A Solution to the Simultaneous Localization and Map Building (SLAM) Problem

M. W. M. Gamini Dissanayake, *Member, IEEE,* Paul Newman, *Member, IEEE,* Steven Clark, Hugh F. Durrant-Whyte, *Member, IEEE,* and M. Csorba

***Abstract*—The simultaneous localization and map building (SLAM) problem asks if it is possible for an autonomous vehicle to start in an unknown location in an unknown environment and then to incrementally build a map of this environment while simultaneously using this map to compute absolute vehicle location. Starting from the estimation-theoretic foundations of this problem developed in [1]–[3], this paper proves that a solution to the SLAM problem is indeed possible. The underlying structure of the SLAM problem is first elucidated. A proof that the estimated map converges monotonically to a relative map with zero uncertainty is then developed. It is then shown that the absolute accuracy of the map and the vehicle location reach a lower bound defined only by the initial vehicle uncertainty. Together, these results show that it is possible for an autonomous vehicle to start in an unknown location in an unknown environment and, using relative observations only, incrementally build a perfect map of the world and to compute simultaneously a bounded estimate of vehicle location. This paper also describes a substantial implementation of the SLAM algorithm on a vehicle operating in an outdoor environment using millimeter-wave (MMW) radar to provide relative map observations. This implementation is used to demonstrate how some key issues such as map management and data association can be handled in a practical environment. The results obtained are cross-compared with absolute locations of the map landmarks obtained by surveying. In conclusion, this paper discusses a number of key issues raised by the solution to the SLAM problem including suboptimal map-building algorithms and map management.**

***Index Terms*—Autonomous navigation, millimeter wave radar, simultaneous localization and map building.**

I. Introduction

THE solution to the simultaneous localization and map building (SLAM) problem is, in many respects, a "Holy Grail" of the autonomous vehicle research community. The ability to place an autonomous vehicle at an unknown location in an unknown environment and then have it build a map, using only relative observations of the environment, and then to use this map simultaneously to navigate would indeed make such a robot "autonomous". Thus the main advantage of SLAM is that it eliminates the need for artificial infrastructures or *a priori* topological knowledge of the environment. A solution to the SLAM problem would be of inestimable value in a range of applications where absolute position or precise map information is unobtainable, including, amongst others, autonomous planetary exploration, subsea autonomous vehicles, autonomous air-borne vehicles, and autonomous all-terrain vehicles in tasks such as mining and construction.

The general SLAM problem has been the subject of substantial research since the inception of a robotics research community and indeed before this in areas such as manned vehicle navigation systems and geophysical surveying. A number of approaches have been proposed to address both the SLAM problem and also more simplified navigation problems where additional map or vehicle location information is made available. Broadly, these approaches adopt one of three main philosophies. The most popular of these is the estimation-theoretic or Kalman filter based approach. The popularity of this approach is due to two main factors. First, it directly provides both a recursive solution to the navigation problem and a means of computing consistent estimates for the uncertainty in vehicle and map landmark locations on the basis of statistical models for vehicle motion and relative landmark observations. Second, a substantial corpus of method and experience has been developed in aerospace, maritime and other navigation applications, from which the autonomous vehicle community can draw. A second philosophy is to eschew the need for absolute position estimates and for precise measures of uncertainty and instead to employ more qualitative knowledge of the relative location of landmarks and vehicle to build maps and guide motion. This general philosophy has been developed by a number of different groups in a number of different ways; see [4]–[6]. The qualitative approach to navigation and the general SLAM problem has many potential advantages over the estimation-theoretic methodology in terms of limiting the need for accurate models and the resulting computational requirements, and in its significant "anthropomorphic appeal". The third, very broad philosophy does away with the rigorous Kalman filter or statistical formalism while retaining an essentially numerical or computational approach to the navigation and SLAM problem. Such approaches include the use of iconic landmark matching [7], global map registration [8], bounded regions [9] and other measures to describe uncertainty. Notable are the work by Thrun *et al.* [10] and Yamauchi *et al.* [11]. Thrun *et al.* use a bayesian approach to map building that does not assume Gaussian probability distributions as required by the Kalman filter. This technique, while very effective for localization with respect to maps, does not lend itself to provide an incremental solution to SLAM where a map is gradually built as information is received from sensors. Yamauchi *et al.*

Manuscript received March 22, 1999; revised September 15, 2000. This paper was recommended for publication by Associate Editor N. Xi and Editor V. Lumelsky upon evaluation of the reviewers' comments.

The authors are with the Department of Mechanical and Mechatronic Engineering, Australian Centre for Field Robotics, The University of Sydney, NSW 2006, Australia (e-mail: pnewman@mit.edu).

Publisher Item Identifier S 1042-296X(01)06731-3.

1042–296X/01$10.00 © 2001 IEEE

use a evidence grid approach that requires that the environment is decomposed to a number of cells.

An estimation-theoretic or Kalman filter based approach to the SLAM problem is adopted in this paper. A major advantage of this approach is that it is possible to develop a complete proof of the various properties of the SLAM problem and to study systematically the evolution of the map and the uncertainty in the map and vehicle location. A proof of existence and convergence for a solution of the SLAM problem within a formal estimation-theoretic framework also encompasses the widest possible range of navigation problems and implies that solutions to the problem using other approaches are possible.

The study of estimation-theoretic solutions to the SLAM problem within the robotics community has an interesting history. Initial work by Smith *et al.* [12] and Durrant-Whyte [13] established a statistical basis for describing relationships between landmarks and manipulating geometric uncertainty. A key element of this work was to show that there must be a high degree of correlation between estimates of the location of different landmarks in a map and that indeed these correlations would grow to unity following successive observations. At the same time Ayache and Faugeras [14] and Chatila and Laumond [15] were undertaking early work in visual navigation of mobile robots using Kalman filter-type algorithms. These two strands of research had much in common and resulted soon after in the key paper by Smith, Self and Cheeseman [1]. This paper showed that as a mobile robot moves through an unknown environment taking relative observations of landmarks, the estimates of these landmarks are all necessarily correlated with each other because of the common error in estimated vehicle location. This paper was followed by a series of related work developing a number of aspects of the essential SLAM problem ([2] and [3], for example). The main conclusion of this work was twofold. First, accounting for correlations between landmarks in a map is important if filter consistency is to be maintained. Second, that a full SLAM solution requires that a state vector consisting of all states in the vehicle model *and* all states of every landmark in the map needs to be maintained and updated following each observation if a complete solution to the SLAM problem is required. The consequence of this in any real application is that the Kalman filter needs to employ a huge state vector (of order the number of landmarks maintained in the map) and is in general, computationally intractable. Crucially, this work did not look at the convergence properties of the map or its steady-state behavior. Indeed, it was widely assumed at the time that the estimated map errors would not converge and would instead execute a random walk behavior with unbounded error growth. Given the computational complexity of the SLAM problem and without knowledge of the convergence behavior of the map, a series of approximations to the full SLAM solution were proposed which assumed that the correlations between landmarks could be minimized or eliminated thus reducing the full filter to a series of decoupled landmark to vehicle filters (see Renken [16], Leonard and Durrant-Whyte [3] for example). Also for these reasons, theoretical work on the full estimation-theoretic SLAM problem largely ceased, with effort instead being expended in map-based navigation and alternative theoretical approaches to the SLAM problem.

This paper starts from the original estimation-theoretic work of Smith, Self and Cheeseman. It assumes an autonomous vehicle (mobile robot) equipped with a sensor capable of making measurements of the location of landmarks relative to the vehicle. The landmarks may be artificial or natural and it is assumed that the signal processing algorithms are available to detect these landmarks. The vehicle starts at an unknown location with no knowledge of the location of landmarks in the environment. As the vehicle moves through the environment (in a stochastic manner) it makes relative observations of the location of individual landmarks. This paper then proves the following three results.

1) The determinant of any submatrix of the map covariance matrix decreases monotonically as observations are successively made.
2) In the limit as the number of observations increases, the landmark estimates become fully correlated.
3) In the limit, the covariance associated with any single landmark location estimate is determined only by the initial covariance in the vehicle location estimate.

These three results describe, in full, the convergence properties of the map and its steady state behavior. In particular they show the following.

- The entire structure of the SLAM problem critically depends on maintaining complete knowledge of the cross correlation between landmark estimates. Minimizing or ignoring cross correlations is precisely contrary to the structure of the problem.
- As the vehicle progresses through the environment the errors in the estimates of any pair of landmarks become more and more correlated, and indeed never become less correlated.
- In the limit, the errors in the estimates of any pair of landmarks becomes fully correlated. This means that given the exact location of any one landmark, the location of any other landmark in the map can also be determined with absolute certainty.
- As the vehicle moves through the environment taking observations of individual landmarks, the error in the estimates of the relative location between different landmarks reduces monotonically to the point where the map of relative locations is known with absolute precision.
- As the map converges in the above manner, the error in the absolute location of every landmark (and thus the whole map) reaches a lower bound determined only by the error that existed when the first observation was made.

Thus a solution to the general SLAM problem exists and it is indeed possible to construct a perfectly accurate map and simultaneously compute vehicle position estimates without any prior knowledge of vehicle or landmark locations.

This paper makes three principal contributions to the solution of the SLAM problem. First, it proves three key convergence properties of the full SLAM filter. Second, it elucidates the true structure of the SLAM problem and shows how this can be used in developing consistent SLAM algorithms. Finally, it demonstrates and evaluates the implementation of the full SLAM al-

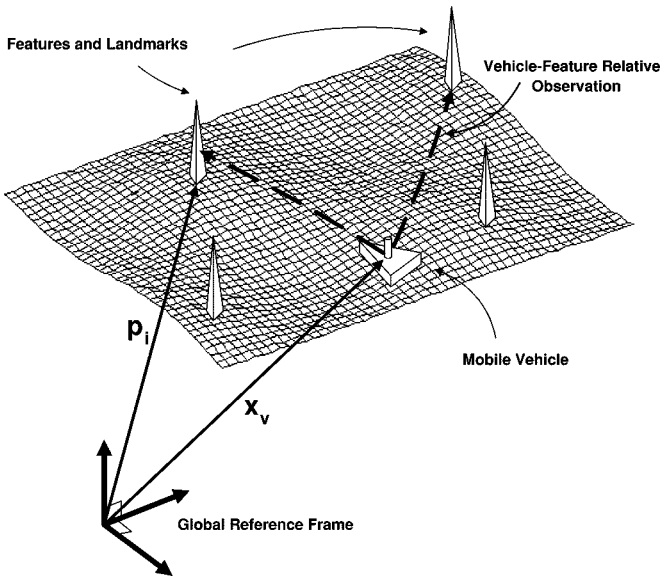

Fig. 1. A vehicle taking relative measurements to environmental landmarks.

gorithm in an outdoor environment using a millimeter wave (MMW) radar sensor.

