

# Euclidean Position Estimation of Static Features Using a Moving Uncalibrated Camera

Nitendra Nath, *Student Member, IEEE*, Darren M. Dawson, *Senior Member, IEEE*, and  
Enver Tatlicioglu, *Member, IEEE*

**Abstract**—In this paper, a novel Euclidean position estimation technique using a single uncalibrated camera mounted on a moving platform is developed to asymptotically recover the 3-D Euclidean position of static object features. The position of the moving platform is assumed to be measurable, and a second object with known 3-D Euclidean coordinates relative to the world frame is considered to be available *a priori*. To account for the unknown camera calibration parameters and to estimate the unknown 3-D Euclidean coordinates, an adaptive least squares estimation strategy is employed based on prediction error formulations and a Lyapunov-type stability analysis. The developed estimator is shown to recover the 3-D Euclidean position of the unknown object features despite the lack of knowledge of the camera calibration parameters. Numerical simulation results along with experimental results are presented to illustrate the effectiveness of the proposed algorithm.

**Index Terms**—Euclidean position estimation, least squares estimation, Lyapunov methods, nonlinear systems, perspective vision systems.

## I. INTRODUCTION

THREE-D RECONSTRUCTION of an object, where the Euclidean coordinates of the features on a moving or fixed object are recovered from a sequence of 2-D images, has received noteworthy attention over the last several years. The recovery of the 3-D Euclidean coordinates is usually done by mounting a camera on a moving vehicle such as an unmanned aerial vehicle (UAV) or a mobile robot which travels through the environment and captures images of static objects or features. 3-D reconstruction or 3-D Euclidean position estimation has significance in several applications such as autonomous vehicle navigation, aerial tracking, path planning, surveillance, etc.

Although, the problem of Euclidean reconstruction is inherently nonlinear, linearization based techniques, such as the extended Kalman filter (EKF) [2], [3], have been used quite frequently.

Manuscript received March 29, 2010; revised October 02, 2010; accepted November 27, 2010. Manuscript received in final form February 21, 2011. Date of publication April 07, 2011; date of current version February 01, 2012. Recommended by Associate Editor Z. Gao. This work was supported in part by a DOE Contract and by a Honda Corporation Grant. A preliminary version of this work has appeared in the Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 2009, pp. 4548–4553.

N. Nath is with Takata-Electronics, Pontiac, MI 48341 USA (e-mail: nnath@g.clemson.edu).

D. M. Dawson is with the Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634 USA (e-mail: ddarren@clemson.edu).

E. Tatlicioglu is with the Department of Electrical and Electronics Engineering, Izmir Institute of Technology, Izmir 35430, Turkey (e-mail: envertatlicioglu@iyte.edu.tr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCST.2011.2120610

However, linearized motion models can cause significant inconsistencies in solutions, as noted in [4]. Moreover, EKF involves *a priori* knowledge of noise distribution. To overcome the shortcomings of the linear model several researchers focused on utilizing nonlinear system analysis and estimation tools to develop nonlinear state observers for depth estimation or 3-D reconstruction [5]–[14]. All these works utilized velocity measurement of the camera or the object in order to estimate the depth where the camera's extrinsic calibration parameters (rotational matrix and translation vector of the camera relative to its mounting frame) were not considered, and the intrinsic calibration parameters (a matrix consisting functions of the camera's internal parameters; namely, focal length, scaling factors, pixel coordinates of the principal point, and camera axes angles) were set to unity to simplify the problem. In practice, these calibration parameters are often required to estimate the depth or the structure of an object. In other words, the aforementioned works require the camera to be calibrated. In [15], a valid intrinsic calibration matrix of the camera was assumed to recover the depth of an object. Our recent work [16] utilized position measurements of the camera to recover the structure of an object. However, this work assumed a calibrated camera, i.e., the intrinsic and the extrinsic camera calibration parameters were assumed to be known. The work in [16] was extended and evaluated experimentally in [17] where it was realized that calibrating a camera is tedious and complicated, especially calibration of the camera's extrinsic parameters.

Hartley and Zisserman in [18] discussed techniques where the essential matrix (when intrinsic calibration parameters are known) can be decomposed to obtain camera's extrinsic calibration parameters. This technique usually results in multiple solutions for the rotation matrix, and the translation vector is found up to an ambiguous scale factor; thus, making it difficult to select the correct solution. If the intrinsic parameters are also unknown along with extrinsic parameters, the fundamental matrix can be obtained. However, the knowledge of just the fundamental matrix is not sufficient to estimate the depth information. It is also noted that recovering the 3-D Euclidean coordinates of an object from a sequence of its 2-D images with a single uncalibrated camera is a very difficult task. In [19], Hartley proposed a multiple-step algorithm for Euclidean reconstruction from uncalibrated camera views. It was noted in [20] that a true 3-D Euclidean scene of an object using a single camera with unconstrained motion and unknown parameters can not be reconstructed. These issues motivated us to develop a simple and an easily implementable estimator for 3-D Euclidean position estimation using a single uncalibrated camera where both, the intrinsic and the extrinsic camera calibration parameters are unknown.

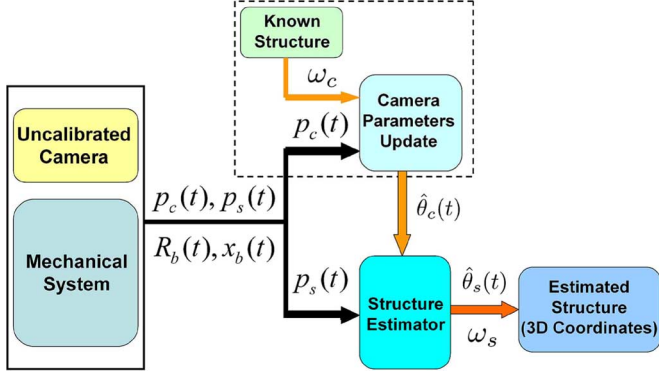


Fig. 1. Illustration of the proposed 3-D Euclidean coordinates estimation technique.

