Proc. of CDC'03 - 42nd IEEE Conference on Decision and Control, December 2003

Minimum-energy State Estimation for Systems with Perspective Outputs and State Constraints

António Pedro Aguiar
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106, USA
aguiar@ece.ucsb.edu

João Pedro Hespanha
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106, USA
hespanha@ece.ucsb.edu

***Abstract*— We address the problem of estimating the state of a system with perspective outputs, whose state is known to satisfy a set of quadratic constraints. We construct a dynamical system that produces an estimate of the state that satisfies the constraints and is "most compatible" with the dynamics, in the sense that it requires the least amount of noise energy to explain the measured output. We apply these results to the estimation of position and orientation of a controlled rigid body, using measurements from a charged-coupled-device (CCD) camera attached to the body. The main contribution of this work is the inclusion of state constraints in the minimum-energy formulation. In the context of our application, these constraints allow us to avoid singularities that previous minimum-energy controllers encountered. The results are validated experimentally by using measurements from a CCD camera mounted on a mobile robot to estimate its position and orientation. These estimates are then used to close the loop and control the robot to a desired position, defined with respect to visual landmarks. The use of state constraints in the estimator allow the system to operate even when the trajectories for the robot do not exhibit "excitation," which was needed in previous minimum-energy estimators.**

I. INTRODUCTION

This paper deals with systems of the form

$$\dot{x} = A(u)x + b(u) + G(u)\mathbf{d}, \tag{1}$$

$$\alpha_j y_j = C_j(u)x + d_j(u) + \mathbf{n}_j, \qquad j \in \{1, 2, \dots, k\}, \tag{2}$$

where $x \in \mathbb{R}^n$ denotes the state of the system, $u \in \mathbb{R}^{n_u}$ its input, $y_j \in \mathbb{R}^{m_j}$ its jth perspective output, $\mathbf{d} \in \mathbb{R}^{n_d}$ an input disturbance that cannot be measured, and $\mathbf{n}_j \in \mathbb{R}^{n_n}$ measurement noise affecting the jth output. Each $\alpha_j \in \mathbb{R}$, $j \in \{1, 2, \dots, k\}$ denotes a scalar that is determined by a normalization constraint such as

$$\|y_j\| = 1 \qquad \text{or} \qquad v_j' y_j = 1, \tag{3}$$

where the $v_j \in \mathbb{R}^{m_j}$ denote constant vectors. We call (1)–(2) a *state-affine system with multiple perspective outputs*, or for short simply a *system with perspective outputs*. This type of systems are inspired by the (single output) perspective systems introduced by Ghosh et al. [8].

Systems with perspective outputs typically arise when charged-coupled-device (CCD) cameras are used to acquire information about the position and orientation of moving rigid bodies. In Section III we will consider the specific problem of estimating the position and orientation of a controlled rigid body using measurements from a CCD camera attached to it. The dynamics of this system can be written as (1)–(2). The reader is referred to [8, 9, 23] for several other examples of perspective systems in the context of motion and shape estimation.

We assume that the state x of the system is constrained to lie on a subset of $\mathbb{R}^n$ specified by n_c quadratic constraints of the form

$$x' S_i x + r_i' x + \rho_i = 0, \qquad \forall i \in \{1, 2, \dots, n_c\}, \tag{4}$$

where the $S_i \in \mathbb{R}^{n\times n}$ are symmetric matrices, the $r_i \in \mathbb{R}^n$ column vectors, and the $\rho_i \in \mathbb{R}$ scalars. We call (1)–(2), (4) a *quadratically constrained state-affine system with multiple perspective outputs*, or for short simply a *constrained system with perspective outputs*. The main reason to consider this type of state constraints is to take into account that some elements of x are known to lie in given manifolds. In the context of rigid body motion, typically part of the state of the system is a rotation matrix that is known to lie in $\mathrm{SO}(3)$. This can be expressed by constraints like (4). State-constraints for systems with perspective outputs also appeared in [23] in the context of motion estimation using a CCD camera and a laser range finder, where the measurements from camera and range finder were related by an algebraic constraint.

This paper addresses state-estimation for constrained systems with perspective outputs. In the last few years, the observability of perspective linear systems was been systematically studied in the literature and [6] provides an elegant algebraic observability test. It should be noted that for perspective linear systems without inputs it is never possible to recover the norm of the state because the system is homogeneous on the initial conditions. Therefore Dayawansa et al. [6] only consider state indistinguishability up a homogeneous scaling of the state. However, as shown in [12, 10], in the presence of inputs it is in principle possible to recover the whole state from projective outputs.

We propose a minimum-energy estimator that produces an estimate for the state of the (1)–(2) that is "most compatible" with system's dynamics and measured outputs. In particular, the optimal state estimate $\hat{x}$ at time t is defined to be the value for the state of (1) that is compatible with the observations (2) collected up to time t and the dynamics (1) for the "smallest" possible noise $\mathbf{n}_j$ and disturbances $\mathbf{d}$, with "smallest" understood in an integral-square sense. This type of estimators were first proposed by Mortensen [19] and further refined by Hijab [11]. Game theoretical versions of these estimators were proposed by McEneaney [18]. It was recently shown by Krener [16] that this type of estimators is globally convergent when the system is observable for *every* input. In [10], it was shown that for projective systems with multiple inputs, convergence can be obtained under less restrictive observability assumptions. In Section II, we improve upon the results in [10] by incorporating state-constraints in the minimum-energy formulation.

