2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Structure Estimation of a Moving Object Using a Moving Camera: An Unknown Input Observer Approach

Ashwin P. Dani, Zhen Kan, Nicholas R. Fischer, Warren E. Dixon

***Abstract*—A state observer is designed for estimating the structure of a moving object with time-varying velocities seen by a moving camera. A nonlinear unknown input observer strategy is used where the object's velocity is considered as an unknown input to the perspective dynamical system. The object is assumed to be moving on a ground plane. The downward-looking camera observing the moving object is also moving (e.g., attached to an air vehicle) with known velocities. The developed method provides the first causal, observer-based structure estimation algorithm for a moving camera viewing a moving object with unknown time-varying object velocities.**

I. Introduction

Recovering the structure of a static scene (i.e., the 3D Euclidean coordinates of feature points) using a moving camera is called 'Structure from Motion (SfM)'. A number of solutions to the SfM problem are given in the form of batch or offline methods [1]–[3] and causal or online methods [4]–[10]. Solutions to the SfM problem (e.g. see [6], [7]) can be used for self-localization and map building of an environment using a moving camera. The fundamental problem behind SfM algorithms is a triangulation problem. Since the object is assumed to be stationary, a moving camera can capture snapshots of the object from two different locations and triangulation is feasible. SfM techniques cannot be used to recover the structure and motion of moving objects using a moving camera because triangulation is not feasible when the object is moving [11].

Recovering the structure of a moving object using a moving camera is termed trajectory triangulation in the pioneering work of [11]. In this paper, the trajectory triangulation problem is called *structure and motion from motion* (SaMfM). In [11], a batch algorithm is applied for points moving in straight lines or conic trajectories given five or nine views, respectively. In [12], a batch algorithm is presented for object motions represented by more general curves. In [13], a factorization-based batch algorithm is proposed where objects are assumed to be moving with constant speed in a straight line, observed by a weak perspective camera. An algebraic geometry approach is presented in [14] to estimate the motion of objects up to a scale given a minimum number of point correspondences. In [15], authors propose a batch algorithm to estimate the structure and motion of objects moving on a ground plane observed by a moving airborne camera. The method relies on a static scene for estimating the projective depth, approximated by the depth of feature points on a static background assuming that one of the feature points of the moving object lies on the static background. In [16], a batch algorithm is developed by approximating the trajectories of a moving object using a linear combination of discrete cosine transform (DCT) basis vectors.

Batch algorithms use an algebraic relationship between 3D coordinates of points in the camera coordinate frame and corresponding 2D projections on the image frame collected over n images to estimate the structure. Hence, batch algorithms are not useful in real-time control algorithms. For visual servo control or video-based surveillance tasks, online structure estimation algorithms are required. The objective of this paper is to estimate the structure of moving objects from a continuous stream of images which can be described using a continuous dynamical model. Instead of algebraic relationships and geometric constraints used by batch algorithms, a rigid body kinematic motion model is used to estimate structure. The use of a dynamical model enables the design of an online/causal algorithm which uses data from images up to the current time step. In [11], authors point out that structure estimation of a moving object using a moving camera can only be obtained if some assumptions are made on the trajectories of the moving object. Recently, a causal algorithm is presented in [17] to estimate the structure and motion of objects moving with constant linear velocities observed by a moving camera with known camera motions. In this paper, efforts are focused on relaxing the assumption of constant object linear velocity. The object is assumed to be moving on a ground plane observed by a downward-looking airborne camera. In the relative rigid body motion dynamics, the moving object's linear velocity can be viewed as an exogenous time-varying disturbance. An unknown input observer (UIO) approach is used to estimate the state of the dynamical system where the moving object's velocity is considered as an unknown input.

Several UIO algorithms are present in literature for estimating the state when an exogenous time-varying unknown input is present in the system. For linear systems UIOs can be found in [18]–[22]. Linear UIO algorithms are extended for various classes of nonlinear systems in [23]–[26]. In [27], an unknown input observer for fault diagnosis is presented.

Ashwin Dani is with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Champaign, IL, USA. Email: adani@illinois.edu.

Zhen Kan, Nicholas Fischer and Warren E. Dixon are with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA. Email: {nic.fischer, kanzhen0322, wdixon}@ufl.edu.

978-1-61284-799-3/11/$26.00 ©2011 IEEE

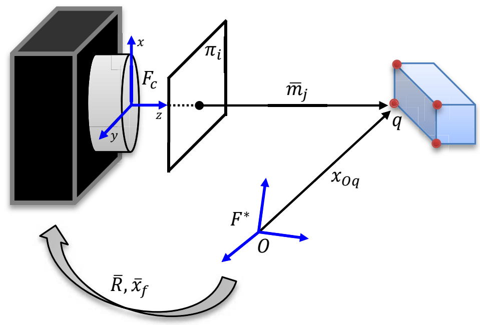

Figure 1. Moving camera looking at a moving object.

