

Estimation of an Affine Motion

Lili Ma, Chengyu Cao, Naira Hovakimyan, Craig Woolsey, and Guoqiang Hu

Abstract—This paper discusses the 3D affine motion estimation problem using two cameras via observations of a single feature point. The unknown parameters to be estimated include the nine rotational parameters, the three translational parameters, and the 3D position. One camera assumes a parabolic projection. The other camera is the conventional camera that has a planar projection surface. The parabolic camera is needed due to its ability of estimating the nine rotational parameters independently of the other camera. The conventional camera is utilized to solve the depth ambiguity problem. A closed-loop nonlinear observer is developed for the affine motion estimation problem. Simulation results are presented to show the robustness of the proposed motion estimation scheme.
Key Words: Motion estimation, affine motion, parabolic projection, stereo vision.

I. INTRODUCTION

The problem of 3D motion estimation has been a central topic in development of vision-based navigation and control algorithms. Solutions to this problem can perceptibly affect applications such as vision-based target tracking, coordination, navigation, and obstacle avoidance. Various solutions to this problem have been proposed over years, such as the nonlinear optimization techniques [1]–[3], linear least-square methods [4], and extended Kalman filter [5], [6], in addition to a class of nonlinear observers that arise from a control view point. These nonlinear observers, typically named as perspective nonlinear observers, study the perspective dynamic systems that result from the perspective projection. In general, a perspective dynamic system is a linear system whose output is observed up to a homogeneous line. It is due to this perspective projection that depth ambiguity occurs.

The problem of 3D motion estimation via perspective nonlinear observers has evolved through two stages in its development. First, range identification was addressed, where all the motion parameters of the moving object have been assumed to be known. The objective was to estimate the range, sometimes called the depth. A variety of nonlinear observers have been proposed, including those in [7]–[16], to name but a few. Next, 3D motion estimation was considered that estimates both the unknown position and all (or some of) the unknown motion parameters, by employing stereo vision [12] or multiple cameras running simultaneously [17], by integrating vision with range sensors [18], by active vision (moving the camera on a rod) [19], or by modifying the behavior of the follower (where the camera is installed) to provide persistent excitation to recover the target's information [20].

Despite the significant amount of work on vision-based motion estimation, the solution to motion estimation of an affine motion using a single feature point is not available,

to the best of the authors' knowledge. For instance, the method proposed in [12] studied the rigid-body motion, where the motion dynamics contain only three rotational parameters. The method proposed in [19] did not consider the rotational dynamics. The algorithms presented in [17], [18] assumed a planar object. In our early study, it has been shown that a parabolic camera can be used to estimate the nine unknown constant rotational parameters under certain conditions, when the translational parameters were assumed known [21]. In this paper, we resort to the stereo pinhole-parabolic system and provide a solution to the affine motion estimation problem.

The contribution of this paper is the closed-loop nonlinear solution to the affine motion estimation problem using one parabolic camera and one conventional camera. Based on the orthographic projection provided by the parabolic camera and the measurements of the two cameras, the motion estimation problem is cast into a nonlinear identifier design problem that can be resolved by applying the existing nonlinear identifier-based-observer (IBO) technique [7], [22], [23]. It is worth mentioning that the motion estimation technique proposed in this paper uses the IBO, but is not limited to the IBO. Other nonlinear observers can be equivalently considered for the same purpose.

The paper is organized as follows. Section II describes the affine motion and formulates the assumptions needed for estimation. The IBO is briefly reviewed in Sec. III. In Sec. IV, the planar and the parabolic projection surfaces are given. Our proposed method for 3D affine motion estimation is presented in Sec. V. Section VI shows simulation results, and finally Section VII concludes the paper.

II. AFFINE MOTION

One model for the relative motion of a point in the camera's field of view is given by the following linear system

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \\ \dot{Z}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \\ Z(t) \end{bmatrix} + \begin{bmatrix} b_1(t) \\ b_2(t) \\ b_3(t) \end{bmatrix}, \quad (1)$$

where the matrix $[a_{ij}]$ presents the rotational dynamics, the vector $[b_j]$ corresponds to the translational motion, and $[X(t), Y(t), Z(t)]^\top$ are the coordinates of the point in the camera frame at time instance t . The affine motion dynamics introduced in (1) describes an object motion that undergoes a rotation, translation, and linear deformations [24]. In general, a_{ij} can be time-varying functions. However, in this paper we limit the discussion to constant parameters a_{ij} .

Let

$$\begin{aligned} \mathbf{a} &= [a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}]^\top, \\ \mathbf{b}(t) &= [b_1(t), b_2(t), b_3(t)]^\top. \end{aligned} \quad (2)$$

We make the following assumptions:

- The motion parameters \mathbf{a} and $\mathbf{b}(t)$ are bounded by some known constants.
- $Z(t)$ is bounded away from zero $Z(t) > \epsilon > 0$ for all $t \geq 0$, where ϵ is a constant.
- $X(t)$ and $Y(t)$ are bounded for all $t \geq 0$.
- Similar to that in [12] (pp. 2087), it is assumed that $\dot{\mathbf{b}}(t)$ can be written as

$$\dot{\mathbf{b}}(t) = p(\boldsymbol{\theta}(t), t) \mathbf{b}(t), \quad (3)$$

where

$$\begin{aligned} \boldsymbol{\theta}(t) &= [b_1(t)x_4(t), b_2(t)x_4(t), b_3(t)x_4(t), \mathbf{a}^\top]^\top, \\ x_4(t) &= \frac{1}{\sqrt{X^2(t) + Y^2(t) + Z^2(t)} - Z(t)}. \end{aligned} \quad (4)$$

Further, the function $p(\boldsymbol{\theta}(t), t) \in \mathbb{R}^{3 \times 3}$ and its first time derivative are piecewise smooth and uniformly bounded.