In this paper, our goal is to develop an estimator to identify the 3-D Euclidean coordinates of features on a stationary object by mounting an uncalibrated camera on a mobile platform whose position is measurable. To achieve this goal, first a geometric model (see Fig. 2) is developed to relate the fixed features on the object with the moving camera where the 3-D Euclidean coordinates of the other object is considered to be available *a priori*. The model is then parameterized for known and unknown object features. A prediction error formulation is then presented along with an auxiliary prediction error that allow us to utilize nonlinear estimation theory to design two adaptive least-squares estimators to compensate for the unknown camera calibration parameters and to estimate the structure. We show that the developed structure estimator (see Fig. 1) identifies the Euclidean coordinates of the object features upon satisfaction of a persistency of excitation (PE) condition, and is not dependent on an accurate estimation of the unknown camera calibration parameters. The proposed estimation technique can be useful in places where a known object exists, and the 3-D Euclidean coordinates of another object has to be estimated. Also, when a camera on a mechanical system is replaced or orientation of the camera is changed, it is not required to recalibrate the camera if a known object is present. The developed estimator provides good results and is robust to noise as demonstrated by the simulation results. The validity of the proposed estimation technique is demonstrated by experimental results.

## II. GEOMETRIC MODEL

To develop a geometric relationship between a perspective moving camera and features on known, and unknown static objects, we define an orthogonal coordinate frame, denoted by  $\mathcal{C}$ , whose origin coincides with the optical center of the camera, an inertial coordinate frame, denoted by  $\mathcal{W}$ , and an orthogonal coordinate frame, denoted by  $\mathcal{B}$  (see Fig. 2). In Fig. 2,  $\mathcal{F}_{si} \forall i = 1, \dots, n$  denotes the  $i$ th object feature whose unknown 3-D Euclidean coordinates relative to the world frame  $\mathcal{W}$  are

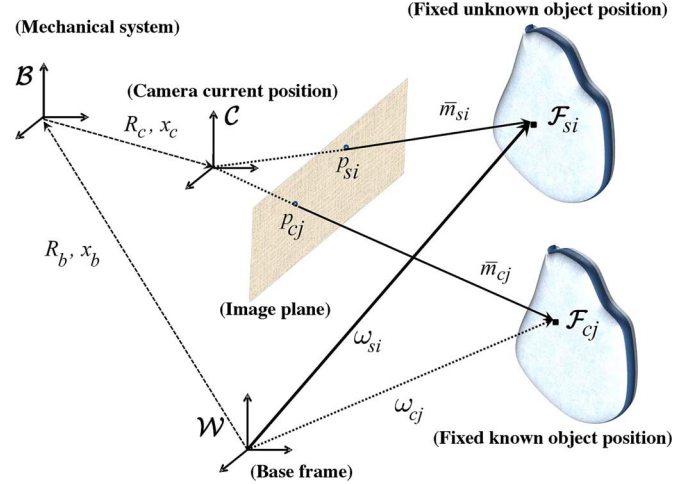


Fig. 2. Geometric relationships between the fixed objects, mechanical system, and the camera.

denoted as the constant  $\omega_{si} \in \mathbb{R}^3$ , and  $\mathcal{F}_{cj} \forall j = 1, \dots, m$  represents the  $j$ th object feature<sup>1</sup> whose corresponding 3-D Euclidean coordinates relative to the world frame  $\mathcal{W}$  are known *a priori* and denoted by the constant  $\omega_{cj} \in \mathbb{R}^3$ . The 3-D coordinates of the object features relative to the camera frame  $\mathcal{C}$ , denoted by  $\bar{m}_{si/cj}(t) \in \mathbb{R}^3$ , is defined as follows<sup>2</sup>

$$\bar{m}_{si/cj} \triangleq [x_{si/cj} \quad y_{si/cj} \quad z_{si/cj}]^T. \quad (1)$$

In the subsequent development, it is assumed that both the objects are always in the field of view of the camera; hence, the distances from the origin of  $\mathcal{C}$  to all the features are always positive and bounded. To relate the coordinate systems, let  $R_b(t) \in SO(3)$  and  $x_b(t) \in \mathbb{R}^3$  denote the measurable rotation matrix and the measurable translation vector, respectively, from  $\mathcal{B}$  to  $\mathcal{W}$ , expressed in  $\mathcal{W}$ . Let  $R_c \in SO(3)$  and  $x_c \in \mathbb{R}^3$  denote the unknown rotation matrix and the unknown translation vector, respectively, from  $\mathcal{C}$  to  $\mathcal{B}$ , expressed in  $\mathcal{B}$ . The pixel coordinates of the object features projected on the image plane, denoted by  $p_{si/cj}(t) \in \mathbb{R}^3$ , is defined as follows:

$$p_{si/cj} \triangleq [u_{si/cj} \quad v_{si/cj} \quad 1]^T \quad (2)$$

where  $u_{si/cj}(t), v_{si/cj}(t) \in \mathbb{R}$ . The projected pixel coordinates of the features are related to the Euclidean coordinates by the pin-hole camera model [21] such that

$$p_{si/cj} = \frac{1}{z_{si/cj}} A \bar{m}_{si/cj} \quad (3)$$

<sup>1</sup>Throughout the paper the subscript  $si$  denotes the  $i$ th feature whose 3-D coordinates relative to the frame  $\mathcal{W}$  is unknown, and  $cj$  denotes the  $j$ th feature whose 3-D coordinates relative to the frame  $\mathcal{W}$  is known.

