

# State Estimation for Systems with Implicit Outputs for the Integration of Vision and Inertial Sensors

A. Pedro Aguiar and João P. Hespanha

**Abstract**—This paper addresses the state estimation of a system with implicit outputs. We formulate the problem in the so-called deterministic  $H_\infty$  filtering setting by computing the value of the state that minimizes the induced  $\mathcal{L}_2$ -gain from disturbances to estimation error, while remaining compatible with the past observations. To avoid weighting the distant past as much as the present, a forgetting factor is also introduced. We show that, under appropriate observability assumptions, the optimal estimate converges globally asymptotically to the true value of the state in the absence of noise and disturbance. In the presence of noise, the estimate converges to a neighborhood of the true value of the state. We apply these results to the estimation of position and attitude of an autonomous vehicle using measurements from an inertial measurement unit (IMU) and a monocular charged-coupled-device (CCD) camera attached to the vehicle.

## I. INTRODUCTION

During the last few decades there has been an extensive study on the design of observers for nonlinear systems. We consider a new class of systems with nonlinear output equations, called *systems with implicit outputs*. These systems arise when one needs to estimate the pose (position and attitude) of autonomous vehicles using measurements from an inertial measurement unit (IMU) and a monocular charged-coupled-device (CCD) camera attached to the vehicle.

Consider a continuous-time system described by

$$\dot{x} = A(x, u) + G(u)w, \quad (1)$$

$$E(x, v)y = C(x, u) + v, \quad (2)$$

where  $A(x, u)$ ,  $C(x, u)$ ,  $E(x, v)$  are affine functions in  $x$ ,  $x \in \mathbb{R}^n$  denotes the state of the system,  $u \in \mathbb{R}^m$  its control input,  $y \in \mathbb{R}^q$  its measured output,  $w \in \mathbb{R}^{n_w}$  an input disturbance that cannot be measured, and  $v \in \mathbb{R}^p$  measurement noise. The initial condition  $x(0)$  of (1) and the signals  $w$  and  $v$  are assumed deterministic but unknown. The measured output  $y$  is only defined implicitly through (2) and the map  $E(x, \cdot)$  is such that

$$\text{Image } E(x, \cdot) = \left\{ E_0(x) + \sum_{i=1}^k \alpha_i E_i : \alpha_i \in \mathbb{R} \right\} \quad (3)$$

Research supported by NSF grants ECS-0242798 and CCR-0311084. Part of this research was done while the first author was with the Center for Control Engineering and Computation, University of California, Santa Barbara, CA 93106-9560.

A. Pedro Aguiar is with the ISR/IST Institute for Systems and Robotics, Instituto Superior Técnico, Torre Norte 8, Av. Rovisco Pais, 1049-001 Lisboa, Portugal. E-mail: pedro@isr.ist.utl.pt

João P. Hespanha is with Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560, USA. E-mail: hespanha@ece.ucsb.edu

where  $E_i \in \mathbb{R}^{p \times q}$  and  $E_0(x)$  is affine in  $x$ . We call (1)–(3) a *state-affine system with implicit outputs*, or for short simply a *system with implicit outputs*. These type of systems are motivated by applications in dynamic vision such as the estimation of the motion of the camera from a sequence of images. System (1)–(3) can be seen as a generalization of perspective systems introduced by Ghosh et al. [1]. The reader is referred to [2], [3] for several other examples of perspective systems in the context of motion and shape estimation. See also [4], [5] that address the observability problem of perspective linear systems. The system with implicitly defined outputs described in [6] and the state-affine systems with multiple perspective outputs considered in [7] (see also [8]–[10]) are also special cases of (1)–(3).

In this paper we design a state-estimator for (1)–(3) using a deterministic  $H_\infty$  approach. Given an initial estimate and the past controls and observations collected up to time  $t$ , the optimal state estimate  $\hat{x}$  at time  $t$  is defined to be the value that minimizes the induced  $\mathcal{L}_2$ -gain from disturbances to estimation error. To avoid weighting the distant past as much as the present, a forgetting factor  $\lambda$  is also introduced.

Over the last two decades the  $H_\infty$  criterion has been applied to filtering problems, cf., e.g., [11]–[15]. See also [16], where  $H_\infty$  filtering techniques are discussed and similarities of the type error bounds with those obtained for the stochastic risk-sensitive filter are pointed out. Closely related to  $H_\infty$  filtering are the minimum-energy estimators, which were first proposed by Mortensen [17] and further refined by Hijab [18]. Game theoretical versions of these estimators were proposed by McEneaney [19]. In [7], minimum-energy estimators were derived for systems with perspective outputs and input-to-state stability like properties with respect to disturbances were presented. These results were also extended to solve the estimation problem when the measured outputs are transmitted through a network, that is, when the measurements arrive at discrete-time instants, are time-delayed, noisy, and may not be complete.

It is worth pointing out that in general, for nonlinear systems, either minimum-energy, or  $H_\infty$  state estimators approaches lead to infinite dimensional observers with state evolving according to a first-order nonlinear PDE of Hamilton-Jacobi type driven by the observations. Due to the particular structure of (1)–(3) in this paper we obtain a closed-form solution that resembles a Kalman-Bucy filter. Furthermore, under appropriate observability assumptions, we show that the state-estimator proposed has the desirable property that the state estimate converges asymptotically to the true value of the state in the absence of noise and

disturbance. In the presence of noise, the estimate converges to a neighborhood of the true value of the state. We can therefore use this state-estimator to design output-feedback controllers by using the estimated state to drive state-feedback controllers.