The observer design relies on a coordinate transformation and solving a parametric Lyapunov equation for computing an observer gain. However, no systematic method exists for solving a parametric Lyapunov equation. In [24], a nonlinear UIO is presented where the design procedure is based on H_∞ optimization. The observer is called a dynamic UIO which provides an extra degree of design freedom but increases the order of the system. In [26], a nonlinear UIO is presented for a class of nonlinear systems based on a linear matrix inequality (LMI) approach. In this paper, a nonlinear UIO is presented for a more general class of nonlinear systems than considered in [26]. The observer synthesis, i.e., finding the observer gain matrices, is achieved by solving a LMI feasibility problem or a LMI eigenvalue optimization problem.

The contribution of this work is to provide a causal algorithm for estimating the structure of a moving object using a moving camera with relaxed assumptions on the object's motion. The object is assumed to be moving on a ground plane with arbitrary velocities observed by a downward looking camera with arbitrary linear motion in 3D space. No assumptions are made on the minimum number of points or minimum number of views required to estimate the structure. Feature point data and camera velocity data from each image frame is required. Estimating the structure of a moving object is recast into an unknown input observer design problem. Another contribution of this paper is to extend the nonlinear UIO design for a more general class of nonlinear systems inspired by the UIO design in [18], [26] and to develop sufficient conditions for the existence of the UIO.

II. Euclidean to Image Space Mapping

Consider a scenario depicted in Fig. 1 where a moving camera views moving point objects (such as feature points on a rigid object). In Fig. 1, an inertial reference frame is denoted by $\mathcal{F}^*$[1]. After the initial time, a reference frame $\mathcal{F}_c$ attached to a pinhole camera undergoes some rotation $\bar{R}(t) \in SO(3)$ and translation $\bar{x}_f(t) \in \mathbb{R}^3$ away from $\mathcal{F}^*$.

[1] w.l.o.g. $\mathcal{F}^*$ can be attached to the camera at the location corresponding to an initial point in time t_0 where the object is in the camera field of view (FOV) and $\mathcal{F}^*$ is identical to $\mathcal{F}_c(t_0)$.

The Euclidean coordinates $\bar{m}_j(t) \in \mathbb{R}^3$ (where $j = \{1, 2, ..., n\}$ denotes a point number) of points observed by a camera expressed in the camera frame $\mathcal{F}_c$ and the respective normalized Euclidean coordinates $m_j(t) \in \mathbb{R}^3$ are defined as

$$\bar{m}_j(t) = \begin{bmatrix} X_j(t), & Y_j(t), & Z_j(t) \end{bmatrix}^T, \quad (1)$$

$$m_j(t) = \begin{bmatrix} \frac{X_j(t)}{Z_j(t)}, & \frac{Y_j(t)}{Z_j(t)}, & 1 \end{bmatrix}^T. \quad (2)$$

Consider a closed and bounded set $\mathcal{Y} \subset \mathbb{R}^3$. To facilitate the subsequent development, the state vector $x_j(t) = [x_{1j}(t), x_{2j}(t), x_{3j}(t)]^T \in \mathcal{Y}$ is constructed from (2) as

$$x_j = \begin{bmatrix} \frac{X_j}{Z_j}, & \frac{Y_j}{Z_j}, & \frac{1}{Z_j} \end{bmatrix}^T. \quad (3)$$

Using projective geometry, the normalized Euclidean coordinates $m_j(t)$ can be related to the pixel coordinates in the image space as

$$q_j = A_c m_j \quad (4)$$

where $q_j(t) = \begin{bmatrix} u_j(t) & v_j(t) & 1 \end{bmatrix}^T$ is a vector of the image-space feature point coordinates $u_j(t), v_j(t) \in \mathbb{R}$ defined on the closed and bounded set $\mathcal{I} \subset \mathbb{R}^3$, and $A_c \in \mathbb{R}^{3\times3}$ is a constant, known, invertible camera calibration matrix [28]. Since A_c is known, the expression in (4) can be used to recover $m_j(t)$, which can be used to partially reconstruct the state $x_j(t)$ so that the first two components of $x_j(t)$ can be determined.

Assumption 1: The relative Euclidean distance $Z_j(t)$ between the camera and the feature points observed on the object is upper and lower bounded by some known positive constants (i.e., the object remains within some finite distance away from the camera). Therefore, the definition in (3) can be used to assume that

$$\bar{x}_3 \geq x_{3j}(t) \geq \underline{x}_3 \quad (5)$$

where $\bar{x}_3, \underline{x}_3 \in \mathbb{R}$ denote known positive bounding constants. Likewise, since the image coordinates are constrained (i.e., the object is assumed to remain in the camera field of view (FOV)), the relationships in (2)-(4) along with the fact that A_c is invertible can be used to conclude that

$$\bar{x}_1 \geq |x_{1j}(t)| \geq \underline{x}_1 \qquad \bar{x}_2 \geq |x_{2j}(t)| \geq \underline{x}_2$$

where $\bar{x}_1$, $\bar{x}_2$, $\underline{x}_1$, $\underline{x}_2 \in \mathbb{R}$ denote known positive bounding constants.