<sup>2</sup>The notation  $Y_{si/cj}$  implies  $Y_{si}$  or  $Y_{cj}$  throughout the paper. If the left-hand side of any expression is considered with the subscript  $si$  then the right-hand side of the expression is with  $si$ , and similarly for  $cj$ .

where  $A \in \mathbb{R}^{3 \times 3}$  is the unknown constant intrinsic camera calibration matrix of the following form [22]:

$$A \triangleq \begin{bmatrix} fk_u & fk_u \cot \phi & u_0 \\ 0 & \frac{fk_v}{\sin \phi} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

where  $k_u, k_v \in \mathbb{R}$  denote camera scaling factors,  $u_0, v_0 \in \mathbb{R}$  represent the pixel coordinates of the principal point,  $\phi \in \mathbb{R}$  is the angle between the camera axes, and  $f \in \mathbb{R}$  is the camera focal length.

### III. PARAMETERIZATION OF THE MODEL

In this section, the parameterization of  $A\bar{m}_{si/cj}(t)$  and the depth variable  $z_{si/cj}(t)$  are presented. The nonlinear static model given in (3) is parameterized for two cases: case 1, where the 3-D coordinates of the features relative to  $\mathcal{W}$  are unknown (i.e., for  $si$ ), and case 2, where the 3-D coordinates of the features relative to  $\mathcal{W}$  are known *a priori* (i.e., for  $cj$ ).

From Fig. 2,  $\bar{m}_{si/cj}(t)$  can be written as follows [16]:

$$\bar{m}_{si/cj} = R_c^T [R_b^T(\omega_{si/cj} - x_b) - x_c]. \quad (5)$$

After substituting (5) into (3), the pixel coordinates for the object features can be written as follows:

$$p_{si/cj} = \frac{1}{z_{si/cj}} AR_c^T [R_b^T(\omega_{si/cj} - x_b) - x_c] \quad (6)$$

and the corresponding depth  $z_{si/cj}(t)$  can be written as follows:

$$z_{si/cj} = R_{c3}^T [R_b^T(\omega_{si/cj} - x_b) - x_c] \quad (7)$$

where  $R_{c3}^T \in \mathbb{R}^{1 \times 3}$  is the last row of  $R_c^T$ .

*Case 1:* For the object features that have unknown 3-D coordinates relative to  $\mathcal{W}$ ,  $p_{si}(t)$  is parameterized as follows:

$$p_{si} = \frac{1}{\Pi(X, \theta_{c2})\theta_{si}} W(X, \theta_{c1})\theta_{si} \quad \forall i = 1, \dots, n \quad (8)$$

where

$$\Pi(\cdot)\theta_{si} = z_{si} = R_{c3}^T [R_b^T(\omega_{si} - x_b) - x_c] \quad (9)$$

$$W(\cdot)\theta_{si} = AR_c^T [R_b^T(\omega_{si} - x_b) - x_c]. \quad (10)$$

In (8)–(10),  $\theta_{c1} \in \mathbb{R}^{12}$  and  $\theta_{c2} \in \mathbb{R}^4$  are the constant vectors containing all the camera parameters [23], the variable  $X(t)$  is a matrix of appropriate dimension containing the combinations of the elements of the measurable signals  $R_b(t)$  and  $x_b(t)$ ,  $\Pi(\cdot) \in \mathbb{R}^{1 \times 4}$ ,  $W(\cdot) \in \mathbb{R}^{3 \times 4}$  are regression matrices, and  $\theta_{si} \in \mathbb{R}^4$  is an unknown constant parameter vector, which is defined as

$$\theta_{si} \triangleq [\omega_{si1} \quad \omega_{si2} \quad \omega_{si3} \quad 1]^T \quad (11)$$

where  $\omega_{sir} \in \mathbb{R} \forall r = 1, 2, 3$ , is the unknown Euclidean coordinate of the  $i$ th feature for the unknown object relative to  $\mathcal{W}$ .

*Case 2:* For known 3-D coordinates of the features relative to  $\mathcal{W}$ ,  $p_{cj}(t)$  is parameterized as follows [23]:

$$p_{cj} = \frac{1}{\Pi_{xj}(\bar{X})\theta_{c2}} W_{xj}(\bar{X})\theta_{c1} \quad \forall j = 1, \dots, m \quad (12)$$

where  $W_{xj}(\cdot) \in \mathbb{R}^{3 \times 12}$ ,  $\Pi_{xj}(\cdot) \in \mathbb{R}^{1 \times 4}$  are known regression matrices ( $\bar{X}(t)$  is a matrix of appropriate dimension containing combinations of the measurable signals and structure information of the known object).

It should be noted that  $z_{si/cj}(t)$  is assumed to satisfy the following inequalities:

$$\rho_{si/cj}(\bar{m}_{si/cj}) \geq z_{si/cj}(t) = \text{denum}(p_{si/cj}) \geq \varepsilon_{si/cj} > 0 \quad (13)$$

where  $\text{denum}(\cdot)$  denotes the denominator of  $(\cdot)$ ,  $\rho_{si/cj}(\bar{m}_{si/cj}) \in \mathbb{R}$  is a positive function and  $\varepsilon_{si/cj} \in \mathbb{R}$  is a positive constant  $\forall i, j$ . The objective of this work is to identify the 3-D Euclidean coordinates of the features on an unknown object  $\theta_{si}$  in the presence of unknown camera parameters  $\theta_{c1}, \theta_{c2}$ .<sup>3</sup> It is to be noted that our novel structure estimation technique guarantees the estimation of the 3-D Euclidean coordinates upon the satisfaction of a PE condition, and it does not require an accurate estimation of the camera parameters.