Section II of this paper introduces the mathematical structure of the estimation-theoretic SLAM problem. Section III then proves and explains the three convergence results. Section IV provides a practical demonstration of an implementation of the full SLAM algorithm in an outdoor environment using MMW radar to provide relative observations of landmarks. An algorithm addressing pertinent issues of map initialization and management is also presented. The algorithm outputs are shown to exhibit the convergent properties derived in Section III. Section V discusses the many remaining problems with obtaining a practical, large scale solution to the SLAM problem including the development of suboptimal solutions, map management, and data association.

II. The Estimation-Theoretic SLAM Problem

This section establishes the mathematical framework employed in the study of the SLAM problem. This framework is identical in all respects to that used in Smith *et al.* [1] and uses the same notation as that adopted in Leonard and Durrant-Whyte [3].

A. Vehicle and Landmark Models

The setting for the SLAM problem is that of a vehicle with a known kinematic model, starting at an unknown location, moving through an environment containing a population of features or landmarks. The terms feature and landmark will be used synonymously. The vehicle is equipped with a sensor that can take measurements of the relative location between any individual landmark and the vehicle itself as shown in Fig. 1. The absolute locations of the landmarks are not available. Without prejudice, a linear (synchronous) discrete-time model of the evolution of the vehicle and the observations of landmarks is adopted. Although vehicle motion and the observation of landmarks is almost always nonlinear and asynchronous in any real navigation problem, the use of linear synchronous models does not affect the validity of the proofs in Section III other than to require the same linearization assumptions as those normally employed in the development of an extended Kalman filter. Indeed, the implementation of the SLAM algorithm described in Section IV uses nonlinear vehicle models and nonlinear asynchronous observation models. The state of the system of interest consists of the position and orientation of the vehicle together with the position of all landmarks. The state of the vehicle at a time step k is denoted $\mathbf{x}_v(k)$. The motion of the vehicle through the environment is modeled by a conventional linear discrete-time state transition equation or process model of the form

$$\mathbf{x}_v(k+1) = \mathbf{F}_v(k)\mathbf{x}_v(k) + \mathbf{u}_v(k+1) + \mathbf{v}_v(k+1) \tag{1}$$

where

- $\mathbf{F}_v(k)$ state transition matrix;
- $\mathbf{u}_v(k)$ vector of control inputs; and
- $\mathbf{v}_v(k)$ vector of temporally uncorrelated process noise errors with zero mean and covariance $\mathbf{Q}_v(k)$ (see [17] and [18] for further details).

The location of the ith landmark is denoted $\mathbf{p}_i$. The "state transition equation" for the ith landmark is

$$\mathbf{p}_i(k+1) = \mathbf{p}_i(k) = \mathbf{p}_i \tag{2}$$

since landmarks are assumed stationary. Without loss of generality the number of landmarks in the environment is arbitrarily set at N. The vector of all N landmarks is denoted

$$\mathbf{p} = [\,\mathbf{p}_1^{\mathrm{T}} \quad \dots \quad \mathbf{p}_N^{\mathrm{T}}\,]^{\mathrm{T}} \tag{3}$$

where T denotes the transpose and is used both inside and outside the brackets to conserve space. The augmented state vector containing both the state of the vehicle and the state of all landmark locations is denoted

$$\mathbf{x}(k) = [\,\mathbf{x}_v^{\mathrm{T}}(k) \quad \mathbf{p}_1^{\mathrm{T}} \quad \dots \quad \mathbf{p}_N^{\mathrm{T}}\,]^{\mathrm{T}}. \tag{4}$$

The augmented state transition model for the complete system may now be written as

$$\begin{bmatrix} \mathbf{x}_v(k+1) \\ \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_N \end{bmatrix} = \begin{bmatrix} \mathbf{F}_v(k) & 0 & \dots & 0 \\ 0 & \mathbf{I}_{p_1} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{I}_{p_N} \end{bmatrix} \begin{bmatrix} \mathbf{x}_v(k) \\ \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_N \end{bmatrix} + \begin{bmatrix} \mathbf{u}_v(k+1) \\ \mathbf{0}_{p_1} \\ \vdots \\ \mathbf{0}_{p_N} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_v(k+1) \\ \mathbf{0}_{p_1} \\ \vdots \\ \mathbf{0}_{p_N} \end{bmatrix} \tag{5}$$

$$\mathbf{x}(k+1) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{u}(k+1) + \mathbf{v}(k+1) \tag{6}$$

where $\mathbf{I}_{p_1}$ is the $\dim(p_i) \times \dim(p_i)$ identity matrix and $\mathbf{0}_{p_1}$ is the $\dim(p_i)$ null vector.

The case in which landmarks $\mathbf{p}_i$ are in stochastic motion may easily be accommodated in this framework. However, doing so offers little insight into the problem and furthermore the convergence properties presented by this paper do not hold.

B. The Observation Model

The vehicle is equipped with a sensor that can obtain observations of the relative location of landmarks with respect to the vehicle. Again, without prejudice, observations are assumed to be linear and synchronous. The observation model for the ith landmark is written in the form

$$\mathbf{z}_i(k) = \mathbf{H}_i\mathbf{x}(k) + \mathbf{w}_i(k) \tag{7}$$
$$= \mathbf{H}_{p_i}\mathbf{p} - \mathbf{H}_v\mathbf{x}_v(k) + \mathbf{w}_i(k) \tag{8}$$

where $\mathbf{w}_i(k)$ is a vector of temporally uncorrelated observation errors with zero mean and variance $\mathbf{R}_i(k)$. The term $\mathbf{H}_i$ is the observation matrix and relates the output of the sensor $\mathbf{z}_i(k)$ to the state vector $\mathbf{x}(k)$ when observing the ith landmark. It is important to note that the observation model for the ith landmark is written in the form

$$\mathbf{H}_i = [-\mathbf{H}_v, \mathbf{0}\dots\mathbf{0}, \mathbf{H}_{p_i}, \mathbf{0}\dots\mathbf{0}]. \tag{9}$$

This structure reflects the fact that the observations are "relative" between the vehicle and the landmark, often in the form of relative location, or relative range and bearing (see Section IV).

C. The Estimation Process

In the estimation-theoretic formulation of the SLAM problem, the Kalman filter is used to provide estimates of vehicle and landmark location. We briefly summarize the notation and main stages of this process as a necessary prelude to the developments in Section III and IV of this paper. Detailed descriptions may be found in [17], [18] and [3]. The Kalman filter recursively computes estimates for a state $\mathbf{x}(k)$ which is evolving according to the process model in (5) and which is being observed according to the observation model in (7). The Kalman filter computes an estimate which is equivalent to the conditional mean $\hat{\mathbf{x}}(p|q) = E[\mathbf{x}(p)|\mathbf{Z}^q]$ $(p \geq q)$, where $\mathbf{Z}^q$ is the sequence of observations taken up until time q. The error in the estimate is denoted $\tilde{\mathbf{x}}(p|q) = \hat{\mathbf{x}}(p|q) - \mathbf{x}(p)$. The Kalman filter also provides a recursive estimate of the covariance $\mathbf{P}(p|q) = E\left[\tilde{\mathbf{x}}(p|q)\tilde{\mathbf{x}}(p|q)^{\mathrm{T}}|\mathbf{Z}^q\right]$ in the estimate $\hat{\mathbf{x}}(p|q)$. The Kalman filter algorithm now proceeds recursively in three stages:

- Prediction: Given that the models described in (5) and (7) hold, and that an estimate $\hat{\mathbf{x}}(k|k)$ of the state $\mathbf{x}(k)$ at time k together with an estimate of the covariance $\mathbf{P}(k|k)$ exist, the algorithm first generates a prediction for the state estimate, the observation (relative to the ith landmark) and the state estimate covariance at time $k+1$ according to

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{F}(k)\hat{\mathbf{x}}(k|k) + \mathbf{u}(k) \tag{10}$$
$$\hat{\mathbf{z}}_i(k+1|k) = \mathbf{H}_i(k)\hat{\mathbf{x}}(k+1|k) \tag{11}$$
$$\mathbf{P}(k+1|k) = \mathbf{F}(k)\mathbf{P}(k|k)\mathbf{F}^{\mathrm{T}}(k) + \mathbf{Q}(k) \tag{12}$$

  respectively.
- Observation: Following the prediction, an observation $\mathbf{z}_i(k+1)$ of the ith landmark of the true state $\mathbf{x}(k+1)$ is made according to (7). Assuming correct landmark association, an innovation is calculated as follows:

$$\nu_i(k+1) = \mathbf{z}_i(k+1) - \hat{\mathbf{z}}_i(k+1|k) \tag{13}$$

  together with an associated innovation covariance matrix given by

$$\mathbf{S}_i(k+1) = \mathbf{H}_i(k)\mathbf{P}(k+1|k)\mathbf{H}_i^{\mathrm{T}}(k) + \mathbf{R}_i(k+1). \tag{14}$$

- Update: The state estimate and corresponding state estimate covariance are then updated according to

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{W}_i(k+1)\nu_i(k+1) \tag{15}$$
$$\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) - \mathbf{W}_i(k+1) \times \mathbf{S}(k+1) \times \mathbf{W}_i^{\mathrm{T}}(k+1) \tag{16}$$

  where the gain matrix $\mathbf{W}_i(k+1)$ is given by

$$\mathbf{W}_i(k+1) = \mathbf{P}(k+1|k)\mathbf{H}_i^{\mathrm{T}}(k)\mathbf{S}_i^{-1}(k+1). \tag{17}$$

  The update of the state estimate covariance matrix is of paramount importance to the SLAM problem. Understanding the structure and evolution of the state covariance matrix is the key component to this solution of the SLAM problem.