For the remainder of this paper, the feature point subscript j is omitted to streamline the notation.

III. Camera Motion Model and State Space Dynamics

Consider a moving camera viewing a moving point q. As shown in Fig. 1, the point q can be expressed in the coordinate system $\mathcal{F}_c$ as

$$\bar{m} = \bar{x}_f + \bar{R} x_{oq} \quad (6)$$

where x_{oq} is a vector from the origin of coordinate system $\mathcal{F}^*$ to the point q expressed in the coordinate system $\mathcal{F}^*$. Differentiating (6), the relative motion of q as observed in the camera coordinate system can be expressed by the following kinematics [28]

$$\dot{\bar{m}} = [\omega]_\times \bar{m} + v_r \tag{7}$$

where $\bar{m}(t)$ is defined in (1), $[\omega]_\times \in \mathbb{R}^{3\times 3}$ denotes a skew symmetric matrix formed from the angular velocity vector of the camera $\omega(t) = \begin{bmatrix} \omega_1 & \omega_2 & \omega_3 \end{bmatrix}^T \in \mathcal{W}$, and $v_r(t)$ represents the relative velocity of the camera with respect to the moving point, defined as

$$v_r = v_c - v_p. \tag{8}$$

In (8), $v_c(t)$ denotes the camera velocity in the camera reference frame given by $v_c(t) = \begin{bmatrix} v_{cx} & v_{cy} & v_{cz} \end{bmatrix}^T \in \mathcal{V}_c$ and $v_p(t)$ denotes the velocity of the point in the camera reference frame given by $v_p(t) = \begin{bmatrix} v_{px} & v_{py} & v_{pz} \end{bmatrix}^T \in \mathcal{V}_p$. The sets $\mathcal{W}$, $\mathcal{V}_c$ and $\mathcal{V}_p$ are closed and bounded sets such that $\mathcal{W} \subset \mathbb{R}^3, \mathcal{V}_c \subset \mathbb{R}^3$ and $\mathcal{V}_p \subset \mathbb{R}^3$. Let the linear and angular camera velocities be denoted by $u = \begin{bmatrix} v_c & \omega \end{bmatrix}^T$.

The states defined in (3) contain unknown structure information of the object. To recover the 3D structure, the state $x(t)$ should be estimated. Using (3) and (7), the dynamics of the state vector $x(t)$ are expressed as

$$\begin{aligned} \dot{x}_1 &= \Omega_1 + f_1 - v_{px}x_3 + x_1 v_{pz} x_3, \\ \dot{x}_2 &= \Omega_2 + f_2 - v_{py}x_3 + x_2 v_{pz} x_3, \\ \dot{x}_3 &= -v_{cz}x_3^2 - (x_2\omega_1 - x_1\omega_2)x_3 + v_{pz}x_3^2, \\ y &= \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T. \end{aligned} \tag{9}$$

where $\Omega_1(u,y)$, $\Omega_2(u,y)$, $f_1(u,x)$, $f_2(u,x)$, $f_3(u,x) \in \mathbb{R}$ are defined as

$$\begin{aligned} \Omega_1(u,y) &\triangleq -x_1x_2\omega_1 + \omega_2 + x_1^2\omega_2 - x_2\omega_3, \\ \Omega_2(u,y) &\triangleq -\omega_1 - x_2^2\omega_1 + x_1x_2\omega_2 + x_1\omega_3, \\ f_1(u,x) &\triangleq (v_{cx} - x_1 v_{cz})x_3, \\ f_2(u,x) &\triangleq (v_{cy} - x_2 v_{cz})x_3, \\ f_3(u,x) &\triangleq -v_{cz}x_3^2 - (x_2\omega_1 - x_1\omega_2)x_3. \end{aligned}$$

Assumption 2: The camera velocities $\omega(t)$, and $v_c(t)$, the object velocity $v_p(t)$, and the feature points $y(t)$ are assumed to be upper bounded by constants.

Assumption 3: The linear velocity of the moving object along the Y and Z-directions of the camera coordinate system is zero, i.e., $v_{py}(t) = v_{pz}(t) = 0$, or the linear velocity of the moving object along only the Z-direction of the camera coordinate system is zero; i.e., $v_{pz}(t) = 0$, $\forall t > 0$. Non-zero linear velocities of the object can be time-varying and unknown.

