H_∞ ESTIMATION OF SYSTEMS WITH IMPLICIT OUTPUTS

An application to Pose Estimation of Autonomous Vehicles

A. Pedro Aguiar * **João P. Hespanha** **

* *ISR/IST Institute for Systems and Robotics, Instituto Superior Técnico, Torre Norte 8, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.*
pedro@isr.ist.utl.pt

** *Department of Electrical and Computer Engineering, Univ. of California, Santa Barbara, CA 93106-9560*
hespanha@ece.ucsb.edu

Abstract: This paper addresses the problem of nonlinear filter design to estimate the relative position and attitude of an autonomous vehicle with respect to a desired coordinate system defined by visual landmarks using measurements from an inertial measurement unit (IMU) and a monocular charged-coupled-device (CCD) camera mounted on-board. We formulate the problem in the framework of state estimation of a state-affine system with implicit outputs. Resorting to dynamic programming, we derive a H_∞ estimator which produces an estimate of the state that is compatible with the dynamics and ensures a prescribed bound γ on the discounted induced L_2-gain from disturbances and noise to estimation error. In our formulation we take directly into account that the measurements arrive at discrete-time instants, are time-delayed, and may not be complete. In this way, we can deal with usual problems in vision systems such as noise as well as latency and intermittency of observations. The convergence of the proposed observer system is analyzed and simulations results are presented and discussed.

Keywords: Observers for nonlinear systems; Estimation; Visual servo control; Autonomous Vehicles;

1. INTRODUCTION

The problem of estimating the position and attitude 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., [Matthies, Kanade and Szeliski, 1989; Jankovic and Ghosh, 1995; Soatto, Frezza and Perona, 1996; Kaminer, Pascoal, Kang and Yakimenko, 2001; Chiuso, Favaro, Jin and Soatto, 2002; Rehbinder and Ghosh, 2003] and references therein). Interesting algorithms are the ones that are filtering-like or iterative that continuously improve the estimates as more data (i.e., images) are acquired and that are robust with respect to measurement noise. Soatto et al. [1996] formulates 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 [Rehbinder and Ghosh, 2003], rigid-body pose estimation using inertial sensors and a monocular camera is considered. A local convergent observer where the states evolve on $SO(3)$ (the rotation estimation is decoupled from the position estimation) is proposed. In the area of wheeled mobile robots, Ma, Kosecka and Sastry [1999] address the problem of tracking an arbitrarily shaped continuous ground curve by formulating it as controlling the shape of the curve in the image plane. Observability of the curve dynamics is studied and an extended Kalman filter is proposed to dynamically estimate the image quantities needed for feedback control from the actual noisy images. An application for landing an unmanned air vehicle using vision in the control loop is described in [Shakernia, Ma, Koo and Sastry, 1999]. In [Kaminer et al., 2001], based on measurements provided by airborne vision and inertial sensors, the problem of navigation system design for autonomous aircraft landing 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 [Hespanha, Yakimenko, Kaminer and Pascoal, 2001] to deal with the so-called out-of-frame events.

This paper addresses the problem of estimating the relative position and attitude of an autonomous vehicle with respect to a desired coordinate system defined by visual landmarks. The measurements are provided by an inertial measurement unit (IMU) and a monocular charged-

[1] Research supported in part by the Portuguese FCT (Foundation for Science and Technology) POSI Programme under framework QCA III, by project MAYA-Sub of the AdI, and NSF Grants ECS-0242798 and CCR-0311084.

coupled-device (CCD) camera mounted on-board that observes the apparent motion os stationary points. More precisely, given a desired inertial coordinate frame defined by visual landmarks $\{\mathcal{V}\}$ and a body-fixed coordinate frame $\{\mathcal{B}\}$ whose origin is located e.g., at the center of mass of the vehicle. The IMU provides the vehicle's linear velocity $v \in \mathbb{R}^3$, angular velocity $\omega \in \mathbb{R}^3$, and pose (position and attitude) $(p, R) \in \mathrm{SE}(3)$ of $\{\mathcal{B}\}$ with respect to some inertial coordinate frame $\{\mathcal{I}\}$. The camera attached to the vehicle sees N points $Q_i \in \mathbb{R}^3$, $i = 1, 2, \ldots, N$ whose coordinates expressed in $\{\mathcal{V}\}$ are assumed to be known. The objective is to determine the position ${}^{\mathcal{V}}P_{\mathcal{B}} \in \mathbb{R}^3$ and orientation ${}^{\mathcal{V}}_{\mathcal{B}}R \in \mathrm{SO}(3)$ of the vehicle with respect to the visual coordinate system $\{\mathcal{V}\}$. It is assumed that the position and orientation of $\{\mathcal{I}\}$ with respect $\{\mathcal{V}\}$ are unknown.