A fundamental problem in autonomous vehicles is the determination of position and attitude with respect to an inertial coordinate system. A promising solution is to combine the measurements provided by an inertial measurement unit (IMU) and a monocular charged-coupled-device (CCD) camera mounted on-board. 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 (cf., e.g., [20]–[25] and references therein). 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 are especially desirable. Soatto et al. [22] formulated the visual motion estimation problem in terms of identification of nonlinear implicit systems with parameters on a topological manifold, and propose a dynamic solution either in the local coordinates, or in the embedding space of the parameter manifold. In [25], rigid-body pose estimation using inertial sensors and a monocular camera is considered. A local convergent observer where the states evolve on  $SO(3)$  is proposed (the rotation estimation is decoupled from the position estimation). In the area of wheeled mobile robots, Ma et al. [26] addressed the problem of tracking an arbitrarily shaped continuous ground curve by formulating it as controlling the shape of the curve in the image plane. An application for landing an unmanned air vehicle using vision in the control loop is described in [27]. In [23], the autonomous aircraft landing problem based on measurements provided by airborne vision and inertial sensors is addressed. The authors cast the problem in a linear parametrically varying framework and solve it using tools that borrows from the theory of linear matrix inequalities. These results are extended in [28] to deal with the so-called out-of-frame events.

In Section V we formulate the problem of estimating the pose of an autonomous vehicle using measurements from an IMU and a monocular CCD camera attached to the vehicle. The problem is then solved by using the  $H_\infty$  state-estimators derived in the previous sections. One of the main contributions is that, contrary to what happens with most previous algorithms, the ones proposed here are globally convergent provided that suitable observability assumptions are satisfied. Global convergence to the correct position and orientation is achieved in the absence of noise. When there is noise, the magnitude of the estimation error is essentially proportional to the amount of noise.

Due to space limitations, only outlines of the proofs are provided.

## II. PROBLEM FORMULATION

This section formulates the state estimation problem using a  $H_\infty$  deterministic approach.

Consider the system with implicit outputs (1)–(3). Our goal is to design an observer which estimates the state vector  $x(t)$  given an initial estimate  $\hat{x}_0$  and the past controls and observations  $\{(u(\tau), y(\tau)) : 0 \leq \tau \leq t\}$ , and minimize the induced  $\mathcal{L}_2$ -gain from disturbances to estimation error. In particular, for a given  $\gamma > 0$ , the estimate  $\hat{x}$  should satisfy

$$\int_0^t \|x(\tau) - \hat{x}(\tau)\|^2 d\tau \leq \gamma^2 \left( (x(0) - \hat{x}_0)' P_0 (x(0) - \hat{x}_0) + \int_0^t \|w(\tau)\|^2 + \|v(\tau)\|^2 d\tau \right)$$

where  $P_0 > 0$ ,  $\hat{x}_0$  encode a-priori information about the state. To avoid the problem of weighting the distant past as much as the present, we introduce a forgetting factor  $\lambda > 0$ . More specifically, we address the following deterministic optimization problem:

*Problem 1 ( $H_\infty$  state estimation):* Given an initial estimate  $\hat{x}_0$ , an input  $u$  and a measured output  $y$  defined on an interval  $[0, t)$ , compute the estimate  $\hat{x}(t)$  of the state at time  $t$  defined by

$$\hat{x}(t) := \arg \min_{z \in \mathbb{R}^n} J(z, t), \quad (4)$$

where

$$\begin{aligned} J(z; t) := & \min_{w, v} \left\{ \gamma^2 e^{-2\lambda t} (x(0) - \hat{x}_0)' P_0 (x(0) - \hat{x}_0) \right. \\ & + \gamma^2 \int_0^t e^{-2\lambda(t-\tau)} (\|w(\tau)\|^2 + \|v(\tau)\|^2) d\tau \\ & - \int_0^t e^{-2\lambda(t-\tau)} \|x(\tau) - \hat{x}(\tau)\|^2 d\tau : \\ & x(t) = z, \dot{x} = A(x, u) + G(u)w, \\ & \left. E(x, v)y = C(x, u)x + v \right\}. \end{aligned} \quad (5)$$

□

The negative of  $J(z, t)$  is the *information state* introduced in [29], [30] and can be interpret as a measure of the likelihood of state  $x = z$  at time  $t$ .

## III. THE OBSERVER EQUATIONS

We propose the following observer for (1)–(3):

$$\begin{aligned} \dot{Q} = & (\nabla A(u) + \lambda I)Q + Q(\nabla A(u) + \lambda I)' \\ & - \gamma^2 Q(\Psi - \gamma^{-2}I)Q + \gamma^{-2}GG', \quad Q(0) = P_0^{-1} \end{aligned} \quad (6)$$

$$\dot{\hat{x}} = A(\hat{x}, u) - \gamma^2 Q \Psi \hat{x} - \gamma^2 Q \psi, \quad \hat{x}(0) = \hat{x}_0 \quad (7)$$

where

$$\begin{aligned} \Psi(t) := & (\mathbf{J}E_0(y) - \nabla C)'(I - YY^\perp)' \times \\ & (I - YY^\perp)(\mathbf{J}E_0(y) - \nabla C), \\ \psi(t) := & (\mathbf{J}E_0(y) - \nabla C)'(I - YY^\perp) \times \\ & (I - YY^\perp)(E_0(0)y - C(0, u)), \end{aligned}$$

$Y := [E_1 y | E_2 y | \dots | E_k y]$ ,  $Y^\perp$  is its pseudo-inverse,  $\nabla A(u)$  and  $\mathbf{J}E_0(y)$  are respectively the gradient of  $A(x, u)$  and the Jacobian of  $E_0(x)y$  both with respect to  $x$ . Since  $A(x, u)$

and  $E_0(x)$  are affine in  $x$ , it follows that  $\nabla A(\cdot)$  and  $\mathbf{J}E_0(\cdot)$  only depend on  $u$  and  $y$ , respectively.