Remark 1. Assumption 3 is satisfied in many practical scenarios such as an object moving along a straight line with a time-varying unknown velocity or an object moving on a ground plane seen by a downward looking camera attached to an airborne UAV. Consider a scenario where the Z-axis of the world coordinate system is pointed upwards perpendicular to the ground plane and the X, Y axes are in the ground plane. Since the object is moving in the ground plane, the linear velocity of the object in the Z-axis of the world coordinate system is zero and the velocities in X, and Y directions are time-varying unknowns. The Z-axis of the camera coordinate system is pointing downwards, hence, for an object moving in a straight line, $v_{py}(t) = v_{pz}(t) = 0$, and for object moving in a plane $v_{pz}(t) = 0$.

IV. Structure and Motion Estimation

A. Structure and Motion from Motion (SaMfM) Objective

The objective of SaMfM is to recover the structure (i.e., Euclidean coordinates with a scaling factor) and motion (i.e., velocities) of moving objects observed by a moving camera, assuming that all camera velocities are known. In this section, an observer is presented which estimates the structure of the moving object with respect to the moving camera. It is assumed that one or more feature points on the object are tracked in each image frame and camera velocities are recorded using sensors such as an IMU. The camera is assumed to be internally calibrated. Estimating the structure of an object is equivalent to estimating the state $x(t)$ of the feature points on the object in each image frame. Based on the definition of the state in (3), the structure of the moving object can be estimated by scaling $\hat{x}_1(t)$ and $\hat{x}_2(t)$ by $\hat{x}_3(t)$.

B. Nonlinear Unknown Input Observer

In this section a nonlinear unknown input observer is developed for a class of nonlinear system in the following form

$$\begin{aligned} \dot{x} &= f(x,u) + g(y,u) + Dd \\ y &= Cx \end{aligned} \tag{10}$$

where $x(t) \in \mathbb{R}^n$ is a state of the system, $u(t) \in \mathbb{R}^m$ is a measurable control input, $d(t) \in \mathbb{R}^q$ is an unmeasurable input, $y(t) \in \mathbb{R}^p$ is output of the system, the function $f(x,u)$ is nonlinear in $x(t)$, and $u(t)$ and satisfies Lipschitz condition $||f(x,u) - f(\hat{x},u)|| \le \gamma_1 ||x - \hat{x}||$ where $\gamma_1 \in \mathbb{R}^+$. The system given by (9) can be represented in the form of (10) with $f(x,u) = \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix}^T$, $g(y,u) = \begin{bmatrix} \Omega_1 & \Omega_2 & 0 \end{bmatrix}^T$, $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $n = 3$, $m = 6$, $p = 2$. For the object moving along straight line $d(t) = v_{px}x_3$, $D = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, $q = 1$ and for the object moving on a plane $d(t) = \begin{bmatrix} v_{px}x_3 & v_{py}x_3 \end{bmatrix}^T$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^T$, $q = 2$. A general n-dimensional model described by (10) is used for the subsequent development of the nonlinear UIO. The UIO can be used for the specific SaMfM dynamics in (9).

Remark 2. The dynamics in (9) are not observable if: a) the camera is stationary, i.e., $u = 0$, or b) the camera moves along the ray projected by the feature point on the image, i.e., $(v_{cx} - v_{px} - x_1(v_{cz} - v_{pz})) = (v_{cy} - v_{py} - x_2(v_{cz} - v_{pz})) = 0$.

The system in (10) can be written in the following form

$$\begin{aligned} \dot{x} &= Ax + \bar{f}(x,u) + g(y,u) + Dd \\ y &= Cx \end{aligned} \tag{11}$$

where $\bar{f}(x,u) = f(x,u) - Ax$, and $A \in \mathbb{R}^{n\times n}$. The function $\bar{f}(x,u)$ satisfies the Lipschitz condition [29], [30]

$$||f(x,u) - f(\hat{x},u) - A(x-\hat{x})|| \leq (\gamma_1+\gamma_2)\,||x-\hat{x}|| \quad (12)$$

where $\gamma_2 \in \mathbb{R}^+$. In this section, the goal is to design an asymptotically converging state observer to estimate $x(t)$ in the presence of an unknown input $d(t)$ (i.e., the moving object's velocity).

Remark 3. If γ_1 is large, the UIO is stable even in the presence of fast moving nonlinear dynamics in $f(x,u)$. For the SaMfM problem, larger values of γ_1 means camera can move with faster velocities.

An unknown input reduced order state observer for the system (11) is designed as

$$\begin{aligned} \dot{z} &= Nz + Ly + M\bar{f}(\hat{x},u) + Mg(y,u) \\ \hat{x} &= z - Ey \end{aligned} \quad (13)$$

where $\hat{x}(t) \in \mathbb{R}^n$ is an estimate of the unknown state $x(t)$, $z(t) \in \mathbb{R}^n$ is an auxiliary signal, the matrices $N \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{n\times p}$, $E \in \mathbb{R}^{n\times p}$, $M \in \mathbb{R}^{n\times n}$ are designed as [18]

$$\begin{aligned} M &= I + EC \\ N &= MA - KC \\ L &= K(I+CE) - MAE \end{aligned} \quad (14)$$

where $K \in \mathbb{R}^{n\times p}$ is a gain matrix, and E is subsequently designed.