The objective is to estimate $[X(t), Y(t), Z(t)]^\top$ and the motion parameters \mathbf{a} and $\mathbf{b}(t)$, using a stereo vision system consisting of a conventional camera and a parabolic camera.

III. REVIEW OF IBO

We review the IBO in this section, which is used for the estimation task. The IBO is designed for a class of nonlinear systems with the following structure [7]:

$$\begin{cases} \dot{\mathbf{x}}_1(t) = w^\top(\mathbf{x}_1(t), \mathbf{u}(t))\mathbf{x}_2(t) + \phi(\mathbf{x}_1(t), \mathbf{u}(t)), \\ \dot{\mathbf{x}}_2(t) = g(\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{u}(t)), \\ \mathbf{y}(t) = \mathbf{x}_1(t), \end{cases} \quad (5)$$

where $\mathbf{x}_1(t) \in \mathbb{X}_1 \subset \mathbb{R}^{n_1}$, $\mathbf{x}_2(t) \in \mathbb{X}_2 \subset \mathbb{R}^{n_2}$, and $\mathbf{u}(t) \in \mathbb{U} \subset \mathbb{R}^k$. The $n_1 \times n_2$ matrix $w^\top(\mathbf{x}_1(t), \mathbf{u}(t))$ and the vector $g(\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{u}(t))$ are general nonlinear functions of their parameters. Let $\mathbf{x}(t) = [\mathbf{x}_1^\top(t), \mathbf{x}_2^\top(t)]^\top$ and $n = n_1 + n_2$ so that $\mathbf{x}(t) \in \mathbb{X} \subset \mathbb{R}^n$, where $\mathbb{X} = \mathbb{X}_1 \oplus \mathbb{X}_2$.

Following [7], we introduce the following assumptions:

Assumption 3.1:

- 1) Let $\mathbf{x}(t)$ be bounded: $\|\mathbf{x}(t)\| < M$, where $M > 0$. Let $\Omega = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < M\}$. Further, for some fixed constant $\gamma > 1$, let $\Omega_\gamma = \{\mathbf{x}(t) \in \mathbb{R}^n : \|\mathbf{x}(t)\| < \gamma M\}$. Assume that the function $g(\mathbf{x}_1(t), \mathbf{x}_2(t), \mathbf{u}(t))$ is locally Lipschitz in Ω_γ with respect to $\mathbf{x}_2(t)$, i.e., there exists a positive constant α such that

$$\|g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{u}) - g(\mathbf{x}_1, \mathbf{z}_2, \mathbf{u})\| < \alpha \|\mathbf{x}_2 - \mathbf{z}_2\|, \quad (6)$$

for all $\mathbf{x}_2(t), \mathbf{z}_2(t) \in \Omega_\gamma \cap \mathbb{X}_2$, uniformly in $\mathbf{x}_1(t) \in \Omega_\gamma \cap \mathbb{X}_1$ and $\mathbf{u}(t) \in \mathbb{U}$.

- 2) Let the regressor matrix $w^\top(\mathbf{x}_1(t), \mathbf{u}(t))$ and its first time derivative be piecewise smooth and uniformly

bounded. Further, assume that there exist positive constants L_1, L_2, η , and μ such that

$$\|w^\top(\mathbf{x}_1, \mathbf{u})\| < L_1, \quad \left\| \frac{dw^\top(\mathbf{x}_1, \mathbf{u})}{dt} \right\| < L_2, \quad (7a)$$

$$\int_t^{t+\mu} w(\mathbf{x}_1(\tau), \mathbf{u}(\tau)) w^\top(\mathbf{x}_1(\tau), \mathbf{u}(\tau)) d\tau > \eta \mathbb{I}, \quad (7b)$$

for all $t \geq 0$, for all trajectories that originate in \mathbb{X} and for all $\mathbf{u}(t) \in \mathbb{U}$, while \mathbb{I} denotes the identity matrix of appropriate dimension.