The following result solves Problem 1.

*Theorem 1 ( $H_\infty$  state estimator):* Assuming that (6)–(7) has a solution on  $[0, T]$ ,  $T \in [0, \infty]$ , then the  $H_\infty$  state estimate defined by (4)–(5) can be obtained from (6)–(7). Furthermore, the cost function  $J(z; t)$  defined in (5) is quadratic and can be written as

$$J(z; t) = (z - \hat{x}(t))' P(t) (z - \hat{x}(t)) + c(t), \quad (8)$$

where  $c(t)$  satisfies an appropriate ODE (cf. (13) below).

*Proof:* The function  $J(z; t)$ ,  $z \in \mathbb{R}^n$ ,  $t \geq 0$  defined in (5) can be viewed as a cost-to-go and computed using dynamic programming. After some algebraic manipulation and exploring the fact that  $E(z, v)$  satisfies (3), we conclude that (see [31])

$$\begin{aligned} J_t(z; t) = & -\frac{1}{4\gamma^2} \|G'(u)J'_z(z; t)\|^2 + \gamma^2 (E_0(z)y - C(z, u))' \\ & (I - YY^\perp)' (I - YY^\perp) (E_0(z)y - C(z, u)) \\ & - \|z - \hat{x}\|^2 - J_z(z; t)A(z, u) - 2\lambda J(z; t), \end{aligned} \quad (9)$$

where  $J_t$  and  $J_z$  denote the partial derivatives of  $J$  with respect to  $t$  and  $z$ , respectively. The value of  $J(z; t)$  can be determined from the linear partial differential equation (9) with initial condition

$$J(z; 0) = (z - \hat{x}_0)' P_0 (z - \hat{x}_0), \quad z \in \mathbb{R}^n. \quad (10)$$

It turns out that there exists a solution to (9)–(10) which is differentiable with respect to  $z$  and can be written as (8) for appropriately defined signals  $\hat{x}(t)$  and  $c(t)$ . The signal  $\hat{x}$  is then precisely the estimate for the state  $x$  of (1)–(3). Moreover, matching (10) with (8) we conclude that  $P(0) = P_0$ ,  $\hat{x}(0) = \hat{x}_0$ , and  $c(0) = 0$ . To verify that the solution to (9)–(10) can indeed be written as (8), we substitute it into (9) and using the fact that  $A(z, u) = \nabla A(u)z + A(0, u)$  and  $E_0(z)y = \mathbf{J}E_0(y)z + E_0(0)y$ , we get

$$\begin{aligned} z' \left[ \dot{P} + \frac{1}{\gamma^2} PGG'P + P\nabla A + \nabla A'P + 2\lambda P - \gamma^2 \Psi + I \right] z \\ + 2z' \left[ -P\dot{\hat{x}} - \dot{P}\hat{x} - \frac{1}{\gamma^2} PGG'P\hat{x} - \nabla A'P\hat{x} + PA(0, u) \right. \\ \left. - 2\lambda P\hat{x} - \gamma^2 \psi - \hat{x} \right] + \dot{c} + 2\hat{x}' P\dot{\hat{x}} + \hat{x}' \dot{P}\hat{x} + \frac{1}{\gamma^2} \hat{x}' PGG'P\hat{x} \\ - \hat{x}' PA(0, u) + 2\lambda \hat{x}' P\hat{x} + 2\lambda c - \gamma^2 (E_0(0)y - C(0, u))' \times \\ (I - YY^\perp)' (I - YY^\perp) (E_0(0)y - C(0, u)) + \|\hat{x}\|^2 = 0. \end{aligned}$$

Since this equation must hold for all  $z \in \mathbb{R}^n$  we conclude that

$$\begin{aligned} \dot{P} + \frac{1}{\gamma^2} PGG'P + P\nabla A + \nabla A'P + 2\lambda P \\ - \gamma^2 \Psi + I = 0 \end{aligned} \quad (11)$$

$$\begin{aligned} -P\dot{\hat{x}} - \dot{P}\hat{x} - \frac{1}{\gamma^2} PGG'P\hat{x} - \nabla A'P\hat{x} \\ + PA(0, u) - 2\lambda P\hat{x} - \gamma^2 \psi - \hat{x} = 0 \end{aligned} \quad (12)$$

$$\begin{aligned} \dot{c} + 2\hat{x}' P\dot{\hat{x}} + \hat{x}' \dot{P}\hat{x} + \frac{1}{\gamma^2} \hat{x}' PGG'P\hat{x} \\ - \hat{x}' PA(0, u) + 2\lambda \hat{x}' P\hat{x} + 2\lambda c - \gamma^2 (E_0(0)y \\ - C(0, u))' (I - YY^\perp)' (I - YY^\perp) \times \\ (E_0(0)y - C(0, u)) + \|\hat{x}\|^2 = 0 \end{aligned} \quad (13)$$