We formulate this problem (see Section 5) in the framework of state estimation of a system with implicit outputs of the form

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

$$E_j(x, v_j)y_j = C_j(x, u) + v_j, \tag{2}$$

$j \in \mathcal{I} := \{1, 2, \ldots N\}$, where $x \in \mathbb{R}^n$ denotes the state of the system, $u \in \mathbb{R}^m$ its control input, $y_j \in \mathbb{R}^{q_j}$ its jth measured output, $w \in \mathbb{R}^r$ an input disturbance that cannot be measured, and $v_j \in \mathbb{R}^{p_j}$ measurement noise affecting the jth output. The functions $A(x, u)$, $C_j(x, u)$, and $E_j(x, v_j)$ are affine in x. The initial condition $x(0)$ of (1) and the signals w and v_j are assumed deterministic but unknown. Each measured output y_j is only defined implicitly through (2) and the map $E_j(x, \cdot)$ is such that

$$\mathrm{Image}\, E_j(x, \cdot) = \Big\{E_{0j}(x) + \sum_{i=1}^{\ell_j} \alpha_{ij}E_{ij} : \alpha_{ij} \in \mathbb{R}\Big\} \tag{3}$$

where $E_{ij} \in \mathbb{R}^{p_j \times q_j}$ and $E_{0j}(x)$ are affine in x. Note that although the implicit representation (2) is affine in x, an explicit representation would generally be nonlinear. We call (1)–(3) a *state-affine system with implicit outputs*, or for short simply a *system with implicit outputs*. These type of systems were introduced in [Aguiar and Hespanha, 2005] and can be seen as a generalization of perspective systems introduced by Ghosh et al. [Ghosh, Jankovic and Wu, 1994]. The reader is referred to [Ghosh and Loucks, 1995; Takahashi and Ghosh, 2001] for several other examples of perspective systems in the context of motion and shape estimation. The system with implicitly defined outputs described in [Matveev, Hu, Frezza and Rehbinder, 2000] and the state-affine systems with multiple perspective outputs considered in [Aguiar and Hespanha, 2006b] are also special cases of (1)–(3).

In this paper, we also suppose that the measurements are acquired only at discrete times t'_i, $i = 0, 1, \ldots, k$, with $t'_0 < t'_1 < \ldots < t'_k$, and that we only have access to them after a time-delay τ_i. Our sequence of measurements from t'_0 to time $t \geq t'_0$ is therefore given by

$$\big\{t'_i,\ \bar{y}_j(t_i),\ j \in \mathcal{I}_i\big\}_{i=0,\ldots,k} \tag{4}$$

where k is the number of arrived measurements from t'_0 to time t (i.e., $t_k \leq t$), $\bar{y}_j(t_i) := y_j(t'_i) = y_j(t_i - \tau_i)$ denotes the time-delay observed variable, and $t_i = t'_i + \tau_i$. Note that the measurements may not be complete, that is, at time t'_i only the outputs $y_j \in \mathbb{R}^{q_j}$ with $j \in \mathcal{I}_i$ are available, where $\mathcal{I}_i \subseteq \mathcal{I}$, and the inclusion may be strict when some measurements are missing.

The problem under consideration is *to design an observer which estimates the continuous-time state vector $x(t)$ governed by (1), given the discrete time-delay measurements $\bar{y}_j(t_i)$ expressed by the output equation*

$$\begin{aligned} E_j\big(x(t_i - \tau_i), v(t_i - \tau_i)\big)\bar{y}_j(t_i) \\ = C_j\big(x(t_i - \tau_i), u(t_i - \tau_i)\big) + v_j(t_i - \tau_i), \end{aligned} \tag{5}$$

Using a H_∞ deterministic approach, we propose an observer that estimates the state vector $x(t)$ given an initial estimate $\hat{x}_0$, the past controls $\{u(\sigma) : 0 \leq \sigma \leq t\}$ and the observations (4), and minimize the induced $\mathcal{L}_2$-gain from disturbances to estimation error. In particular, for the simple case of $\tau_i = 0$, for a gain level $\gamma > 0$, the estimate $\hat{x}$ should satisfy

$$\int_0^t \|x(\sigma) - \hat{x}(\sigma)\|^2\, d\sigma \leq \gamma^2\Big((x(0) - \hat{x}_0)' P_0 (x(0) - \hat{x}_0) + \int_0^t \|w(\sigma)\|^2\, d\sigma + \sum_{i=0}^{k}\sum_{j \in \mathcal{I}_i} \|v_j(t_i)\|^2\Big)$$

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 $\lambda > 0$.

Over the last two decades the H_∞ criterion has been applied to filtering problems, cf., e.g., [Başar and Bernhard, 1995; Nagpal and Khargonekar, 1991; Xie, Soh and de Souza, 1994; Krener, 1997; Sayed, 2001; Boel, James and Petersen, 2002; Aguiar and Hespanha, 2005]. In [Aguiar and Hespanha, 2005], a state-estimator for (1)–(3) was designed using a deterministic H_∞ 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 was defined to be the value that minimizes the induced $\mathcal{L}_2$-gain from disturbances to estimation error. Closely related to H_∞ filtering are the minimum-energy estimators, which were first proposed by Mortensen [Mortensen, 1968] and further refined by Hijab [Hijab, 1980]. Game theoretical versions of these estimators were proposed by McEneaney [McEneaney, 1998]. In [Aguiar and Hespanha, 2006b], minimum-energy estimators were derived for systems with perspective outputs and input-to-state stability like properties of the estimation error with respect to disturbances were presented.

In general, both minimum-energy and H_∞ state estimators for nonlinear systems lead to infinite dimensional observers with state evolving according to a first-order nonlinear PDE of Hamilton-Jacobi type driven by the observations. We present a closed-form solution that is filtering-like and iterative-algorithm that continuously improve the estimates as more measurements are acquired. Under appropriate observability assumptions, we show that the optimal 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.

The paper is organized as follows: In Section 2 we formulate the state estimation problem using a H_∞ deterministic approach. Section 3 presents the equations for the optimal observer and in Section 4 we determine under what conditions does the state estimate $\hat{x}$ converges to the true state x. In Section 5 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_∞ state-estimators derived in the previous sections. Concluding remarks are given in Section 6.