III. Structure of the SLAM Problem

In this section proofs for the three key results underlying the structure of the SLAM problem are provided. The implications of these results will also be examined in detail. The appendix provides a summary of the key properties of positive semidefinite matrices that are invoked implicitly in the following proofs.

A. Convergence of the Map Covariance Matrix

The state covariance matrix may be written in block form as

$$\mathbf{P}(i|j) = \begin{bmatrix} \mathbf{P}_{vv}(i|j) & \mathbf{P}_{vm}(i|j) \\ \mathbf{P}_{vm}^{\mathrm{T}}(i|j) & \mathbf{P}_{mm}(i|j) \end{bmatrix}$$

where

- $\mathbf{P}_{vv}(i|j)$ error covariance matrix associated with the vehicle state estimate;
- $\mathbf{P}_{mm}(i|j)$ map covariance matrix associated with the landmark state estimates; and
- $\mathbf{P}_{vm}(i|j)$ cross-covariance matrix between vehicle and landmark states.

Theorem 1: The determinant of any submatrix of the map covariance matrix decreases monotonically as successive observations are made.

The algorithm is initialized using a positive semidefinite (*psd*) state covariance matrix $\mathbf{P}(0|0)$. The matrices $\mathbf{Q}$ and $\mathbf{R}_i$ are both *psd*, and consequently the matrices $\mathbf{P}(k+1|k)$, $\mathbf{S}_i(k+1)$, $\mathbf{W}_i(k+1)\mathbf{S}_i(k+1)\mathbf{W}_i^{\mathrm{T}}(k+1)$ and $\mathbf{P}(k+1|k+1)$ are all *psd*. From (16), and for any landmark i

$$\det\mathbf{P}(k+1|k+1) = \det\left[\mathbf{P}(k+1|k) - \mathbf{W}_i(k+1)\mathbf{S}_i(k+1)\mathbf{W}^{\mathrm{T}}(k+1)\right] \leq \det\mathbf{P}(k+1|k). \tag{18}$$

The determinant of the state covariance matrix is a measure of the volume of the uncertainty ellipsoid associated with the state estimate. Equation (18) states that the total uncertainty of the state estimate does not increase during an update.

Any principal submatrix of a *psd* matrix is also *psd* (see Appendix 1). Thus, from (18) the map covariance matrix also has the property

$$\det \mathbf{P}_{mm}(k+1|k+1) \leq \det \mathbf{P}_{mm}(k+1|k) \tag{19}$$

From (12), the full state covariance prediction equation may be written in the form

$$\begin{bmatrix} \mathbf{P}_{vv}(k+1|k) & \mathbf{P}_{vm}(k+1|k) \\ \mathbf{P}_{vm}^{\mathrm{T}}(k+1|k) & \mathbf{P}_{mm}(k+1|k) \end{bmatrix} = \mathbf{Y}_1$$

where

$$\mathbf{Y}_1 = \begin{bmatrix} \mathbf{F}_v \mathbf{P}_{vv}(k|k) {\mathbf{F}_v}^{\mathrm{T}} + \mathbf{Q}_{vv} & \mathbf{F}_v \mathbf{P}_{vm}(k|k) \\ \mathbf{P}_{vm}^{\mathrm{T}}(k|k) {\mathbf{F}_v}^{\mathrm{T}} & \mathbf{P}_{mm}(k|k) \end{bmatrix}$$

Thus, as landmarks are assumed stationary, no process noise is injected in to the predicted map states. Consequently, the map covariance matrix and any principal submatrix of the map covariance matrix has the property that

$$\mathbf{P}_{mm}(k+1|k) = \mathbf{P}_{mm}(k|k) \tag{20}$$

Note that this is clearly not true for the full covariance matrix as process noise is injected in to the vehicle location predictions and so the prediction covariance grows during the prediction step.

It follows from (19) and (20) that the map covariance matrix has the property that

$$\det \mathbf{P}_{mm}(k+1|k+1) \leq \det \mathbf{P}_{mm}(k|k) \tag{21}$$

Furthermore, the general properties of *psd* matrices ensure that this inequality holds for *any* submatrix of the map covariance matrix. In particular, for any diagonal element σ_{ii}^2 of the map covariance matrix (state variance),

$$\sigma_{ii}^2(k+1|k+1) \leq \sigma_{ii}^2(k|k)$$

Thus the error in the estimate of the absolute location of every landmark also diminishes.

Theorem 2: In the limit the landmark estimates become fully correlated.

As the number of observations taken tends to infinity a lower limit on the map covariance limit will be reached such that

$$\lim_{k\to\infty} [\mathbf{P}_{mm}(k+1|k+1)] = \mathbf{P}_{mm}(k|k) \tag{22}$$

Writing $\mathbf{P}(k+1|k)$ as $\mathbf{P}^{\ominus}$ and $\mathbf{P}(k+1|k+1)$ as $\mathbf{P}^{\ominus}$ for notational clarity, the SLAM algorithm update stage can be written as

$$\begin{aligned} \mathbf{P}^{\ominus} =& \mathbf{P}^{\ominus} - \mathbf{W}_i(k+1)\mathbf{S}_i{\mathbf{W}_i}^{\mathrm{T}}(k+1) \\ =& \mathbf{P}^{\ominus} - \mathbf{P}^{\ominus}{\mathbf{H}_i}^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{H}_i\mathbf{P}^{\ominus} \\ =& \mathbf{P}^{\ominus} - \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{bmatrix} \mathbf{S}_i^{-1} [{\mathbf{M}_1}^{\mathrm{T}} \quad {\mathbf{M}_2}^{\mathrm{T}}] \\ =& \mathbf{P}^{\ominus} - \begin{bmatrix} \mathbf{M}_1\mathbf{S}_i^{-1}{\mathbf{M}_1}^{\mathrm{T}} & \mathbf{M}_1\mathbf{S}_i^{-1}{\mathbf{M}_2}^{\mathrm{T}} \\ \mathbf{M}_2\mathbf{S}_i^{-1}{\mathbf{M}_1}^{\mathrm{T}} & \mathbf{M}_2\mathbf{S}_i^{-1}{\mathbf{M}_2}^{\mathrm{T}} \end{bmatrix} \end{aligned} \tag{23}$$

where

$$\begin{aligned} \mathbf{M}_1 =& -\mathbf{P}_{vv}^{\ominus}{\mathbf{H}_v}^{\mathrm{T}} + \mathbf{P}_{vm}^{\ominus}{\mathbf{H}_{pi}}^{\mathrm{T}} \\ \mathbf{M}_2 =& -\mathbf{P}_{vm}^{\ominus}{\mathbf{H}_v}^{\mathrm{T}} + \mathbf{P}_{mm}^{\ominus}{\mathbf{H}_{pi}}^{\mathrm{T}} \end{aligned} \tag{24}$$

The update of the map covariance matrix $\mathbf{P}_{mm}$ can be written as

$$\begin{aligned} \mathbf{P}_{mm}(k+1|k+1) =& \mathbf{P}_{mm}(k+1|k) - \mathbf{M}_2\mathbf{S}_i^{-1}{\mathbf{M}_2}^{\mathrm{T}} \\ =& \mathbf{P}_{mm}(k|k) - \mathbf{M}_2\mathbf{S}_i^{-1}{\mathbf{M}_2}^{\mathrm{T}} \end{aligned} \tag{25}$$

Equation (22) and (25) imply that the matrix $\mathbf{M}_2\mathbf{S}_i^{-1}{\mathbf{M}_2}^{\mathrm{T}} = \mathbf{0}$. The inverse of the innovation covariance matrix $\mathbf{S}_i^{-1}$ is always *psd* because the observation noise covariance $\mathbf{R}_i$ is *psd*, therefore (22) requires that $\mathbf{M}_2 = \mathbf{0}$

$$\mathbf{P}_{mm}(k|k){\mathbf{H}_{pi}}^{\mathrm{T}} = \mathbf{P}_{vm}(k|k){\mathbf{H}_v}^{\mathrm{T}} \quad \forall\, i \tag{26}$$

Equation (26) holds for all i and therefore the block columns of $\mathbf{P}_{mm}$ are linearly dependent. A consequence of this fact is that in the limit the determinant of the covariance matrix of a map containing more than one landmark tends to zero.

$$\lim_{k\to\infty} [\det \mathbf{P}_{mm}(k|k)] = \mathbf{0} \tag{27}$$

This implies that the landmarks become progressively more correlated as successive observations are made. In the limit then, given the exact location of one landmark the location of all other landmarks can be deduced with absolute certainty and the map is fully correlated.

Consider the implications of (26) upon the estimate $\mathbf{d}$ of the relative position between any two landmarks $\mathbf{p}_i$ and $\mathbf{p}_j$ of the same type.

$$\mathbf{d} = \hat{\mathbf{p}}_i - \hat{\mathbf{p}}_j = \mathbf{G}_{ij}\mathbf{x}$$

The covariance $\mathbf{P}_d$ of $\mathbf{d}$ is given by

$$\mathbf{P}_d = \mathbf{G}_{ij}\mathbf{P}{\mathbf{G}_{ij}}^{\mathrm{T}} = \mathbf{P}_{ii} + \mathbf{P}_{jj} - \mathbf{P}_{ij} - {\mathbf{P}_{ij}}^{\mathrm{T}}$$

With similar landmark types, $\mathbf{H}_{pi} = \mathbf{H}_{pj}$ and so (26) implies that the block columns of $\mathbf{P}_{mm}$ are identical. Furthermore, because $\mathbf{P}_{mm}$ is symmetric it follows that

$$\mathbf{P}_{ii} = \mathbf{P}_{jj} = \mathbf{P}_{ij} = {\mathbf{P}_{ij}}^{\mathrm{T}} \quad \forall\, i,j \tag{28}$$

Therefore, in the limit, $\mathbf{P}_d = \mathbf{0}$ and the relationship between the landmarks is known with complete certainty. It is important to note that this result does not mean that the determinants of the landmark covariance matrices tend to zero. In the limit the absolute location of landmarks may still be uncertain.