### IV. EUCLIDEAN STRUCTURE ESTIMATION

In this section, a prediction error formulation for the parameterized model given in (8) is presented along with an auxiliary prediction error formulation which accounts for the unknown camera parameters. A structure estimator is then presented along with the stability analysis.

#### A. Prediction Error Formulation

To proceed with the error development, the term  $\Pi(\cdot)\theta_{si}$  is multiplied to the both sides of (8) that results in the following expression:

$$p_{si}\Pi\theta_{si} = W\theta_{si}. \quad (14)$$

The estimate of (14) can be written as follows:

$$\hat{p}_{si}\hat{\Pi}\hat{\theta}_{si} = \hat{W}\hat{\theta}_{si} \quad (15)$$

where  $\hat{p}_{si}(t) \in \mathbb{R}^3$  and  $\hat{\theta}_{si}(t) \in \mathbb{R}^4$  are the estimates of  $p_{si}(t)$  and  $\theta_{si}$ , respectively,  $\hat{\Pi}$  and  $\hat{W}$  denote  $\Pi(X, \hat{\theta}_{c2})$  and  $W(X, \hat{\theta}_{c1})$ , respectively, where  $\hat{\theta}_{c1}(t) \in \mathbb{R}^{12}$ ,  $\hat{\theta}_{c2}(t) \in \mathbb{R}^4$  are the estimates of  $\theta_{c1}$  and  $\theta_{c2}$ , respectively. To facilitate the development, the terms  $\hat{W}\theta_{si}$  and  $\hat{p}_{si}\Pi\theta_{si}$  are added and subtracted from the right-hand side and the left-hand side of (14), respectively, that yields the following expression:

$$[p_{si} - \hat{p}_{si}]\Pi\theta_{si} + \hat{p}_{si}\Pi\theta_{si} = \hat{W}\theta_{si} + [W - \hat{W}]\theta_{si}. \quad (16)$$

After subtracting (15) from (16), the following expression is obtained:

$$\tilde{p}_{si}\Pi\theta_{si} + \hat{p}_{si}\Pi\theta_{si} - \hat{p}_{si}\hat{\Pi}\hat{\theta}_{si} = \hat{W}\tilde{\theta}_{si} + \tilde{W}\theta_{si} \quad (17)$$

where  $\tilde{p}_{si}(t) \triangleq p_{si} - \hat{p}_{si} \in \mathbb{R}^3$  is the prediction error for the  $i$ th feature point  $\forall i = 1, \dots, n$ ,  $\tilde{\theta}_{si}(t) \triangleq \theta_{si} - \hat{\theta}_{si} \in \mathbb{R}^4$  is the structure estimation error, and  $\tilde{W}$  denotes  $W(X, \tilde{\theta}_{c1})$ , where  $\tilde{\theta}_{c1}(t) \triangleq \theta_{c1} - \hat{\theta}_{c1} \in \mathbb{R}^{12}$ . After adding and subtracting the

<sup>3</sup>The estimates of  $A$ ,  $R_c$ , and  $x_c$  can be obtained from the estimates of  $\theta_{c1}$ ,  $\theta_{c2}$  (see [23] for a detailed explanation).

term  $\hat{p}_{si}\hat{\Pi}\theta_{si}$  to the left-hand side of (17) and simplifying, the following expression can be obtained:

$$\tilde{p}_{si} = \frac{1}{\hat{\Pi}\theta_{si}} \left\{ [\hat{W} - \hat{p}_{si}\hat{\Pi}]\tilde{\theta}_{si} + [\tilde{W} - \hat{p}_{si}\tilde{\Pi}]\theta_{si} \right\} \quad (18)$$

where  $\tilde{\Pi}$  denotes  $\Pi(X, \tilde{\theta}_{c2})$  and  $\tilde{\theta}_{c2}(t) \triangleq \theta_{c2} - \hat{\theta}_{c2} \in \mathbb{R}^4$ . To ease the subsequent analysis, we combine these individual vectors  $\forall i$  to obtain their respective compact forms. The combined form of the pixel coordinates and their estimates for all the feature points on the unknown object, denoted by  $p_s(t), \hat{p}_s(t) \in \mathbb{R}^{3n}$ , respectively, are defined as follows:

$$p_s \triangleq [p_{s1}^T \ p_{s2}^T \ \dots \ p_{sn}^T]^T; \hat{p}_s \triangleq [\hat{p}_{s1}^T \ \hat{p}_{s2}^T \ \dots \ \hat{p}_{sn}^T]^T \quad (19)$$

and the prediction error  $\tilde{p}_s(t) \in \mathbb{R}^{3n}$  is defined as follows:

$$\tilde{p}_s \triangleq p_s - \hat{p}_s = [\tilde{p}_{s1}^T \ \tilde{p}_{s2}^T \ \dots \ \tilde{p}_{sn}^T]^T. \quad (20)$$