This paper builds upon and extends previous results by the authors [Aguiar and Hespanha, 2006b; Aguiar and Hespanha, 2005; Aguiar and Hespanha, 2006c]. Due to space limitations, all the proofs are omitted. These can be found in [Aguiar and Hespanha, 2006a].

2. PROBLEM STATEMENT

This section formulates the state estimation problem using a H_∞ deterministic approach. Consider the system with implicit outputs (1), (5). From (1), $x(t_i)$ satisfies

$$x(t_i) = \Phi(t_i, t_i - \tau_i)x(t_i - \tau_i) + \int_{t_i-\tau_i}^{t_i} \Phi(t_i,\sigma)\big[A(0,u(\sigma)) + G(u(\sigma))w(\sigma)\big]\, d\sigma, \tag{6}$$

where $\Phi(t, t_0)$ is the transition matrix of system (1) satisfying the linear time-varying differential equation

$$\dot{\Phi} = \nabla A(u)\Phi. \tag{7}$$

In (7), $\nabla A(u)$ denote the gradient of $A(x, u)$ with respect to x. Since $A(x, u)$ is affine in x, it follows that $\nabla A(\cdot)$ only depend on u. From (6) we obtain

$$x(t_i - \tau_i) = \Phi^{-1}(t_i, t_i - \tau_i)x(t_i) - \Phi^{-1}(t_i, t_i - \tau_i) \times \int_{t_i-\tau_i}^{t_i} \Phi(t_i,\sigma)\big[A(0,u(\sigma)) + G(u(\sigma))w(\sigma)\big]\, d\sigma.$$

Substituting this equation in (5) and exploring the fact that $E_j(x, v)$ and $C_j(x, u)$ are affine functions in x which implies that $E_j(x, v)y = \mathbf{J}E_j(y, v)x + E_j(0, v)y$ and $C_j(x, u) = \nabla C_j(u)x + C_j(0, u)$, where $\mathbf{J}E_j(y, v)$ is the Jacobian of $E_j(x, v)y$ with respect to x, we obtain

$$\bar{E}_{ji}(x(t_i), v_j(t_i - \tau_i))\bar{y}_j(t_i) = \bar{C}_j(x(t_i), u) + \bar{v}_j(t_i), \tag{8}$$

where the definitions of $\bar{E}_{ji}$, $\bar{C}_j(x(t_i), u)$ and $\bar{v}_j(t_i)$ can be found in [Aguiar and Hespanha, 2006a].

The estimation problem can now be stated as follows:

Problem 1. Consider the continuous-time state equation (1) together with the discrete-time implicit output equation (8). For a small gain level $\gamma > 0$, given an initial estimate $\hat{x}_0$, an input u defined on an interval $[0, t)$, and measured outputs $\bar{y}_j(t_i)$, $j \in \mathcal{I}_i$ with $i = 0, 1, \dots k$, $t_0 := 0 \le t_1 \le \cdots \le t_k \le t \le t_{k+1}$, compute the estimate $\hat{x}(t)$ of the state at time t defined as

$$\hat{x}(t) := \arg\min_{z\in\mathbb{R}^n} J(z, t), \tag{9}$$

where $J(z, t)$ is defined in (10), at the top of the next page, $P_0 > 0$, and $\lambda \ge 0$ is a forgetting factor. □

In a broad sense, for a given gain level $\gamma > 0$, the optimal state $\hat{x}$ at time t is defined to be the value for the state that is compatible with the initial estimate $\hat{x}_0$, the past controls and the observations collected up to time t, and the dynamics of the system that ensures the prescribed bound γ on the discounted induced L_2-gain from disturbances and noise to estimation error. The negative of $J(z, t)$ can also be viewed as the *information state* introduced in [James, Baras and Elliott, 1993; James, Baras and Elliott, 1994] and interpreted as a measure of the likelihood of state $x = z$ at time t.

3. THE OBSERVER EQUATIONS

We propose the following observer and will shortly show that it solves Problem 1.

a) Initial condition

$$t_0 = 0, \quad P(t_0) = P_0, \quad \hat{x}(t_0) = \hat{x}_0 \tag{11}$$

b) Dynamic equations for $t \in [t_i, t_{i+1})$, $i = 0, 1, \dots, k$

$$\dot{P}(t) = -P(t)(\nabla A(u) + \lambda I) - (\nabla A(u) + \lambda I)'P(t) - \gamma^{-2}(P(t)G(u)G(u)'P(t) + \gamma^2 I), \tag{12}$$

$$\dot{\hat{x}}(t) = A(\hat{x}, u), \tag{13}$$

with $P(t_i) = P_i$, and $\hat{x}(t_i) = \hat{x}_i$.

c) Impulse equations at $t = t_{i+1}$, $i = 0, 1, \dots, k-1$

$$P(t_{i+1}) = P(t_{i+1}^-) + \gamma^2 \sum_{j\in\mathcal{I}_{i+1}} \Psi_j(t_{i+1}), \tag{14}$$

$$\hat{x}(t_{i+1}) = \hat{x}(t_{i+1}^-) - P(t_{i+1})^{-1}\gamma^2 \sum_{j\in\mathcal{I}_{i+1}} \big[\Psi_j(t_{i+1})\hat{x}(t_{i+1}^-) + \psi_j(t_{i+1})\big] \tag{15}$$

where

$$\Psi_j(t_i) := \Big(\mathbf{J}E_{0j}(\bar{y}_j(t_i))\Phi(t_i - \tau_i, t_i) - \nabla\bar{C}_j\Big)'\Big(I - Y_{ji}Y_{ji}^{\perp}\Big)' \times \Big(I - Y_{ji}Y_{ji}^{\perp}\Big)\Big(\mathbf{J}E_{0j}(\bar{y}_j(t_i))\Phi(t_i - \tau_i, t_i) - \nabla\bar{C}_j\Big),$$