Theorem 3: In the limit, the lower bound on the covariance matrix associated with any single landmark estimate is determined only by the initial covariance in the vehicle estimate $\mathbf{P}_{0v}$ at the time of the first sighting of the first landmark.

As described previously, the covariance of the landmark location estimates decrease as successive observations are made. The best estimates are obviously obtained when the covariance matrices of the vehicle process noise $\mathbf{Q}$ and the observation noise $\mathbf{R}_i$ are small. The limiting value and hence lower bound of the state covariance matrix can be obtained when the vehicle is

stationary giving $\mathbf{Q} = \mathbf{0}$. Under these circumstances it is convenient to use the information form of the Kalman filter to examine the behavior of the state covariance matrix. Continuing with the single landmark environment, the state covariance matrix after this solitary landmark has been observed for k instances can be written as

$$\mathbf{P}(k|k)^{-1} = \mathbf{P}(k|k-1)^{-1} + \begin{bmatrix} -\mathbf{H}_v{}^{\mathrm{T}} \\ \mathbf{H}_{p1}{}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \mathbf{R}_1{}^{-1} \begin{bmatrix} -\mathbf{H}_v & \mathbf{H}_{p1} \end{bmatrix} \quad (29)$$

where because $\mathbf{Q} = \mathbf{0}$,

$$\mathbf{P}(k|k-1)^{-1} = \mathbf{P}(k-1|k-1)^{-1} \quad (30)$$

Using (29) and (30) $\mathbf{P}(k|k)^{-1}$ can now be written as

$$\mathbf{P}(k|k)^{-1} = \begin{bmatrix} \mathbf{P}_{0v}{}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \mathbf{Y}_2 \quad (31)$$

where

$$\mathbf{Y}_2 = \begin{bmatrix} k\mathbf{H}_v{}^{\mathrm{T}}\mathbf{R}_1{}^{-1}\mathbf{H}_v & -k\mathbf{H}_v{}^{\mathrm{T}}\mathbf{R}_1{}^{-1}\mathbf{H}_{p1} \\ -k\mathbf{H}_{p1}{}^{\mathrm{T}}\mathbf{R}_1{}^{-1}\mathbf{H}_v & k\mathbf{H}_{p1}{}^{\mathrm{T}}\mathbf{R}_1{}^{-1}\mathbf{H}_{pi} \end{bmatrix}$$

Invoking the matrix inversion lemma for partitioned matrices,

$$\mathbf{P}(k|k) = \begin{bmatrix} \mathbf{P}_{0v} & \mathbf{P}_{0v}\mathbf{H}_v{}^{\mathrm{T}}\left[\mathbf{H}_{p1}{}^{\mathrm{T}}\right]^{-1} \\ \mathbf{H}_{p1}^{-1}\mathbf{H}_v\mathbf{P}_{0v} & \mathbf{H}_{p1}^{-1}\mathbf{H}_v\mathbf{P}_{0v}\mathbf{H}_{p1}^{-1}\mathbf{H}_v{}^{\mathrm{T}} + \mathbf{Y}_3 \end{bmatrix} \quad (32)$$

where

$$\mathbf{Y}_3 = \frac{\mathbf{H}_{p1}^{-1}\mathbf{R}_1\left[\mathbf{H}_{p1}^{-1}\right]^{\mathrm{T}}}{k}.$$

Examination of the lower right block matrix of (32) shows how the landmark uncertainty estimate stems only from the initial vehicle uncertainty $\mathbf{P}_{0v}$. To find the lower bound on the state covariance matrix and in turn the landmark uncertainty the number of observations of the landmark is allowed to tend to infinity to yield in the limit

$$\lim_{k\to\infty} \mathbf{P}(k|k) = \begin{bmatrix} \mathbf{P}_{0v} & \mathbf{P}_{0v}\mathbf{H}_v{}^{\mathrm{T}}\left[\mathbf{H}_{p1}{}^{\mathrm{T}}\right]^{-1} \\ \mathbf{H}_{p1}^{-1}\mathbf{H}_v\mathbf{P}_{0v} & \mathbf{H}_{p1}^{-1}\mathbf{H}_v\mathbf{P}_{0v}\mathbf{H}_{p1}^{-1}\mathbf{H}_v{}^{\mathrm{T}} \end{bmatrix} \quad (33)$$

Equation (33) gives the lower bound of the solitary landmark state estimate variance as $\mathbf{H}_{p1}^{-1}\mathbf{H}_v\mathbf{P}_{0v}\mathbf{H}_{p1}^{-1}\mathbf{H}_v{}^{\mathrm{T}}$. Examine now the case of an environment containing N landmarks, The smallest achievable uncertainty in the estimate of the ith landmark when the landmark has been observed at the exclusion of all other landmarks is $\mathbf{H}_{pi}^{-1}\mathbf{H}_v\mathbf{P}_{0v}\mathbf{H}_{pi}^{-1}\mathbf{H}_v{}^{\mathrm{T}}$. If more than one landmark is observed as $k \to \infty$, as will be the case in any nontrivial navigation problem, then it is possible for the landmark uncertainty to be further reduced by theorem 1. In the limit the lower bound on the uncertainty in the ith landmark state is written as

$$\mathbf{P}_{p_i,p_i}{}^{\infty} = \min_{i\in[1,N]} \left\{ \mathbf{H}_{pi}^{-1}\mathbf{H}_v\mathbf{P}_{0v}\mathbf{H}_{pi}^{-1}\mathbf{H}_v{}^{\mathrm{T}} \right\} \quad (34)$$

and is determined only by the initial covariance in the vehicle location estimate $\mathbf{P}_{0v}$. Note that because $\mathbf{Q}$ was set to zero in search of the lower bound the vehicle uncertainty remains unchanged at $\mathbf{P}_{0v}$ as $k \to \infty$. In the simple case where $\mathbf{H}_{pi}$ and $\mathbf{H}_v$ are identity matrices in the limit the certainty of each landmark estimate achieves a lower bound given by the initial uncertainty of the vehicle.

When the process noise is not zero the two competing effects of loss of information content due to process noise and the increase in information content through observations, determine the limiting covariance. The problem is now analytically intractable, although the limiting covariance of the map can never be below the limit given by the above equation and will be a function of $\mathbf{P}_{0v}, \mathbf{Q}$ and $\mathbf{R}$.

It is important to note that the limit to the covariance applies because all the landmarks are observed and initialized solely from the observations made from the vehicle. The covariances of landmark estimates can not be further reduced by making additional observations to previously unknown landmarks. However, incorporation of external information, for example using an observation made to a landmark whose location is available through external means such as GPS, will reduce the limiting covariance.

In summary, the three theorems derived above describe, in full, the convergence properties of the map and its steady state behavior. As the vehicle progresses through the environment the total uncertainty of the estimates of landmark locations reduces monotonically to the point where the map of relative locations is known with absolute precision. In the limit, errors in the estimates of any pair of landmarks become fully correlated. This means that given the exact location of any one landmark, the location of any other landmark in the map can also be determined with absolute certainty. As the map converges in the above manner, the error in the absolute location estimate of every landmark (and thus the whole map) reaches a lower bound determined only by the error that existed when the first observation was made.

Thus a solution to the general SLAM problem exists and it is indeed possible to construct a perfectly accurate map describing the relative location of landmarks and simultaneously compute vehicle position estimates without any prior knowledge of landmark or vehicle locations.

IV. Implementation of the Simultaneous Localization and Map Building Algorithm

This section describes a practical implementation of the simultaneous localization and map building (SLAM) algorithm on a standard road vehicle. The vehicle is equipped with a millimeter wave radar (MMWR), which provides observations of the location of landmarks with respect to the vehicle. The implementation is aimed at demonstrating key properties of the SLAM algorithm; convergence, consistency and boundedness of the map error.

The implementation also serves to highlight a number of key properties of the SLAM algorithm and its practical development. In particular, the implementation shows how generally nonlinear vehicle and observation models may be incorporated in the algorithm, how the issue of data association can be dealt

Fig. 2. The test vehicle, showing mounting of the MMWR and GPS systems.

Fig. 3. Test vehicle at the test site. A radar point landmark can be seen in the left foreground.

with, and how landmarks are initialized and tracked as the algorithm proceeds.

The implementation described here is, however, only a first step in the realization and deployment of a fully autonomous SLAM navigation system. A number of substantive further issues in landmark extraction, data association, reduced computation and map management are discussed further in following sections.

A. Experimental Setup

Fig. 2 shows the test vehicle; a conventional utility vehicle fitted with a MMWR system as the primary sensor used in the experiments. Encoders are fitted to the drive shaft to provide a measure of the vehicle speed and a linear variable differential transformer (LVDT) is fitted to the steering rack to provide a measure of vehicle heading. A differential GPS system and an inertial measurement unit are also fitted to the vehicle but are not used in the experiments described below. In the environment used for the experiments, DGPS was found to be prone to large errors due to the reflections caused by nearby large metal structures (usually known as "multipath" errors). The radar employed in the experiments is a 77–GHz FMCW unit. The radar beam is scanned 360° in azimuth at a rate of 1–3 Hz. After signal processing the radar provides an amplitude signal (power spectral density), corresponding to returns at different ranges, at angular increments of approximately 1.5°. This signal is thresholded to provide a measurement of range and bearing to a target. The radar employs a dual polarization receiver so that even and odd bounce specularities can be distinguished. The radar is capable of providing range measurements to 250 m with a resolution of 10 cm in range and 1.5° in bearing. A detailed description of the radar and its performance can be found in [19]. Fig. 3 shows the test vehicle moving in an environment that contains a number of radar reflectors. These reflectors appear as omni-directional point landmarks in the radar images. These, together with a number of natural landmark point targets, serve as the landmarks to be estimated by the SLAM algorithm.