Substituting (11) in (12), we obtain

$$\begin{aligned} -\dot{P} = P(\nabla A + \lambda I) + (\nabla A + \lambda I)' P \\ + \gamma^{-2} PGG'P - \gamma^2 \Psi + I \end{aligned} \quad (14)$$

$$P\dot{\hat{x}} = PA(\hat{x}, u) - \gamma^2 \Psi \hat{x} - \gamma^2 \psi \quad (15)$$

Since the solution  $Q$  to (6) is positive definite (cf. Lemma 2 in Appendix), then  $Q^{-1}$  is defined on  $[0, T]$ . Using the fact that  $\dot{Q}^{-1} = -Q^{-1}\dot{Q}Q^{-1}$ , it is straightforward to show that both  $Q^{-1}$  and  $P$  satisfy (14). Since  $Q(0) = P_0^{-1}$ , by unicity of solution  $Q(t) = P^{-1}(t)$ ,  $\forall t \in [0, T]$ , and therefore (14)–(15) and (6)–(7) are equivalent. ■

To guarantee that the  $H_\infty$  state estimate has a global solution ( $T = \infty$ ), the value of  $\gamma$  should be sufficiently large. In particular, a sufficient condition for this is given by the following result.

*Lemma 1 (Observability condition):* The  $H_\infty$  estimator (6)–(7) has a global solution and

$$P(t) \geq \delta I > 0, \quad \forall t \geq 0, \quad (16)$$

for some  $\delta > 0$ , if there exists a sufficiently large  $\gamma > 0$  such that the following condition

$$\gamma^2 W_0(t) \geq \int_0^t \Phi(\tau, t)' \Phi(\tau, t) d\tau + \delta I \quad \forall t \geq 0 \quad (17)$$

holds, where

$$W_0(t) := \int_0^t \Phi(\tau, t)' \Psi(\tau) \Phi(\tau, t) d\tau, \quad (18)$$

and  $\Phi(t, \tau)$  denotes the state transition matrix of  $\dot{z} = (\nabla A + \gamma^{-2}GG'P + \lambda I)z$ .

*Proof:* Due to space limitations, the proof is omitted. It can be found in [31]. ■

#### IV. ESTIMATOR CONVERGENCE

We are now interesting in determining under what conditions does the state estimate  $\hat{x}$  converges to the true state  $x$ .

*Theorem 2 (Convergence):* Assuming that the solutions to the system with implicit outputs (1)–(3) and to state estimator (6)–(7) exist on  $[0, T]$ ,  $T \in [0, \infty]$ , and  $P(t) \geq \delta I$ , then there exist positive constants  $c, \kappa, \gamma_w, \gamma_v$  such that

$$\begin{aligned} \|\tilde{x}(t)\| \leq ce^{-\kappa t} \|\tilde{x}(0)\| + \gamma_w \sup_{\tau \in [0, t]} \|w(\tau)\| \\ + \gamma_v \sup_{\tau \in [0, t]} \|v(\tau)\|, \quad \forall t \in [0, T] \end{aligned} \quad (19)$$

where  $\tilde{x}(t) := \hat{x}(t) - x(t)$  denotes the state estimation error.

*Proof:* From (1) and (7) we conclude that

$$\begin{aligned} \dot{\tilde{x}} = (\nabla - \gamma^2 Q\Psi)\tilde{x} - \gamma^2 Q[(\mathbf{J}E_0(y) - \nabla C)'(I - YY^\perp)' \times \\ (I - YY^\perp)v - Gw \end{aligned} \quad (20)$$

Defining  $V(\tilde{x}) := \tilde{x}'P\tilde{x}$ ,  $P := Q^{-1}$ , computing its time-derivative, and using (14), we get

$$\begin{aligned} \dot{V} = & -\tilde{x}' \left( \frac{1}{\gamma^2} PGG'P + 2\lambda P + \gamma^2 \Psi + I \right) \tilde{x} - 2\tilde{x}' PGw \\ & - 2\gamma^2 (\mathbf{J}E_0(y) - \nabla C)' (I - YY^\perp)' (I - YY^\perp) v \end{aligned}$$

By completing the squares, we further conclude that

$$\dot{V} \leq -2(\lambda - \theta)V - \frac{1}{\lambda_{max}(P)}V, \quad V \geq \frac{\gamma^2}{2\theta} \|w\|^2 + \frac{\gamma^2}{2\theta} \|v\|^2,$$

where  $\theta \in (0, \lambda)$ . It is now straightforward to conclude that the input-to-state stability bound (19) holds with  $c = \sqrt{\frac{\lambda_{max}(P_0)}{\delta}}$ ,  $\kappa = \lambda - \theta + \frac{1}{2\lambda_M}$ ,  $\gamma_w = \gamma_v = \frac{\gamma}{\sqrt{2\delta\theta}}$ , where  $\lambda_M = \sup_{\tau \in [0, T]} \lambda_{max}(P(\tau))$ . ■

## V. AUTONOMOUS VEHICLES MOTION ESTIMATION USING CCD CAMERAS AND INERTIAL SENSORS

In this section we show how one can estimate the position and attitude of an autonomous vehicle with respect to an inertial coordinate frame defined by visual landmarks using both measurements from an inertial measurement unit (IMU) and a monocular charged-coupled-device (CCD) camera mounted on-board. We do this by reducing the problem to the estimation of the state of a system with implicit outputs of the form (1)–(3).