$$\psi_j(t_i) := \Big(\mathbf{J}E_{0j}(\bar{y}_j(t_i))\Phi(t_i - \tau_i, t_i) - \nabla\bar{C}_j\Big)'\Big(I - Y_{ji}Y_{ji}^{\perp}\Big) \times \Big(I - Y_{ji}Y_{ji}^{\perp}\Big)\Big(E_{0j}(0)\bar{y}_j(t_i) - \mathbf{J}E_{0j}(\bar{y}_j(t_i)) \times \Phi(t_i - \tau_i, t_i)\int_{t_i-\tau_i}^{t_i} \Phi(t_i,\sigma)A(0,u(\sigma))\, d\sigma - \bar{C}_j(0,u)\Big),$$

$$Y_{ji} := \Big[E_{1j}\bar{y}_j(t_i)|E_{2j}\bar{y}_j(t_i)|\cdots|E_{\ell_j j}\bar{y}_j(t_i)\Big], \tag{16}$$

$Y_{ji}^{\perp}$ denotes the pseudo-inverse of Y_{ji}, $\nabla A(u)$ and $\mathbf{J}E_{0j}(y)$ are respectively the gradient of $A(x, u)$ and the Jacobian of $E_{0j}(x)y$ both with respect to x. The following result solves Problem 1.

Theorem 1. The H_∞ state estimate $\hat{x}(t)$ defined by (9)–(10) can be obtained from the *impulse system* (11)–(15). Furthermore, the cost function $J(z; t)$ defined in (10) is quadratic and can be written as

$$J(z; t) = \big(z - \hat{x}(t)\big)'P(t)\big(z - \hat{x}(t)\big) + c(t), \tag{17}$$

where $c(t)$ satisfies an appropriate impulse equation. □

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

$$J(z;t) := \min_{\substack{w:[0,t),\\ \bar v_j(t_i)\ i=0,1,\dots k}} \Big\{ \gamma^2 e^{-2\lambda t}(x(0)-\hat x_0)' P_0 (x(0)-\hat x_0) + \gamma^2 \int_0^t e^{-2\lambda(t-\sigma)} \|w(\sigma)\|^2 \, d\sigma + \gamma^2 \sum_{i=0}^{k} \sum_{j\in\mathcal{I}_i} e^{-2\lambda(t_k - t_i)} \|\bar v_j(t_i)\|^2$$

$$- \int_0^t e^{-2\lambda(t-\sigma)} \|x(\sigma) - \hat x(\sigma)\|^2 \, d\sigma : x(t) = z, \dot x = A(x,u) + G(u)w, \bar E_{ji} \bar y_j(t_i) = \bar C_j(x(t_i), u) + \bar v_j(t_i) \Big\} \quad (10)$$

Lemma 1. The H_∞ estimator (11)–(15) has a global solution and

$$P(t) \geq \delta I > 0, \qquad \forall t \geq 0, \qquad (18)$$

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

$$\gamma^2 W_0(t) \geq \int_0^t \bar\Phi(\tau,t)' \bar\Phi(\tau,t) d\tau + \delta I \quad \forall t \geq 0 \qquad (19)$$

holds, where

$$W_0(t) := \sum_{i=1}^{k} \sum_{j\in\mathcal{I}_i} \bar\Phi(t_i,t)' \Psi_j(t_i) \bar\Phi(t_i,t), \qquad (20)$$

$$\bar\Phi(\tau,\sigma) := \begin{cases} \bar\Phi_i(\tau,\sigma), & i=j \\ \bar\Phi_i(\tau,t_{i+1})\bar\Phi_{i+1}(t_{i+1},t_{i+2})\cdots\bar\Phi_j(t_j,\sigma), & i<j \end{cases}$$

$\forall \tau \in [t_i, t_{i+1})$ and $\forall \sigma \in [t_j, t_{j+1})$, and $\bar\Phi_i(t,\tau)$ denotes the state transition matrix of $\dot z = (\nabla A + \gamma^{-2} G G' P + \lambda I) z$, $\forall \tau \in [t_i, t_{i+1})$. □

4. ESTIMATOR CONVERGENCE

We are now interesting in determining under what conditions does the state estimate $\hat x$ converges to the true state x. As in [Aguiar and Hespanha, 2006b], the following technical assumption is needed:

Assumption 1. Let $\text{Num}(t,\sigma)$, $0 \leq \sigma < t$ denote the number of time instants at which measurement arrive in the open interval (σ, t). There exist finite positive constants τ_D and N_0, for which the following condition holds:

$$\text{Num}(t,\sigma) \leq N_0 + \frac{t-\sigma}{\tau_D}.$$

The constant τ_D is called the *average dwell-time* and N_0 the *chatter bound.* □

This assumption roughly speaking guarantees that the average interval between consecutive arrival of measurements is no less than τ_D. In this way, the summation in (10) will not grow unbounded due to "too frequent" measurements.

The following result establishes the convergence of the state estimate.

