

On The Problem of Simultaneous Localization, Map Building, and Servoing of Autonomous Vehicles

A. Bicchi, F. Lorussi, P. Murrieri, and V. Scordio

Centro Interdipartimentale di Ricerca “Enrico Piaggio”.
E-mail: bicchi@ing.unipi.it, lorussi@piaggio.cci.unipi.it,
p.murrieri@ing.unipi.it, scordio@piaggio.cci.unipi.it .
URL: www.piaggio.cci.unipi.it

Abstract. In this chapter, we consider three of the main problems that arise in the navigation of autonomous vehicles in partially or totally unknown environments, i.e. building a map of the environment, self-localizing, and servoing the robot so as to achieve given goals based on sensorial information. As compared to most part of the existing literature on SLAM, we privilege here a system-theoretic view to the problem, which allows the localization and mapping problems to be cast in a unified framework with the control problem. The chapter is an overview of existing results in this vein, and of some interesting directions for research in the field.

1 Introduction

Autonomous vehicles have a wide range of applications, both in indoor and outdoor environments, and represent one of the areas with largest potential for advanced robotics. A very important trend in research related to mobile robots is concerned with their sensorization, and in particular with the tradeoffs between effectiveness and cost of different possible sensorial equipments.

Three of the main technical difficulties in applying mobile robots to partially or totally unstructured environments are indeed sensor-related: the localization of the vehicle with respect to the environment, the construction of a map of the environment itself, and the control of the vehicle to desired postures relative to the environment. Naturally, the three problems are closely interconnected. While the acronym SLAM (for Simultaneous Localization And Map building) has been gaining wide acceptance in the robotics literature ([5,29,41]) to indicate the composition of the first two aspects, the connection to control is less frequently addressed. Indeed, in the SLAM literature, vehicles are often commanded in open loop. On the other hand, in the rather extensive literature on control of autonomous robots, localization is often simply taken for granted. Such is the case e.g. in many papers dealing with set-point stabilization of wheeled vehicles, which assume full state information (viz. [6,7,13,45,34,8,12]). In practical applications of automated vehicle control, however, one is confronted with the problem of estimating the current position and orientation of the vehicle only through indirect, noisy measurements by available sensors. Although much work has been done on techniques for vehicle localization based on combinations of sensory information (odometry, laser range finders, cameras, etc.),

very little is known about the real time connection of a localization algorithm and a feedback control law.

In this chapter, we consider the problem of simultaneous localization, mapping, and servoing (SLAMS) from a unified system-theoretic viewpoint, and report on work towards integrating solutions allowing an autonomous vehicle to navigate in an unknown environment. The chapter is organized as follows: in section 2 we formulate the problem under consideration, and in section 3 we provide a brief survey of the state of the art. In section 4 we discuss aspects related to the existence of solutions to the SLAM problem, and to the choice of optimal exploratory paths to elicit SLAM information. In section 5 we report on the problem of simultaneous localization and servoing, before concluding in section 6.

2 Modeling of the SLAMS problem

Let us consider a system comprised of a vehicle moving in an environment with the aim of localizing itself and the environment features. For simplicity, we assume that features are distinctive 3D points in the environment where the vehicle moves (more general features are described e.g. in [39]). The vehicle is endowed with sensors, such as a radial laser rangefinder or video cameras. Both the vehicle initial position and orientation, and the feature positions, are unknown or, more generally, known up to some a priori probability distribution. A particular pose of the vehicle, or set of poses, is regarded as the goal. Sensor readings corresponding to the goal pose are known (by e.g. recording them in a preliminary learning phase). Among the features that the sensor head detects in the robot environment, we will distinguish between those belonging to objects with unknown positions (which we shall call *targets*), and those belonging to objects whose absolute position is known (which will be referred to as *markers*). Indeed, as it can be argued, this distinction is only useful for simplicity of description, as in general the case is that there exist features that are more or less uncertain.

The vehicle dynamics are supposed to be slow enough to be neglected (dynamics do not add much to the problem structure, while increasing formal complexity). Kinematics of wheeled vehicles can usually be written as a nonlinear system of the type $\dot{x} = G(x)u$, where $x \in \mathbb{R}^{n_v}$ is the robot pose (typically, $n_v = 3$ for a vehicle moving in a plane with an orientation), and $u \in \mathbb{R}^m$ are the input velocities. It is often the case where the system velocities are affected by disturbances μ (such as e.g. slippage of the wheels), and the model is accordingly modified to include process noise as $\dot{x} = G(x)(u + \mu)$.

Let the i -th target absolute coordinates be denoted by $p_i \in \mathbb{R}^d$, with $d = 2$ for planar features and $d = 3$ in case of 3D environments, and use $p \in \mathbb{R}^{dn_f}$ to denote the collection of all features. According to the sensor equipment specifics, the relative position of the vehicle and of the features form sensor readings, or *observables*, described by the map $h : \mathbb{R}^{n_v} \times \mathbb{R}^{dn_f} \rightarrow \mathbb{R}^q$, $(x, p) \mapsto y = h(x, p)$. Measurement noise ν adds to this as $y = h(x, p) + \nu$.

In system-theoretic terms, the three problems in SLAMS can be described by referring to the input-state-output system

$$\begin{aligned} \begin{bmatrix} \dot{x} \\ \dot{p} \end{bmatrix} &= f(x, p, u, \mu) = \begin{bmatrix} G(x) \\ 0 \end{bmatrix} (u + \mu) \\ y &= h(x, p) + \nu. \end{aligned} \quad (1)$$

In this framework, localization and mapping are *observability* problems, dealing with the reconstruction of the present pose x and feature map p , respectively, from current and past observables, from model and input knowledge, and from statistics on process noise μ and measurement noise ν . Servoing is a *stabilization* problem, aiming at devising what inputs u are to be given to the system so as to reach the desired pose, based on available data. Should the current pose x be known exactly at all times, servoing would amount to find a state feedback law in the form $u(x, t)$, such that $\dot{x} = G(x)u(x, t)$ asymptotically converged to the desired pose. However, such knowledge is not available in general, because typically $q < n_v + dn_f$ and, even when this inequality would be reversed (such as e.g. when using absolute landmarks and a trinocular stereo camera head), because of measurement noise. Servoing in SLAMS should therefore be regarded in general as an *output stabilization* problem, whereby a new dynamic system must be designed in the additional states $w \in \mathbb{R}^r$ as

$$\begin{aligned} \dot{w} &= S(w, y) \\ u &= F(w, y) \end{aligned} \quad (2)$$

such that, when connected to system (1), asymptotic stability of the compound $n_v + dn_f + r$ states can be achieved. It is often (but not always) the case that the auxiliary system (2) includes an *estimator* of the system (1), i.e. its design is aimed at achieving the convergence of $w(t)$ to the pose $x(t)$ (the prevailing design for the estimator is based on Extended Kalman Filters, see below). According to this approach, a design is often attempted for the control in the form of a state-feedback stabilizer $u(w, t)$, where w is used in place of x . Naturally, convergence of the estimator and of the state-feedback law separately are only necessary conditions in order for their composition to provide a stable and satisfactory behavior.