### A. Kinematic equations of motion

Let  $\{\mathcal{V}\}$  be an inertial coordinate frame defined by visual landmarks and  $\{\mathcal{B}\}$  a body-fixed coordinate frame whose origin is located e.g. at the center of mass of the vehicle. The configuration of the vehicle  $({}^V_B R, {}^V P_B)$  or for simplicity of notation  $(R, p)$ , is an element of the Special Euclidean group  $SE(3) := SO(3) \times \mathbb{R}^3$ , where  $R \in SO(3) := \{R \in \mathbb{R}^{3 \times 3} : RR' = I_3, \det(R) = +1\}$  is a rotation matrix that describes the orientation of the vehicle by mapping body coordinates into inertial coordinates, and  $p \in \mathbb{R}^3$  is the position of the origin of  $\{\mathcal{B}\}$  in  $\{\mathcal{V}\}$ . Denoting by  $v \in \mathbb{R}^3$  and  $\omega \in \mathbb{R}^3$  the linear and angular velocities of the vehicle relative to  $\{\mathcal{V}\}$  expressed in  $\{\mathcal{B}\}$ , respectively, the following kinematic relations apply:

$$\dot{p} = Rv, \quad (21)$$

$$\dot{R} = RS(\omega), \quad (22)$$

where

$$S(x) := \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \quad \forall x := (x_1, x_2, x_3) \in \mathbb{R}^3.$$

The objective is to determine the position  $p$  and orientation  $R$  with respect to the visual coordinate system  $\{\mathcal{V}\}$ .

### B. Sensor Measurements

We consider that the IMU provides the vehicle's linear velocity  $v$ , angular velocity  $\omega$ , and pose (position and attitude) with respect to some inertial coordinate frame  $\{\mathcal{I}\}$ . It is assumed that the position and orientation of  $\{\mathcal{I}\}$  with

respect to the visual coordinate frame  $\{\mathcal{V}\}$  are unknown. The measurements are denoted by

$$\begin{aligned} \zeta_1 &= v, \\ \zeta_2 &= \omega, \\ \zeta_3 &= {}^I P_B, \\ \zeta_4 &= {}^I R, \end{aligned} \quad (23)$$

$$\zeta_4 = {}^I R, \quad (24)$$

where  $\zeta_1 \in \mathbb{R}^3, \zeta_2 \in \mathbb{R}^3, \zeta_3 \in \mathbb{R}^3, \zeta_4 \in SO(3)$ , and  $({}^I_B R, {}^I P_B)$  denotes the configuration of the frame  $\{\mathcal{B}\}$  with respect to frame  $\{\mathcal{I}\}$ .

We also suppose that there is a camera attached to the vehicle that sees  $N$  points  $Q_i = (x_i, y_i, z_i)'$ ,  $i = 1, 2, \dots, N$  whose coordinates are known for the visual coordinate system. Denoting by  $\zeta_{i+4} \in \mathbb{R}^3$  the homogeneous image coordinates provided by the camera of the point  $Q_i$ , the following relationships apply:

$$\mu_{i+4} \zeta_{i+4} = F {}^C Q_i, \quad (25)$$

$$[0 \ 0 \ 1] \zeta_{i+4} = 1, \quad \forall i \in \{1, 2, \dots, N\} \quad (26)$$

where  ${}^C Q_i$  is the position of  $Q_i$  expressed in the camera's frame,  $\mu_{i+4} \in \mathbb{R}$  captures the depth of the point  ${}^C Q_i$  (which is unknown), and  $F$  is an upper triangular matrix with the camera's intrinsic parameters, of the form

$$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ 0 & f_{22} & f_{23} \\ 0 & 0 & 1 \end{bmatrix},$$

where each  $f_{ij}$  denotes a scalar [32, Chapter 3].

Given the measurements  $\zeta_i$ ,  $i = 1, \dots, N + 4$ , we now proceed with the formulation of a system with implicit outputs.

### C. System with implicit outputs

Let  ${}^V Q_1$  and  ${}^B Q_1$  be the coordinates of a point  $Q_1$  in the frames  $\{\mathcal{V}\}$  and  $\{\mathcal{B}\}$ , respectively. Then, we have that

$${}^V Q_1 = {}^V P_B + {}^V_B R {}^B Q_1. \quad (27)$$

From this and (21)–(22), we obtain the state equations

$${}^B \dot{Q}_1 = {}^V_B R' {}^V \dot{Q}_1 - S(\omega) {}^B Q_1 - v, \quad (28)$$

$${}^V_B \dot{R} = -S(\omega) {}^V_B R'. \quad (29)$$

To obtain the output equations for the vision system, we first note that if  ${}^V Q_j$  and  ${}^B Q_j$  denote the coordinates of another point  $Q_j$  in the frames  $\{\mathcal{V}\}$  and  $\{\mathcal{B}\}$ , respectively, we conclude that

$$\begin{aligned} {}^B Q_j &= {}^V_B R' {}^V Q_j - {}^V_B R' {}^V P_B \\ &= {}^V_B R' ({}^V Q_j - {}^V Q_1) + {}^B Q_1. \end{aligned}$$

Using now (25) and the fact that  ${}^C Q_i = {}^C P_B + {}^C_B R {}^B Q_i$ , we obtain the output equations

$$\begin{aligned} \mu_{i+4} \zeta_{i+4} &= F \left( {}^C P_B + {}^C_B R {}^V_B R' ({}^V Q_i - {}^V Q_1) \right. \\ &\quad \left. + {}^C_B R {}^B Q_1 \right), \quad \forall i \in \{1, 2, \dots, N\} \end{aligned} \quad (30)$$

where  $({}^C_B R, {}^C P_B) \in SE(3)$  denotes the configuration of the frame  $\{\mathcal{B}\}$  with respect to the camera's frame  $\{\mathcal{C}\}$ .