Based on (18), the prediction error  $\tilde{p}_s(t)$  can be written as

$$\tilde{p}_s = B\bar{W}_s\tilde{\theta}_s + BW_s\theta_s \quad (21)$$

where  $\bar{W}_s(t) \in \mathbb{R}^{3n \times 4n}$  is a measurable signal defined as follows:

$$\bar{W}_s \triangleq \begin{bmatrix} \hat{W} - \hat{p}_{s1}\hat{\Pi} & 0_{3 \times 4} & \dots & 0_{3 \times 4} \\ 0_{3 \times 4} & \hat{W} - \hat{p}_{s2}\hat{\Pi} & \dots & 0_{3 \times 4} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{3 \times 4} & 0_{3 \times 4} & \dots & \hat{W} - \hat{p}_{sn}\hat{\Pi} \end{bmatrix} \quad (22)$$

and  $W_s(t) \in \mathbb{R}^{3n \times 4n}$  is an auxiliary matrix defined as

$$W_s \triangleq \begin{bmatrix} \tilde{W} - \hat{p}_{s1}\tilde{\Pi} & 0_{3 \times 4} & \dots & 0_{3 \times 4} \\ 0_{3 \times 4} & \tilde{W} - \hat{p}_{s2}\tilde{\Pi} & \dots & 0_{3 \times 4} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{3 \times 4} & 0_{3 \times 4} & \dots & \tilde{W} - \hat{p}_{si}\tilde{\Pi} \end{bmatrix} \quad (23)$$

where  $0_{3 \times 4} \in \mathbb{R}^{3 \times 4}$  is a zero matrix. In (21),  $\theta_s \triangleq [\theta_{s1}^T \ \theta_{s2}^T \ \dots \ \theta_{sn}^T]^T \in \mathbb{R}^{4n}$ , and  $B(t) \in \mathbb{R}^{3n \times 3n}$  is an auxiliary matrix defined as

$$B \triangleq \text{diag}\{\varphi_1, \varphi_1, \varphi_1, \dots, \varphi_n, \varphi_n, \varphi_n\} \quad (24)$$

where  $\varphi_i \triangleq (1/\hat{\Pi}\theta_{si}) \forall i = 1, \dots, n$ . The combined form of the structure estimation errors of the features, denoted by  $\tilde{\theta}_s(t) \in \mathbb{R}^{4n}$ , is defined as follows:

$$\tilde{\theta}_s \triangleq \theta_s - \hat{\theta}_s = [\tilde{\theta}_{s1}^T \ \tilde{\theta}_{s2}^T \ \dots \ \tilde{\theta}_{sn}^T]^T. \quad (25)$$

The expression given in (21), can be rewritten as follows:

$$\tilde{p}_s = B\bar{W}_s\tilde{\theta}_s + BS \quad (26)$$

where  $S(t) \triangleq W_s\theta_s \in \mathbb{R}^{3n}$ . After utilizing (18),  $S_i(t) \in \mathbb{R}^3$  for the  $i$ th object feature can be written as follows:

$$S_i = [W(X, \tilde{\theta}_{c1}) - \hat{p}_{si}\Pi(X, \tilde{\theta}_{c2})]\theta_{si} \quad (27)$$

which can be further written as follows:<sup>4</sup>

$$S_i = [W_{xi}(X, \theta_{si}) - \hat{p}_{si}\Pi_{xi}(X, \theta_{si})]\tilde{\theta}_c \quad \forall i = 1, \dots, n \quad (28)$$

where  $W_{xi}(\cdot) \in \mathbb{R}^{3 \times 12}$ ,  $\Pi_{xi}(\cdot) \in \mathbb{R}^{1 \times 4}$ , and  $\tilde{\theta}_c \triangleq [\theta_{c1}^T \ \theta_{c2}^T]^T \in \mathbb{R}^{16}$ .

In order to make the structure estimation error  $\tilde{\theta}_s(t)$  go to zero, we seek to make the prediction error  $\tilde{p}_s(t)$  given in (26) go to zero.

### B. Auxiliary Prediction Error Formulation

To further facilitate the development and to account for the unknown intrinsic and extrinsic camera parameters, both sides of (12) are multiplied with the term  $\Pi_{xj}(\cdot)\theta_{c2}$  to obtain the following expression:

$$p_{cj}\Pi_{xj}(\cdot)\theta_{c2} = W_{xj}(\cdot)\theta_{c1} \quad \forall j = 1, \dots, m. \quad (29)$$

The estimate form of (29) is written as follows:

$$\hat{p}_{cj}\Pi_{xj}(\cdot)\hat{\theta}_{c2} = W_{xj}(\cdot)\hat{\theta}_{c1} \quad (30)$$

where  $\hat{p}_{cj}(t) \in \mathbb{R}^3$  is the pixel estimate of the  $j$ th feature. After subtracting (30) from (29), and then adding and subtracting the term  $\hat{p}_{cj}\Pi_{xj}\theta_{c2}$  to the left-hand side results in the following expression:

$$\tilde{p}_{cj} = \frac{1}{\Pi_{xj}(\cdot)\theta_{c2}} [W_{xj}(\cdot)\tilde{\theta}_{c1} - \hat{p}_{cj}\Pi_{xj}(\cdot)\tilde{\theta}_{c2}] \quad (31)$$

where  $\tilde{p}_{cj}(t) \triangleq p_{cj} - \hat{p}_{cj} \in \mathbb{R}^3$  is the prediction error for the  $j$ th feature point  $\forall j = 1, \dots, m$ . Based on (31), the combined form of the prediction errors, denoted by  $\tilde{p}_c \triangleq [\tilde{p}_{c1}^T \ \tilde{p}_{c2}^T \ \dots \ \tilde{p}_{cm}^T]^T \in \mathbb{R}^{3m}$ , can be written as follows:

$$\tilde{p}_c = F[W_x\Pi_x] \begin{bmatrix} \tilde{\theta}_{c1}^T \\ \tilde{\theta}_{c2}^T \end{bmatrix}^T. \quad (32)$$