Letting $\hat{\mathbf{x}}_1(t)$ and $\hat{\mathbf{x}}_2(t)$ be the estimates of $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$, respectively, the IBO is introduced as:

$$\begin{cases} \dot{\hat{\mathbf{x}}}_1(t) = GA_m(\hat{\mathbf{x}}_1 - \mathbf{x}_1) + w^\top(\mathbf{x}_1, \mathbf{u})\hat{\mathbf{x}}_2 + \phi(\mathbf{x}_1, \mathbf{u}), \\ \dot{\hat{\mathbf{x}}}_2(t) = -G^2 w(\mathbf{x}_1, \mathbf{u})P(\hat{\mathbf{x}}_1 - \mathbf{x}_1) + g(\mathbf{x}_1, \hat{\mathbf{x}}_2, \mathbf{u}), \\ \hat{\mathbf{x}}(t_i^+) = M \frac{\hat{\mathbf{x}}(t_i^-)}{\|\hat{\mathbf{x}}(t_i^-)\|}, \end{cases} \quad (8)$$

where G is a scalar constant and A_m is an $n_1 \times n_1$ Hurwitz matrix. The matrix P is the positive definite solution of the Lyapunov equation $A_m^\top P + PA_m = -Q$ for some $Q > 0$. The sequence t_i is defined as follows:

$$t_i = \min\{t : t > t_{i-1} \text{ and } \|\hat{\mathbf{x}}(t)\| \geq \gamma M\}, \quad t_0 = 0, \quad (9)$$

where γ is a fixed constant.

Theorem 3.1: Subject to Assumption 3.1, there exists a positive constant G_0 such that the estimation errors $e_1(t) = \mathbf{x}_1(t) - \hat{\mathbf{x}}_1(t)$ and $e_2(t) = \mathbf{x}_2(t) - \hat{\mathbf{x}}_2(t)$ converge to zero exponentially, if the constant G in (8) is chosen larger than G_0 [7].

IV. PLANAR AND PARABOLIC PROJECTION

To provide a solution to the 3D affine motion estimation problem, two cameras are used: a conventional camera (with planar projection surface) and a parabolic camera (with parabolic projection surface). We briefly review these two projections in this section.

Conventional Camera: The conventional camera has a planar projection surface. It is supposed that the observed position on the planar surface is defined by

$$\mathbf{y}^*(t) = [y_1^*(t), y_2^*(t)]^\top = \left[\frac{X(t) - m}{Z(t)}, \frac{Y(t) - n}{Z(t)} \right]^\top, \quad (10)$$

where m and n are known constants.

Parabolic Camera: The parabolic camera assumes parabolic projection, which refers to the projection induced by a parabolic mirror onto an image plane. The parabolic projection of a point P in space is the orthographic projection of the intersection of the line connecting the point P with the paraboloid's focus, and the paraboloid [25]. Consider a paraboloid placed in such a way that its symmetric axis is the Z -axis and its focus is at the origin, as shown in Fig. 1. Further, assume that the paraboloid's focal length f equals $1/2$, without loss of generality.

The function of the parabolic surface is

$$\frac{1}{2}(x_1^2 + x_2^2 - 1) = x_3. \quad (11)$$

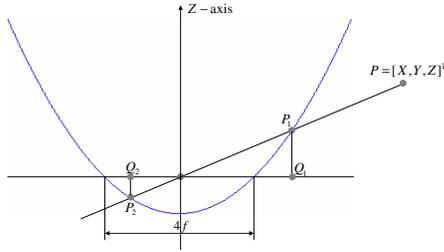


Fig. 1. 2D illustration of a parabolic projection.

The projection of a point $P = [X, Y, Z]^T$ onto the paraboloid surface can be described by $[X, Y, Z]^T / L$ with $L = \pm\sqrt{X^2 + Y^2 + Z^2} - Z$. Let P_1 be the intersection point (as light ray enters the paraboloid), as shown in Fig. 1. Correspondingly, $L = \sqrt{X^2 + Y^2 + Z^2} - Z$ [15], [16], [25].

Let

$$\begin{aligned} \mathbf{x}(t) &= [x_1(t), x_2(t), x_3(t), x_4(t)]^T, \\ &= \frac{1}{L(t)} [X(t), Y(t), Z(t), 1]^T, \end{aligned} \quad (12)$$

where

$$L(t) = \sqrt{X^2(t) + Y^2(t) + Z^2(t)} - Z(t). \quad (13)$$

The system (1) with output observations (12) is equivalent to the following system:

$$\left\{ \begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} &= \begin{bmatrix} b_1 + \Omega_0(t)x_1 \\ b_2 + \Omega_0(t)x_2 \\ b_3 + \Omega_0(t)x_3 \end{bmatrix} x_4 + \begin{bmatrix} \left(1 - \frac{x_1^2}{1+x_3}\right) \sum_{j=1}^3 a_{1j}x_j \\ -\frac{x_1x_2}{1+x_3} \sum_{j=1}^3 a_{1j}x_j \\ -\frac{x_1x_3}{1+x_3} \sum_{j=1}^3 a_{1j}x_j \end{bmatrix} \\ &+ \begin{bmatrix} -\frac{x_1x_2}{1+x_3} \sum_{j=1}^3 a_{2j}x_j + \frac{x_1}{1+x_3} \sum_{j=1}^3 a_{3j}x_j \\ \left(1 - \frac{x_2^2}{1+x_3}\right) \sum_{j=1}^3 a_{2j}x_j + \frac{x_2}{1+x_3} \sum_{j=1}^3 a_{3j}x_j \\ -\frac{x_2x_3}{1+x_3} \sum_{j=1}^3 a_{2j}x_j + \left(1 + \frac{x_3}{1+x_3}\right) \sum_{j=1}^3 a_{3j}x_j \end{bmatrix}, \\ \dot{x}_4(t) &= \left(\sum_{j=1}^3 a_{3j}x_j - \frac{1}{1+x_3} \sum_{i=1}^3 \sum_{j=1}^3 a_{ij}x_jx_i \right) x_4 + \Omega_0(t)x_4^2, \end{aligned} \right. \quad (14)$$