A fundamental problem in mobile robotics is the determination of the position and orientation of a robot with respect to an inertial coordinate system. A promising solution to this problem is to utilize a camera mounted on the robot that observes the apparent motion on the image of stationary points. The linear and angular velocities of the camera can be assumed known in its own coordinate system (possibly with errors due to noise) but not in the inertial coordinate system. This is quite reasonable in mobile robotics where the motion of the camera is determined by the applied control signals. The problem of estimating the position and orientation of a camera mounted on a rigid body from the apparent motion of point features has a long tradition in the computer vision literature. We are interested here in filtering-like or iterative algorithms that continuously improve the estimates as more data (i.e., images) are acquired and that are robust with respect to measurement noise. Following the lead of other researchers (cf., e.g., [17, 13, 22, 14, 5, 21] and references therein) in Section III we formulate this as a state-estimation problem and utilize the minimum-energy state estimator to solve it. The estimator proposed here inherits the convergence properties of the one proposed in [10], when the trajectory exhibits sufficient excitation. Moreover, it may still converge when the observability assumptions in [10] are violated by making use of the constraint that the rotation matrix must be an element of $\mathrm{SO}(3)$. This constraint is also used to avoid singularities that often arise when all points being observed are coplanar.

The theoretical results were experimentally validated by applying them to estimate the position and orientation of a mobile robot using measurements from an on-board CCD camera. We then use these estimates to close the feedback loop and control the robot to a desired position, defined with respect to visual landmarks. The results obtained are discussed in Section IV.

II. State estimation with constraints

In [10] we proposed to formulate state estimation for the system with perspective outputs (1)–(2), as a deterministic optimization problem in which the estimate $\hat{x}(t)$ of the state at time $t \geq 0$ is the value for which the measured outputs can be made compatible with the system dynamics (1)–(2) for the "smallest" possible noise $\mathbf{n}_j$ and disturbance $\mathbf{d}$. Formally, given an input u and measured outputs y_j defined on an interval $[0, t)$, we define

$$\hat{x}(t) := \arg \min_{z \in \mathbb{R}^n} J(z, t), \tag{5}$$

where

$$\begin{aligned} J(z;t) := \min_{\mathbf{d}, \mathbf{n}_j, \alpha_j} \Big\{ &(x(0) - \hat{x}_0)' P_0 (x(0) - \hat{x}_0) \\ &+ \int_0^t \Big(\|\mathbf{d}\|^2 + \sum_j \|\mathbf{n}_j\|^2 \Big) d\tau : \\ x(t) = z,\ &\dot{x} = A(u)x + b(u) + G(u)\mathbf{d}, \\ &\alpha_j y_j = C_j(u)x + d_j(u) + \mathbf{n}_j \Big\}, \end{aligned} \tag{6}$$

with $P_0 > 0$ and $\hat{x}_0$ encodes a-priori information about the state. The following result from [10] solves the state-estimation problem defined above.

Theorem 1: The cost J defined in (6) is given by

$$J(z;t) = \big(z - \bar{x}(t)\big)' P(t) \big(z - \bar{x}(t)\big) + J_0(t),$$

where

$$-\dot{P} = PA + A'P + PGG'P - W, \qquad P(0) = P_0 \tag{7}$$

$$P\dot{\bar{x}} = PA\bar{x} + Pb - W\bar{x} - w, \qquad \bar{x}(0) = \hat{x}_0 \tag{8}$$

$$\dot{J}_0 = \bar{x}'W\bar{x} + \sum_j d_j' \Big(I - \frac{y_j y_j'}{\|y_j\|^2} \Big) d_j, \quad J_0(0) = 0, \tag{9}$$

and

$$W(t) := \sum_j C_j'(u) \ \ I - \frac{y_j y_j'}{\|y_j\|^2} \ \ C_j(u),$$

$$w(t) := \sum_j C_j'(u) \ \ I - \frac{y_j y_j'}{\|y_j\|^2} \ \ d_j(u), \quad t \geq 0.$$

Moreover, $P(t) > 0\ \forall t \geq 0$ and the estimate $\hat{x}(t)$ defined by (5) is uniquely defined and equal to $\bar{x}(t)$.