In (32),  $\Pi_x(t) \in \mathbb{R}^{3m \times 4}$ , and  $W_x(\bar{X}) \in \mathbb{R}^{3m \times 12}$  are defined as follows:

$$\Pi_x \triangleq - [(\hat{p}_{c1}\Pi_{x1})^T (\hat{p}_{c2}\Pi_{x2})^T \dots (\hat{p}_{cm}\Pi_{xm})^T]^T \\ W_x \triangleq [W_{x1}(\bar{X})^T W_{x2}(\bar{X})^T \dots W_{xm}(\bar{X})^T]^T \quad (33)$$

and  $F(t) \in \mathbb{R}^{3m \times 3m}$  is an auxiliary matrix defined as follows:

$$F \triangleq \text{diag}\{\bar{\varphi}_1, \bar{\varphi}_1, \bar{\varphi}_1, \dots, \bar{\varphi}_n, \bar{\varphi}_n, \bar{\varphi}_n\} \quad (34)$$

where  $\bar{\varphi}_j \triangleq (1/\Pi_{xj}\theta_{c2}) \forall j = 1, \dots, m$ . The expression given in (32) can be further simplified as follows:

$$\tilde{p}_c = FW_c\tilde{\theta}_c \quad (35)$$

where  $W_c(t) \in \mathbb{R}^{3m \times 16}$  is defined as follows:

$$W_c \triangleq [W_x \ \Pi_x]. \quad (36)$$

<sup>4</sup>The reader is referred to [24] for the matrix property.

To account for the unknown camera parameters, the following update law is designed [23]:

$$\dot{\hat{\theta}}_c \triangleq \text{Proj} \{ \alpha_c \Gamma_c W_c^T \tilde{p}_c \} \quad (37)$$

where  $\text{Proj}\{\cdot\}$  ensures the positiveness of the term  $\Pi_{x_j}(\cdot)\hat{\theta}_{c2}(t)$  (see [25] for a detailed description) and  $\alpha_c(t) \in \mathbb{R}$  is a positive function defined as follows:

$$\alpha_c \triangleq 1 + \frac{1}{\bar{\varepsilon}_c} \bar{\rho}_c(\cdot) \quad (38)$$

where  $\bar{\rho}_c(\cdot) \triangleq \max_j \{ \rho_{c_j}^2(\cdot) \} \in \mathbb{R}$  is a positive function and  $\bar{\varepsilon}_c \triangleq \min_j \{ \varepsilon_{c_j} \} \in \mathbb{R}$  is a positive constant. In (37),  $\Gamma_c(t) \in \mathbb{R}^{16 \times 16}$  is the covariance matrix, designed as follows:<sup>5</sup>

$$\frac{d}{dt} \{ \Gamma_c^{-1}(t) \} = 2W_c^T W_c, \quad \Gamma_c(t_m^+) = \Gamma_c(t_0) = \eta_0 I_{16} \quad (39)$$

where  $t_m \in \mathbb{R}^+$  is the time instant at which the minimum eigenvalue of  $\Gamma_c(t)$  is less than or equal to  $\eta_1$  (i.e., the covariance matrix is reset each time when its minimum eigenvalue becomes less than  $\eta_1$ ) and  $\eta_0, \eta_1 \in \mathbb{R}$  are positive constants satisfying the inequality  $\eta_0 > \eta_1$ .

*Remark 1:* It is to be noted that due to resetting,  $\Gamma_c(t)$  is guaranteed to be positive definite for all  $t \geq 0$ . At the resetting time  $t_m$ ,  $\Gamma_c(t_m^+) = \Gamma_c(t_0) = \eta_0 I_{16}$ ; therefore,  $\Gamma_c^{-1}(t_0) = \eta_0^{-1} I_{16}$  and between the discontinuities  $(d/dt)\Gamma_c^{-1}(t) \geq 0$  (i.e.,  $\Gamma_c^{-1}(t_2) - \Gamma_c^{-1}(t_1) \geq 0, \forall t_2 \geq t_1 \geq 0$ ). Thus, it can be easily inferred that  $\Gamma_c^{-1} \geq \eta_0^{-1} I_{16}, \forall t \geq 0$ . Because of the resetting,  $\Gamma_c(t)$  is always lower bounded by  $\eta_1 I_{16}, \forall t \geq 0$ ; therefore, the following inequalities are always guaranteed [26]:

$$\eta_0 I_{16} \geq \Gamma_c(t) \geq \eta_1 I_{16}, \quad \eta_1^{-1} I_{16} \geq \Gamma_c^{-1}(t) \geq \eta_0^{-1} I_{16}. \quad (40)$$

### C. Euclidean Structure Estimator

Based on the subsequent stability analysis, the following update law is designed to estimate the Euclidean coordinates of the unknown object features relative to  $\mathcal{W}$

$$\dot{\hat{\theta}}_s \triangleq \text{Proj} \{ \alpha \Gamma \bar{W}_s^T \tilde{p}_s \} \quad (41)$$

where  $\text{Proj}\{\cdot\}$  ensures the positiveness of the term  $\hat{\Pi}(\cdot)\hat{\theta}_{si}(t)$  (see [25] for a detailed description), and  $\alpha(t) \in \mathbb{R}$  is a positive scalar function defined as follows:

$$\alpha \triangleq 1 + \frac{1}{\bar{\varepsilon}_s} \bar{\rho}_s(\cdot) \quad (42)$$

where  $\bar{\rho}_s(\cdot) \in \mathbb{R}$  is a positive function defined as

$$\bar{\rho}_s(\cdot) \triangleq \max_i \{ \rho_{si}^2(\cdot) \} \quad (43)$$

and  $\bar{\varepsilon}_s \in \mathbb{R}$  is a positive constant defined as follows:

$$\bar{\varepsilon}_s \triangleq \min_i \{ \varepsilon_{si} \}. \quad (44)$$

<sup>5</sup>Throughout the paper,  $I_q$  will be used to denote a  $q \times q$  standard identity matrix.