The vehicle is driven manually. Radar range and bearing measurements are logged together with encoder and steer information by an on-board computer system. In the evaluation of the SLAM algorithm, this information is employed without any *a priori* knowledge of landmark location to deduce estimates for both vehicle position and landmark locations.

To evaluate the SLAM algorithm, it is necessary to have some idea of the true vehicle track and true landmark locations that can be compared with those estimated by the SLAM algorithm. For this reason the true landmark locations were accurately surveyed for comparison with the output of the map building algorithm. A second navigation algorithm that employs knowledge of beacon locations is then run on the same data set as used to generate the map estimates. (This algorithm is very similar to that described in [20].) This algorithm provides an accurate estimate of true vehicle location which can be used for comparison with the estimate generated from the SLAM algorithm.

In the following, the vehicle state is defined by $\mathbf{x}_v = [x, y, \varphi]^{\mathrm{T}}$ where x and y are the coordinates of the centre of the rear axle of the vehicle with respect to some global coordinate frame and φ is the orientation of the vehicle axis. The landmarks are modeled as point landmarks and represented by a cartesian pair such that $\mathbf{p}_i = [x_i, y_i]$, $i = 1 \ldots N$. Both vehicle and landmark states are registered in the same frame of reference.

1) The Process Model: Fig. 4 shows a schematic diagram of the vehicle in the process of observing a landmark. The following kinematic equations can be used to predict the vehicle state from the steering γ and velocity inputs V:

$$\begin{aligned}\dot{x} &= V\cos(\varphi)\\ \dot{y} &= V\sin(\varphi)\\ \dot{\varphi} &= \frac{V\tan(\gamma)}{L},\end{aligned}$$

where L is the wheel-base length of the vehicle. These equations can be used to obtain a discrete-time vehicle process model in the form

$$\begin{bmatrix} x(k+1)\\ y(k+1)\\ \varphi(k+1)\end{bmatrix} = \begin{bmatrix} x(k)+\Delta T V(k)\cos(\varphi(k))\\ y(k)+\Delta T V(k)\sin(\varphi(k))\\ \varphi(k)+\frac{\Delta T V(k)\tan(\gamma(k))}{L}\end{bmatrix} \tag{35}$$

for use in the prediction stage of the vehicle state estimator. The landmarks in the environment are assumed to be stationary point targets. The landmark process model is thus

$$\begin{bmatrix} x_i(k+1)\\ y_i(k+1)\end{bmatrix} = \begin{bmatrix} x_i(k)\\ y_i(k)\end{bmatrix} \tag{36}$$

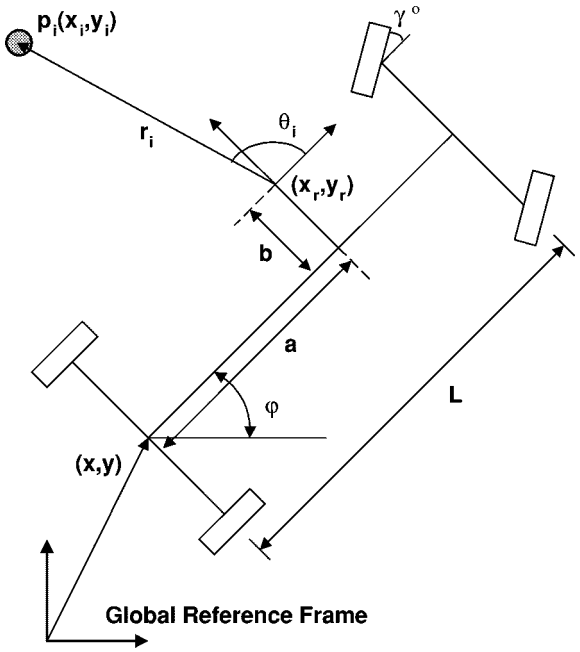

Fig. 4. Vehicle and observation kinematics.

for all landmarks $i = 1 \ldots N$. Together, (35) and (36) define the state transition matrix $\mathbf{F}(\cdot)$ for the system.

2) The Observation Model: The MMWR used in the experiments returns the range $r_i(k)$ and bearing $\theta_i(k)$ to a landmark i. Referring to Fig. 4, the observation model can be written as

$$\begin{aligned} r_i(k) &= \sqrt{(x_i - x_r(k))^2 + (y_i - y_r(k))^2} + w_r(k) \\ \theta_i(k) &= \arctan\left(\frac{y_i - y_r(k)}{x_i - x_r(k)}\right) - \varphi(k) + w_\theta(k) \end{aligned} \tag{37}$$

where w_r and w_θ are the noise sequences associated with the range and bearing measurements, and $[x_r(k), y_r(k)]$ is the location of the radar given, in global coordinates, by

$$\begin{aligned} x_r(k) &= x(k) + a\cos(\varphi(k)) - b\sin(\varphi(k)) \\ y_r(k) &= y(k) + a\sin(\varphi(k)) + b\cos(\varphi(k)) \end{aligned}$$

Equation (37) defines the observation model $\mathbf{H}_i(\cdot)$ for a specific landmark.

3) Estimation Equations: The theoretical developments in this paper employed only linear models of vehicle and landmark kinematics. This was necessary to develop the necessary proofs of convergence. However, the implementation described here requires the use of nonlinear models of vehicle and landmark kinematics $\mathbf{f}(\cdot)$ and nonlinear models of landmark observation $\mathbf{h}(\cdot)$.

Practically an extended Kalman filter (EKF) rather than a simple linear Kalman filter is employed to generate estimates. The EKF uses linearized kinematic and observation equations for generating state predictions. The use of the EKF in vehicle navigation and the necessary assumptions needed for successful operation is well known (see, for example, the development in [20]), and is thus not developed further here.

4) Map Initialization and Management: In any SLAM algorithm the number and location of landmarks is not known *a priori*. Landmark locations must be initialized and inferred from observations alone. The radar receives reflections from many objects present in the environment but only the observations resulting from reflections from stationary point landmarks should be used in the estimation process. Fig. 5 shows a typical test run and the locations that correspond to all the reflections received by the radar. Only about 30% of the radar observations corresponded to identifiable point landmarks in the environment. In addition to these, a large freight vehicle and a number of nearby buildings also reflect the radar beam and produced range and bearing observations. This data set illustrates the importance of correct landmark identification, initialization and subsequent data association. In this implementation a simple measure of landmark quality is employed to initialize and track potential landmarks. Landmark quality implicitly tests whether the landmark behaves as a stationary point landmark. Range and bearing measurements which exhibit this behavior are assigned a high quality measure and are incorporated as a landmark. Those that do not are rejected.

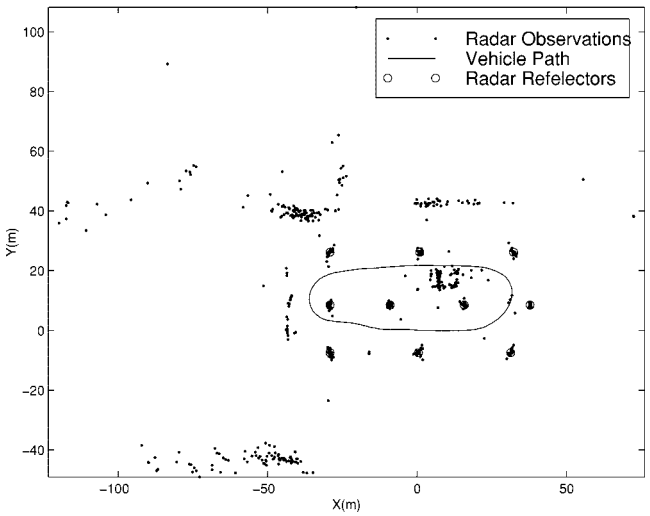

Fig. 5. A typical test run. Only 30% of radar observations correspond to identifiable landmarks.

The landmark quality algorithm is described in detail in Appendix 2. The algorithm uses two landmark lists to record "tentative" and "confirmed" targets. A tentative landmark is initialized on receipt of a range and bearing measurement. A tentative target is promoted to a confirmed landmark when a sufficiently high quality measure is obtained. Once confirmed, the landmark is inserted into the augmented state vector to be estimated as part of the SLAM algorithm. The landmark state location and covariance is initialized from observation data obtained when the landmark is promoted to confirmed status. Fig. 6 shows the computed landmark quality obtained at the end of the test run described in this paper.

B. Experimental Results

The vehicle starts at the origin, remaining stationary for approximately 30 s and then executing a series of loops at speeds up to 10 m/s. Fig. 7 shows the overall results of the SLAM algorithm. The estimated landmark locations are designated by stars. The actual (true) surveyed landmark locations are designated by circles. The vehicle path is shown with a solid line. In

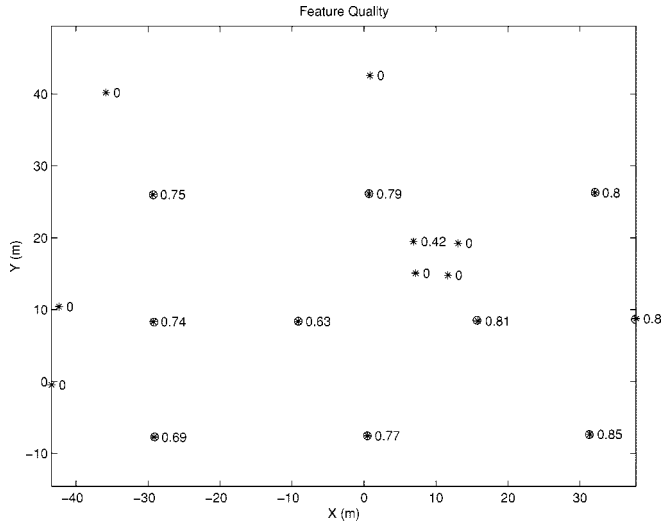

Fig. 6. Calculated landmark qualities at the end of the test run. Landmark quality ranges from 1 (for highest quality) to 0 (for lowest quality). Landmarks are marked with a star. Landmarks which are also circled are the artificial landmarks whose locations were surveyed so that the accuracy of the map building algorithm can be evaluated.