The model in (1) is sometimes referred to as *world-centric*. It is rather obvious that, unless geographic markers or other equivalent information (from compass, GPS, etc.) are present, reconstruction of absolute robot position and orientation is impossible. A different description of the same problem can hence be given in coordinates relative to the vehicle (a *robot-centric* model), which would be written in the form

$$\begin{aligned} {}^v\dot{p} &= \ell({}^v p, u, \mu) \\ y &= \hat{h}({}^v p) + \nu, \end{aligned} \quad (3)$$

Such a model is applicable for instance to the case where a camera is mounted on the vehicle, with the output map $\hat{h}(\cdot)$ representing the projection of 3D features to the image plane of the camera. Output feedback control of (3) amounts then to what is commonly referred to as image-based visual servoing of the vehicle. In this case, explicit estimation of the robot pose is clearly unnecessary.

In the rest of this chapter, we will discuss these different aspects of the SLAMS problem in more detail, emphasizing the insight that an integrated system theoretic approach can bring to the field.

3 Approaches to the SLAM Problem

As shown in the former section, the Simultaneous Localization and Map Building problem — also known in the literature as CML (Concurrent Mapping and Localization)— is characterized by two sources of uncertainty: the vehicle model (because of both uncertain parameters appearing in the dynamics and process noise) and sensor noise.

Uncertainty can be dealt with in basically two ways, i.e. deterministically or by using probabilistic models. The first approach assumes that all uncertainty sources may generate errors that are unknown but bounded, and seeks for bounds on how these error can propagate through the reconstruction process. Naturally, the problem tends to be overly complex from the computational and memory-occupation viewpoints, hence efficient algorithms to approximate the worst-case bounds are in order. An application of this approach to robot localization is reported in [18], where an efficient, recursive algorithm to approximate the set of robot poses compatible with present and past measurements is presented.

Deterministic algorithms tend to suffer from excessive conservativeness, and are typically not very suited to take into account the existence of large, sporadic errors in sensor readings (*outliers*), which are common in some types of sensors used in SLAMS (e.g. spurious reflections of lasers or sonars, feature mismatch, etc.). When an excess of conservatism is not justified by particularly risk-sensitive applications, it is often preferred to adopt probabilistic models of uncertainty.

The basis for virtually all probabilistic methods is Bayesian theory of inference, which assumes that the statistical properties of the *data space* and of the *model space* are well defined. These are the vector spaces, of suitable dimension, where observables y and unknowns (and estimates thereof, denoted for brevity as x) take their values, and where a probability density function (p.d.f.) is defined for the variables of interest. The *a priori* state of information consists in a p.d.f. defined over the model space X , $f_{prior}(x)$, which models any knowledge one may have on the system model parameters independently from the present act of measurement, due e.g. to physical insight or to independent measurements carried out previously.

In the formation of estimates, two information sources are to be considered, i.e. the forward solution of the physical model, and the act of measuring itself. The state of information on the experimental uncertainties in measurement outputs can be modelled by means of a p.d.f. $f_{exp}(y)$ over the data space Y (this should be provided by the instrument supplier), while modelling errors (due to imperfection of (1), or to process noise) can be represented by a conditional p.d.f. $f_{mod}(y|x)$ in the data space Y (or, more generally, by a joint p.d.f. $f_{mod}(y, x)$ over $X \times Y$).

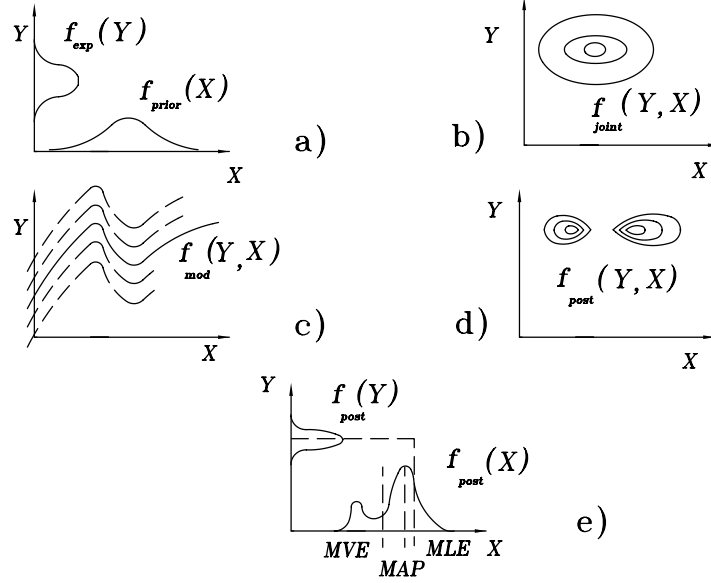


Fig. 1. The process of Bayesian inference. (a) A priori information on the model space, $f_{prior}(x)$, and information on experimental data $f_{exp}(y)$ are independent, and combine in the joint p.d.f. $f_{joint}(y, x)$, (b). Information on modelling is represented by $f_{mod}(y, x)$ (c). The conjunction of $f_{joint}(y, x)$ and $f_{mod}(y, x)$ is $f_{post}(y, x)$ (d). The marginal p.d.f.'s $f_{post}(x)$ and $f_{post}(y)$ (e) can be obtained directly from $f_{post}(y, x)$. Different estimators can be applied to these results, as illustrated in (e).

Fusing the different information in an estimate of x leads to a *posterior* p.d.f over X , that is described by Bayes' formula

$$f_{post}(x) = f(x|y) = \alpha_b f_{prior}(x) \int_Y f_{exp}(y) f_{mod}(y|x) dy, \quad (4)$$

where α_b is a normalization factor such that $\int_X f_{post}(x) dx = 1$. The process of information fusion is described in figure 1 (adapted from [44]), with reference to the case where the measurement equation forming observables y from unknowns x is nonlinear (such as it actually is in SLAM). Although the posterior p.d.f. on the model space represents the most complete description of the state of information on the quantity to be measured one may wish, a final decision on what is the "best" estimate of x needs usually be taken. Several possibilities arise in general, such as the *maximum a posteriori* estimate (MAP), *maximum likelihood* estimates (MLE, which coincides with MAP if no priors are available), the *minimum variance* estimate (MVE) alias *minimum mean square* (MMSE). Figure 1-e illustrates these estimates. While very little can be said in general about the performance of such estimators, well known particularizations apply under certain assumptions on the prior distributions. Thus, if a normal distribution (an order-2 Gaussian) can be assumed for all prior information, the MAP estimate enjoys many useful properties:

first (and perhaps most importantly for the problem at hand), since the convolution in (4) of two Gaussian distributions is Gaussian, the modelling and experimental errors in measurements simply combine by addition of the covariance matrices of experimental and modelling errors, $C_Y = C_{exp} + C_{mod}$. Roughly speaking, errors in the model knowledge (kinematic model of the systems and odometry errors) can be ignored, provided that experimental measurement errors in y are suitably increased. This result holds for nonlinear sensor models as well. For linearized measurement models ($y = Hx$), the a posteriori p.d.f. would also be Gaussian, the MVE and MAP estimates would coincide and evaluate to

$$\begin{aligned}\hat{x} &= C_{post}(H^T C_Y^{-1} y + C_{prior}^{-1} x_{prior}), \\ C_{post} &= (\mathcal{F} + C_{prior}^{-1})^{-1},\end{aligned}\tag{5}$$

where \mathcal{F} , the *Fisher information matrix* for the linear case at hand, is defined as

$$\mathcal{F} = H^T C_Y^{-1} H.\tag{6}$$

As a final remark, the Gauss-Markov theorem [38] ensures that the estimate (5) is the best linear unbiased estimate (BLUE) in the minimum-variance sense even for non-Gaussian a priori distributions. This result may seem to indicate some “absolute optimality” of the least-squares estimate. However, the MVE of a non-Gaussian distribution may not be a significant estimate, as apparent in figure 1-e. This is the case for instance when a few measurements are grossly in error (*outliers*): the MVE in this case can provide meaningless results. This fact is sometimes used to point out the *lack of robustness* of the MVE.

In the literature on mobile robot localization and mapping, methods to evaluate an estimate of the posterior p.d.f. over the space of unknown robot poses and targets have been studied extensively. While for an exhaustive review the reader is referred to [41], we limit ourselves to point out that methods proposed so far can be roughly classified in two main groups: batch and recursive.

Batch methods attempt as accurate a solution of the posterior as possible, by taking into account that often in SLAM the posterior p.d.f. is a complex multimodal distribution. To such complexity contribute different factors, among which the non-linearity of dynamics and measurement equations (1), and the fact that measurement noise in different measurements is statistically correlated, because errors in control accumulate over time, and they affect how subsequent measurements are interpreted ([41]). A crucial aspect of SLAM is indeed that, when features are not distinctive, multiple correspondences are possible, a problem also known as *data association*. The correspondence problem, consisting in determining if sensor measurements taken at different times correspond to the same physical object in the world, is very hard to be tackled, since the number of possible hypotheses can grow exponentially over time. A family of methods recently introduced to deal with these problems, which is based on Dempster’s Expectation Maximization Algorithms (EM) [24,41], represent the current state-of-art in this regard. However, since EM have to process data multiple times they are not suitable to real-time implementation, as needed e.g. to interface with servoing algorithms.

On the other hand, most often new updates of model estimates are needed in real-time, without referring to the whole history of sensed data. To cope with this requirement, further simplifications are usually done: for instance, assuming a Gaussian posterior distribution, the given record of data can be completely described by the mean vector and the covariance matrix. When a new datum is available, all prior information can be extracted from those statistics. A method that does not use prior information explicitly, but through its statistics only, is called *recursive*. The Kalman filter is one such recursive method, implementing the optimal minimum variance observer for a linear system subject to uncorrelated, zero-mean, Gaussian white noise disturbances.

Unfortunately, these assumptions are unfulfilled in SLAM applications. Hence, different simplifying assumptions and approximations are employed. Filters resulting from repeated approximate linearization of (1) are commonly referred to Extended Kalman Filters (EKF). Although extended Kalman filters for the SLAM problem do not guarantee any optimality property, they remain the most widely used filters in SLAM. EKF maintain all information on the estimated posteriors in the vector of means and in a covariance matrix, whose update at each step is a costly operation (quadratic with the number of features). In practical implementations, a key limitation of EKF is the low number of features it can deal with.

Algorithms have been recently proposed to overcome this limitation. The Fast-Slam [33] algorithm is based on the assumption that the knowledge of the robot path renders measurements of individual markers independent, so that the problem of determining the position of K features could be decomposed into K estimation problems, one for each feature [33]. Compressed EKF (CEKF), see [20], stores and maintains all the information gathered in a local area with a cost proportional to the square of the number of landmarks in the area. This information can then be transferred to the rest of the global map with a cost that is similar to full SLAM, but in only one iteration. Sparse Extended Information Filter (SEIF), see [43], is an algorithm whose updates require constant time, independent of the number of features in the map. It exploits the particular form of the information matrix, i.e. the inverse of the covariance matrix. Since the information matrix is sparse, it possesses a large number of elements whose values, when normalized, are near zero and can be neglected in the updating process. Some algorithms, see [17,16,26], based on incremental update of uncertain maps, use a fuzzy logic approach to manage uncertainty on obstacle poses and successively implement obstacle avoidance strategies.

An interesting possibility in SLAM is the possibility of using multiple vehicles in a cooperative way in order to perform tasks more quickly and robustly than a single vehicle can do. In [15,42], the problem of performing concurrent mapping and localization with a team of cooperating autonomous vehicles is considered, and the advantages of such a multiagent cooperation are illustrated.

One of the most challenging topics in SLAM is the optimization of autonomous robotic exploration. Indeed, it is often the case that robots have degrees of freedom in the choice of the path to follow, which should be used to maximize the information that the system can gather on the environment. The problem is clearly of great

relevance to many tasks, such as e.g. surveillance or exploration. However, it is in general a difficult problem, as several quantities have to be traded off, such as the expected gain in map information, the time and energy it takes to gain this information, the possible loss of pose information along the way, and so on. This problem is considered in detail in the next section.

4 Solvability and Optimization of SLAM

As already mentioned, simultaneous localization and mapping amounts to estimating the state of system (1) through integration of input velocities (odometry) and knowledge of the observations y . Input velocities and observables are affected by process and measurement noise, respectively.

We start by observing that system (1) is nonlinear in an intrinsic way, in the sense that approximating the system with a linear time-invariant model destroys the very property of observability: this entails that elementary theory and results on linear estimation do not hold in this case.

