References

[1] A. Bemporad, M. Di Marco, and A. Tesi, "Sonar-based wall-following control of mobile robots," *J. of Dynamic Systems, Measurement, and Control*, vol. 122, no. 1, pp. 226–229, 2000.

[2] 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 *Advances in Control of Articulated and Mobile Robots*, B. Siciliano, A. De Luca, C. Melchiorri, and G. Casalino, Eds. Berlin, Heidelberg, Germany: Springer-Verlag, 2004, vol. 10, pp. 191–222.

[3] J. Borenstein, H. Everett, L. Feng, and D. Wehe, "Mobile robot positioning: Sensors and techniques," *J. of Robotic Systems*, vol. 14, no. 4, pp. 231–249, 1997.

[4] G. Ippoliti, L. Jetto, A. La Manna, and S. Longhi, "Improving the robustness properties of robot localization procedures with respect to environment features uncertainties," in *Proc. of the IEEE International Conf. on Robotics and Automation*, Barcelona, Spain, 2005, pp. 1463–1470.

[5] J. Borenstein and Y. Koren, "Histogramic in-motion mapping for mobile robot obstacle avoidance," *IEEE Trans. on Robotics and Automation*, vol. 7, no. 4, pp. 535–539, 1991.

[6] G. Oriolo, G. Ulivi, and M. Vendittelli, "Real-time map building and navigation for autonomous robots in unknown environments," *IEEE Trans. on Systems, Man and Cybernetics – Part B: Cybernetics*, vol. 28, no. 3, pp. 316–333, 1998.

[7] M. Montemerlo and S. Thrun, "Simultaneous localization and mapping with unknown data association using FastSLAM," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA '03)*, vol. 2, 2003, pp. 1985–1991.

[8] A. Bicchi, F. Lorussi, P. Murrieri, and V. Scordio, "On the problem of simultaneous localization, map building, and servoing of autonomous vehicles," in *Advances in Control of Articulated and Mobile Robots*, B. Siciliano, A. De Luca, C. Melchiorri, and G. Casalino, Eds. Berlin, Heidelberg, Germany: Springer-Verlag, 2004, vol. 10, pp. 223–242.

[9] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, 2001.

[10] S. Panzieri, F. Pascucci, and G. Ulivi, "Vision based navigation using Kalman approach for SLAM," in *Proc. of the 11th International Conference on Advanced Robotics*, 2003.

[11] G. Zunino and H. Christensen, "Simultaneous localization and mapping in domestic environments," in *Proc. of the International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2001)*, 2001, pp. 67–72.

[12] C. Wang, "Localization estimation and uncertainty analysis for mobile robots," in *Proc. of the Int. Conf. on Robotics and Automation*, 1988, pp. 1230–1235.

[13] 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. on Robotics and Automation*, vol. 15, pp. 219–229, 1999.

[14] N. Ayache and O. Faugeras, "Maintaining representations of the environment of a mobile robot," *IEEE Trans. on Robotics and Automation*, vol. 5, pp. 804–819, 1989.

[15] A. Kak, "Depth perception for robots," in *Handbook of industrial robotics*, S. Nof, Ed. New York: Wiley, 1985, pp. 272–319.

[16] A. Bonci, T. Leo, and S. Longhi, "Data fusion for visual and ultrasonic map-building," in *Proc. of the 6th IFAC Symp. on Robot Control (SYROCO 2000)*, Wien, Austria, 2000, pp. 241–256.

[17] P. Hough, *Methods and means for recognising complex patterns.* U.S. Patent, 1962, vol. 3,069,654.

[18] I. Sobel, *Camera models and machine perception.* Stanford University CA: PhD Thesis, Electrical Engineering Department, 1970.

[19] 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," in *Proc. of the IEEE Int. Conference on Systems, Man and Cybernetics (SMC02)*, Hammamet, Tunisia, 2002.

[20] A. Angeloni, T. Leo, S. Longhi, and R. Zulli, "Real time collision avoidance for mobile robots," in *Proc. of the 6th Symposium on Measurement and Control in Robotics, ISMCR'96*, Brussels, Belgium, 1996, pp. 239–244.

[21] E. Bender, *Mathematical Methods in Artificial Intelligence.* Wiley-IEEE Computer Society Press, 1996.