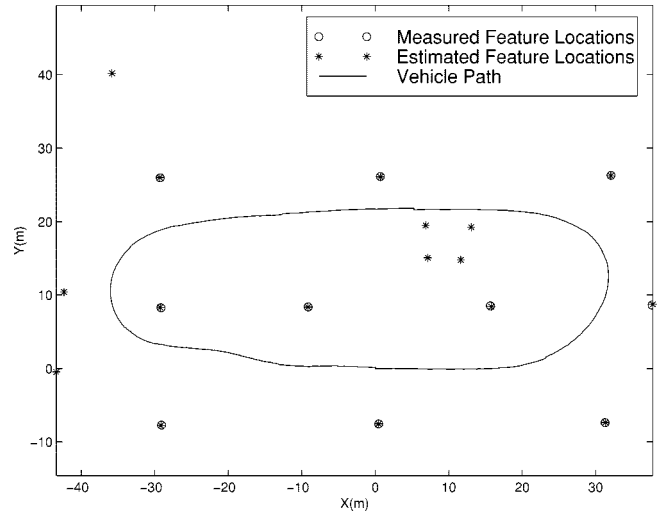

Fig. 7. The "true" vehicle path together with surveyed (circles) and estimated (starred) landmark locations.

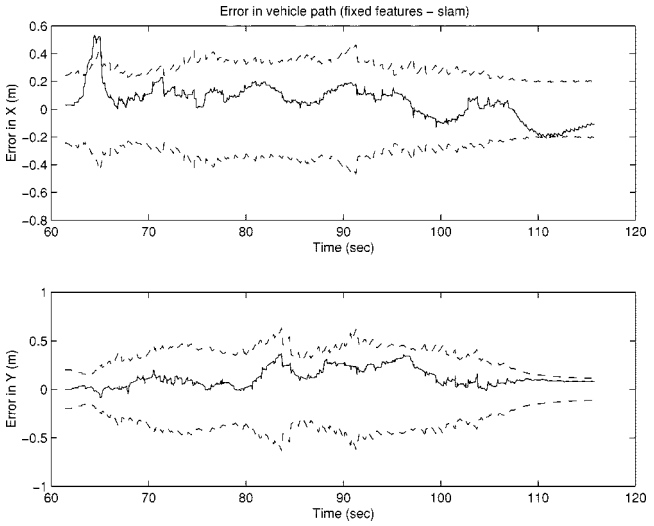

Fig. 8. Actual error in vehicle location estimate in x and y (solid line), together with 95% confidence bounds (dotted lines) derived from estimated location errors. The reduction in the estimated location errors around 110 s is due to the vehicle slowing down and stopping.

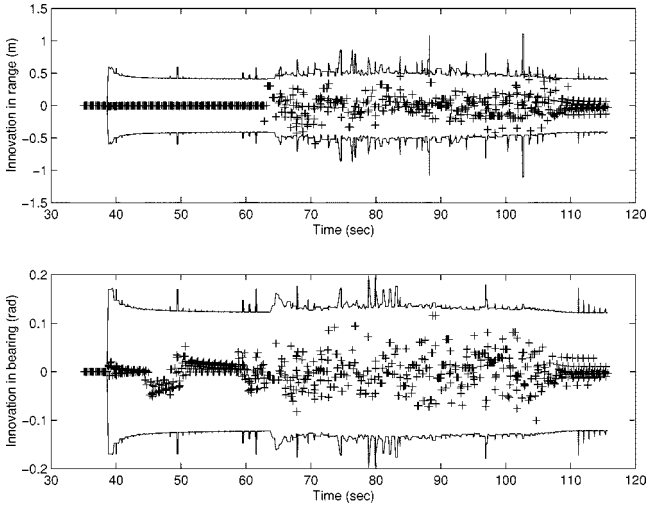

Fig. 9. Range and bearing innovations together with associated 95% confidence bounds.

order to compute the "true" vehicle path, surveyed locations of the artificial point landmarks were used for running a map based estimation algorithm [20]. The absolute accuracy of the vehicle path obtained using this algorithm was computed to be approximately 5 cm.

1) Vehicle Localization Results: The differences between the "true" vehicle path shown in Fig. 7 and the path estimated from the SLAM algorithm are too small to be seen on the scale used in Fig. 7. Fig. 8 shows the error between true and SLAM estimated position in more detail. The figure shows the actual error in estimated vehicle location in both x and y as a function of time (the central solid line). The figure also shows 95% (2σ) confidence limits in the estimate error. These confidence bounds are derived from the state estimate covariance matrix and represent the *estimated* vehicle error.

The actual vehicle error is clearly bounded by the confidence limits of estimated vehicle error. The estimate produced by the SLAM algorithm is thus consistent (and indeed is conservative). The estimated vehicle error as defined by the confidence bounds does not diverge so the estimates produced are stable. The jump in x error near the start of the run is caused by the vehicle accelerating when the model implicitly assumes a constant velocity model. The slight oscillation in errors and estimated errors are due to the vehicle cornering and thus coupling long-travel with lateral errors. Selecting a suitable model for the vehicle motion requires giving consideration to the tradeoff between the filter accuracy and the model complexity. It is possible to improve the results obtained by assuming a constant acceleration model, and relaxing the nonholonomic constraint used in the derivation of the vehicle kinematic (35) by incorporating vehicle slip. This clearly increases the computational complexity of the algorithm.

The SLAM algorithm thus generates vehicle location estimates which are consistent, stable and have bounded errors.

2) Map Building Results: Fig. 9 shows the innovations in range and bearing observations together with the estimated

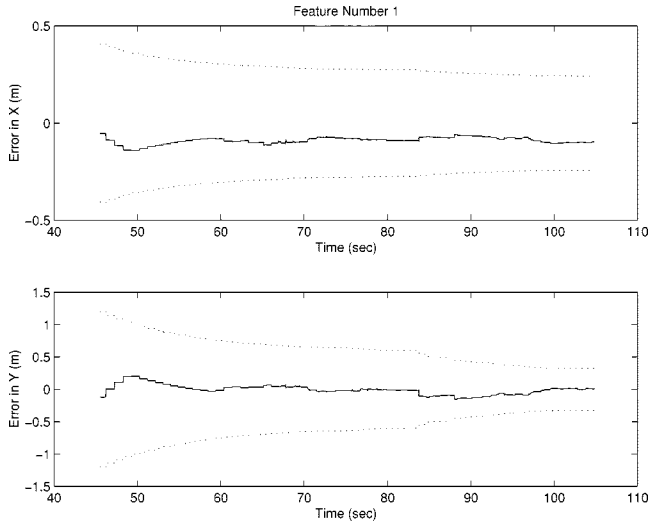

Fig. 10. Difference between the actual and estimated location for landmark 1. The 95% confidence limit of the difference is shown by dotted lines.

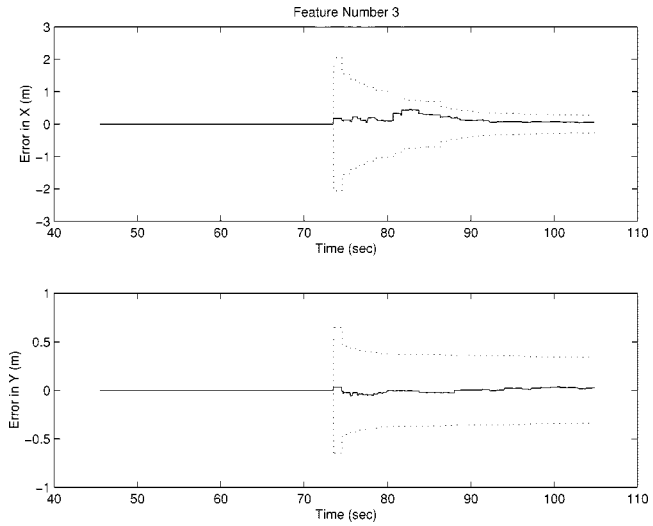

Fig. 11. Difference between the actual and estimated location for landmark 3. The 95% confidence limit of the difference is shown by dotted lines.

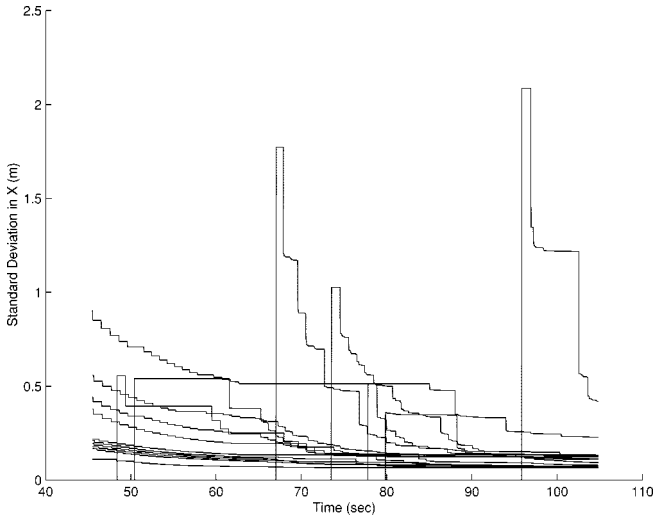

Fig. 12. Decreasing uncertainty in landmark location estimates.