To quantify the estimation objective an estimation error is defined as

$$e(t) \triangleq \hat{x}(t) - x(t) = z - Ey - x. \quad (15)$$

Taking the time derivative of the estimation error and using (11) and (13) yields

$$\begin{aligned} \dot{e} &= \dot{z} - (I+EC)\,\dot{x}, \\ \dot{e} &= Nz + Ly + M\bar{f}(\hat{x},u) - (I+EC)Ax \\ &\quad -(I+EC)\bar{f}(x,u) - (I+EC)Dd. \end{aligned} \quad (16)$$

Using (14) and (15), the error system in (16) can be written as

$$\begin{aligned} \dot{e} &= Ne + N(I+EC)\,x + LCx \\ &\quad + M\left(\bar{f}(\hat{x},u) - \bar{f}(x,u)\right) - MAx - MDd, \\ &= Ne + (NM + LC - MA)\,x \\ &\quad + M\left(\bar{f}(\hat{x},u) - \bar{f}(x,u)\right) - MDd. \end{aligned}$$

Using (14) the equality $NM + LC - MA = 0$ is satisfied, and if E is chosen such that

$$MD = (I+EC)\,D = 0, \quad (17)$$

then the error dynamics can be written as

$$\dot{e} = Ne + M\left(\bar{f}(\hat{x},u) - \bar{f}(x,u)\right). \quad (18)$$

The condition in (17) can be written as

$$ECD = -D.$$

A solution exists for matrix E if $rank(CD) = q$ and the solution is given in a generalized form by [18]

$$E = F + YG \quad (19)$$

where $Y \in \mathbb{R}^{n\times p}$ can be chosen arbitrarily, F and G are given by

$$F \triangleq -D(CD)^{\dagger}, \;\; G \triangleq \left(I_p - (CD)(CD)^{\dagger}\right)$$

and $(CD)^{\dagger}$ denotes the generalized pseudo inverse of the matrix CD given by

$$(CD)^{\dagger} = \left((CD)^T(CD)\right)^{-1}(CD)^T.$$

C. Stability Analysis

Since E can be computed using (19), the only unknowns in (14) are the matrices K and Y. The following theorem gives a condition for choosing K and Y such that the observation error $e(t)$ converges to zero.

Theorem: The nonlinear unknown input observer in (13) is exponentially stable such that

$$\|e(t)\| \leq \|e(t_0)\|\, exp(-\lambda t)$$

where $\lambda \in \mathbb{R}^+$ is a constant, provided Assumptions 1-3 and following sufficient conditions are satisfied

$$N^TP + PN + \left(\gamma_1^2+\gamma_2^2\right)PMM^TP + 2I \quad < \quad 0 \quad (20)$$

where $P \in \mathbb{R}^{n\times n}$ is a positive definite, symmetric matrix.

Proof: Consider a Lyapunov candidate function $V : \mathbb{R}^n \to \mathbb{R}$ defined as

$$V = e^TPe. \quad (21)$$

The Lyapunov function satisfies

$$\lambda_{min}(P)\,\|e\|^2 \leq V \leq \lambda_{max}(P)\,\|e\|^2 \quad (22)$$

where λ_{min} and λ_{max} are the min and max eigen values of the matrix P. Taking the time derivative of (21) along the trajectories of (18) yields

$$\begin{aligned} \dot{V} &= e^T\left(N^TP + PN\right)e + 2e^TPM\left(\bar{f}(\hat{x},u) - \bar{f}(x,u)\right) \\ \dot{V} &= e^T\left(N^TP + PN\right)e + 2e^TPM\left(f(\hat{x},u) - f(x,u)\right) \\ &\quad -2e^TPMA\,(\hat{x}-x) \\ \dot{V} &\leq e^T\left(N^TP + PN\right)e + 2\left\|e^TPM\right\|\|A\|\,\|e\| \\ &\quad +2\left\|e^TPM\right\|\|f(\hat{x},u) - f(x,u)\| \\ \dot{V} &\leq e^T\left(N^TP + PN\right)e + 2\left\|e^TPM\right\|\gamma_1\|e\| \\ &\quad +2\left\|e^TPM\right\|\gamma_2\|e\| \end{aligned}$$

where $\gamma_1 \in \mathbb{R}^+$ is a Lipschitz constant, and $\gamma_2 \in \mathbb{R}^+$ is norm of the matrix A. Using the norm inequality

$$2\gamma_i\left\|e^TPM\right\|\|e\| \quad \leq \quad \gamma_i^2\left\|e^TPM\right\|^2 + \|e\|^2, \;\; \forall i = \{1,2\}$$

the upper bound on $\dot{V}$ is given by

$$\begin{aligned} \dot{V} &\leq e^T\left(N^TP + PN\right)e \\ &\quad + \left(\gamma_1^2+\gamma_2^2\right)e^TPMM^TPe + 2e^Te \\ \dot{V} &\leq e^T\left(N^TP + PN + \left(\gamma_1^2+\gamma_2^2\right)PMM^TP + 2I\right)e. \end{aligned}$$