In (41),  $\Gamma(t) \in \mathbb{R}^{4n \times 4n}$  is the least-squares estimation gain matrix, designed as follows:

$$\frac{d}{dt} \{ \Gamma^{-1}(t) \} = 2\bar{W}_s^T \bar{W}_s. \quad (45)$$

The Euclidean structure estimator given in (41) is run simultaneously with the adaptive update law given in (37). The latter updates the camera parameters for  $\bar{W}_s(t)$  which is used in the structure estimator. See Fig. 1 for an illustration of the estimation technique.

*Remark 2:* It should be noted that if  $\Gamma^{-1}(t_0)$  is selected to be positive definite and symmetric then  $\Gamma(t_0)$  is also positive definite and symmetric. Therefore, it follows that both  $\Gamma^{-1}(t)$  and  $\Gamma(t)$  are positive definite and symmetric. The following expression can be obtained from (45):

$$\dot{\Gamma} = -2\Gamma \bar{W}_s^T \bar{W}_s \Gamma. \quad (46)$$

It can be easily seen from (46) that  $\dot{\Gamma}(t)$  is negative semidefinite; therefore,  $\Gamma(t)$  is always constant or decreasing; hence, it follows that  $\Gamma(t)$  is bounded (for more details, the reader is referred to [27]).

*Remark 3:* The projection algorithm utilized in (37), and (41) ensures that  $\tilde{p}_{si}(t)$  is bounded  $\forall i = 1, \dots, n$ . Furthermore, it satisfies the following inequalities (see [17] for a detailed description):

$$-\tilde{\theta}_s \Gamma^{-1} \text{Proj} \{ \tau_s \} \leq -\tilde{\theta}_s \Gamma^{-1} \tau_s \quad (47)$$

$$-\tilde{\theta}_c \Gamma_c^{-1} \text{Proj} \{ \tau_c \} \leq -\tilde{\theta}_c \Gamma_c^{-1} \tau_c \quad (48)$$

where  $\tau_s = \alpha \Gamma \bar{W}_s^T \tilde{p}_s$  and  $\tau_c = \alpha_c \Gamma_c W_c^T \tilde{p}_c$ .

### D. Stability Analysis

*Theorem 1:* The update law defined in (41) ensures that  $\|\tilde{\theta}_s(t)\| \rightarrow 0$  as  $t \rightarrow +\infty$  provided that the following PE condition [26] holds:

$$\gamma_1 I_{4n} \leq \int_{t_0}^{t_0+T} \bar{W}_s^T(\tau) \bar{W}_s(\tau) d\tau \leq \gamma_2 I_{4n} \quad (49)$$

where  $\gamma_1, \gamma_2, T \in \mathbb{R}$  are positive constants.

*Proof:* See [28].

*Remark 4:* The parameter vector  $\hat{\theta}_{si}(t)$  provides a scaled estimate of the Euclidean coordinates of the object features relative to the world frame. Since the last element in the unknown constant parameter vector is equal to 1 as defined in (11), the scale factor can be computed as

$$\lambda_i = \hat{\theta}_{si4} \quad (50)$$

where  $\lambda_i(t) \in \mathbb{R}$  is the scale factor for the  $i$ th feature and  $\hat{\theta}_{si4}(t) \in \mathbb{R}$  is the last entry of  $\hat{\theta}_{si}(t)$ . It should be noted that  $\hat{\theta}_{si4}(t)$  is always nonzero which is guaranteed by the projection algorithm introduced in (37). The estimates of the Euclidean coordinates of the  $i$ th feature can now be recovered as follows:

$$\hat{\omega}_{sih} = \frac{1}{\lambda_i} \hat{\theta}_{sih} \quad (51)$$

where  $\hat{\theta}_{s_{ih}}(t) \forall h = 1, 2, 3$ , is the  $h$ th element of the estimated parameter vector for  $i$ th feature.

*Remark 5:* It should be noted that the camera parameters' update may affect the convergence speed of the Euclidean structure estimator; however, subsequently presented simulations and experiments confirm that as long as the second term in (26) goes to zero, the structure estimation is purely dependent on the prediction error and the structure estimator.

*Remark 6:* The PE condition given in (49) can be fulfilled by guaranteeing that the signals in the regression matrix vary in a sufficient independent manner within a time-window. The reader is referred to [29] for a detailed description.

## V. SIMULATION AND EXPERIMENTAL RESULTS

The reader is referred to [28] for detailed simulation and experimental results.

## VI. CONCLUSION

A novel technique for estimation of 3-D Euclidean coordinates of features on a static object with an uncalibrated camera mounted on a moving mechanical system was presented. The position information of the mechanical system was available and information of a second object was assumed to be known. Two adaptive update laws were presented utilizing formulations of a prediction error and an auxiliary prediction error, respectively, that facilitated the 3-D Euclidean coordinates estimation and compensation for the unknown camera calibration parameters. It was proven that the Euclidean distance estimation error signals were driven to zero.