We will regard  $\zeta_1$  and  $\zeta_2$  as inputs to the implicit output system. The dynamics of (23)–(24) can be written as

$$\dot{P}_I = -S(\omega)P_I, \quad (31)$$

$${}^V_I\dot{R}' = 0, \quad (32)$$

with the output equations

$$\zeta_3 = {}^V_B R' V Q_1 - {}^B Q_1 - P_I, \quad (33)$$

$$\zeta_4 {}^V_I R' = {}^V_B R', \quad (34)$$

where  $P_I := {}^V_B R' V P_I$ . Thus, our implicit output system is composed by (31)–(32), (28)–(29), (33)–(34) and (30). We now need to rewrite it in the form (1)–(3).

To proceed we use the following notation: Given an  $m \times n$ -matrix  $M$ , we denote by  $\text{stack}(M)$  the  $mn$ -vector obtained from stacking the columns of  $M$  one on top of each other, with the first column on top. Given two matrices  $M_i \in \mathbb{R}^{m_i \times n_i}$ ,  $i \in \{1, 2\}$  we denote by  $M_1 \otimes M_2 \in \mathbb{R}^{m_1 m_2 \times n_1 n_2}$  the Kronecker product of  $M_1$  by  $M_2$ . Using the fact that given three matrices  $A, B, X$  with appropriate dimensions,  $\text{stack}(AXB) = (B' \otimes A) \text{stack}(X)$  [33], we can rewrite (31)–(32), (28)–(29), (30), (33)–(34) as follows:

$$\dot{P}_I = -S(\omega)P_I, \quad (35)$$

$$\text{stack}({}^V_I \dot{R}') = 0_{9 \times 1} \quad (36)$$

$${}^B \dot{Q}_1 = -S(\omega) {}^B Q_1 - v + {}^V \dot{Q}'_1 \otimes I_{3 \times 3} \text{stack}({}^V_B R'), \quad (37)$$

$$\text{stack}({}^V_B \dot{R}') = I_{3 \times 3} \otimes -S(\omega) \text{stack}({}^V_B R'), \quad (38)$$

$$\zeta_3 = {}^V \dot{Q}'_1 \otimes I_{3 \times 3} \text{stack}({}^V_B R') - {}^B Q_1 - P_I, \quad (39)$$

$${}^V_I R' \otimes I_{3 \times 3} \text{stack}(\zeta_4) = \text{stack}({}^V_B R'), \quad (40)$$

$$\mu_{i+4} \zeta_{i+4} = F {}^C P_B + {}^V Q_i - V Q_1' \otimes F {}^C_B R \text{stack}({}^V_B R') + F {}^C_B R {}^B Q_1 \quad (41)$$

Thus, defining the vectors  $x \in \mathbb{R}^{24}$ ,  $y \in \mathbb{R}^{12+N}$ , and  $u \in \mathbb{R}^6$  as

$$x := \begin{bmatrix} P_I \\ \text{stack}({}^V_I R') \\ {}^B Q_1 \\ \text{stack}({}^V_B R') \end{bmatrix}, \quad y := \begin{bmatrix} \zeta_3 \\ \text{stack}(\zeta_4) \\ \zeta_5 \\ \vdots \\ \zeta_{4+N} \end{bmatrix}, \quad u := \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix},$$

it follows that system (35)–(41) can be expressed in the form (1)–(3) with

$$A(x, u) := \begin{bmatrix} -S(\omega) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -S(\omega) & {}^V \dot{Q}'_1 \otimes I_{3 \times 3} \\ 0 & 0 & 0 & I_{3 \times 3} \otimes -S(\omega) \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ -v \\ 0 \end{bmatrix},$$

$$C(x, u) := \begin{bmatrix} -I & 0 & -I & {}^V \dot{Q}'_1 \otimes I_{3 \times 3} \\ 0 & 0 & 0 & I \\ 0 & 0 & F {}^C_B R & 0 \\ 0 & 0 & F {}^C_B R & ({}^V Q_2 - V Q_1)' \otimes F {}^C_B R \\ 0 & 0 & F {}^C_B R & ({}^V Q_3 - V Q_1)' \otimes F {}^C_B R \\ \dots & \dots & \dots & \dots \\ 0 & 0 & F {}^C_B R & ({}^V Q_{2+N} - V Q_1)' \otimes F {}^C_B R \end{bmatrix} x + \begin{bmatrix} 0 \\ F {}^C P_B \\ \vdots \\ F {}^C P_B \end{bmatrix}$$

To determine  $E(x, v)$  we introduce additive noise to (25), i.e.,

$$\mu_{i+4} \zeta_{i+4} = F {}^C Q_i + v_i. \quad (42)$$

Note that noise does not destroy the constraint (26).