If $Q \triangleq N^T P + PN + \left(\gamma_1^2 + \gamma_2^2\right) PMM^T P + 2I < 0$ then $\dot{V}$ can be expressed as

$$\dot{V} \quad \leq \quad e^T Q e. \tag{23}$$

Using (22) and (23), the following upper bound for $V(t)$ can be developed

$$V(t) \leq V(t_0)\, exp\left(-\lambda t\right)$$

where $\lambda = \frac{\lambda_{max}(Q)}{\lambda_{min}(P)}$ and the upper bound for the estimation error is given by

$$\|e(t)\| \leq \zeta \, \|e(t_0)\| \, exp\left(-\lambda t\right)$$

where $\zeta = \frac{\lambda_{max}(P)}{\lambda_{min}(P)}$.

Remark 4. Model uncertainties can be represented by an additive disturbance term $d_1(t) \in \mathbb{R}^n$ in (10). The estimation error will be uniformly ultimately bounded in the presence of model uncertainties.

D. Sufficient Condition

The inequality in (20) is satisfied if the pair (MA, C) is observable [18] and the following condition is satisfied [31]

$$\underset{\omega \in \mathbb{R}^+}{min}\, \sigma_{min}\left(MA - KC - j\omega I_3\right) > \sqrt{\gamma_3}\left(\gamma_1 + \gamma_2\right) \tag{24}$$

where $\sigma_{min}(\cdot)$ denotes the minimum singular value of a matrix, and $\gamma_3 \triangleq \lambda_{max}\left(MM^T\right)$. If the pair (MA, C) is observable then the gain matrix K can be selected so that $N = MA - KC$ is Hurwitz. Since $rank(CD) = rank(D) = q$ the condition

$$rank \begin{bmatrix} sI_n - A & D \\ C & 0 \end{bmatrix} = n + q, \;\; \forall s \in \mathbb{C} \tag{25}$$

implies that the pair (MA, C) is observable [18]. Thus, the matrix A in (11) should be chosen such that the condition in (25) is satisfied. Another criteria on the selection of A is to minimize the Lipschitz constant in (12). In the following section, the condition in (20) is reformulated as a LMI feasibility problem.

E. LMI Formulation

The matrices P, K and Y should be selected such that the sufficient condition for the observer error stability in (20) is satisfied. Substituting N and M from (14) into (20) yields

$$\begin{aligned} &(MA - KC)^T P + P\left(MA - KC\right) + 2I \\ &+ \left(\gamma_1^2 + \gamma_2^2\right) P\left(I + EC\right)\left(I + EC\right)^T P \quad < \quad 0. \end{aligned} \tag{26}$$

After using (19), the inequality in (26) can be expressed as

$$\begin{aligned} &A^T\left(I + FC\right)^T P + P\left(I + FC\right) A \\ &+ A^T C^T G^T P_Y^T + P_Y GCA - C^T P_K^T - P_K C \\ &+ 2I + \left(\gamma_1^2 + \gamma_2^2\right)\left(P + PFC + P_Y GC\right) \quad < \quad 0 \\ &\left(P + PFC + P_Y GC\right)^T \end{aligned} \tag{27}$$

where $P_Y = PY$ and $P_K = PK$. For the observer synthesis, the matrices Y, K and $P > 0$ should be computed such that the matrix inequality in (27) is satisfied. Since $P > 0$, P^{-1} exists and Y, and K can be computed using $Y = P^{-1}P_Y$, and $K = P^{-1}P_K$. Using Schur's complement, the inequality in (27) can be transformed into the matrix inequality

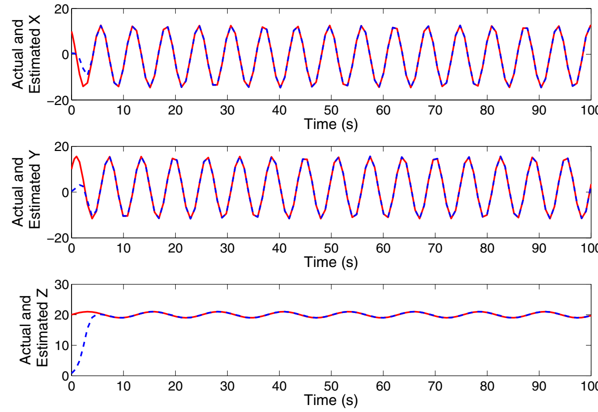

Figure 2. Comparison of the actual and estimated X,Y and Z positions of a moving object with respect to a moving camera.