Theorem 2. Assuming that the solutions to the system with implicit outputs (1), (5) exists on $[0,T)$, $T \in [0,\infty]$, the solution to the impulse state estimator (11)–(15) also exists on $[0,T)$. Moreover, if $P(t) \geq \delta I$, $\forall t \in [0,T)$, $\delta > 0$, then there exist positive constants $c > 0, r < 1, \gamma_w, \gamma_1, \dots, \gamma_N$ such that

$$\|\tilde x(t_k)\| \leq c\, r^k \|\tilde x(0)\| + \gamma_w \sup_{\sigma\in(0,t_k)} \|w(\sigma)\| + \sum_{j=1}^{N} \gamma_j \sup_{\sigma\in(0,t_k)} \|\bar v_j(\sigma)\|, \quad t_k > 0 \quad (21)$$

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

Combining Theorem 2 and Lemma 1 we obtain the following:

Corollary 1. When Assumption 1 holds, and there exist constants $\mathbf{N}$, ϵ such that the persistence of excitation condition (19) holds, the state-estimate $\hat x$ converges to the state x in the absence of disturbance input and measurement noise. When the disturbance and noise are bounded but nonzero, $\hat x$ converges to a neighborhood of the true state x. □

5. POSE ESTIMATION OF AUTONOMOUS VEHICLES 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).

5.1 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 $\left({}^{\mathcal{V}}_{\mathcal{B}}R,\ {}^{\mathcal{V}}P_{\mathcal{B}}\right)$ is an element of the Special Euclidean group $\text{SE}(3) := \text{SO}(3) \times \mathbb{R}^3$, where ${}^{\mathcal{V}}_{\mathcal{B}}R \in \text{SO}(3) := \{R \in \mathbb{R}^{3\times 3} : RR' = I_3, \det(R) = +1\}$ is a rotation matrix that describes the orientation of the vehicle by mapping body coordinates into $\{\mathcal{V}\}$, and ${}^{\mathcal{V}}P_{\mathcal{B}} \in \mathbb{R}^3$ is the position of the origin of $\{\mathcal{B}\}$ in $\{\mathcal{V}\}$. Denoting by ${}^{\mathcal{B}}v \in \mathbb{R}^3$ and ${}^{\mathcal{B}}\omega \in \mathbb{R}^3$ the linear and angular velocities of the vehicle relative to $\{\mathcal{V}\}$ expressed in $\{\mathcal{B}\}$, respectively, the following kinematic relations apply:

$${}^{\mathcal{V}}\dot P_{\mathcal{B}} = {}^{\mathcal{V}}_{\mathcal{B}}R\,{}^{\mathcal{B}}v, \qquad {}^{\mathcal{V}}_{\mathcal{B}}\dot R = {}^{\mathcal{V}}_{\mathcal{B}}R\,S({}^{\mathcal{B}}\omega), \qquad (22)$$

where $S(x) := \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}$, $\forall x := (x_1, x_2, x_3) \in \mathbb{R}^3$. The objective is to determine the position and attitude of the vehicle with respect to the visual coordinate system $\{\mathcal{V}\}$, that is, to estimate ${}^{\mathcal{V}}P_{\mathcal{B}}$ and ${}^{\mathcal{V}}_{\mathcal{B}}R$.

5.2 Sensor Measurements

We consider that the IMU provides the vehicle's linear velocity ${}^{\mathcal{B}}v$, angular velocity ${}^{\mathcal{B}}\omega$, 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

$$\zeta_1 = {}^{\mathcal{B}}v,\ \zeta_2 = {}^{\mathcal{B}}\omega,\ \zeta_3 = {}^{\mathcal{I}}P_{\mathcal{B}},\ \zeta_4 = {}^{\mathcal{I}}_{\mathcal{B}}R, \quad (23)$$

where $\zeta_1 \in \mathbb{R}^3, \zeta_2 \in \mathbb{R}^3, \zeta_3 \in \mathbb{R}^3$, $\zeta_4 \in \text{SO}(3)$, and $\left({}^{\mathcal{I}}_{\mathcal{B}}R,\ {}^{\mathcal{I}}P_{\mathcal{B}}\right)$ 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 $Q_i = (x_i, y_i, z_i)'$,

$i = 1, 2, \ldots, N$ whose coordinates expressed in the visual coordinate system are assumed to be known. 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:

$$\mu_{i+4}\zeta_{i+4} = F\,{}^{\mathcal{C}}Q_i, \tag{24}$$

$$\left[\, 0\ 0\ 1 \,\right]\zeta_{i+4} = 1, \quad \forall i \in \{1, 2, \ldots, N\} \tag{25}$$

where ${}^{\mathcal{C}}Q_i$ is the position of Q_i expressed in the camera's frame, $\mu_{i+4} \in \mathbb{R}$ captures the depth of the point ${}^{\mathcal{C}}Q_i$ (which is unknown), and F is an upper triangular matrix with the camera's intrinsic parameters.
Given the measurements ζ_i, $i = 1, \ldots, N+4$, we now proceed with the formulation of a system with implicit outputs.

5.3 System with implicit outputs

Let ${}^{\mathcal{V}}Q_1$ and ${}^{\mathcal{B}}Q_1$ be the coordinates of a point Q_1 in the frames $\{\mathcal{V}\}$ and $\{\mathcal{B}\}$, respectively. Then, we have that ${}^{\mathcal{V}}Q_1 = {}^{\mathcal{V}}P_{\mathcal{B}} + {}^{\mathcal{V}}_{\mathcal{B}}R\,{}^{\mathcal{B}}Q_1$. From this and (22), we obtain the state equations

$${}^{\mathcal{B}}\dot{Q}_1 = {}^{\mathcal{V}}_{\mathcal{B}}R'\,{}^{\mathcal{V}}\dot{Q}_1 - S(\omega)\,{}^{\mathcal{B}}Q_1 - v, \tag{26}$$