From (42), (26) and (30) we conclude that

$$\mu_{i+4} = [0 \ 0 \ 1] F [{}^C P_B + {}^C_B R {}^V_B R' ({}^V Q_i - V Q_1) + {}^C_B R {}^B Q_1] + v_i, \quad \forall i \in \{1, 2, \dots, N\}$$

which leads to

$$E(x, v) := \begin{bmatrix} I & 0 & \dots & \dots & \dots & 0 \\ 0 & {}^V_I R' \otimes I_{3 \times 3} & 0 & \dots & \dots & 0 \\ 0 & 0 & \mu_3 & 0 & \dots & 0 \\ 0 & \dots & 0 & \mu_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & \mu_{2+N} \end{bmatrix}.$$

The image of  $E(x, v)$  satisfies (3) with

$$E_0(x) := \begin{bmatrix} I & 0 & \dots & \dots & \dots & 0 \\ 0 & {}^V_I R' \otimes I_{3 \times 3} & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & 0 \end{bmatrix}$$

$$E_1 = \text{diag}\{0_{3 \times 3}, 0_{9 \times 9}, 1, 0, \dots, 0\},$$

$$E_i = \text{diag}\{0_{3 \times 3}, 0_{9 \times 9}, 0, \dots, 1, \dots, 0\},$$

$$E_N = \text{diag}\{0_{3 \times 3}, 0_{9 \times 9}, 0, \dots, 0, 1\}.$$

We can now use the results given in the previous sections to compute estimates  $\hat{P}_I$ ,  ${}^V_I \hat{R}'$ ,  ${}^B \hat{Q}_1$  and  ${}^V_B \hat{R}'$  for  $P_I$ ,  ${}^V_I R'$ ,  ${}^B Q_1$  and  ${}^V_B R'$ , respectively. From  ${}^B \hat{Q}_1$  and  ${}^V_B \hat{R}'$ ,  $p$  can also be estimated using

$$\hat{p} = V Q_1 - {}^V_B \hat{R}' {}^B \hat{Q}_1.$$

## VI. CONCLUSIONS

We considered the problem of estimating the state of a system with implicit outputs. We designed estimators using a deterministic  $H_\infty$  approach that are globally convergent under appropriate observability assumptions and can therefore, be used to design output-feedback controllers. We apply these results to the estimation of position and attitude of an autonomous vehicle using measurements from an inertial measurement unit and a monocular charged-coupled-device camera attached to the vehicle. Future work will address experimental validation of these results. The estimation problem in the presence of latency and intermittency of the observations is a topic of current research.

## APPENDIX

*Lemma 2:* Assuming that (6) has a solution on  $[0, T)$ ,  $T \in [0, \infty]$ , then  $Q(t)$  is positive definite for all  $t \in [0, T)$ .

*Proof:* Due to space limitations, the proof is omitted. It can be found in [31]. ■

## REFERENCES

- [1] B. K. Ghosh, M. Jankovic, and Y. T. Wu, "Perspective problems in system theory and its application in machine vision," *J. Math. Syst. Estimation Contr.*, vol. 4, no. 1, pp. 3–38, 1994.
- [2] B. K. Ghosh and E. P. Loucks, "A perspective theory for motion and shape estimation in machine vision," *SIAM J. Contr. Optimization*, vol. 33, no. 5, pp. 1530–1559, 1995.
- [3] S. Takahashi and B. K. Ghosh, "Motion and shape parameters identification with vision and range," in *Proc. of the 2001 Amer. Contr. Conf.*, vol. 6, June 2001, pp. 4626–4631.

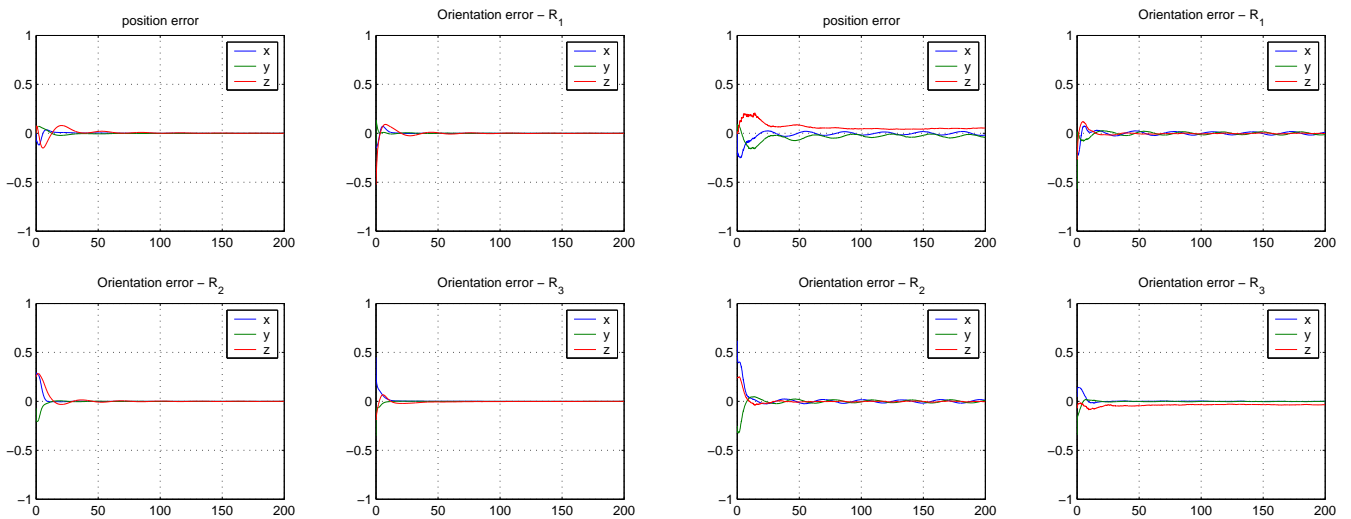


Fig. 1. Estimation errors for a simulation in which a robot follows a circular path with a camera looking up at four non-coplanar points. The four plots on the left correspond to no measurement noise, whereas the ones on the right correspond to Gaussian measurement noise with standard deviation equal to roughly 5% of the measurements. The orientation errors labelled  $R_1$ ,  $R_2$ , and  $R_3$  correspond to the estimation errors for the first, second, and third columns of  $R$ , respectively.