The intrinsic nonlinear nature of the problem can be illustrated directly by the simple example in Fig.(2) of a planar vehicle ($n_v = 3$) with M markers and N targets (hence $dn_f = 2N$). Outputs in this examples would be the $q = M + N$

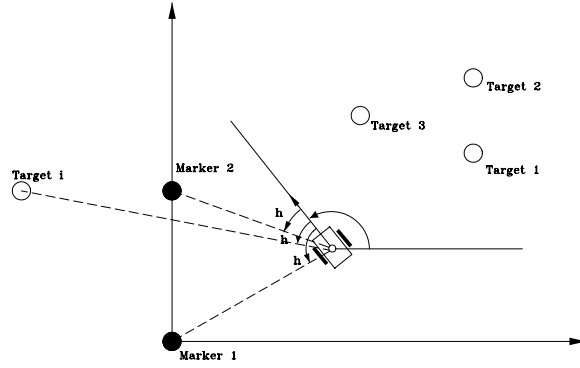


Fig. 2. A vehicle in an unknown environment with markers and targets.

angles formed by the rover's fore axis with lines through the sensor head and the M markers and N targets. The linear approximation of system (1) at any equilibrium $x = x_0, p = p_0, u = 0$, would indeed have a null dynamic matrix

$$A = \left. \frac{\partial f(\cdot)}{\partial (x, p)} \right|_{eq.} = 0 \in \mathbf{R}^{(2N+3) \times (2N+3)},$$

and output matrix

$$C = \left. \frac{\partial h(\cdot)}{\partial (x, p)} \right|_{eq.} \in \mathbf{R}^{(M+N) \times (2N+3)}.$$

Hence, in any nontrivial case (i.e., whenever there is at least one targets ($N \neq 0$) or there are less than three known markers ($M < 3$)) the linearized system is unobservable.

On the other hand, it is intuitively clear (and everyday's experience in surveyors' work) that simple triangulation calculations using two or more measurements from different positions would allow the reconstruction of all the problem unknowns, except at most for singular configurations. Analytically, complete observability of system (1) can be checked, as an exercise in nonlinear system theory, by computing the dimension of $\langle f(\cdot) | \text{span} \{dh(\cdot)\} \rangle$, the smallest codistribution that contains the output one-forms and is invariant under the control vector fields (see [2] for details on calculations). By such nonlinear analysis, it is also possible to notice that observability can be destroyed by choosing particular input functions, the so-called "bad inputs". A bad input for our example is the trivial input $u = 0$: the vehicle cannot localize itself nor the targets without moving. Other bad inputs are illustrated in Fig.(3).

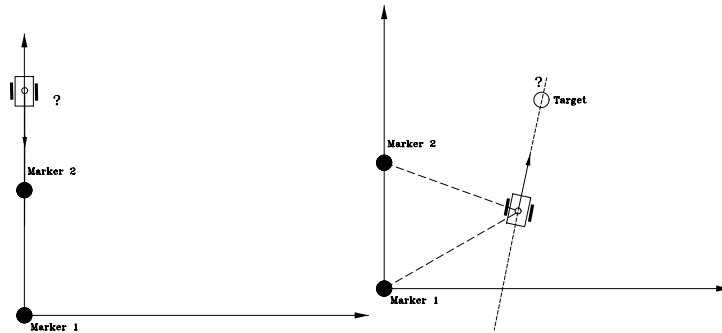


Fig. 3. A vehicle triangulating with two markers cannot localize itself if the inputs are such that it remains aligned with the markers; it cannot localize a target if it aims at the target directly.

In order to drive a rover to explore its environment, it is clear that bad inputs should be avoided. Indeed, the very fact that there exist bad inputs suggests that there should also be "good", and possibly optimal, inputs. To find such optimal exploratory strategies, however, the differential geometric analysis tools such as those introduced above are not well suited, as they only provide topological criteria for observability. What is needed instead is a metric information on the "distance" of a system from unobservability, and to how maximize it. More generally, it is to be expected that different trajectories will elicit different amounts of information: a complete SLAM system should not only provide estimates of the vehicle and feature positions, but also as precise as possible a description of the statistics of those estimates as random variables, so as to allow evaluation of confidence intervals on possible decisions.

To provide a better understanding of how two different states can be distinguished via dynamic measurements, let us consider the output $y(t) = h(x, p) = y(x_o, u, t)$ as a function of the initial conditions x_o and of the inputs u . Let x_o^o and x_o' denote two different initial conditions, with $\|x_o^o - x_o'\| < \epsilon$, and let us consider

$$y(x_o', u, t) - y(x_o^o, u, t) = \left. \frac{\partial y}{\partial x_o} \right|_{x_o=x_o^o} (x_o' - x_o^o) + O^2(\epsilon) \quad (7)$$

i.e. a linear measurement equation of the form

$$\tilde{y}(t) + \delta_y = M(t)\tilde{x} \quad (8)$$

where $\tilde{x} = (x_o' - x_o^o)$ is unknown, \tilde{y} comes from measurements, and the perturbation term δ_y accounts for measurement noise and approximation errors. Notice explicitly that the linear operator $M = \left. \frac{\partial y}{\partial x_o} \right|_{x_o=x_o^o} \mathcal{F}$ depends in general on applied inputs, as only for very special systems (in particular, linear) superposition of effects of initial states and inputs holds. By premultiplying both sides of (8) by $M^T W$, with $W > 0$ a suitable positive definite matrix weighing accuracy of different sensors, and by integrating from time 0 to T , we obtain

$$\mathcal{Y} + \Delta_y = \mathcal{F}\tilde{x}, \quad (9)$$

where $\mathcal{Y} = \int_0^T M^T(t)W\tilde{y}(t)dt$, and $\mathcal{F} = \int_0^T M^T(t)WM(t)dt$ is the *Fisher Information Matrix* for our system.

Singularity of \mathcal{F} (for some input choice) clearly implies that distinct initial values of the state exist which provide exactly the same measurements over the time interval, hence is tantamount to unobservability of the system.

A different argument to support the same conclusion can be derived from Kalman estimation theory. Indeed, in the linear case, for the covariance matrix P of a Kalman filter, the Cramèr-Rao inequalities [38] hold:

$$[\mathcal{F} + \mathcal{N}^{-1}]^{-1} \leq P \leq \mathcal{F}^{-1} + \mathcal{N} \quad (10)$$

where \mathcal{F} is the Fisher Information Matrix (defined in (9) in this framework), and \mathcal{N} , the covariance matrix of process noise, is assumed to be independent of the trajectory. According to this, minimization of \mathcal{F}^{-1} can be considered as an instrument to minimize P . This is further justified by the fact that, in the absence of process noise and of prior information, the Riccati equation solution for the filter is exactly $P(t) = \mathcal{F}^{-1}(t)$. Cramèr-Rao bounds can be extrapolated to estimate covariance for non-linear systems (see e.g. [27]) (although, in the context of non-linear systems, minimum-variance estimates do not enjoy the properties that make them desirable for linear systems, and MVE-based optimal sensor design is often questionable [1]).