Theorem 1 solves the state estimation problem for the unconstrained system (1)–(2). However, even when the system satisfy the constraints (4), there is no guarantee that the estimates $\hat{x}(t) = \bar{x}(t)$, $t \geq 0$ will satisfy these constraints. To address this, we will use the estimate

$$\hat{x}(t) := \arg \min_{\substack{z \in \mathbb{R}^n : z' S_i z + r_i' z + \rho_i = 0 \\ \forall i \in \{1, \ldots, n_c\}}} J(z, t),$$

instead of (5). We solve this optimization problem using Lagrange multiplier theory [4], which leads to the following sufficient conditions for optimality:

$$P(\hat{x} - \bar{x}) + \big(\bar{S}(\hat{x}) + \bar{R}\big)\lambda = 0, \tag{10}$$

$$(\bar{S}(\hat{x}) + 2\bar{R})'\hat{x} + \bar{\rho} = 0, \tag{11}$$

$$z'\bar{P}z > 0, \qquad \forall z \in \mathbb{R}^n : (\bar{S}(\hat{x}) + 2\bar{R})'z = 0, \tag{12}$$

where $\lambda \in \mathbb{R}^{n_c}$ is the Lagrange multiplier vector and

$$\bar{P} := P + \sum_{i=1}^{n_c} \lambda_i S_i, \qquad \bar{S}(\hat{x}) := [S_1\hat{x} \quad \cdots \quad S_{n_c}\hat{x}],$$

$$\bar{R} := \frac{1}{2} [r_1 \quad \cdots \quad r_{n_c}], \qquad \bar{\rho} := [\rho_1 \quad \cdots \quad \rho_{n_c}]'.$$

Note that (11) is simply a vectorized reformulation of the constraints $\hat{x}'S_i\hat{x} + r_i'\hat{x} + \rho_i = 0$, $\forall i \in \{1, \ldots, n_c\}$. The following result utilizes these conditions to construct an asymptotic estimate for x.

Theorem 2: Suppose that along trajectories of the system we have $\bar{P} > 0$ and $\bar{S}(\hat{x}) + \bar{R}$ remains full column rank. Setting

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}} \\ \dot{\lambda} \end{bmatrix} = &\begin{bmatrix} \bar{P} & (\bar{S}(\hat{x}) + \bar{R}) \\ (\bar{S}(\hat{x}) + \bar{R})' & 0 \end{bmatrix}^{-1} \\ &\left(\begin{bmatrix} -\dot{P}\hat{x} + P\dot{\bar{x}} + \dot{P}\bar{x} \\ 0 \end{bmatrix} - \mu \begin{bmatrix} P(\hat{x} - \bar{x}) + (\bar{S}(\hat{x}) + \bar{R})\lambda \\ (\frac{1}{2}\bar{S}(\hat{x}) + \bar{R})'\hat{x} + \frac{1}{2}\bar{\rho} \end{bmatrix} \right), \end{aligned} \tag{13}$$

where μ is a positive constant, the sufficient conditions for optimality (10)–(12) hold asymptotically. Moreover, in the absence of numerical errors and setting $\hat{x}(0) = \hat{x}_0$, $\lambda(0) = 0$ equations (10)–(12) hold exactly for all times.

Before proving Theorem 2, note that $\bar{P}$ is indeed positive definite while λ is small, because P is always positive definite. Also, the rank condition on $\bar{S}(\hat{x}) + \bar{R}$ is the standard condition in Lagrange multiplier theory for constraints independence. Note also that one can implement (13) without differentiation since it follows from (7)–(8) that

$$-\dot{P}\hat{x} + P\dot{\bar{x}} + \dot{P}\bar{x} = P(A\hat{x} + b) - W\hat{x} + w \\ + (A + GG'P)'P(\hat{x} - \bar{x}).$$

Proof: We start by noting that as long as $\bar{P} > 0$ and $\bar{S}(\hat{x}) + \bar{R}$ is full column rank, the matrix

$$\begin{bmatrix} \bar{P} & (\bar{S}(\hat{x})+\bar{R}) \\ (\bar{S}(\hat{x})+\bar{R})' & 0 \end{bmatrix}$$

is nonsingular and therefore the right-hand-side of (13) is well defined. Defining

$$V(\hat{x}, \lambda) := \frac{1}{2}\|P(\hat{x} - \bar{x}) + \bar{S}(\hat{x}) + \bar{R} \ \lambda\|^2 \\ + \frac{1}{2}\|(\bar{S}(\hat{x}) + 2\bar{R})'\hat{x} + \bar{\rho}\|^2,$$

we obtain

$$\begin{aligned} \dot{V} = {} & P(\hat{x} - \bar{x}) + \bar{S}(\hat{x}) + \bar{R} \ \lambda \ ' \\ & \bar{P}\dot{\hat{x}} + \bar{S}(\hat{x}) + \bar{R})\dot{\lambda} - P\dot{\bar{x}} + \dot{P}\hat{x} - \dot{P}\bar{x} \\ & + (\bar{S}(\hat{x}) + 2\bar{R})'\hat{x} + \bar{\rho} \ '(2\bar{S}(\hat{x}) + 2\bar{R})'\dot{\hat{x}} \\ = {} & \begin{matrix} P(\hat{x} - \bar{x}) + (\bar{S}(\hat{x}) + \bar{R})\lambda \\ 2(\bar{S}(\hat{x}) + 2\bar{R})'\hat{x} + 2\bar{\rho} \end{matrix} \ ' \\ & \begin{matrix} \bar{P} & (\bar{S}(\hat{x}) + \bar{R}) \\ (\bar{S}(\hat{x}) + \bar{R})' & 0 \end{matrix} \ \begin{matrix} \dot{\hat{x}} \\ \dot{\lambda} \end{matrix} + \begin{matrix} \dot{P}\hat{x} - P\dot{\bar{x}} - \dot{P}\bar{x} \\ 0 \end{matrix} \\ = {} & -2\mu V, \end{aligned}$$

from which we conclude that $V \to 0$ as $t \to \infty$ and therefore (10) and (11) hold asymptotically. Equation (12) is a consequence of the fact that $\bar{P} > 0$. When $\hat{x}(0) = \hat{x}_0$ and $\lambda(0) = 0$, V is initially zero and will remain so for all times. ∎