$$\begin{bmatrix} P_1 & \beta R \\ \beta R^T & -I \end{bmatrix} < 0 \tag{28}$$

where

$$\begin{aligned} P_1 &= A^T\left(I + FC\right)^T P + P\left(I + FC\right) A + \\ &\quad A^T C^T G^T P_Y^T + P_Y GCA - C^T P_K^T - P_K C \\ &\quad +2I, \\ R &= P + PFC + P_Y GC, \\ \beta &= \sqrt{\gamma_1^2 + \gamma_2^2}. \end{aligned}$$

The matrix inequality in (28) is an LMI in variables P, P_Y, and P_K. The LMI feasibility problem can be solved using standard LMI algorithms [32]. The LMI problem in (28) can also be seen as a problem of finding P, P_Y and P_K such that β is maximized. Maximizing β is equivalent to maximizing γ_1 which means the observer can be designed for nonlinear functions with a larger Lipschitz constant.

V. Simulation

Consider a moving camera observing an object moving along a straight line. Camera velocities are given by $v_c(t) = \begin{bmatrix} 2 & 1 & 0.5cos(t/2) \end{bmatrix}^T$ and $\omega(t) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. The object is assumed to be moving with velocity $v_p(t) = \begin{bmatrix} 0.5 & 0 & 0 \end{bmatrix}^T$. The camera calibration matrix is chosen as

$$A_c = \begin{bmatrix} 720 & 0 & 320 \\ 0 & 720 & 240 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrices A, C and D are given by

$$A = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

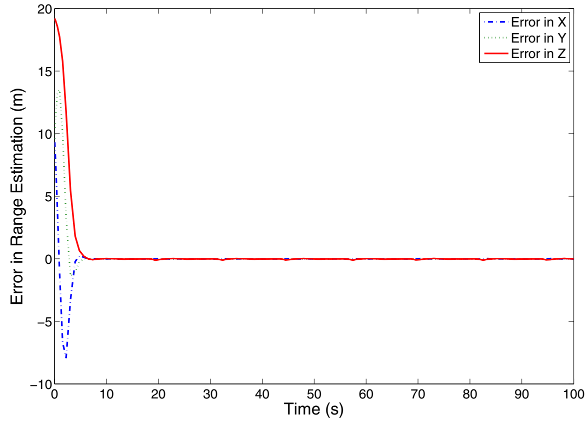

Figure 3. Error in the range estimation of the moving object.

The matrix Y and the gain matrix K are computed using the LMI feasibility command '$feasp$' in Matlab and are given by

$$K = \begin{bmatrix} 0.8278 & 0 \\ 0 & 0.8278 \\ -1.5374 & 0 \end{bmatrix}, \ Y = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & -1.5374 \end{bmatrix}$$

Fig. 2 shows comparison of the actual and estimated X, Y and Z coordinates of the object in the camera coordinate frame. Fig. 3 shows the range estimation error between the moving object and the moving camera.

VI. Conclusion

A nonlinear observer is developed to solve the SaMfM problem. The proposed algorithm estimates the structure of a moving object using a moving camera with less restrictive assumptions on the object motion. The object motion is assumed to be along a straight line or in a plane observed by a moving airborne camera. The algorithm improves on our previous work in [17] by relaxing the constant object velocity assumption to arbitrary object motion in a straight line or in a plane. The observer-based approach is causal and does not assume a minimum number of views or feature points. The structure estimation is insensitive to the object motion in the sense that the state estimation is completely decoupled from the object motion which acts as an exogenous disturbance input.

References

[1] F. Kahl and R. Hartley, "Multiple-view geometry under the L_∞-norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 9, pp. 1603–1617, Sept. 2008.

[2] J. Oliensis, "A critique of structure-from-motion algorithms," *Comput. Vis. Image. Understand.*, vol. 80, pp. 172–214, 2000.

[3] P. Sturm and B. Triggs, "A factorization based algorithm for multi-image projective structure and motion," *Lect. Notes Comput. Sci.*, vol. 1065, pp. 709–720, 1996.

[4] S. Soatto and P. Perona, "Reducing "structure from motion": A general framework for dynamic vision, part 1: Modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 9, pp. 933–942, 1998.

[5] O. Dahl, F. Nyberg, and A. Heyden, "Nonlinear and adaptive observers for perspective dynamic systems," in *Proc. Am. Control Conf.*, New York City, USA, July 2007, pp. 966–971.

[6] M. Jankovic and B. Ghosh, "Visually guided ranging from observations points, lines and curves via an identifier based nonlinear observer," *Syst. Contr. Lett.*, vol. 25, no. 1, pp. 63–73, 1995.

[7] W. E. Dixon, Y. Fang, D. M. Dawson, and T. J. Flynn, "Range identification for perspective vision systems," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2232–2238, Dec. 2003.