From the above considerations on state reconstruction and on Cramèr-Rao inequalities, it is clear that the information matrix can provide the desired notion of “distance” from unobservability, that is, a merit figure for different inputs (hence trajectories) of the exploring rover. Indeed, the smallest eigenvalue $E = \lambda_{\min}(\mathcal{F}) = 1/\|\mathcal{F}^{-1}\|_2$, the determinant index $D = {}^{(n_v + dn_f)} \sqrt{\det \mathcal{F}}$ the

trace index $T = \frac{\text{trace}(\mathcal{F})}{n_v + dn_f}$, and the average-variance index $A = \frac{n_v + dn_f}{\text{trace}(\mathcal{F}^{-1})}$, are among the most often used such criteria (known as E-, D-, T-, and A-criterion, respectively).

Notice that information based criteria do not reflect any particular choice in the estimator or filter adopted in the actual localization procedure, rather it is intrinsic to the reconstructibility of the state from the given trajectory. This is a very useful property, in view of the fact that several different estimators and filters can be applied to the SLAM problem.

The problem of choosing exploratory paths of fixed length L to maximize SLAM information can be formalized (in the E-criterion sense) as an *optimal control* problem, i.e.

$$\text{maximize } J(u) = \lambda_{\min}(\mathcal{F}), \quad (11)$$

subject to the constraints

$$L = \int_0^T \sqrt{(\dot{x}_1^2 + \dot{x}_2^2)} dt,$$

$$\dot{x} = G(x)u; \quad x(0) = x_o,$$

$$y = h(x).$$

Solving this problem can be expected to be quite difficult in general. Using system-theoretic tools, an analytic solution was given in [31] for the simplified case of an omnidirectional vehicle moving in a planar environment with only two markers. Extremal paths for the functional J were shown to be contained in the pencil of curves spanned by the parameter α as

$$[\cos(\alpha) \sin(\alpha)] \left(\frac{\partial y}{\partial x_o} \right)^T \frac{\partial y}{\partial x_o} - \frac{\partial y}{\partial x_o} \Big|_{x=x_o}^T \frac{\partial y}{\partial x_o} \Big|_{x=x_o} \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix} = 0, \quad (12)$$

where the actual value of α depends on L . It can be easily seen that the obtained pencil is a set of conics (some examples of optimal exploratory paths, for different lengths, are represented in Fig.(4)).

Extensions of the analytic solutions to nonholonomically constrained vehicles with unknown target features are feasible (work in this direction is undergoing). However, to obtain solutions in most general cases, efficient numerical methods are in order. In a recent overview [41], where the importance of the SLAM optimization problem is acknowledged, currently available solutions are reported to be mostly limited to heuristic, greedy algorithms. Furthermore, most known methods often disregard the nonlinear character of the SLAM problem, which on the contrary is of large momentum, as we discussed.

The main limitation of gradient-descent methods in this framework is of course the presence of local minima in the information return function: application of methods from receding-horizon optimal control theory in this context can be expected to

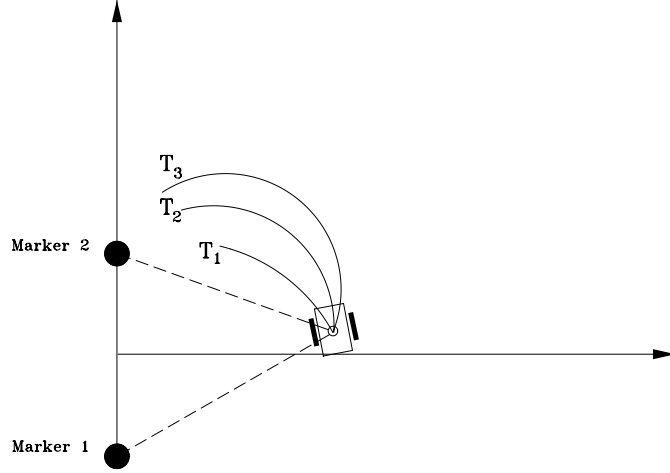


Fig. 4. Optimal trajectories for three different path lengths $T_1 = 1sec$, $T_2 = 2sec$, and $T_3 = 3sec$

offer a substantial edge. In the following, we illustrate application of such techniques to a few examples of on-line, numeric SLAM trajectory optimization.

To apply numerical methods, continuous-time system equations (1) or (3) are first discretized, so that the information matrix is rewritten as the sum of products

$$\mathcal{F} = \sum_{i=0}^k \left. \frac{\partial y}{\partial x_o} \right|_{x_i}^T \left. \frac{\partial y}{\partial x_o} \right|_{x_i} \quad (13)$$

evaluated at each point of a candidate trajectory. Using techniques developed in [36], we furthermore introduce a quantization of the input space (i.e., the set of possible incremental moves of the vehicle), thus inducing a discretization of the configuration space. It can be shown that, for vehicles with chained-form kinematics, the reachable set is indeed a lattice in this case, which is a very convenient structure to apply numerical search methods to.

If d is the cardinality of the input set, there are d^k paths of length k stemming from a generic configuration, for which the contribution to information is given by (13). An exhaustive search of the most informative path is possible for moderate values of d and k . The receding-horizon optimal control policy consists then in applying only the first control of the locally optimizing sequence, to recompute the next optimizing sequence, thus proceeding iteratively. The method can be easily used in conjunction with other techniques for e.g. obstacle avoidance. How practical the method is depend very much on the affordable horizon length for which real-time computations are feasible, hence choices concerning time and input quantization, information representation, etc., are an important area of research.

Simulation results reported in Fig.(5) compare performance of a greedy algorithm with the receding-horizon method. Walls are considered here as pure obstacles,

i.e. they are detected iff the vehicle “bumps” into them, while information for self localization and mapping is only extracted from measurements relative to two markers (black circles) and to four target features. Results show how the receding-horizon methods collects richer information in this case.