with output

$$\mathbf{y}(t) = [y_1(t), y_2(t), y_3(t)]^T = \frac{1}{L(t)} [X(t), Y(t), Z(t)]^T, \quad (15)$$

where $\Omega_0(t)$ in (14) is

$$\Omega_0(t) = b_3 - \frac{1}{1+x_3} \sum_{i=1}^3 b_i x_i. \quad (16)$$

In the next section, estimation of the affine motion in (1) will be performed using the visual measurements $\mathbf{y}^*(t)$ and $\mathbf{y}(t)$ in equations (10) and (15), respectively.

V. MOTION ESTIMATION

Motion estimation of the rotational parameters is achieved via the parabolic camera. Estimation of the unknown depth/position and the translational parameters is obtained appealing to the stereo vision setup.

A. Estimation of Rotational Parameters \mathbf{a}

Estimation of the rotational parameters \mathbf{a} is achieved via the parabolic camera using visual measurements $\mathbf{y}(t)$. Let

$$\boldsymbol{\theta}(t) = \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \vdots \\ \theta_{12}(t) \end{bmatrix} = \begin{bmatrix} b_1(t)x_4(t) \\ b_2(t)x_4(t) \\ b_3(t)x_4(t) \\ \mathbf{a} \end{bmatrix}, \quad (17)$$

where $x_4(t)$ is given in (4). Using the assumption in equation (3), the system (14) can be rewritten as

$$\left\{ \begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} &= \mathbf{w}_1^T(x_1(t), x_2(t), x_3(t)) \boldsymbol{\theta}(t), \\ \dot{\boldsymbol{\theta}}(t) &= \begin{bmatrix} g_1(x_1(t), x_2(t), x_3(t), \boldsymbol{\theta}(t)) & \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \end{bmatrix} \\ \mathbf{0}_{9 \times 1} \end{bmatrix} \\ &+ \begin{bmatrix} p(\boldsymbol{\theta}(t), t) & \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \end{bmatrix} \\ \mathbf{0}_{9 \times 1} \end{bmatrix}, \end{aligned} \right. \quad (18)$$

where

$$\begin{aligned} \mathbf{w}_1^T(x_1(t), x_2(t), x_3(t)) &= \\ &\begin{bmatrix} 1 - \frac{x_1^2}{1+x_3} & -\frac{x_1x_2}{1+x_3} & \frac{x_1}{1+x_3} & \left(1 - \frac{x_1^2}{1+x_3}\right)[x_1, x_2, x_3] \\ -\frac{x_1x_2}{1+x_3} & 1 - \frac{x_2^2}{1+x_3} & \frac{x_2}{1+x_3} & -\frac{x_1x_2}{1+x_3}[x_1, x_2, x_3] \\ -\frac{x_1x_3}{1+x_3} & -\frac{x_2x_3}{1+x_3} & 1 + \frac{x_3}{1+x_3} & -\frac{x_1x_3}{1+x_3}[x_1, x_2, x_3] \\ -\frac{x_1x_2}{1+x_3}[x_1, x_2, x_3] & \frac{x_1}{1+x_3}[x_1, x_2, x_3] & & \\ \left(1 - \frac{x_2^2}{1+x_3}\right)[x_1, x_2, x_3] & \frac{x_2}{1+x_3}[x_1, x_2, x_3] & & \\ -\frac{x_2x_3}{1+x_3}[x_1, x_2, x_3] & \left(1 + \frac{x_3}{1+x_3}\right)[x_1, x_2, x_3] & & \end{bmatrix} \end{aligned} \quad (19a)$$

and

$$\begin{aligned} g_1(x_1(t), x_2(t), x_3(t), \boldsymbol{\theta}(t)) &= \\ &= \frac{x_1^2\theta_4 + x_1x_2\theta_5 + x_1x_3\theta_6 + x_1x_2\theta_7 + x_2^2\theta_8 + x_2x_3\theta_9}{1+x_3} \\ &+ \frac{x_1\theta_{10} + x_2\theta_{11} + x_3\theta_{12}}{1+x_3} + \theta_3 - \frac{x_1\theta_1 + x_2\theta_2 + x_3\theta_3}{1+x_3}. \end{aligned} \quad (19b)$$

The system (18) exhibits the structure given in (5), to which IBO may be applied, by treating $\mathbf{x}_1(t) = [x_1(t), x_2(t), x_3(t)]^T$, $\mathbf{x}_2(t) = \boldsymbol{\theta}(t)$, and $\phi(\mathbf{x}_1(t), \mathbf{u}(t)) = 0$. To apply the IBO, we need the following assumption for the system (18):

Assumption 5.1: There do not exist constants κ_i (for $i = 1, 2, 3, 4$) with $\sum_{i=1}^4 \kappa_i^2 \neq 0$ such that

$$\kappa_1 x_1(\tau) + \kappa_2 x_2(\tau) + \kappa_3 x_3(\tau) + \kappa_4 = 0, \quad (20)$$

for all $\tau \in [t, t + \mu]$, where μ is an arbitrarily small positive constant.