$${}^{V}_{B}\dot{R}' = -S(\omega)\,{}^{\mathcal{V}}_{\mathcal{B}}R'. \tag{27}$$

To obtain the output equations of the vision subsystem, we first note that if ${}^{\mathcal{V}}Q_j$ and ${}^{\mathcal{B}}Q_j$ denote the coordinates of another point Q_j in the frames $\{\mathcal{V}\}$ and $\{\mathcal{B}\}$, respectively, we conclude that ${}^{\mathcal{B}}Q_j = {}^{\mathcal{V}}_{\mathcal{B}}R'\,{}^{\mathcal{V}}Q_j - {}^{\mathcal{V}}_{\mathcal{B}}R'\,{}^{\mathcal{V}}P_{\mathcal{B}} = {}^{\mathcal{V}}_{\mathcal{B}}R'\big({}^{\mathcal{V}}Q_j - {}^{\mathcal{V}}Q_1\big) + {}^{\mathcal{B}}Q_1$. Using now (24) and the fact that ${}^{\mathcal{C}}Q_i = {}^{\mathcal{C}}P_{\mathcal{B}} + {}^{\mathcal{C}}_{\mathcal{B}}R\,{}^{\mathcal{B}}Q_i$, we obtain the output equations

$$\mu_{i+4}\zeta_{i+4} = F\Big({}^{\mathcal{C}}P_{\mathcal{B}} + {}^{\mathcal{C}}_{\mathcal{B}}R\,{}^{\mathcal{V}}_{\mathcal{B}}R'\big({}^{\mathcal{V}}Q_i - {}^{\mathcal{V}}Q_1\big) + {}^{\mathcal{C}}_{\mathcal{B}}R\,{}^{\mathcal{B}}Q_1\Big), \quad \forall i \in \{1, 2, \ldots, N\} \tag{28}$$

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

We will regard ζ_1 and ζ_2 as inputs to the implicit output system. The dynamics of (23) can be written as

$$\overbrace{{}^{\mathcal{V}}_{\mathcal{B}}R'\ {}^{\mathcal{V}}P_{\mathcal{I}}}^{\cdot} = -S({}^{\mathcal{B}}\omega)\,{}^{\mathcal{V}}_{\mathcal{B}}R'\ {}^{\mathcal{V}}P_{\mathcal{I}}, \tag{29}$$

$${}^{V}_{I}\dot{R}' = 0, \tag{30}$$

with the output equations

$${}^{\mathcal{I}}_{\mathcal{B}}R'\ {}^{\mathcal{I}}P_{\mathcal{B}} = {}^{\mathcal{V}}_{\mathcal{B}}R'\,{}^{\mathcal{V}}Q_1 - {}^{\mathcal{B}}Q_1 - P_I, \tag{31}$$

$${}^{\mathcal{I}}_{\mathcal{B}}R'\,{}^{\mathcal{V}}_{\mathcal{I}}R' = {}^{\mathcal{V}}_{\mathcal{B}}R'. \tag{32}$$

Thus, our implicit output system is composed by (29)–(30), (26)–(27), (31)–(32) and (28). 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 $\mathrm{stack}(M)$ 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, $\mathrm{stack}(A\,X\,B) = (B' \otimes A)\,\mathrm{stack}(X)$ [Horn and Johnson, 1994], we can rewrite (29)–(30), (26)–(27), (28), (31)–(32) as (1)–(3) where

$x := \mathrm{stack}({}^{\mathcal{V}}_{\mathcal{B}}R'\ {}^{\mathcal{V}}P_{\mathcal{I}}, \mathrm{stack}({}^{\mathcal{V}}_{\mathcal{I}}R'), {}^{\mathcal{B}}Q_1, \mathrm{stack}({}^{\mathcal{V}}_{\mathcal{B}}R'))$, $y := \mathrm{stack}(\zeta_4'\zeta_3, \mathrm{stack}(\zeta_4), \zeta_{4+1}, \ldots, \zeta_{4+N})$, $u := \mathrm{stack}(\zeta_1, \zeta_2)$,

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

$$C_1(x,u) := \left[\, -I\ 0\ -I\ {}^{\mathcal{V}}Q_1'\otimes I_{3\times 3} \,\right] x, \quad C_2(x,u) := \left[\, 0\ 0\ 0\ I \,\right] x$$

$$C_{2+i}(x,u) := \left[\, 0\ 0\ F\,{}^{\mathcal{C}}_{\mathcal{B}}R\ ({}^{\mathcal{V}}Q_i - {}^{\mathcal{V}}Q_1)'\otimes F\,{}^{\mathcal{C}}_{\mathcal{B}}R \,\right] x + F\,{}^{\mathcal{C}}P_{\mathcal{B}},$$