[8] A. P. Dani, N. Fischer, and W. E. Dixon, "Single camera structure and motion," *IEEE Trans. Autom. Contr.*, 2011 to appear.

[9] A. P. Dani and W. E. Dixon, *Single Camera Structure and Motion Estimation*. Springer, Lec. Notes in Contr. Inform. Sci., 2010, vol. 401, ch. Visual Servoing via Advanced Numerical Methods, pp. 209–229.

[10] A. P. Dani, K. Rifai, and W. E. Dixon, "Globally exponentially convergent observer for vision-based range estimation," in *IEEE Multi-Conf. Sys. and Contr.*, Yokohama, Japan, September 2010, pp. 801–806.

[11] S. Avidan and A. Shashua, "Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 4, pp. 348–357, Apr 2000.

[12] J. Kaminski and M. Teicher, "A general framework for trajectory triangulation," *J. Math. Imag. Vis.*, vol. 21, no. 1, pp. 27–41, 2004.

[13] M. Han and T. Kanade, "Reconstruction of a scene with multiple linearly moving objects," *Int. J. Comput. Vision*, vol. 59, no. 3, pp. 285–300, 2004.

[14] R. Vidal, Y. Ma, S. Soatto, and S. Sastry, "Two-view multibody structure from motion," *Int. J. Comput. Vision*, vol. 68, no. 1, pp. 7–25, 2006.

[15] C. Yuan and G. Medioni, "3D reconstruction of background and objects moving on ground plane viewed from a moving camera," in *Comput. Vision Pattern Recongnit.*, vol. 2, 2006, pp. 2261 – 2268.

[16] H. Park, T. Shiratori, I. Matthews, and Y. Sheikh, "3D reconstruction of a moving point from a series of 2D projections," in *Euro. Conf. on Comp. Vision*, vol. 6313, 2010, pp. 158–171.

[17] A. P. Dani, Z. Kan, N. Fischer, and W. E. Dixon, "Structure and motion estimation of a moving object using a moving camera," in *Proc. Am. Control Conf.*, Baltimore, MD, 2010.

[18] M. Darouach, M. Zasadzinski, and S. Xu, "Full-order observers for linear systems with unknown inputs," *IEEE Trans. Autom. Control*, vol. 39, no. 3, pp. 606–609, Mar. 1994.

[19] Y. Guan and M. Saif, "A novel approach to the design of unknown input observers," *IEEE Trans. Autom. Control*, vol. 36, no. 5, pp. 632–635, May 1991.

[20] M. Hou and P. Muller, "Design of observers for linear systems with unknown inputs," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 871–875, Jun. 1992.

[21] C.-C. Tsui, "A new design approach to unknown input observers," *IEEE Trans. Autom. Control*, vol. 41, no. 3, pp. 464–468, Mar. 1996.

[22] S. Bhattacharyya, "Observer design for linear systems with unknown inputs," *IEEE Trans. Autom. Control*, vol. 23, no. 3, pp. 483–484, Jun. 1978.

[23] J. Moreno, "Unknown input observers for SISO nonlinear systems," in *IEEE Conf. Decision and Control*, vol. 1, 2000, pp. 790–801.

[24] A. Pertew, H. Marquez, and Q. Zhao, "Design of unknown input observers for Lipschitz nonlinear systems," in *Proc. of American Control Conf.*, 2005, pp. 4198–4203.

[25] W. Chen and M. Saif, "Fault detection and isolation based on novel unknown input observer design," in *Proc. of American Control Conf.*, 2006.

[26] ——, "Unknown input observer design for a class of nonlinear systems: an LMI approach," in *Proc. of American Control Conf.*, 2006.

[27] D. Koenig and S. Mammar, "Design of a class of reduced order unknown inputs nonlinear observer for fault diagnosis," in *Proc. of American Control Conf.*, vol. 3, 2001, pp. 2143–2147.

[28] Y. Ma, S. Soatto, J. Kosecká, and S. Sastry, *An Invitation to 3-D Vision*. Springer, 2004.

[29] E. Yaz and A. Azemi, "Observer design for discrete and continuous nonlinear stochastic systems," *Int. J. Syst. Sci.*, vol. 24, no. 12, pp. 2289–2302, 1993.

[30] L. Xie and P. P. Khargonekar, "Lyapunov-based adaptive state estimation for a class of nonlinear stochastic systems," in *Proc. of American Controls Conf.*, Baltimore, MD, 2010, pp. 6071–6076.

[31] R. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Trans. Autom. Control*, vol. 43, no. 3, pp. 397–401, Mar. 1998.

[32] P. Gahinet, A. Nemirovskii, A. Laub, and M. Chilali, "The LMI control toolbox," in *IEEE Conf. Decision and Control*, vol. 3, 1994, pp. 2038–2041.