It is straightforward to verify that $w_1(x_1, x_2, x_3)$ and its first time derivative are piecewise smooth and uniformly bounded. It can be seen from (19a) that rearranging the columns of $w_1^\top(x_1(t), x_2(t), x_3(t))$ one can arrive at

$$\bar{w}_1^\top(x_1(t), x_2(t), x_3(t)) = \underbrace{\begin{bmatrix} 1 - \frac{x_1^2}{1+x_3} & \frac{-x_1x_2}{1+x_3} & \frac{x_1}{1+x_3} \\ -\frac{x_1x_2}{1+x_3} & 1 - \frac{x_2^2}{1+x_3} & \frac{x_2}{1+x_3} \\ -\frac{x_1x_3}{1+x_3} & -\frac{x_2x_3}{1+x_3} & 1 + \frac{x_3}{1+x_3} \end{bmatrix}}_A \otimes \underbrace{[1, x_1, x_2, x_3]}_B, \quad (21)$$

where \otimes denotes the Kronecker product. Using straightforward algebraic manipulations, it can be shown that Assumption 5.1 ensures that all columns in the matrices A and B are linearly independent. Thus, the twelve columns in $w_1^\top(x_1(t), x_2(t), x_3(t))$ are also linearly independent, which leads to (7b), as shown below.

Let $v_i(t)$ be the i^{th} column of $w_1^\top(x_1, x_2, x_3)$ in (19a) (for $i = 1, \dots, 12$). From Assumption 5.1, there do not exist constants κ_i (for $i = 1, \dots, 12$) with $\sum_{i=1}^{12} \kappa_i^2 \neq 0$ such that $\sum_{i=1}^{12} \kappa_i v_i(t) = 0$. Therefore, for any nonzero 12×1 vector ν with $\|\nu\| = 1$, we have $\nu^\top w_1(x_1, x_2, x_3) w_1^\top(x_1, x_2, x_3) \nu > \varepsilon \|\nu\|^2 = \varepsilon$. Thus $w_1(x_1, x_2, x_3) w_1^\top(x_1, x_2, x_3) > \varepsilon \mathbb{I}$ and (7b) is satisfied. Assumption 5.1 says that the observed image data $(x_1(t), x_2(t), x_3(t))$ should not stay in the neighborhood of any fixed plane during a very short time interval.

Let

$$\hat{\theta}(t) = [\hat{\theta}_1(t), \hat{\theta}_2(t), \dots, \hat{\theta}_{12}(t)]^\top \quad (22)$$

and

$$\begin{aligned} e_1(t) &= \hat{x}_1(t) - x_1(t), \\ e_2(t) &= \hat{x}_2(t) - x_2(t), \\ e_3(t) &= \hat{x}_3(t) - x_3(t). \end{aligned} \quad (23)$$

The following observer is considered for (18):

$$\left\{ \begin{aligned} \begin{bmatrix} \dot{\hat{x}}_1(t) \\ \dot{\hat{x}}_2(t) \\ \dot{\hat{x}}_3(t) \end{bmatrix} &= GA_{3 \times 3} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} + w_1^\top(x_1(t), x_2(t), x_3(t)) \hat{\theta}(t), \\ \dot{\hat{\theta}}(t) &= -G^2 w_1(x_1(t), x_2(t), x_3(t)) P_{3 \times 3} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \\ &+ \begin{bmatrix} g_1(x_1(t), x_2(t), x_3(t), \hat{\theta}(t)) & \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \hat{\theta}_3(t) \end{bmatrix} \\ \mathbf{0}_{9 \times 1} \end{bmatrix} \\ &+ \begin{bmatrix} p(\hat{\theta}(t), t) & \begin{bmatrix} \hat{\theta}_1(t) \\ \hat{\theta}_2(t) \\ \hat{\theta}_3(t) \end{bmatrix} \\ \mathbf{0}_{9 \times 1} \end{bmatrix}, \end{aligned} \right. \quad (24)$$

where G is a scalar constant and $A_{3 \times 3}$ is a 3×3 Hurwitz matrix. The matrix $P_{3 \times 3}$ is the positive definite solution of the Lyapunov equation $A_{3 \times 3}^\top P_{3 \times 3} + P_{3 \times 3} A_{3 \times 3} = -Q$ for some $Q > 0$.

According to Theorem 3.1, there exists a positive constant G_0 , such that choosing $G > G_0$ ensures the estimation errors

$$[e_1(t), e_2(t), e_3(t), \hat{\theta}^\top(t) - \theta^\top(t)]^\top$$

converge to zero exponentially. Clearly, this gives $\hat{\alpha}(t)$:

$$\hat{\alpha}(t) = [\hat{\theta}_4(t), \hat{\theta}_5(t), \dots, \hat{\theta}_{12}(t)]^\top, \quad (25)$$

where $\hat{\theta}_i(t)$ (for $i = 4, \dots, 12$) denotes the i^{th} element of $\hat{\theta}(t)$ obtained in (24).