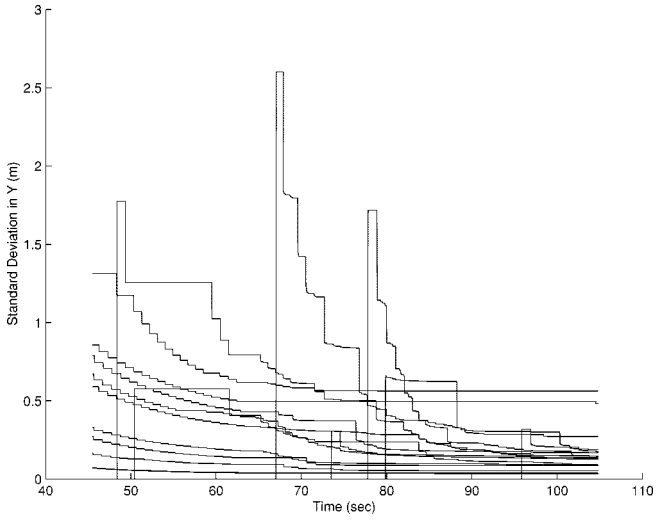

Fig. 13. Standard deviation of the landmark location estimate for all detected landmarks. As predicted, the uncertainty in landmark location estimates decrease monotonically.

95% confidence limits. The innovations are the only available measure for analysing on-line filter behavior when true vehicle states are unavailable. The innovations here indicate that the filter and models employed are consistent.

In addition to the 10 radar reflectors placed in the environment, the map building algorithm recognizes a further seven natural landmarks as suitable point landmarks. These can be seen in Fig. 7 as stars (confirmed landmarks) without circles (without a survey location). Some of these natural landmarks correspond to the legs of a large cargo moving vehicle parked in the test site. The others do not correspond to any obvious identifiable point landmarks, but were recognized as landmarks simply because they returned consistent point-like radar reflections. The landmark qualities Q calculated at the end of the test run are shown in Fig. 6.

Figs. 10 and 11 show the error between the actual and estimated landmark locations for two of the radar reflectors. One of these landmarks (landmark 1) was observed from the initial vehicle location at the start of the test run. The second landmark (landmark 3) was first observed about 30 s later into the run. Figs. 10 and 11 also show the associated 95% confidence limits in the location estimates. As before, these are calculated using the *estimated* landmark location covariances. The landmark location estimates are thus consistent (and conservative) with actual landmark location errors being smaller than the estimated error.

It can be seen that the initial variance of landmark 3 is much greater than that of landmark 1. This is due to the fact that the uncertainty in the vehicle location is small when landmark 1 is initialized, whereas landmark 3 is initialized while the vehicle is in motion and uncertainty in vehicle location is relatively high. These figures also show that there is some bias in the landmark location estimates. However, this bias is well within the accuracy of the true measurement (estimated to be ± 0.1 m).

Figs. 12 and 13 show the estimated standard deviations in x and y of all landmark location estimates produced by the filter (for graphical purposes, the variances are set to zero until the landmark is confirmed). As predicted by theory, the estimated

errors in landmark location decrease monotonically; and thus the overall error in the map reduces at each observation. Visually, the errors in landmark location estimates reach a common lower bound. As predicted by theory, this lower bound corresponds to the initial uncertainty in vehicle location.

V. Discussion and Conclusions

The main contribution of this paper is to demonstrate the existence of a nondivergent estimation theoretic solution to the SLAM problem and to elucidate upon the general structure of SLAM navigation algorithms. These contributions are founded on the three theoretical results proved in Section II; that uncertainty in the relative map estimates reduces monotonically, that these uncertainties converge to zero, and that the uncertainty in vehicle and absolute map locations achieves a lower bound.

Propagation of the full map covariance matrix is essential to the solution of the SLAM problem. It is the cross-correlations in this map covariance matrix which maintain knowledge of the relative relationships between landmark location estimates and which in turn underpin the exhibited convergence properties. Omission of these cross correlations destroys the whole structure of the SLAM problem and results in inconsistent and divergent solutions to the map building problem.

However, the use of the full map covariance matrix at each step in the map building problem causes substantial computational problems. As the number of landmarks N increases, the computation required at each step increases as N^2, and required map storage increases as N^2. As the range over which it is desired to operate a SLAM algorithm increases (and thus the number of landmarks increases), it will become essential to develop a computationally tractable version of the SLAM map building algorithm which maintains the properties of being consistent and nondivergent. There are currently two approaches to this problem; the first uses bounded approximations to the estimation of correlations between landmarks, the second method exploits the structure of the SLAM problem to transform the map building process into a computationally simpler estimation problem.

Bounded approximation methods use algorithms which make worst-case assumptions about correlatedness between two estimates. These include the covariance intersect method [21], and the bounded region method [22]. These algorithms result in SLAM methods which have constant time update rules (independent of the number of landmarks in the map), and which are statistically consistent. However, the conservative nature of these algorithms means that observation information is not fully exploited and consequently convergence rates for the SLAM method are often impracticably slow (and in some cases divergent).

Transformation methods attempt to re-frame the map building problem in terms of alternate map or landmark representations which particularly have relative independence properties. For example, it makes sense that landmarks which are distant from each other should have estimates that are relatively independent and so do not need to be considered in the same estimation problem. One example of a transformation method is the relative filter [23], [24] which directly estimates the relative, rather than absolute, location of landmarks. The relative landmark location errors may be considered independent thus resulting in a map building algorithm which has constant time complexity regardless of the number of landmarks. More generally, a number of approaches are being developed for constructing "local maps" and embedding these in a map management process. Here, consistent (full filter) local maps are linked by conservative transformations between local maps to generate and maintain larger scale maps. This embodies the idea that local landmarks are more important to immediate navigation needs than distant landmarks, and that landmarks can naturally be grouped into localized sets. Such transformation methods can exploit relatively low degrees of correlation between landmark elements to generate relatively decoupled submaps. The advantage of these transformation methods is that they highlight the real issue of large-scale map management.

The theoretical results described in this paper are essential in developing and understanding these various approaches to map building. As discussed in the introduction there are a number of existing approaches to the localization and map building problem. The important contribution of this paper is the proof that a solution exists. Furthermore, it presents an algorithm that is efficient in the sense that it makes optimal use of the observations of relative location of landmarks for estimating landmark and vehicle locations, a property inherent in the Kalman filter. The value of all other alternative real-time SLAM algorithms that use similar information can be evaluated with respect to this "full" solution. This is particularly true in the case where simplifications are made to SLAM algorithms in order to increase the computational efficiency.

The most effective algorithm for SLAM depends much on the operating environment. For example, the relative sparseness of occupied regions in the environment used for the example shown in Section IV, or the grid based approaches proposed in [11] and [10]. In an indoor environment the use of only point landmarks as in Section IV will be inefficient as the information such as the ranges to walls will not be utilized. The framework proposed in this paper, however, can incorporate geometric features such as lines (for example, see [25]). In environments where geometric features are difficult to detect, for example in an underground mine, the proposed strategy will not be feasible. Many navigation systems used in outdoor environments rely on exogenous systems such as GPS. Clearly if external information is available these can be incorporated to the framework proposed in this paper. This is particularly useful in the situations where the exogenous sensor is unreliable, for example GPS not being able to observe sufficient numbers of satellites or multi-path errors caused when operating in cluttered environments.

The implementation described in this paper is relatively small scale. It does, however, serve to illustrate a range of practical issues in landmark extraction, landmark initialization, data association, maintenance and validation of the SLAM algorithm. The implementation and deployment of a large-scale SLAM system, capable of vehicle localization and map building over large areas, will require further development of these practical issues as well as a solution to the map management problem. However, such a substantial deployment would represent a

major step forward in the development of autonomous vehicle systems.

APPENDIX I
PROPERTIES OF POSITIVE SEMIDEFINITE MATRICES

1) Diagonal entries of a *psd* matrix are nonnegative.
2) If $A \in R^{n\times n}$ is *psd* then for any matrix $B \in R^{n\times m}$, BAB^{T} is *psd*.
3) If $A \in R^{n\times n}$ and $C \in R^{n\times n}t$ are both *psd* then $\det(A+C) \geq \det(A)$.
4) All principal submatrices of *psd* matrices are *psd*.
5) For $A \in R^{m\times m}$, $B \in R^{n\times m}$ and $C \in R^{n\times n}$, if

$$D = \begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix}$$

is *psd*, then

$$\det(D) \leq \det(A)\det(C)$$

APPENDIX II
LANDMARK INITIALIZATION ALGORITHM

The following describes the procedure used to initialize the landmark locations and associate observations to particular landmarks. This procedure also evaluates the quality of the landmarks. This algorithm is an essential precursor to the estimation process. It is not specific to the radar sensor but can be generalized to any sensor capable of observing landmarks in the environment.

Two landmark lists are maintained. One list stores landmarks that are confirmed to be valid $\mathbf{p}_i$, $i = 1 \ldots N$ and the other stores potential landmarks yet to be validated $\mathbf{q}_i$, $i = 1 \ldots M$. Initially both lists are empty. The map management algorithm proceeds as follows:

1) Given an observation $[r_f, \theta_f]$ at time instant k from the radar, the location of the landmark possibly responsible for this observation $\mathbf{p}_f = [x_f, y_f]^{\mathrm{T}} = g(x, y, \varphi, r_f, \theta_f)$ and its covariance $\mathbf{P}_f$ is calculated using the following relationship.

$$\begin{bmatrix} x_f \\ y_f \end{bmatrix} = \begin{bmatrix} x(k) \\ y(k) \end{bmatrix} + \begin{bmatrix} x_{vf}(k) \\ y_{vf}(k) \end{bmatrix}$$

where

$$\begin{bmatrix} x_{vf}(k) \\ y_{vf}(k) \end{bmatrix} = \begin{bmatrix} a\cos(\varphi(k)) - b\sin(\varphi(k)) + r_f\cos(\varphi(k)+\theta_f) \\ a\sin(\varphi(k)) + b\cos(\varphi(k)) + r_f\sin(\varphi(k)+\theta_f) \end{bmatrix}$$

and

$$\mathbf{P}_f = \nabla g_{xy\varphi}\mathbf{P}_v\nabla g_{xy\varphi}^{\mathrm{T}} + \nabla g_{r_f\theta_f}\mathbf{R}\nabla g_{r_f\theta_f}^{\mathrm{T}}$$

where $\mathbf{P}_v$ is the covariance matrix of the vehicle location estimate extracted from the state covariance matrix $\mathbf{P}(k|k)$ and $\mathbf{R}$ is the measurement noise covariance.