III. Rigid body motion estimation using CCD cameras

In this section we illustrate how the result from the previous section can be used to estimate the position and orientation of a mobile robot using a CCD camera mounted on the robot that observes the apparent motion on the image of stationary points. The setup is adapted from [10], except that by enforcing state constraints we can avoid some of the singularities encountered in [10].

Consider a coordinate frame $\{b\}$ attached to a rigid body that moves with respect to an inertial frame $\{i\}$. We denote[1] by $(p_{ib}, R_{ib}) \in \mathrm{SE}(3)$ the configuration of the frame $\{b\}$ with respect to $\{i\}$. Thus, if q_1^i and q_1^b denote the coordinates of a point Q_1 in the frames $\{i\}$ and $\{b\}$, respectively, we have that

$$q_1^i = p_{ib} + R_{ib} q_1^b. \tag{14}$$

[1]We denote by $\mathrm{SE}(3)$ the Cartesian product of $\mathbb{R}^3$ with the group $\mathrm{SO}(3)$ of 3×3 rotation matrices; and by $\mathrm{se}(3)$ the Cartesian product of $\mathbb{R}^3$ with the space $\mathrm{so}(3)$ of 3×3 skew-symmetric matrices (cf.,e.g., [20]).

Moreover, if q_j^i and q_j^b denote the coordinates of another point Q_j in the frames $\{i\}$ and $\{b\}$, respectively, we conclude that

$$q_j^b = R_{ib}' q_j^i - R_{ib}' p_{ib} = R_{ib}'(q_j^i - q^i) + q_1^b.$$

We denote by $(v_{ib}^b, \Omega_{ib}^b) \in \mathrm{se}(3)$ the twist that defines the velocity of frame $\{b\}$ with respect to $\{i\}$, expressed in the frame $\{b\}$, i.e.,

$$\Omega_{ib}^b := R_{ib}'\dot{R}_{ib}, \qquad v_{ib}^b := R_{ib}'\dot{p}_{ib}.$$

¿From this and (14), we obtain

$$\dot{q}_1^b = -\Omega_{ib}^b q_1^b - v_{ib}^b + R_{ib}'\dot{q}_1^i, \qquad \dot{R}_{ib} = R_{ib}\Omega_{ib}^b.$$

Suppose now that a camera attached to the body frame $\{b\}$ sees k points $Q_1, Q_2, \ldots, Q_k$. Denoting by $y_j \in \mathbb{R}^3$ the homogeneous image coordinates provided by the camera of the point Q_j , the dynamics of the system can be described by the following system with k perspective outputs:

$$\dot{q}_1^b = -\Omega_{ib}^b q_1^b + R_{ib}'\dot{q}_1^i - v_{ib}^b, \tag{15}$$

$$\dot{R}_{ib}' = -\Omega_{ib}^b R_{ib}', \tag{16}$$

$$\alpha_j y_j = F\big(p_{cb} + R_{cb} q_1^b + R_{cb} R_{ib}'(q_j^i - q_1^i)\big), \tag{17}$$

$\forall j \in \{1, 2, \ldots, k\}$, where $(p_{cb}, R_{cb}) \in \mathrm{SE}(3)$ denotes the configuration of the frame $\{b\}$ with respect to the camera's frame $\{c\}$, and F an upper triangular matrix with the cameras intrinsic parameters of the form

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ 0 & f_{22} & f_{21} \\ 0 & 0 & 1 \end{bmatrix},$$

where each f_{ij} denotes a scalar. For ideal pin-hole cameras $f_{12} = f_{13} = f_{21} = 0$ and $f_{11} = f_{22} = f$, where f is the camera's focal length. For the general case, we refer the reader *e.g.* to [7] for details on the physical meaning of the entries of F. Note that F and (p_{cb}, R_{cb}) can be time-varying in case the camera is allowed to zoom or pan and tilt, which is often needed to get good visual information. The normalization constraints (3) are given by

$$[0 \quad 0 \quad 1]\, y_j = 1, \quad \forall j \in \{1, 2, \ldots, k\}.$$

To put this system in the form of (1)–(2) and use the results in Section II, we simply need to define x to be a 12-dimensional vector whose first 3 entries are the entries of q_1^b and the remaining 9 entries are the columns of R_{ib} stacked on top of each other. Since R_{ib} is a rotation matrix, it must satisfy the following constraints

$$r_{ib_x}' r_{ib_x} = r_{ib_y}' r_{ib_y} = 1, \ r_{ib_x}' r_{ib_y} = 0, \ r_{ib_z} = r_{ib_x} \times r_{ib_y}, \tag{18}$$

where r_{ib_x}, r_{ib_y}, and r_{ib_z} denotes the columns of R_{ib}. Note that (18) are indeed quadratic constraints as in (4).