B. Estimation of $Z(t)$ and $\mathbf{b}(t)$

Estimation of the unknown depth $Z(t)$ and the translational parameters $\mathbf{b}(t)$ is obtained using observations of the feature point from two cameras, the parabolic camera and the conventional camera.

From equations (10) and (15), we have

$$\frac{y_1(t)}{y_3(t)} = \frac{X(t)}{Z(t)} = \frac{X(t) - m + m}{Z(t)} = y_1^*(t) + \frac{m}{Z(t)}.$$

Thus, $Z(t)$ can be computed as $Z(t) = \frac{m}{\frac{y_1(t)}{y_3(t)} - y_1^*(t)}$. Similarly, we can have $Z(t) = \frac{m}{\frac{y_2(t)}{y_3(t)} - y_2^*(t)}$. Hence, the estimate of $Z(t)$, denoted by $\hat{Z}(t)$, can be computed from the visual measurements $\mathbf{y}(t)$ and $\mathbf{y}^*(t)$ as

$$\hat{Z}(t) = \sqrt{\left(\frac{m}{\frac{y_1(t)}{y_3(t)} - y_1^*(t)}\right)^2 + \left(\frac{m}{\frac{y_2(t)}{y_3(t)} - y_2^*(t)}\right)^2}, \quad (26a)$$

where $(y_1(t), y_2(t))$ and $(y_1^*(t), y_2^*(t))$ are the visual measurements from the parabolic and the conventional camera, respectively.

Similarly, the estimates of $X(t)$ and $Y(t)$, denoted by $\hat{X}(t)$ and $\hat{Y}(t)$, can be computed as

$$\begin{aligned} \hat{X}(t) &= \sqrt{\left(\frac{y_1(t)}{y_3(t)} \hat{Z}(t)\right)^2 + \left(y_1^*(t) \hat{Z}(t) + m\right)^2}, \\ \hat{Y}(t) &= \sqrt{\left(\frac{y_2(t)}{y_3(t)} \hat{Z}(t)\right)^2 + \left(y_2^*(t) \hat{Z}(t) + m\right)^2}. \end{aligned} \quad (26b)$$

Denote

$$\hat{L}(t) = \sqrt{\hat{X}^2(t) + \hat{Y}^2(t) + \hat{Z}^2(t)} - \hat{Z}(t).$$

It follows from equation (17) that the estimate of $\mathbf{b}(t)$, denoted by $\hat{\mathbf{b}}(t)$, can be obtained by

$$\hat{\mathbf{b}}(t) = \hat{L}(t) [\hat{\theta}_1(t), \hat{\theta}_2(t), \hat{\theta}_3(t)]^\top, \quad (26c)$$

where $\hat{\theta}_i(t)$ (for $i = 1, 2, 3$) denotes the i^{th} element of $\hat{\theta}(t)$.

In summary, 3D affine motion estimation is obtained via equations (25) and (26). It is clear that estimation of \mathbf{a} is achieved using the parabolic camera only. The other quantities are obtained resorting to the stereo vision system. The proposed method is feasible due to the following factors:

- 1) The nonlinearity introduced by the parabolic projection gives the possibility of estimating all the rotational parameters a_{ij} (see the regressor matrix $w_1^\top(x_1(t), x_2(t), x_3(t))$ in (19a), where all columns of the regressor matrix can be linearly independent to satisfy the observability condition (7b)).
- 2) The assumption that the rotational parameters are unknown constants reduces the resulting perspective dynamic system to a linearly parameterized structure, such that existing perspective nonlinear observers can be applied (see (18)).
- 3) The depth ambiguity is resolved via stereo vision.

VI. NUMERICAL SIMULATION

In this section, numerical simulation results are provided to illustrate the performance of the motion estimation technique. The feature point has the following motion dynamics:

$$\begin{bmatrix} \dot{X}(t) \\ \dot{Y}(t) \\ \dot{Z}(t) \end{bmatrix} = \begin{bmatrix} 0.3 & -4 & 0.3 \\ 4 & -0.2 & 0.4 \\ -0.6 & -0.5 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} 2\pi \sin(2\pi t) \\ 2\pi \cos(2\pi t) \\ 0 \end{bmatrix},$$

$$(X_0, Y_0, Z_0) = (1, 1.5, 1), \quad x_0 = (X_0, Y_0, Z_0)/L_0, \quad (27)$$

where $L_0 = \sqrt{X_0^2 + Y_0^2 + Z_0^2} - Z_0$. The observer's initial states are chosen to be $[\hat{x}_1(0), \hat{x}_2(0), \hat{x}_3(0), \hat{x}_4(0)] = [X_0/L_0, Y_0/L_0, Z_0/L_0, 1]$ and $\hat{\mathbf{a}}(0) = 0$. The observer parameters are chosen to be $M = 15$, $A_m = \mathbb{I}_3$, $P = -1/2 \times \mathbb{I}_3$ and $G = 10$, where \mathbb{I}_3 denotes the 3×3 identity matrix.

In the simulations, 1% uniform noise was injected into the measurable image-space signals $\mathbf{y}^*(t)$ and $\mathbf{y}(t)$ via the `randn()` function in Matlab. A step size of 0.02 second is utilized to simulate a typical real-time image processing rate of 30Hz.