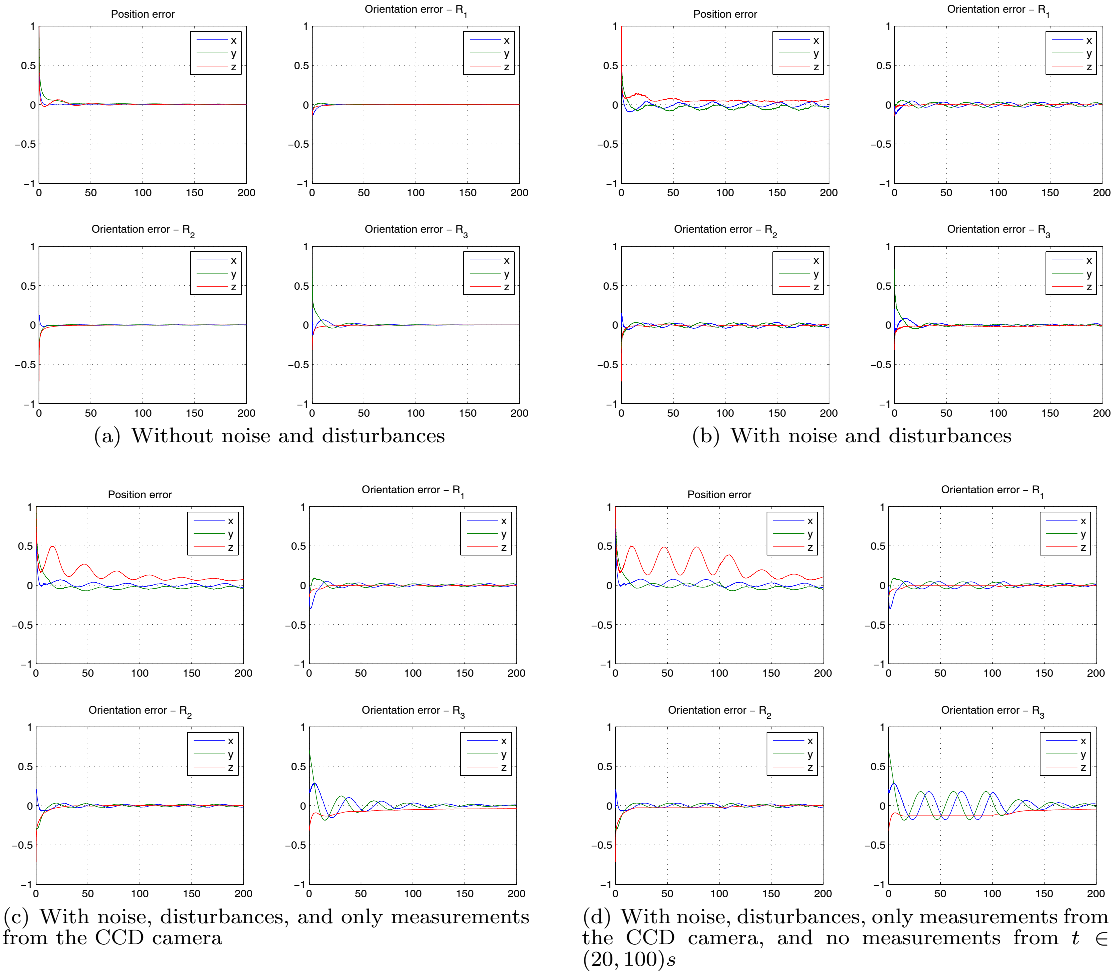
$\forall i \in \{1, \ldots, N\}$. By introducing additive noise to the output equations (31)–(32) and (24) we conclude that $E_1(x, v_1) := I$, $E_2(x, v_2) := {}^{\mathcal{V}}_{\mathcal{I}}R \otimes I_{3\times 3}$, $E_{2+i}(x, v_{2+i}) := \left[\, 0\ 0\ 1 \,\right] F\big[\,{}^{\mathcal{C}}P_{\mathcal{B}} + {}^{\mathcal{C}}_{\mathcal{B}}R\,{}^{\mathcal{V}}_{\mathcal{B}}R'\big({}^{\mathcal{V}}Q_i - {}^{\mathcal{V}}Q_1\big) + {}^{\mathcal{C}}_{\mathcal{B}}R\,{}^{\mathcal{B}}Q_1\big] + {}^{\mathcal{B}}v_{2+i}$, $\forall i \in \{1, \ldots, N\}$. The image of $E_j(x, v_j)$ satisfies (3) with $E_{01}(x) := I$, $\ell_1 = 0$, $E_{02}(x) := {}^{\mathcal{V}}_{\mathcal{I}}R \otimes I_{3\times 3}$, $\ell_2 = 0$, $E_{0,2+i}(x) := 0$, $\ell_{2+i} = 1$, $E_{1,2+i} := 1\ \forall i \in \{1, \ldots, N\}$. We can now use the results given in the previous sections to compute $\hat{x}$. From ${}^{\mathcal{B}}\hat{Q}_1$ and ${}^{V}_{B}\hat{R}'$, the position ${}^{\mathcal{V}}P_{\mathcal{B}}$ can also be estimated using ${}^{\mathcal{V}}\hat{P}_{\mathcal{B}} = {}^{\mathcal{V}}Q_1 - {}^{\mathcal{V}}_{\mathcal{B}}\hat{R}\,{}^{\mathcal{B}}\hat{Q}_1$.

We now illustrate the performance of the proposed estimator through computer simulation. The autonomous vehicle starts at the origin ${}^{\mathcal{V}}P_{\mathcal{B}} = 0$ with orientation ${}^{\mathcal{V}}_{\mathcal{B}}R = I$ and follows a circular path with a camera looking up at four noncoplanar points. The linear velocity is ${}^{\mathcal{B}}v = [0.3, 0, 0]'\,m/s$ and the angular velocity is ${}^{\mathcal{B}}\omega = [0, 0, 0.2]'\,rad/s$. The vision sampling interval is $T_{CCD} = 0.4\,s$ and the time-delay is $\tau_{CCD} = 0.05\,s$. The IMU sampling interval is $T_{IMU} = 0.1\,s$ and there is no time-delay. The estimator was initialized with ${}^{\mathcal{V}}\hat{P}_{\mathcal{B}} = [1, 1, 1]'\,m/s$ and ${}^{\mathcal{V}}_{\mathcal{B}}\hat{R} = \left[\begin{smallmatrix} 0.9814 & -0.0179 & 0.1913 \\ -0.1246 & 0.6983 & 0.7049 \\ -0.1462 & -0.7156 & 0.6831 \end{smallmatrix}\right]$. The measurements were corrupted with additive Gaussian noise with standard deviation equal to roughly 5% of the measurements.