## REFERENCES

- [1] N. Nath, E. Tatlicioglu, and D. M. Dawson, "Euclidean position estimation of static features using a moving uncalibrated camera," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2009, pp. 4548–4553.
- [2] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. Comput. Vision*, vol. 3, no. 3, pp. 209–238, 1989.
- [3] A. Chiuso, P. Favaro, H. Jin, and S. Soatto, "Structure from motion casually integrated over time," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 523–535, Apr. 2002.
- [4] S. J. Julier and J. K. Uhlmann, "A counter example to the theory of simultaneous localization and map building," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2001, pp. 4238–4243.
- [5] X. Chen and H. Kano, "A new state observer for perspective systems," *IEEE Trans. Autom. Control*, vol. 47, no. 4, pp. 658–663, Apr. 2002.
- [6] X. Chen and H. Kano, "State observer for a class of nonlinear systems and its application to machine vision," *IEEE Trans. Autom. Control*, vol. 49, no. 11, pp. 2085–2091, Nov. 2004.
- [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] L. Ma, Y. Chen, and K. L. Moore, "Range identification for perspective dynamic system with single homogeneous observation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2004, pp. 5207–5211.
- [9] X. Hu and T. Ersson, "Active state estimation of nonlinear systems," *Automatica*, vol. 40, no. 12, pp. 2075–2082, 2004.
- [10] R. Abdursal, H. Inaba, and B. Ghosh, "Nonlinear observers for perspective time-invariant linear systems," *Automatica*, vol. 40, no. 3, pp. 481–490, 2004.
- [11] D. Karagiannis and A. Astolfi, "A new solution to the problem of range identification in perspective vision systems," *IEEE Trans. Autom. Control*, vol. 50, no. 12, pp. 2074–2077, Dec. 2005.
- [12] S. Gupta, D. Aiken, G. Hu, and W. E. Dixon, "Lyapunov-based range and motion identification for a nonaffine perspective dynamic system," in *Proc. Amer. Control Conf.*, 2006, pp. 4471–4476.
- [13] O. Dahl, F. Nyberg, and A. Heyden, "Nonlinear and adaptive observers for perspective dynamic systems," in *Proc. Amer. Control Conf.*, 2007, pp. 1966–971.
- [14] A. De Luca, G. Oriolo, and P. R. Giordano, "On-line estimation of feature depth for image-based visual servoing schemes," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007, pp. 2823–2828.
- [15] D. Braganza, D. M. Dawson, and T. Hughes, "Euclidean position estimation of static features using a moving camera with known velocities," in *Proc. IEEE Int. Conf. Dec. Control*, 2007, pp. 2695–2700.
- [16] N. Nath, D. Braganza, and D. M. Dawson, "Position based structure from motion using a moving calibrated camera," in *Proc. Amer. Control Conf.*, 2008, pp. 1764–1769.
- [17] N. Nath, D. Braganza, D. M. Dawson, and T. Burg, "Range identification for perspective vision systems: A position based approach," Clemson Univ. CRB, Clemson, SC, Tech. Rep. CU/CRB/7/18/08/1, Jul. 2008. [Online]. Available: <http://www.ces.clemson.edu/ece/crb/publicn/tr.htm>
- [18] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [19] R. I. Hartley, *Euclidean Reconstruction From Uncalibrated Views, Applications of Invariance in Computer Vision—Lecture Notes in Computer Science*. Berlin, Germany: Springer, 1994, vol. 825, pp. 235–256.
- [20] Y. F. Li and R. S. Lu, "Uncalibrated Euclidean 3-D reconstruction using an active vision system," *IEEE Trans. Robot. Autom.*, vol. 20, no. 1, pp. 15–25, Feb. 2004.
- [21] O. Faugeras, *Three-Dimensional Computer Vision*. Cambridge, MA: MIT Press, 1993.
- [22] E. Malis and F. Chaumette, "2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," *Int. J. Comput. Vision*, vol. 37, no. 1, pp. 79–97, 2000.
- [23] A. Kapadia, D. Braganza, D. M. Dawson, and M. L. McIntyre, "Adaptive camera calibration with measurable position of fixed features," in *Proc. Amer. Control Conf.*, 2008, pp. 3869–3874.
- [24] S. Nicosia and P. Tomei, "Robot control by using only joint position measurements," *IEEE Trans. Autom. Control*, vol. 35, no. 9, pp. 1058–1061, Sep. 1990.
- [25] E. Tatlicioglu, D. M. Dawson, and B. Xian, "Adaptive visual servo regulation control for camera-in-hand configuration with a fixed-camera extension," *Int. J. Robot. Autom.*, vol. 24, no. 4, pp. 346–355, 2009.
- [26] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [27] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.
- [28] N. Nath, D. M. Dawson, and E. Tatlicioglu, "Euclidean position estimation of static features using a moving uncalibrated camera," Clemson University CRB, Clemson, SC, Tech. Rep. CU/CRB/3/31/09/1, Mar. 2009. [Online]. Available: <http://www.ces.clemson.edu/ece/crb/publicn/tr.htm>
- [29] H. F. Grip, L. Imsland, T. A. Johansen, T. I. Fossen, J. C. Kalkkuhl, and A. Suissa, "Nonlinear vehicle side-slip estimation with friction adaptation," *Automatica*, vol. 44, no. 3, pp. 611–622, 2008.
- [30] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*, 1981, pp. 674–679.
- [31] S. Birchfield, "KLT: An implementation of the Kanade-Lucas-Tomasi feature tracker," Clemson Univ., Clemson, SC, 2007. [Online]. Available: <http://www.ces.clemson.edu/~stb/klf>