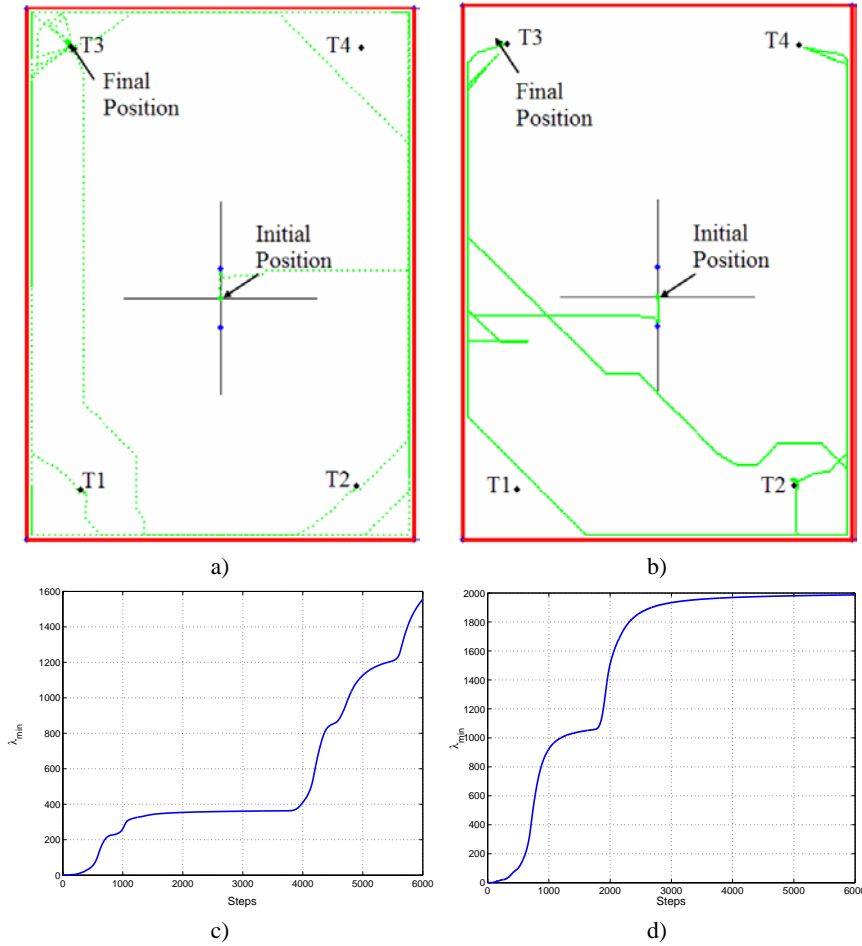


Fig. 5. Trajectory of a vehicle during the exploration of a rectangular environment with 2 markers and 4 target features, using gradient-descent (a) and a 3-steps receding horizon (b), respectively. Time evolutions of the corresponding information return function $E = \lambda_{min}(\mathcal{F})$ are reported in c) and d).

More simulation results relative to different environments are reported in Fig.(6). While these results show how the method is quite versatile in navigating in a cluttered environment fetching for information where that is available, it is of course an open research issue to provide a provable, quantitative assessment of the advantages of

this method with respect to others, and to design the numerous parameters that play an important role in its implementation.

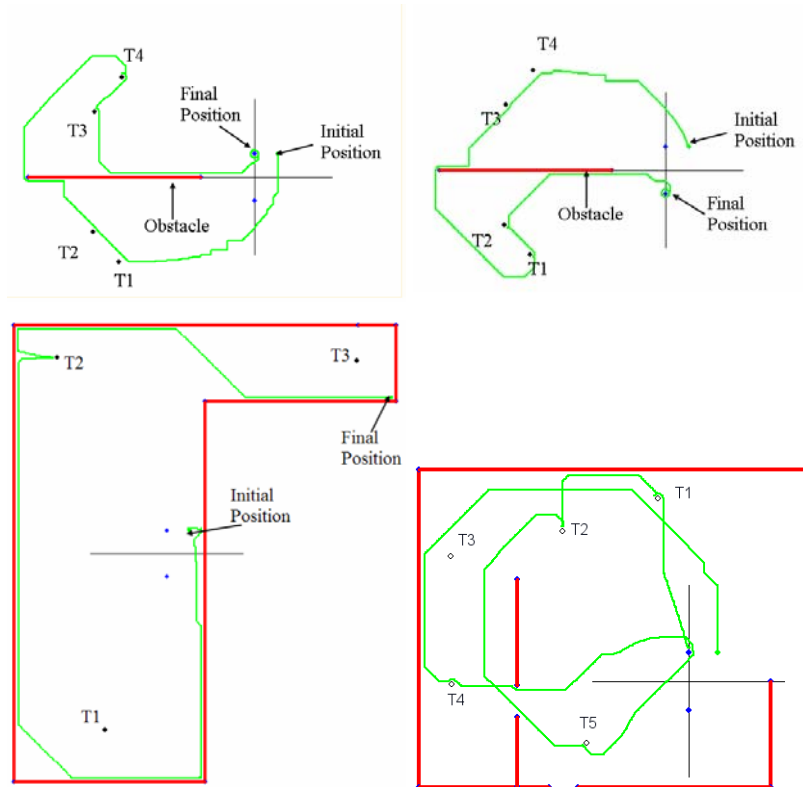


Fig. 6. Receding-horizon optimal trajectories in different environments, whereby the task of maximizing the information return function leads the vehicle to cover target areas. Observe how slightly different initial conditions may lead to completely different exploration strategies (upper right and left), however with similar characteristics. More complex environments are also dealt with satisfactorily (bottom left and right).

5 Simultaneous Localization and Servoing

As mentioned in the introduction, one of the consequences of the intrinsic nonlinearity of the SLAMS problem is that there exists no guarantee that, even assuming that a converging estimator *and* a stabilizing state-feedback law are available, their connection will provide an overall controller which behaves as expected.

Many different approaches can be taken at addressing this problem, depending on the specifics of the task and of the models at hand. In this section, we will survey two methods which differ in the generality they afford.

5.1 Observer-based SLAMS

A first approach to design simultaneous localization and servoing for system (1) relies on techniques for output-feedback stabilization of nonlinear systems, based on nonlinear observers and extensions of the separation principle. An observer providing an estimate \hat{x} of the current state x is said to enjoy the separation property if, whenever there exists a nominally stabilizing static state feedback $u(x)$, application of the control law $u(\hat{x})$ permits to achieve (local) stability of the closed-loop system.

The widely used extended Kalman filter (see e.g. [5,19,32]) lacks in general provable properties of convergence and separation. Several authors underscored that application of the EKF to localization data is often troublesome. Indeed, the filter convergence properties are very much prone to initialization of filter parameters (e.g., measurement and process covariances). Several alternative schemes of nonlinear observers have been proposed in the literature.

One way of designing an observer is to transform the original nonlinear system into another one for which the design is known. Transformations, which have been proposed in the literature, include system immersion [14], which permits to obtain a bilinear system if the observation space is finite dimensional, and linearization by means of output injection [21,28,30], which assumes that particular differential-geometric conditions on the system vector fields are verified. A nonlinear observer and its practical implementation have been presented in Gauthier et al. [9,4], in which the first step is writing the input affine nonlinear system in a so-called normal observation form. However, this form requires that the trivial input is an universal input [3] for the system. In our problem, this condition is violated by the mobile robot kinematics.

An interesting possibility for an estimator for the localization problem is the extension of the Luenberger filter in a nonlinear setting, by using the time derivatives of the input ([46,40]). In [10], a local nonlinear observer for mobile robot localization was designed, based on the concept of Extended Output Jacobian (EOJ) matrix, which is the collection of the covectors associated to the considered elements of the observability space, i.e. the output and its derivative. Output derivatives are estimated by using high-pass filters. Local practical stability of the observation error dynamics is guaranteed since persistent perturbations introduced by filters can be made arbitrarily small. A singularity-avoidance exploration task is also addressed to deal with the singularity occurrences in the EOJ matrix. For such an observer scheme, [10] showed a local separation property to hold.