Fig. 1 displays the time evolution of the estimation errors. It can be seen that the estimated pose without noise converges to zero (see Fig. 1(a)) and in the presence of noise tend to a small neighborhood of the true value (see Fig. 1(b)). To illustrate the benefit of having measurements from the IMU, Fig. 1(c) shows the time evolution of the estimation errors when there is only measurements from the CCD camera. As expected, although the errors converge to a small neighborhood of the origin, the transients are worst than the ones displayed in Fig. 1(b). In Fig. 1(d) we can also see what happens when the observer does not receive measurements at all from $t = 20\,s$ to $t = 100\,s$.

6. CONCLUSIONS

We considered the problem of estimating the state of a system with implicit outputs, whose measurements arrive at discrete-time instants, are time-delayed, noisy, and may not be complete. We designed a estimator using a deterministic H_∞ approach that is globally convergent under appropriate observability assumptions and can therefore, be used to design output-feedback controllers. These results were applied 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



(a) Without noise and disturbances

(b) With noise and disturbances

(c) With noise, disturbances, and only measurements from the CCD camera

(d) With noise, disturbances, only measurements from the CCD camera, and no measurements from $t \in (20, 100)s$

Fig. 1. Time evolution of the estimation errors in position and orientation. The orientation errors labeled R_1, R_2, and R_3 correspond to the estimation errors for the first, second, and third columns of ${}^{\mathcal{V}}_{\mathcal{B}}R$, respectively.

attached to the vehicle. Future work will address the investigation of easier conditions that satisfy the required observability assumptions and experimental validation of these results.

REFERENCES

Aguiar, A. P. and Hespanha, J. P. [2005]. State estimation for systems with implicit outputs for the integration of vision and inertial sensors, *Proc. of the 44th Conf. on Dec. and Contr.*

Aguiar, A. P. and Hespanha, J. P. [2006a]. H_∞ estimation of systems with implicit outputs, *Technical report*, ISR, IST.

Aguiar, A. P. and Hespanha, J. P. [2006b]. Minimum-energy state estimation for systems with perspective outputs, *IEEE Trans. on Automat. Contr.* **51**(2): 226–241.

Aguiar, A. P. and Hespanha, J. P. [2006c]. State estimation of continuous-time systems with implicit outputs from discrete noisy time-delayed measurements, Submitted to CDC'06.

Başar, T. and Bernhard, P. [1995]. *H^∞ Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd edn, Birkhäuser, Boston, MA.

Boel, R., James, M. and Petersen, I. [2002]. Robustness and risk-sensitive filtering, *IEEE Trans. on Aut. Contr.* **47**(3): 451–461.

Chiuso, A., Favaro, P., Jin, H. and Soatto, S. [2002]. Structure from motion causally integrated over time, *IEEE Trans. on Pattern Anal. Mach. Intell.* **24**(4): 523–535.

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

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

Hespanha, J., Yakimenko, O., Kaminer, I. and Pascoal, A. [2001]. Linear parametrically varying systems with brief instabilities: an application to integrated vision/IMU navigation, *Proc. of the 40th Conf. on Decision and Contr.*

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

Horn, R. A. and Johnson, C. R. [1994]. *Topics in Matrix Analysis*, Cambridge University Press., Cambridge.

James, M., Baras, J. and Elliott, R. [1993]. Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems, *Proc. of the 32nd Conf. on Decision and Contr.*, San Antonio, TX, USA, pp. 3357–3360.

James, M., Baras, J. and Elliott, R. [1994]. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems, *IEEE Trans. on Automat. Contr.* **39**(4): 780–792.

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

Kaminer, I., Pascoal, A., Kang, W. and Yakimenko, O. [2001]. Integrated vision/inertial navigation systems design using nonlinear filtering, *IEEE T. Aeros. and Electr. Syst.* **37**(1): 158–172.

Krener, A. J. [1997]. Necessary and sufficient conditions for nonlinear worst case (H-infinity) control and estimation, *J. Math. Syst. Estimation Contr.* (7): 81-106.

Ma, Y., Kosecka, J. and Sastry, S. [1999]. Vision guided navigation for a nonholonomic mobile robot, *IEEE Trans. Robot. Automat.* **15**(3): 521–536.

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

Matveev, A., Hu, X., Frezza, R. and Rehbinder, H. [2000]. Observers for systems with implicit output, *IEEE Trans. on Automat. Contr.* **45**(1): 168–173.

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

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

Nagpal, K. M. and Khargonekar, P. P. [1991]. Filtering and smoothing in an H_∞ setting, *IEEE T. on Aut. Contr.* **36**(2): 152–166.

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

Sayed, A. [2001]. A framework for state-space estimation with uncertain models, *IEEE T. on Aut. Contr.* **46**(7): 998–1013.

Shakernia, O., Ma, Y., Koo, T. and Sastry, S. [1999]. Landing an unmanned air vehicle: vision based motion estimation and nonlinear control, *Asian J. Contr.* **1**(3): 128–145.

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

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

Xie, L., Soh, Y. C. and de Souza, C. E. [1994]. Robust Kalman filtering for uncertain discrete-time systems, *IEEE Trans. on Automat. Contr.* **29**(6): 1310–1314.