Once we use the results in Section II to compute estimates $\hat{R}_{ib}$ and $\hat{q}_1^b$ for R_{ib} and q_1^b, respectively, we can also estimate p_{ib} using

$$\hat{p}_{ib} = q_1^i - \hat{R}_{ib}\hat{q}_1^b.$$

A. Camera calibration

We can also use the minimum-energy estimator proposed here to calibrate a camera, when its intrinsic parameters F are unknown. To this effect, we re-write the homogeneous image coordinates $y_j \in \mathbb{R}^3$ provided by the camera of the point Q_j, as

$$\alpha_j y_j = F p_{ci} + F R_{ci} q_j^i. \quad \forall j \in \{1, 2, \dots, k\},$$

where $(p_{ci}, R_{ci}) = (p_{cb} - R_{ci} p_{ib}, R_{cb} R'_{ib})$ denotes the configuration of the inertial frame $\{i\}$ with respect to the camera's frame $\{c\}$. Defining $\bar{p} := F p_{ci}$ and $\bar{R} := R'_{ci} F'$ and assuming F constant, we obtain $\forall j \in \{1, 2, \dots, k\}$

$$\dot{\bar{p}} = \bar{R}' v_{ci}^i \quad \dot{\bar{R}}' = \bar{R}' \Omega_{ci}^i \quad \alpha_j y_j = \bar{p} + \bar{R}' q_j^i, \tag{19}$$

where $(v_{ci}^i, \Omega_{ci}^i) \in \mathrm{se}(3)$ denotes the twist that defines the velocity of frame $\{i\}$ with respect to $\{c\}$, expressed in the frame $\{i\}$. This system can be put in the form of (1)–(2) by defining x to be a 12-dimensional vector whose first 3 entries are the entries of $\bar{p}$ and the remaining 9 entries are the columns of $\bar{R}$ stacked on top of each other. ¿From the definition of $\bar{R}$, we conclude that

$$\bar{R}'\bar{R} = FF' = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ 0 & f_{22} & f_{21} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{11} & 0 & 0 \\ f_{12} & f_{22} & 0 \\ f_{13} & f_{21} & 1 \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & 1 \end{bmatrix}.$$

Therefore $\bar{R}$ must satisfy the constraint

$$\bar{r}_z' \bar{r}_z = 1, \tag{20}$$

where $\bar{r}_z$ denotes the last column of $\bar{R}$. .

Once we compute estimates $\hat{p}$ and $\hat{R}$ for $\bar{p}$ and $\bar{R}$, respectively, we can use the QR decomposition to obtain an upper triangular matrix $\hat{F}$ and a rotation matrix $\hat{R}_{ci}$ such that

$$\hat{R}' := \hat{F} \hat{R}_{ci}.$$

These matrix can be viewed as estimates of F and R_{ci}, respectively. Note that (20) guarantees $\hat{F}$ has a one in the last element of the main diagonal. An estimate for p_{ci} can also be obtained from $\hat{p}_{ci} := \hat{F}^{-1}\hat{p}$.

B. Singular configurations

Depending on the configuration of the points $Q_1, Q_2, \dots, Q_k$, the state of (15)–(17) may not be observable. In fact, when all points are coplanar or collinear we can, respectively, find a 9- or 6-dimensional realization for the system. However, it is known that with at least 3 coplanar but non collinear points, it should be possible to recover the position and orientation of the rigid body from the camera measurements. This is possible by exploring the fact that R_{ib} is a rotation matrix and therefore the state of (15)–(17) must satisfy the constraints (18).

In [10], this difficulty was overcome by considering lower dimensional realizations of (15)–(17) for the cases of coplanar points. An unconstrained state-estimator was then constructed for these lower dimensional representations, and an estimate of the rotation matrix was computed from it. This approach however, exhibits two main difficulties: *i)* The lower-dimensional realization was still not minimal for some trajectories, e.g., trajectories for which the optical center of the camera did not translate. For these trajectories the estimation error would not converge to zero. *ii)* Structurally distinct approaches were needed to solve the case of coplanar and noncoplanar points. In practice, both approaches sometimes performed poorly when the point were "almost" coplanar but not quite.