Observer-based approaches are applicable to rather general models of vehicles and sensorial equipment, but has correspondingly some weaknesses. In the first place, convergence and separation can only be proven locally. While this is to be expected with complex nonlinear systems such as those at hand, the drawback shown by laboratory practice is that the large initial estimation errors or displacements from the desired pose often prevent correct functioning of the system.

5.2 Visual servoing

When the SLAMS problem is specialized for a particular class of sensors and vehicles, more powerful techniques can usually be devised. In this paragraph, we report on a particular but important case, where we assume that the sensorial information consists of a video camera mounted on-board the vehicle.

Visual servoing techniques, which have been profitably used in recent years mostly for the control of robot arms ([23]), use visual information either directly, by the computation of an image error signal, or indirectly, by the evaluation of the state of the system. These two approaches were classified by Weiss in 1984 as *Image Based Visual Servoing* (IBVS) and *Position Based Visual Servoing* (PBVS), respectively. Indeed, these two schemes should be regarded as the end-points of a range of different possibilities, whereby the raw sensorial information is gradually abstracted away to a more structured representation using some knowledge of the robot-environment model (a scheme which is roughly half-way between IBVS and PBVS was used e.g. in [11]).

IBVS and other sensor-level control schemes have several advantages, such as robustness (or even insensitivity) to modelling errors and hence suitability to unstructured scenes and environments. On the other hand, PBVS and in general higher-level control schemes also have important attractive features. Using the PBVS approach, for instance, the control law can be synthesized in the usual working coordinates for the robot, and thus usually a simpler synthesis is made possible. Furthermore, abstracting sensor information to a higher level of representation allows using different sensorial sources. In the example of a camera mounted on a mobile robot, for instance, the synergistic use of odometry and visual feedback is only possible if these information can be taken to some common denominator where they can be fused coherently.

Early work on visual servoing of wheeled vehicles include those of [22] and [11]. In the latter papers a feedback control law stabilizing the vehicle posture by using visual information only was solved. More recently, [35] considered the problem under the practically most relevant constraint of keeping tracked features within sight of a limited-aperture camera while the vehicle maneuvers to park. The method proposed in [35] adopts a hybrid control law, that solves the problem by switching among different stabilizing output-feedback laws, depending on conditions triggered by events such as the approach of image boundaries by some tracked features. It is to be noted that, although different sensors (such as some models of laser range finders, or omnidirectional cameras, or pan-tilt heads) may not be affected by view-angle limitations, these are typically some orders of magnitude more expensive than the conventional cameras considered in [35], which are readily available even in the consumer market. Implementation of the visual-servo method of [35] is based on selecting a few target features from an image recorded at the desired configuration, and by comparing their position in the image plane with that obtained in real-time from the robot camera. Some experimental results obtained by application of this method are reported below for illustration.

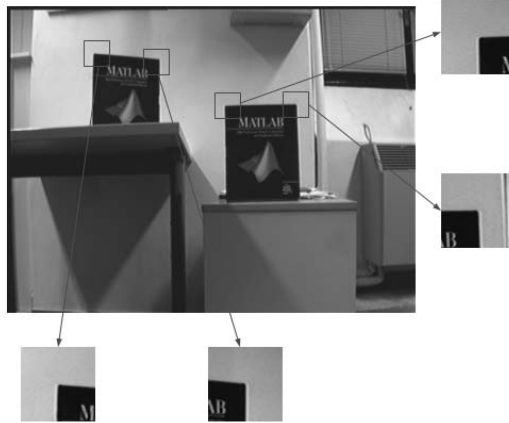


Fig. 7. Image grabbed from the robot camera at the target position, with four selected control features.

6 Conclusion

In this chapter, we have considered the connection of three different problems, localization, map building, and servoing, of mobile vehicles moving in unstructured or partially structured environments. An effort has been paid at casting the three problems within a unique framework, which is that provided by the theory of dynamical control systems with outputs. Although this approach is still to be validated in large-scale applications, where the dimensionality of the space of unknowns and the possible topological complexity of the environment can place formidable obstacles, there seems to be some interesting avenue of development at the confluence of classical computer-science and probabilistic approaches and system theory.

Acknowledgement

Work supported by the following contracts: IST-2001-37170 (RECSYS), ASII/R/124/02 (TEMA).

References

1. A. Bicchi, and G. Canepa, "Optimal Design of Multivariate sensors", *Measurement Science and Technology*, vol. 5, pp. 319–332, IOP Publishing Ltd, 1994.
2. A. Bicchi, D. Pratichizzo, A. Marigo, and A. Balestrino, "On the observability of mobile vehicles localization", *Proc. IEEE Mediterranean Conf. On Control And Systems*, 1998.
3. G. Bonard, F. Celle-Couenne, and G. Gilles, *Nonlinear Systems, Vol.1*, Chapman & Hall, 1995.
4. G. Bonard, and H. Hammouri, "A high gain observer for a class of uniformly observable systems," *Proc. 30th IEEE Conference on Decision and Control*, 1991.



Fig. 8. External views (left column) and subjective images (right column) as taken from the vehicle, in the initial configuration (top row) and in the final configuration (bottom row), after reaching convergence under the visual feedback control scheme of [35]. The bottom right image should be compared with the target image in fig. 7.

5. J. Borenstein, A. Everett, and L. Feng, *Navigating mobile robots*, AK Peters, 1996.
6. G. Campion, B. d'Andr ea–Novel, and G. Bastin, “Controllability and State Feedback Stabilization of Nonholonomic Wheeled Mechanical Systems,” *Advanced Robotic Control*, Springer-Verlag, pp. 106–124, 1991.
7. C. Canudas de Wit, and O.J. Sordalen, “Exponential Stabilization of Mobile Robots with Nonholonomic Constraints,” *IEEE Transactions on Robotics and Automation*, vol. 37, nr. 11, 1992.
8. G. Casalino, M. Aicardi, A. Bicchi, and A. Balestrino, “Closed Loop Steering and Path Following for Unicycle–like Vehicles: a Simple Lyapunov Function based Approach,” *IEEE Robotics and Automation Magazine*, vol. 2, nr. 1, pp. 27–35, 1995.
9. F. Celle, J.P. Gauthier, and D. Kazakos, “Orthogonal representations of nonlinear systems and input-output maps,” *Systems & Control Letters*, nr. 7, pp. 365–372, 1986.
10. F. Conticelli, and A. Bicchi, “Observer Design for Locally Observable Analytic Systems: Convergence, Separation Property, and Redundancy,” *Nonlinear Control in the Year 2000*, Springer-Verlag, nr. 258, pp. 315–330, 2000.
11. F. Conticelli, D. Prattichizzo, F. Guidi, and A. Bicchi, “Vision-based dynamic estimation and set-point stabilization of nonholonomic vehicles,” *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, USA, 2000, pp. 2771–2776.
12. B. d'Andr ea–Novel, G. Campion, and G. Bastin, “Control of nonholonomic wheeled mobile robots by state feedback linearization,” *International Journal of Robotics Research*, vol. 14, nr. 6, pp. 543–559, 1995.

