We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



20

A Daisy-Chaining Visual Servoing Approach with Applications in Tracking, Localization, and Mapping

S. S. Mehta, W. E. Dixon
University of Florida, Gainesville
USA

G. Hu
Kansas State University, Manhattan
USA

N. Gans
University of Texas, Dallas
USA

1. Introduction

Recent advances in image processing, computational technology and control theory are enabling vision-based control, localization and mapping to become more prevalent in autonomous vehicle applications. Instead of relying solely on a global positioning system (GPS) or inertial measurement units (IMU) for navigation, image-based methods are a promising approach to provide autonomous vehicles with position and orientation (i.e., pose) information. Specifically, rather than obtain an inertial measurement of an autonomous vehicle, vision systems can be used to recast the navigation, localization, control and mapping problems in terms of the image space.

Applications involving localization and mapping using camera as a sensor are often described as Visual Simultaneous Localization and Mapping (VSLAM) (Davison et al., 2007; Eustice et al., 2005; Goncalves et al., 2005; Jensfelt et al., 2006; Jung & Lacroix, 2003; Kim & Sukkarieh, 2003; Se et al., 2002), wherein the camera is the main sensor used to estimate the location of a robot in the world, as well as estimate and maintain estimates of surrounding terrain or features. There are many overlapping ways to categorize VSLAM approaches. Some authors (e.g., (Eustice et al., 2005; Jensfelt et al., 2006; Se et al., 2002)) make a distinction between "local VSLAM" and "global VSLAM". Many VSLAM approaches use probabilistic filters (e.g., extended Kalman filter or particle filter) (Davison et al., 2007; Eustice et al., 2005; Jensfelt et al., 2006; Jung & Lacroix, 2003; Kim & Sukkarieh, 2003), typically estimating a state vector composed of the camera/robot position, orientation and velocity, and the 3D coordinates of visual features in the world frame. An option to a filtered based approach is the use of epipolar geometry (Goncalves et al., 2005; Se et al., 2002). A final possible category are methods that

build a true 3D map (i.e., a map that is easily interpreted by a human being such as walls or topography) (Eustice et al., 2005; Jensfelt et al., 2006; Jung & Lacroix, 2003; Kim & Sukkarieh, 2003; Se et al., 2002), and those that build a more abstract map that is designed to allow the camera/robot to accurately navigate and recognize its location, but not designed for human interpretation.

From the navigation perspective, vision-based pose estimation has motivated results such as (Baker & Nayar, 1999; Burschka & Hager, 2001; Chen et al., 2006; Das et al., 2001; Dixon et al., 2001; Fang et al., 2005; Hagar et al., 1998; Kim et al., 2001; Ma et al., 1999; Song & Huang, 2001) and others, where a camera provides feedback information to enable autonomous navigation of a control agent. See (Chen et al., 2006) for a detailed review of these and other related results. Typically these results are focused on the regulation result, and in all the results the targets are static with respect to the moving camera or the camera is stationary and recording images of the moving control agent.

Vision-based cooperative control methods can involve a moving camera supplying regulation/tracking control input to a moving control agent. A practical example application of this scenario is an airborne camera attached to a remote controlled aircraft that is used to determine a desired video of an unmanned ground vehicle (UGV) moving in a terrain, and then another moving camera (which does not have to follow the same trajectory as the previous camera) is used to relate and control the pose of a moving UGV with respect to the recorded video. The challenge here is to account for the relative velocity between the moving camera and the moving UGV. Also, the reference objects (or features) used to evaluate the pose can leave the camera's field-of-view (FOV) while new reference object enters the FOV. In this scenario, the vision-based system should be intelligent to switch from the leaving reference object to a new reference object to provide the pose information to the controller.

This chapter uses a new daisy-chaining method for visual servo tracking control of a rigid-body object, such as an UGV, while providing localization of the moving camera and moving object in the world frame, and mapping the location of static landmarks in the world frame. Hence, this approach can be used in control and local VSLAM of the UGV, with applications toward path planning, real time trajectory generation, obstacle avoidance, multi-vehicle coordination control and task assignment, etc. By using the daisy-chaining strategy, the coordinates of static features out of the FOV can also be estimated. The estimates of static features can be maintained as a map, or can be used as measurements in existing VSLAM methods.

Section 2 introduces the imaging model, geometric model used in this chapter, as well as introduces the daisy-chaining method as applied to the case of controlling a six-DOF planar object through visual data from a moving camera and fixed reference camera. These results are extended to the case of an UGV with nonholonomic constraints and a moving camera and moving reference camera in Section 3. The efforts of previous sections are then brought to bear on a tracking and mapping application, where the UGV is controlled to track a trajectory that takes the vehicle outside of the initial FOV of the camera. The daisy-chaining approach must be extended to allow for new fixed landmarks to enter the FOV and related to previous landmarks and the UGV.