By incorporating the constraints (18) in the state-estimation problem, there is no need to consider lower dimensional realizations for (15)–(17) when the points are coplanar but not collinear. However, when all points are collinear there may a fundamental loss of observability that cannot be overcome without extra information. Indeed, when there is a direction v such that the line segments between all points are all aligned with v, i.e.,

$$q_{j_1}^i - q_{j_2}^i = \beta_j v, \qquad \forall j_1, j_2 \in \{1, 2, \dots, k\},\ \beta_{j_1 j_2} \in \mathbb{R},$$

and the velocity of the points (if any) is also aligned with v, i.e., $\dot{q}_j^i = \gamma_j v$, $\forall j \in \{1, 2, \dots, k\}$,$\gamma_j \in \mathbb{R}$, the matrix $R_v \in \mathrm{SO}(3)$ that corresponds to fixed rotation around v has the property that

$$R_v \dot{q}_1^i = \dot{q}_1^i, \quad R_v(q_j^i - q_1^i) = q_j^i - q_1^i, \quad \forall j \in \{1, 2, \dots, k\}.$$

In this case,

$$\begin{aligned} \dot{q}_1^b &= -\Omega_{ib}^b q_1^b + \bar{R}'_{ib} \dot{q}_1^i - v_{ib}^b, \\ \dot{\bar{R}}'_{ib} &= -\Omega_{ib}^b \bar{R}'_{ib}, \\ \alpha_j y_j &= F\Big(p_{cb} + R_{cb} q_1^b + R_{cb} \bar{R}'_{ib}(q_j^i - q_1^i)\Big), \end{aligned}$$

$\bar{R}'_{ib} := R'_{ib} R_v$, which has exactly the same input-output map as (15)–(17). This means that R_{ib} can only be recovered from the perspective outputs up to a rotation around v.

C. Unknown inertial coordinates

When the inertial coordinates of the points $Q_1, Q_2, \dots, Q_k$ are not known one can still estimate x by using three of the points to define the inertial coordinate frame. To this effect, let

$$S := R'_{ib} \begin{bmatrix} q_2^i - q_1^i & q_3^i - q_1^i & \cdots & q_k^i - q_1^i \end{bmatrix}.$$

We can re-write (15)–(17) as

$$\begin{aligned} \dot{q}_1^b &= -\Omega_{ib}^b q_1^b + R'_{ib} \dot{q}_1^i - v_{ib}^b, \\ \dot{S} &= -\Omega_{ib}^b S, \\ \alpha_j y_j &= F\big(p_{cb} + R_{cb} q_1^b + R_{cb} S e_j\big), \quad \forall j \in \{1, 2, \dots, k\}, \end{aligned}$$

where e_j denotes the jth column of the $(k-1) \times (k-1)$ identity matrix.

To recover an estimate $\hat{R}_{ib}$ of R_{ib} from the estimate $\hat{S}$ of S, we can use the QR decomposition to obtain a rotation matrix $\hat{R}_{ib}$ and an upper triangular matrix $\hat{U}$ such that

$$\hat{S} = \hat{R}'_{ib} \hat{U},$$

and then defining $q_1^i := 0$ and q_j^i, $j \in \{2, 3, \dots, k\}$ equal to the $(j-1)$th column of $\hat{U}$. This corresponds to the following convention to construct the inertial coordinate frame: the origin of $\{i\}$ is the point Q_1; its first axis is defined by the direction from Q_1 to Q_2; its second axis is orthogonal to the first one and lies on the plane defines by Q_1, Q_2, and Q_3; and its third axis is defined by the cross product of the first two.

IV. Experimental results

The theoretical results presented in the previous sections were experimentally validated by applying them to estimate the position and orientation of a mobile robot using measurements from an on-board CCD camera. This section describes the experimental setup and presents the results achieved for two type of experiments: open-loop estimation and output feedback control.

A. Experimental Setup

(a) (b)

Fig. 1. (a) A Pionner 2-DXE with camera. (b) The object seen by the camera used for calibration.

The experiments were carried out on a Pionner 2-DXE mobile robot from ActivMedia [1]. The vehicle, shown in Fig. 1(a), has two rear wheels which are powered by two independent high torque, reversible-DC motors, and one passive rear caster to balance the robot. The robot hardware is controlled by the Pioneer 2 Operating System (P2OS) on-board, which runs on a siemens 88C166-based microcontroller. On-board, there is also a 800 MHz Pentium-based computer with 128 MB memory running RedHat Linux 7.2 that it is connected via a serial tether between its RS232 serial port and the robot's microcontroller. The vehicle is equipped with a Sony EVI D30 pan-tilt-zoom (PTZ) color video camera mounted on the top of the robot with its optical axis oriented towards the forward direction (when pan and tilt are zero). The PTZ robotic camera system is connected to a PC104+ framegrabber attached to the PCI bus of the on-board computer. The CCD camera has a resolution of 768×494 pixels.

The control algorithms operate at a sampling period of $0.1\,s$ and were implemented in C++ using the ARIA [1] software development environment. To detect the image features we used the image processing library VisLib created by ActivMedia [1] together with edge detector algorithm software SUSAN [2]. A resolution of 320×240 was selected to reduce the processing time. In these experiments we used a square black box as a feature and we were interesting in detecting the location of its corners (which correspond to the points q_i in the estimator algorithm). The location of these points were obtained by detecting the edges of the square and then computing their intersections.

B. Open-loop estimation

To validate the minimum-energy state estimator described in Section II and III, several open-loop tests were first carried out. As a simple test we ran the estimator with the vehicle stopped and commanding the pan angle of the camera.