13. A. De Luca, G. Oriolo, and M. Vendittelli, "Control of Wheeled Mobile Robots: An Experimental Overview," in S. Nicosia, B. Siciliano, A. Bicchi, and P. Valigi (Eds.) *RAMSETE. Articulated and Mobile Robots for SERVICES and TEchnology*, Lecture Notes in Control and Information Sciences, vol. 270, pp. 181–226, Springer Verlag, London, 2001.
14. M. Fliess, and I. Kupka, "A finiteness criterion for nonlinear input-output differential systems", *SIAM Journal Control and Optimization*, nr. 21, pp. 712-729, 1983.
15. J. W. Fenwick, P. M. Newman, and J.J. Leonard "Cooperative Concurrent Mapping and Localization," *Proceedings of the 2002 IEEE International Conference on Robotics and Automation*, Washington, pp. 1802-1809, 2002.
16. F. Gambino, G. Oriolo, and G. Ulivi, "A comparison of three uncertainty calculus techniques for ultrasonic map building," *SPIE*, Orlando, pp. 249–260, 1996.
17. F. Gambino, G. Ulivi, and M. Vendittelli, "The transferable belief model in ultrasonic map building," *6th IEEE Conference on Fuzzy Sys*, Barcellona, pp. 601–606, 1997.
18. A. Garulli, and A. Vicino, "Set Membership Localization of Mobile Robots via Angle Measurements," *IEEE Transaction on Robotics and automation*, Vol. 17, nr. 4, AUGUST 2001
19. A. Gelb, "Applied Optimal Estimation", MIT Press, 1974.
20. J. E. Guivant and Eduardo Mario Nebot, "Optimization of the Simultaneous Localization and Map-Building Algorithm for Real-Time," *Proc. IEEE*
21. R. Hermann, and A. Krener, "Nonlinear Controllability and Observability," *TAC*, vol. 22, nr. 5, 1977.
22. K. Hashimoto and T. Noritsugu, "Visual servoing of nonholonomic cart," *Proc. Int. Conf. on Robotics and Automation*, pp. 1719–1724, April 1997.
23. S. A. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robotics and Automation*, vol. 5, nr. 12, pp. 651–670, 1996.
24. A. P. Dempster, A. N. Laird, and D. B. Rubin, "Maximum Likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.
25. A. Isidori, "Nonlinear Control Systems", 3rd ed., Springer Verlag, 1995.
26. L. Jetto, S. Longhi, and D. Vitali, "Localization of a wheeled mobile robot by sensor data fusion based on a fuzzy logic adapted Kalman filter," *Control Engineering Practice*, vol. 7, pp. 763-771, 1999.
27. R.L. Kosut, A. Arbel, and K.M. Messler: "Optimal Sensor Design for State Reconstruction", *IEEE Trans. on Automatic Control*, vol. AC-27, n.1, pp. 242-4, Feb. 1982.
28. A.J. Krener, and H. Schättler, "The Structure of Small Time Reachable Sets in Low Dimensions," *SIAM J. Control Optim.*, vol. 27, nr. 1, pp. 120–147, 1989.
29. J. J. Leonard, H. F. Durrant-White, and I. J. Cox, "Dynamic map building for an autonomous mobile robot," *International Journal of Robotics Research*, vol. 11, pp. 89–96, 1992.
30. J. Levine, and R. Marino, "Nonlinear system immersion, observers and finite dimensional filters," *Systems & Control Letters*, nr. 7, pp. 137-142, 1986.
31. F. Lorussi, A. Marigo, and A. Bicchi, "Optimal exploratory paths for a mobile rover," *Proc. IEEE Int. Conf. on Robotics and Automation*, pp. 2078–2083, 2001.
32. E.A. Misawa, and J.K. Hedrick, "Nonlinear Observers – A state of the art – Survey", *Trans. ASME*, nr. 111, pp. 344–352, 1989.
33. M. Montemerlo, S. Thrun, D. Koller, and Ben Wegbreit, "A Factored Solution to the Simultaneous Localization and Mapping Problem," *Proceedings of the AAAI National Conference on Artificial Intelligence*, Edmonton, 2002.

34. P. Morin, and C. Samson, "Application of Backstepping Techniques to the Time-Varying Exponential Stabilization of Chained Form Systems," *INRIA Research Report*, Sophia-Antipolis, nr. 2792, 1996.
35. P. Murrieri, D. Fontanelli, and A. Bicchi, "Visual-servoed parking with limited view angle," *Experimental Robotics VIII*, Springer Verlag, vol. 5, pp. 254-263, 2002.
36. S. Pancanti, L. Leonardi, L. Pallottino, and A. Bicchi, "Optimal control of quantized input systems," in C. Tomlin and M. Greenstreet, editors, *Hybrid Systems: Computation and Control*, volume LNCS 2289 of Lecture Notes in Computer Science, pages 351-363. Springer-Verlag, Heidelberg, Germany, 2002.
37. F. Piloni, and A. Bicchi, "Navigazione di veicoli autonomi mediante visione artificiale: pianificazione dei moti di esplorazione", Internal Report, Centro Piaggio, University of Pisa, 1994.
38. C. R. Rao: "Linear Statistical Inference and Its Applications", Wiley, New York, 1973.
39. S. Se, D. Lowe, and J. Little, "Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks," *International Journal of Robotics Research*, vol. 21, pp. 735-758, 2002.
40. A.R. Teel, and L. Praly, "Global stabilizability and observability imply semi-global stabilizability by output feedback," *Systems & Control Letters*, nr. 22, pp. 313-325, 1994.
41. S. Thrun, "Robotic Mapping: A Survey," in *Exploring Artificial Intelligence in the New Millenium*, Morgan Kaufmann, 2002.
42. S. Thrun, W. Burgard, and D. Fox, "A Probabilistic Approach to Concurrent Mapping and Localization for Mobile Robots," *Machine Learning and Autonomous Robots*, nr. 31/5, 125, 1998.
43. S. Thrun, D. Koller, Z. Ghahmarani, and H. Durrant-Whyte "SLAM Updates Require Constant Time," Tech. rep., School of Computer Science, Carnegie Mellon University, 2002.
44. A. Tarantola: "Inverse Problem Theory", Elsevier, Amsterdam, 1987.
45. D. Tilbury, and A. Chelouah, "Steering a Three-Input Nonholonomic System using Multi-Rate Controls," *European Control Conference*, pp. 1428-1431, 1993.
46. M. Zeitz, "The extended Luenberger observer for nonlinear systems," *Systems and Control letters*, vol. 9, pp. 149-156, 1987.