Estimation of the rotational parameters \mathbf{a} , the 3D position $[X(t), Y(t), Z(t)]^\top$, and the translational parameters $\mathbf{b}(t)$ are shown in Figs. 2-8, where true values are plotted in solid lines and the corresponding estimates are plotted in dashed lines. Obviously, both position estimation of $[X(t), Y(t), Z(t)]^\top$ and parameter estimations of \mathbf{a} , $\mathbf{b}(t)$ are convergent. Figure 9 shows the trajectory of the observed image data $(x_1(t), x_2(t), x_3(t))$ on the parabolic surface. It can be observed that the image data do not stay within the neighborhood of any fixed plane during any time interval such that Assumption 5.1 is verified.

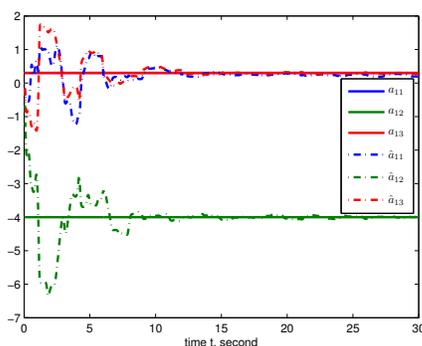


Fig. 2. Estimation of a_{11} , a_{12} , and a_{13} .

VII. CONCLUSION

A nonlinear observer is applied to the affine motion estimation problem, where the motion parameters and the position of a moving object in the space are identified using the perspective observations of a single feature point. Two cameras are used for the estimation task: a conventional camera (with planar projection surface) and a parabolic camera (with parabolic projection). The unknown rotational parameters are estimated via the parabolic camera. Estimation of the position and translational parameters is obtained

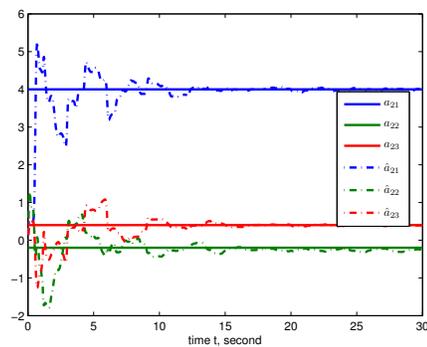


Fig. 3. Estimation of a_{21} , a_{22} , and a_{23} .

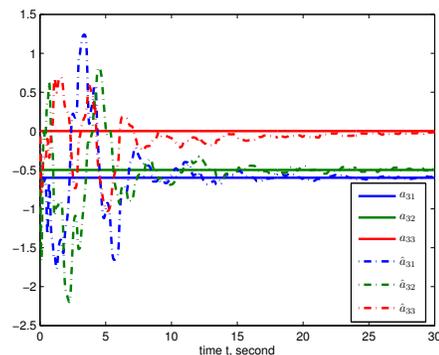


Fig. 4. Estimation of a_{31} , a_{32} , and a_{33} .

resorting to the stereo vision system. Simulation results are presented showing the effectiveness of the proposed estimation technique.

VIII. ACKNOWLEDGMENTS

The work was sponsored by ARO Grant #W911NF-06-1-0330 and AFOSR MURI subcontract F49620-03-1-0401.

REFERENCES

- [1] H. Cho, K. Lee, and S. Lee, "A new robust 3D motion estimation under perspective projection," in *IEEE Conference on Image Processing*, Thessaloniki, Greece, Oct. 2001, pp. 660–663.

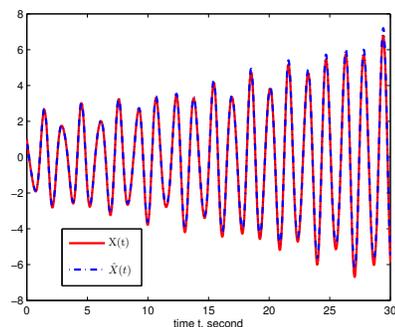


Fig. 5. Estimation of $X(t)$.

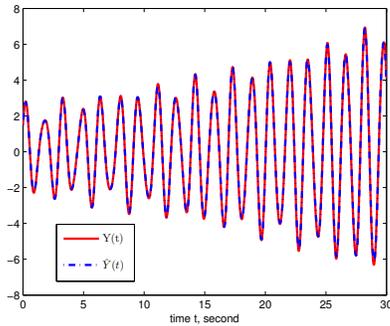


Fig. 6. Estimation of $Y(t)$.

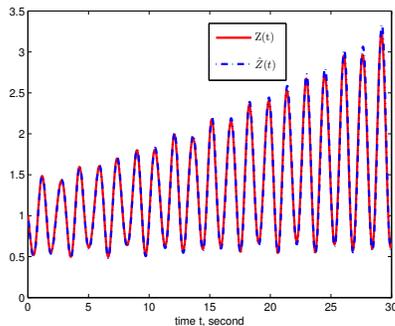


Fig. 7. Estimation of $Z(t)$.