Fig. 2 shows the experimental results of this test. The robot is at position $(x, y) = (-0.8\,m, 0\,m)$ with $\theta = 0\,rad$. The pan angle is set to zero. At $t = 20\,s$ a ramp-like signal is

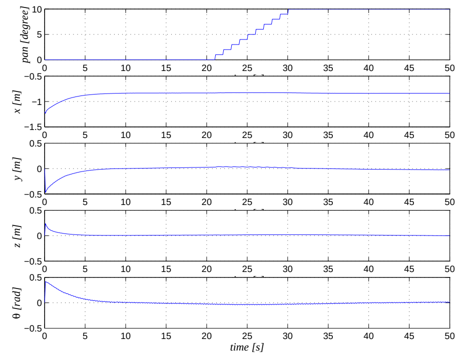


Fig. 2. Open-loop experiment: Time evolution of pan angle and estimator outputs $\hat{p}_{ib}$ and $\hat{\theta}$.

applied in pan until $t = 30\,s$. As shown in the figure, the estimator converges to the true values and is not affected by the camera's pan motion.

C. Output feedback control

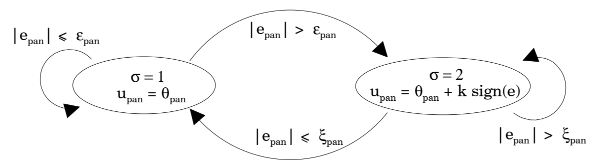


Fig. 3. Schematic of the hybrid pan motion control.

Pan motion control: A hybrid controller was designed and implemented for pan motion. Let σ be defined as a piecewise constant switching signal taking values in the set $\{1, 2\}$ and the error in pan as $e_{pan}(t) = \frac{\sum_{i=1}^{n} x_i(t)}{n} - \frac{W}{2}$, where n is the number of points, x_i is the x-coordinate of the i^{th} point expressed in the camera frame $\{c\}$, and W is the width of the image plane. The control law was designed according to the schematic presented in Fig. 3 and its goal is to keep the points near the center of the image.

Parking experiment: The Pionner 2-DXE is non-holonomic and has as control inputs, the velocities of the left and right wheels, v_1 and v_2. The kinematic model is given by

$$\dot{x} = v\cos(\theta), \quad \dot{y} = v\sin(\theta), \quad \dot{\theta} = \omega,$$

where v and ω denote the linear and angular velocity. These velocities which are the command inputs u_1, u_2 are related to v_1 and v_2 by $v_1 = v - l\omega$, $v_2 = v + l\omega$, where l is half the distance between the two robot wheels. In this section we present experimental results obtained using the minimum-energy state estimator described in Section II and III combined with the pan controller previous described and the point stabilization controller derived in [3]. The stabilization control law used is a piecewise time-invariant continuous feedback law and it is based on a non-smooth state transformation inspired by the polar description of the kinematics of the robot. The special attraction of this controller compared with others proposed in the literature is that its control strategy is intuitive, simple to implement, offers a good performance, and the resulting path of the mobile robot is more natural, *i.e.*, similar to what a human operator would attempt. However, as reported in [15], this controller is inherently sensitive to sensor noise and small perturbations around the origin.

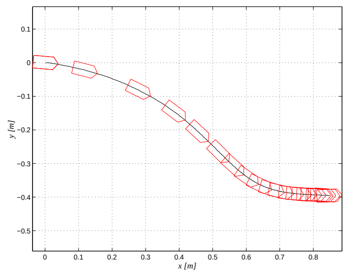


Fig. 4. Parking experiment: Resulting trajectory of the mobile robot in the xy-plane using dead-reckoning data measurements.

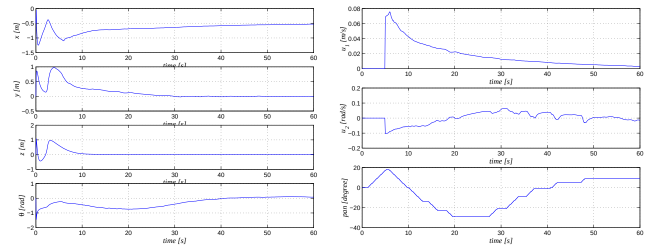
Fig. 4 and Fig. 5 show the experimental results for the case where the vehicle starts stopped with an initial condition $(x, y) = (-0.8\,m, 0.4\,m)$ and with heading $\theta = 0$. The objective was to park the vehicle at position $(x, y) = (-0.5\,m, 0\,m)$ with heading $\theta = 0$. Fig. 4 displays the resulting vehicle trajectory using the dead-reckoning data measurements. Note that dead-reckoning only gives the pose of the mobile robot with respect to the starting initial condition and is therefore not useful to park at a point specified in a referential frame defined by visual landmarks. Fig. 5(a) shows the time evolution of estimator outputs $\hat{p}_{ib}$ and $\hat{\theta}$, and Fig. 5(b) the time evolution of control signals u_1, u_2, and u_{pan}. In this experiment we imposed that the point stabilization control algorithm only starts to operate at time $t = 5\,s$ and the pan controller at time $t = 1\,s$. ¿From the figures it can be seen that the vehicle converges to a very small neighborhood of the desired pose. From Fig. 5(b) it can be observed that the pan controller is indeed able to compensate the motion of the robot in order to keep the features in the center of the image.