2) An observation is associated with a landmark $\mathbf{p}_i$ in the confirmed landmark set if

$$d_{fi} = (\mathbf{p}_f - \mathbf{p}_i)^{\mathrm{T}}(\mathbf{P}_f + \mathbf{P}_i)^{-1}(\mathbf{p}_f - \mathbf{p}_i) < d_{\min}$$

where $\mathbf{P}_i$ is the covariance of the landmark location estimate $\mathbf{p}_i$, extracted from the state covariance matrix $\mathbf{P}(k|k)$. Note that d_{fi} is the Mahalanobis distance between $\mathbf{p}_f$ and $\mathbf{p}_i$, and its probability distribution in this case is that of a χ^2 variable with two degrees of freedom. Therefore, a suitable value for $d_{\min}$ can be selected such that the null hypothesis that $\mathbf{p}_f$ and $\mathbf{p}_i$ are the same is not rejected at some desired confidence level. Also if the above condition results in an observation being associated with more than one landmark $\mathbf{p}$, the observation is rejected. If accepted, the observation is then used to generate a new state estimate.
3) If an observation cannot be associated with any confirmed landmark, then it is checked against the set of potential landmarks for possible association. Mahalanobis distance is again used as the criterion for association. If an association with the potential landmark $\mathbf{q}_j$ is justified the new observation is used to update the location of the potential landmark $\mathbf{q}_j$ and its covariance $\mathbf{P}_j$. In addition, a counter c_j indicating the number of associations with landmark $\mathbf{q}_j$ is also incremented.
4) An observation that is not associated with either a confirmed or a potential landmark can be considered as a new landmark. In this case $\mathbf{p}_f$ is added to the list of potential landmarks as $\mathbf{q}_{M+1}$, a counter c_{M+1} is initialized and the number of the time step k is assigned to a timer t_{M+1}.
5) The potential landmark list $\mathbf{q}_i$, $i = 1 \ldots M$ is then examined against the following criteria.
   (a) If c_j is greater than a predetermined number of associations $c_{\min}$, the landmark j is considered to be sufficiently stable and therefore is transferred to the confirmed landmark list.
   (b) If $(k - t_j)$ is greater than a predetermined $t_{\max}$ then the landmark j has not achieved the desired minimum number of associations over a sufficient length of time. Landmark j is therefore removed from the potential landmark list.
6) The probability density function (PDF) of the observations associated to a given landmark can be used to estimate its "quality". As suggested in [26] the quality Q_i of landmark i is calculated using the following equation:

$$Q_i = \frac{\sum_{j=1}^{l} \frac{1}{2\pi}|\mathbf{S}_j|^{-1/2}\exp(\frac{\nu_j\mathbf{S}_j^{-1}\nu_j^{\mathrm{T}}}{2})}{\sum_{j=1}^{l} \frac{1}{2\pi}|\mathbf{S}_j|^{-1/2}} \tag{38}$$

where l is the number of observations so far associated with the landmark i, where v_j is the innovation of the observation j associated with landmark i observed at time k. $\mathbf{S}_j$ is the innovation covariance as defined by (14). The landmark quality Q_i is the ratio between the sum of the probability densities of the observations and the maximum value of the probability density that is achieved when all the observations coincide with their predicted

values. Therefore landmark qualities Q lie between 0 and 1. At reasonable intervals, landmark qualities can be calculated and landmarks that do not achieve a predetermined $Q_{\min}$ can be deleted from the map.

7) Return to step 2 when the next observation is received.

REFERENCES

[1] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*, I.J. Cox and G.T. Wilfon, Eds, New York: Springer Verlag, 1990, pp. 167–193.

[2] P. Moutarlier and R. Chatlia, "Stochastic multisensor data fusion for mobile robot localization and environment modeling," in *Int. Symp. Robot. Res.*, 1989, pp. 85–94.

[3] J. Leonard and H. Durrant-Whyte, *Directed Sonar Sensing for Mobile Robot Navigation*. Boston, MA: Kluwer Academic, 1992.

[4] R. A. Brooks, "Robust layered control system for a mobile robot," *IEEE J. Robot. Automat.*, vol. RA-2n, pp. 14–23, 1986.

[5] B. J. Kuipers and Y-T Byun, "A robot exploration and mapping strategy based on semantic hierachy of spacial representation," *J. Robot. Autonom. Syst.*, vol. 8, pp. 47–63, 1991.

[6] T.S. Levitt and D.T. Lawton, "Qualitative navigation for mobile robots," *Artif. Intell. J.*, vol. 44, no. 3, pp. 305–360, 1990.

[7] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Computer Vision*, vol. 13, no. 2, pp. 119–152, 1994.

[8] F. Cozman and E. Krotkov, "Automatic mountain detection and pose estimation for teleoperation of lunar rovers," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1997.

[9] U. D. Hanebeck and G. Schmidt, "Set theoretical localization of fast mobile robots using an angle measurement technique," in *Proc. IEEE Int. Conf. Robot. Automat.*, Minneapolis, MN, Apr. 1996, pp. 1387–1394.

[10] S. Thrun, D. Fox, and W. Burgard, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Mach. Learning Autonom. Robots*, vol. 31, pp. 29–53, 1998.

[11] B. Yamauchi, A. Schultz, and W. Adams, "Mobile robot exploration and map building with continuous localization," in *Proc. IEEE Int. Conf. Robot. Automat.*, Leuven, Belgium, May 1998, pp. 3715–3720.

[12] R. Cheeseman and P. Smith, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.*, 1986.

[13] H. F. Durrant-Whyte, "Uncertain geometry in robotics," *IEEE Trans. Robot. Automat.*, vol. 4, pp. 23–31, 1988.

[14] N. Ayache and O. Faugeras, "Maintaining a representation of the environment of a mobile robot," *IEEE Trans. Robot. Automat.*, vol. 5, 1989.

[15] R. Chatila and J-P Laumond, "Position referencing and consistant world modeling for mobile robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1985.

[16] W. D. Renken, "Concurrent localization and map building for mobile robots using ultrasonic sensors," in *Proc. IEEE Int. Conf. Itell. Robots Syst.*, 1993.

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

[18] P. Maybeck, *Stochastic Models, Estimation and Control*, New York: Academic, 1982, vol. 1.

[19] S. Clark and H. F. Durrant-Whyte, "Autonomous land vehicle navigation using millimeter wave radar," in *Int. Conf. Robot. Automat.*, 1998.

[20] H. Durrant-Whyte, "An autonomous guided vehicle for cargo handling applications," *Int. J. Robot. Res.*, vol. 15, no. 5, pp. 407–440, 1996.

[21] J. K. Uhlmann, S. J. Julier, and M. Csorba, "Nondivergent simultaneous map building and localization using covariance intersection," in *SPIE*, Orlando, FL, Apr. 1997, pp. 2–11.

[22] M. Csorba, "Simultaneous localization and map building," Ph.D. dissertation, Univ. Oxford, Robot. Res. Group, 1997.

[23] M. Csorba and H. F. Durrant-Whyte, "New approach to map building using relative position estimates," in *SPIE*, vol. 3087, Orlando, FL, Apr. 1997, pp. 115–125.

[24] P. M. Newman, "On the solution to the simultaneous localization and map building problem," Ph.D. dissertation, Dept. Mechan. Eng., Australian Centre for Field Robotics, Univ. Sydney, Sydney, Australia, 1999.

[25] J. A Castellanos, J.M.M. Montiel, J. Neira, and J.D. Tardos, "Sensor influence in the performance of simultaneous mobile robot localization and map building," in *Proc. 6th Int. Symp. Experiment. Robot.*, Sydney, Australia, Mar. 1999, pp. 203–212.

[26] D. Maksarov and H.F. Durrant-Whyte, "Mobile vehicle navigation in unknown environments: a multiple hypothesis approach," *IEE Proc. Contr. Theory Applicat.*, vol. 142, July 1995.

M. W. M. Gamini Dissanayake received the degree in mechanical/production engineering from the University of Peradeniya, Sri Lanka. He received the M.Sc. degree in machine tool technology and Ph.D. degree in mechanical engineering (robotics) from the University of Birmingham, England, in 1981 and 1985, respectively.

He was a lecturer at the University of Peradeniya and the National University of Singapore before joining the University of Sydney, Australia, where he is currently a senior lecturer. His research interests include localization and map building for mobile robots, navigation systems, dynamics and control of mechanical systems, cargo handling, optimization, and path planning.

Paul Newman (M'97) received the Masters degree in engineering science from Oxford University, England, in 1995 and the Ph.D. degree from The University of Sydney, Australia, in 1999. His thesis examined the structure and solution of the SLAM problem.

Before joining the Massachusetts Institute of Technology, Cambridge, he worked in the industrial subsea navigation domain. He continues to consult in navigation systems architecture and algorithms. Currently, he is a Postdoctoral Associate with the Department of Ocean Engineering at MIT. His research activities include multirobot SLAM and autonomous subsea navigation.

Steven Clark received the Ph.D. degree from Sydney University, Sydney, Australia, in 1999. His dissertation discusses the practical application of millimeter wave radar for field robotics. He is now providing radar equipment and signal processing expertise commercially through Navtech Electronics, Ltd. in the U.K.

Hugh F. Durrant-Whyte (S'84–M'86) received the B.Sc. (Eng.) degree (1st class honors) in mechanical and nuclear engineering from the University of London, England, in 1983, and the M.S.E. and Ph.D. degrees, both in systems engineering, from the University of Pennsylvania, Philadelphia, in 1985 and 1986, respectively.

From 1987 to 1995, he was a Senior Lecturer in Engineering Science with the University of Oxford, England, and a Fellow of Oriel College Oxford. Since July 1995 he has been Professor of Mechatronic Engineering with the Department of Mechanical and Mechatronic Engineering, The University of Sydney, Australia, where he leads the Australian Centre for Field Robotics, a Commonwealth Key Centre of Teaching and Research. His research work focuses on automation in cargo handling, surface and underground mining, defense, unmanned flight vehicles, and autonomous subsea vehicles.

M. Csorba, photograph and biography not available at the time of publication.