- [2] K. Diamantaras, T. Papadimitriou, M. Srinatzis, and M. Roumeliotis, "Total least squares 3-D motion estimation," in *IEEE Conference on Image Processing*, Chicago, IL, Oct. 1998, pp. 923–927.
- [3] K. Diamantaras and M. Srinatzis, "Camera motion parameter recovery under perspective projection," in *IEEE Conference on Image Processing*, Switzerland, Sep. 1996, pp. 807–810.
- [4] T. Papadimitriou, K. Diamantaras, M. Srinatzis, and M. Roumeliotis, "Robust estimation of rigid-body 3-D motion parameters based on point correspondences," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 4, pp. 541–549, June 2000.
- [5] A. Chiuso, P. Favaro, H. Jin, and S. Soatto, "Structure from motion causally integrated over time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 523–535, Apr. 2002.
- [6] S. Soatto, R. Frezza, and P. Perona, "Motion estimation via dynamic vision," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 393–413, Dec. 1996.
- [7] M. Jankovic and B. Ghosh, "Visually guided ranging from observations of points, lines and curves via an identifier based nonlinear

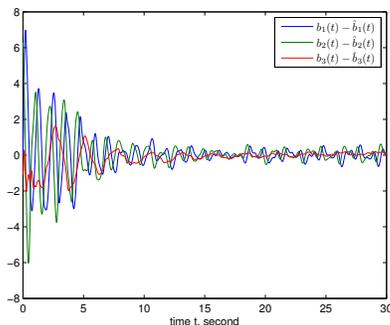


Fig. 8. Estimation of $\mathbf{b}(t)$.

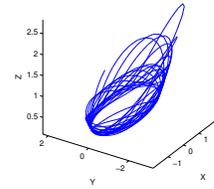


Fig. 9. Trajectory on the parabolic surface.

- observer," *Systems and Control Letters*, vol. 25, pp. 63–73, 1995.
- [8] X. Chen and H. Kano, "A new state observer for perspective systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 658–663, Apr. 2002.
- [9] W. Dixon, Y. Fang, D. Dawson, and T. Flynn, "Range identification for perspective vision systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2232–2238, Dec. 2003.
- [10] L. Ma, Y. Chen, and K. Moore, "Range identification for perspective dynamic systems using linear approximation," in *IEEE Conference on Robotics and Automation*, New Orleans, Apr.-May 2004, pp. 1658–1663.
- [11] D. Karagiannis and A. Astolfi, "A new solution to the problem of range identification in perspective vision systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 2074–2077, Dec. 2005.
- [12] X. Chen and H. Kano, "State observer for a class of nonlinear systems and its application to machine vision," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 2085–2091, Nov. 2004.
- [13] L. Ma, Y. Chen, and K. Moore, "Range identification for perspective dynamic system with a single homogeneous observation," *International Journal of Applied Mathematics and Computer Science*, vol. 15, no. 1, pp. 63–72, 2005.
- [14] O. Dahl, F. Nyberg, J. Holst, and A. Heyden, "Linear design of a nonlinear observer for perspective systems," in *IEEE Conference on Robotics and Automation*, Barcelona, Spain, Apr. 2005, pp. 429–435.
- [15] S. Gupta, D. Aiken, G. Hu, and W. Dixon, "Lyapunov-based range and motion identification for a nonaffine perspective dynamic system," in *American Control Conference*, Minneapolis, MN, June 2006, pp. 4471–4476.
- [16] D. Aiken, S. Gupta, G. Hu, and W. Dixon, "Lyapunov-based range identification for a paracatadioptric system," in *IEEE Conference on Decision and Control*, San Diego, CA, Dec. 2006, pp. 3879–3884.
- [17] B. Ghosh and C. Martin, "Homogeneous dynamical systems theory," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 462–472, Mar. 2002.
- [18] S. Takahashi and B. Ghosh, "Motion and shape identification with vision and range," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1392–1396, Aug. 2002.
- [19] C. Cao, N. Hovakimyan, and J. Evers, "Active control of visual sensor for aerial tracking," in *Proceedings of AIAA Guidance, Navigation and Control Conference*, Keystone, CO, Aug. 2006, AIAA-2006-6610.
- [20] C. Cao and N. Hovakimyan, "Vision-based tracking using intelligent excitation," *International Journal of Control*, 2008, To Appear.
- [21] L. Ma, C. Cao, N. Hovakimyan, W. Dixon, and C. Woolsey, "Range identification in the presence of unknown motion parameters for perspective vision systems," in *American Control Conference*, New York, NY, July 2007, pp. 972–977.
- [22] B. Ghosh, E. Loucks, and M. Jankovic, "An introduction to perspective observability and recursive identification problems in machine vision," in *IEEE Conference on Decision and Control*, 1994, pp. 3229–3234.
- [23] B. Ghosh, M. Jankovic, and E. Loucks, "On the problem of observing motion and shape," in *IEEE Conference on Automation, Robotics, and Computer Vision*, Singapore, Nov. 1994, pp. 653–657.
- [24] R. Tsai and T. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch," *IEEE Transactions on Acoustic, Speech, and Signal Processing*, vol. ASSP-29, no. 6, pp. 1147–1152, Dec. 1981.
- [25] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *European Conference on Computer Vision*, Ireland, June 2000, pp. 445–461.