- [4] W. P. Dayawansa, B. K. Ghosh, C. Martin, and X. Wang, "A necessary and sufficient condition for the perspective observability problem," *Syst. & Contr. Lett.*, vol. 25, pp. 159–166, 1995.
- [5] 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.*, vol. 5, 2000, pp. 5157–5162.
- [6] A. Matveev, X. Hu, R. Frezza, and H. Rehbinder, "Observers for systems with implicit output," *IEEE Trans. on Automat. Contr.*, vol. 45, no. 1, pp. 168–173, Jan. 2000.
- [7] A. P. Aguiar and J. P. Hespanha, "Minimum-energy state estimation for systems with perspective outputs," *IEEE Trans. on Automat. Contr.* To appear, 2005.
- [8] J. P. Hespanha, "State estimation and control for systems with perspective outputs," in *Proc. of the 41st Conf. on Decision and Contr.*, Dec. 2002.
- [9] A. P. Aguiar and J. P. Hespanha, "Pose estimation of autonomous vehicles using visual information: A minimum-energy estimator approach," in *Proc. of IAV2004 - 5th IFAC/EURON Symp. on Intel. Auton. Vehicles*, July 2004.
- [10] —, "State estimation for continuous-time systems with perspective outputs from discrete noisy time-delayed measurements," in *Proc. of the 43rd Conf. on Decision and Contr.*, Dec. 2004.
- [11] T. Başar and P. Bernhard, *H<sup>∞</sup> Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd ed. Boston, MA: Birkhäuser, 1995.
- [12] K. M. Nagpal and P. P. Khargonekar, "Filtering and smoothing in an  $h_\infty$  setting," *IEEE Trans. on Automat. Contr.*, vol. 36, no. 2, pp. 152–166, Feb. 1991.
- [13] L. Xie, Y. C. Soh, and C. E. de Souza, "Robust kalman filtering for uncertain discrete-time systems," *IEEE Trans. on Automat. Contr.*, vol. 29, no. 6, pp. 1310–1314, 1994.
- [14] A. J. Krener, "Necessary and sufficient conditions for nonlinear worst case (h-infinity) control and estimation," *J. Math. Syst. Estimation Contr.*, no. 7, p. 81106, 1997.
- [15] A. H. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Trans. on Automat. Contr.*, vol. 46, no. 7, p. 9981013, July 2001.
- [16] R. K. Boel, M. R. James, and I. R. Petersen, "Robustness and risk-sensitive filtering," *IEEE Trans. on Automat. Contr.*, vol. 47, no. 3, pp. 451–461, Mar. 2002.
- [17] R. E. Mortensen, "Maximum likelihood recursive nonlinear filtering," *J. Opt. Theory and Applications*, vol. 2, pp. 386–394, 1968.
- [18] O. J. Hijab, "Minimum energy estimation," Ph.D. dissertation, University of California, Berkeley, 1980.
- [19] W. M. McEneaney, "Robust/H-infinity filtering for nonlinear systems," *Syst. & Contr. Lett.*, no. 33, pp. 315–325, 1998.
- [20] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. of Comput. Vision*, vol. 3, pp. 209–236, 1989.
- [21] M. Jankovic and Ghosh, "Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear observer," *Syst. & Contr. Lett.*, vol. 25, pp. 63–73, 1995.
- [22] S. Soatto, R. Frezza, and P. Perona, "Motion estimation via dynamic vision," *IEEE Trans. on Automat. Contr.*, vol. 41, no. 3, pp. 393–413, Mar. 1996.
- [23] 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.*, vol. 37, no. 1, pp. 158–172, Jan. 2001.
- [24] A. Chiuso, P. Favaro, H. Jin, and S. Soatto, "Structure from motion causally integrated over time," *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 523–535, Apr. 2002.
- [25] H. Rehbinder and B. K. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," *IEEE Trans. on Automat. Contr.*, vol. 48, no. 2, pp. 186–199, Feb. 2003.
- [26] Y. Ma, J. Kosecka, and S. Sastry, "Vision guided navigation for a nonholonomic mobile robot," *IEEE Trans. Robot. Automat.*, vol. 15, no. 3, pp. 521–536, 1999.
- [27] O. Shakernia, Y. Ma, T. Koo, and S. Sastry, "Landing an unmanned air vehicle: vision based motion estimation and nonlinear control," *Asian J. Contr.*, vol. 1, no. 3, pp. 128–145, 1999.
- [28] J. Hespanha, O. Yakimenko, I. Kaminer, and A. Pascoal, "Linear parametrically varying systems with brief instabilities: an application to integrated vision/IMU navigation," in *Proc. of the 40th Conf. on Decision and Contr.*, Dec. 2001.
- [29] M. James, J. Baras, and R. Elliott, "Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems," in *Proc. of the 32nd Conf. on Decision and Contr.*, San Antonio, TX, USA, 1993, pp. 3357–3360.
- [30] —, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Trans. on Automat. Contr.*, vol. 39, no. 4, p. 780792, 1994.
- [31] A. P. Aguiar and J. P. Hespanha, "State estimation for systems with implicit outputs for the integration of vision and inertial sensors," University of California, Santa Barbara, Tech. Rep., Mar. 2005.
- [32] Y. Ma, J. Kosecka, S. Soatto, and S. Sastry, *An Invitation to 3-D Vision*. Springer Verlag, 2003.
- [33] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge: Cambridge University Press., 1994.