V. Conclusions

We considered the problem of estimating the state of a system with perspective outputs, whose state is known to satisfy a set of quadratic constraints. We designed a dynamical system that produces an estimate of the state that satisfies the constraints and is ”most compatible” with the dynamics, in the sense that it requires the least amount of noise energy to explain the measured output. The use of state constraints in the estimator allows the system to operate even when the trajectories for the robot do not exhibit “excitation,” which was needed by previous minimum-energy estimators. These results were validated experimentally by using measurements from a CCD camera mounted on a mobile robot to estimate its position and orientation. The position estimates are then used to close the loop and control the robot to a desired position, defined with respect to visual landmarks. The experimental results showed good performance.

VI. Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. ECS-0242798. The work of António Aguiar was supported by a Pos-Doc Fellowship PRAXIS XXI from the Portuguese Foundation of Science and Technology.



(a) (b)

Fig. 5. Parking experiment: Time evolution of (a) estimator outputs $\hat{p}_{ib}$ and $\hat{\theta}$; and (b) control signals $u_1(t)$, $u_2(t)$, and $u_{pan}(t)$.

References

[1] ActivMedia. URL: http://www.activmedia.com/.

[2] Susan. URL: http://www.fmrib.ox.ac.uk/ steve/susan/.

[3] M. Aicardi, G. Casalino, A. Bicchi, and A. Balestrino. Closed loop steering of unicycle-like vehicles via Lyapunov techniques. *IEEE Rob. & Autom. Mag.*, 2(1):27–35, 1995.

[4] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999.

[5] A. Chiuso, P. Favaro, H. Jin, and S. Soatto. Structure from motion causally integrated over time. To appear in *IEEE Trans. on Pattern Anal. Mach. Intell.*, 2002.

[6] W. P. Dayawansa, B. K. Ghosh, C. Martin, and X. Wang. A necessary and sufficient condition for the perspective observability problem. *Syst. & Contr. Lett.*, 25:159–166, 1995.

[7] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, Cambridge, MA, 1993.

[8] B. K. Ghosh, M. Jankovic, and Y. T. Wu. Perspective problems in system theory and its application in machine vision. *J. Mathematical Syst. Estimation Contr.*, 4(1):3–38, 1994.

[9] B. K. Ghosh and E. P. Loucks. A perspective theory for motion and shape estimation in machine vision. *SIAM J. Contr. Optimization*, 33 (5):1530–1559, 1995.

[10] J. P. Hespanha. State estimation and control for systems with perspective outputs. In *Proc. of the 41th Conf. on Decision and Contr.*, Dec. 2002.

[11] O. J. Hijab. *Minimum Energy Estimation*. PhD thesis, University of California, Berkeley, 1980.

[12] H. Inaba, A. Yoshida, R. Abdursul, and B. K. Ghosh. Observability of perspective dynamical systems. In *Proc. of the 39th Conf. on Decision and Contr.*, volume 5, pages 5157–5162, 2000.

[13] M. Jankovic and Ghosh. Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer. *Syst. & Contr. Lett.*, 25:63–73, 1995.

[14] I. Kaminer, A. M. Pascoal, W. Kang, and O. Yakimenko. Integrated vision/inertial navigation systems design using nonlinear filtering. *IEEE Trans. Aerospace and Electronic Syst.*, 37(1):158–172, Jan. 2001.

[15] B. Kim and P. Tsiotras. Controllers for unicycle-type wheeled robots: Theoretical results and experimental validation. *IEEE Trans. Robot. Automat.*, 18(3):294–307, 2002.

[16] A. J. Krener. The convergence of the minimum energy estimator. In W. Kang, editor, *New Trends in Nonlinear Dynamics and Control, and Their Applications*, Lecture Notes in Control and Information Sciences. Springer-Verlag, 2003. To appear.

[17] L. Matthies, T. Kanade, and R. Szeliski. Kalman filter-based algorithms for estimating depth from image sequences. *Int. J. of Comput. Vision*, 3:209–236, 1989.

[18] W. M. McEneaney. Robust/H-infinity filtering for nonlinear systems. *Syst. & Contr. Lett.*, (33):315–325, 1998.

[19] R. E. Mortensen. Maximum likelihood recursive nonlinear filtering. *J. Opt. Theory and Applications*, 2:386–394, 1968.

[20] R. M. Murray, Z. Li, and S. S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CDC Press, Boca Raton, Florida, 1994.

[21] H. Rehbinder and B. K. Ghosh. Pose estimation using line-based dynamic vision and inertial sensors. *IEEE Trans. on Automat. Contr.*, 48(2):186–199, Feb. 2003.

[22] S. Soatto, R. Frezza, and P. Perona. Motion estimation via dynamic vision. *IEEE Trans. on Automat. Contr.*, 41(3):393–413, Mar. 1996.

[23] S. Takahashi and B. K. Ghosh. Motion and shape parameters identification with vision and range. In *Proc. of the 2001 Amer. Contr. Conf.*, volume 6, pages 4626–4631, June 2001.