2. Daisy-Chaining Based Tracking Control

In this section, a visual servo tracking controller is developed for a moving six-DOF agent based on daisy-chained image feedback from a moving camera. The objective is to enable a controlled agent to track a desired trajectory determined by a sequence of prerecorded images from a stationary camera. To achieve this result, several technical issues must be resolved

including: discriminating the relative velocity between the moving camera and the moving agent, compensating for the unknown time-varying distance measurement from the camera to the agent, relating the unknown attitude of the control agent to some measurable signals, and using the unit quaternion to formulate the rotation motion and rotation error system. The relative velocity issue is resolved by utilizing multi-view image geometry to daisy-chain homography relationships between the moving camera frame and the moving agent coordinate frames. By using the depth ratios obtained from the homography decomposition, the unknown depth information is related to an unknown constant that can be compensated for by a Lyapunov-based adaptive update law. Lyapunov-based methods are provided to prove the adaptive asymptotic tracking result.

2.1 Problem Scenario

Over the past decade, a variety of visual servo controllers have been addressed for both camera-to-hand and camera-in-hand configurations (e.g., see (Allen et al., 1993; Hager et al., 1995; Hutchinson et al., 1996; Wiiesoma et al., 1993)). Typical camera-to-hand and camera-in-hand visual servo controllers have required that either the camera or the target remain stationary so that an absolute velocity can be determined and used in the control development. For the problem of a moving camera tracking a moving target (i.e. control of relative pose/velocity), integral control or predictive Kalman filters have been used to overcome the unknown target velocity (Bensalah & Chaumette, 1995; Papanikolopoulos et al., 1993). In contrast to these methods, the development in this section and our previous preliminary work in (Hu, Mehta, Gans & Dixon, 2007; Mehta, Dixon, MacArthur & Crane, 2006; Mehta, Hu, Gans & Dixon, 2006) is motivated by the problem when the camera and the target are moving. A practical example application of this scenario is an airborne camera attached to a remote controlled aircraft that is used to determine pose measurements of an UGV and then relay the information to the UGV for closed-loop control.

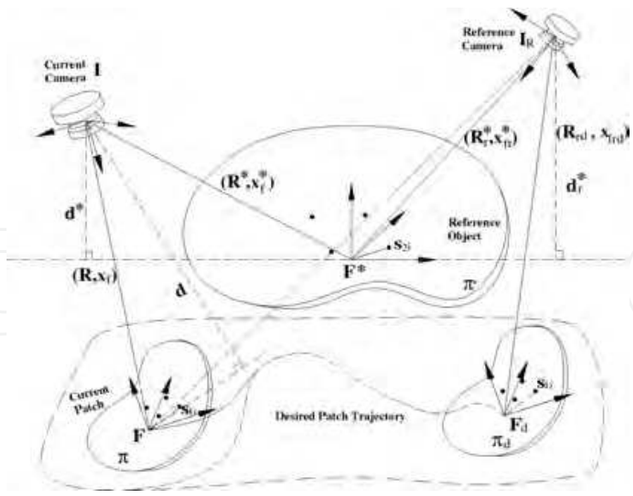


Fig. 1. Geometric model for a moving camera, moving target and stationary reference camera.

The scenario examined in this section is depicted in Fig. 1, where various coordinate frames are defined as a means to develop the subsequent Euclidean reconstruction and control methods. In Fig. 1, a stationary coordinate frame $\mathcal{I}_R$ is attached to a camera and a time-varying

coordinate frame $\mathcal{F}_d$ is attached to some mobile agent (e.g., an aircraft, a ground vehicle, a marine vessel). The agent is identified in an image through a collection of feature points that are assumed (without loss of generality[1]) to be coplanar and non-collinear (i.e., a planar patch of feature points). The camera attached to $\mathcal{I}_R$ a priori records a series of snapshots (i.e., a video) of the motion of the coordinate frame $\mathcal{F}_d$ until $\mathcal{F}_d$ comes to rest. A stationary coordinate frame $\mathcal{F}^*$ is attached to another planar patch of feature points that are assumed to be visible in every frame of the video recorded by the camera. For example, the camera attached to $\mathcal{I}_R$ is on-board a stationary satellite that takes a series of snapshots of the relative motion of $\mathcal{F}_d$ with respect to $\mathcal{F}^*$. Therefore, the desired motion of $\mathcal{F}_d$ can be encoded as a series of relative translations and rotations with respect to the stationary frame $\mathcal{F}^*$ a priori. Spline functions or filter algorithms can then be used to generate a smooth desired feature point trajectory as described in (Chen et al., 2005).

Fig. 1 also depicts a time-varying coordinate frame $\mathcal{I}$ that is attached to another camera (e.g., a camera attached to a remote controlled aircraft), and a time-varying coordinate frame $\mathcal{F}$ that is attached to the current pose of the planar patch. The camera attached to $\mathcal{I}$ captures snapshots of the planar patches associated with $\mathcal{F}$ and $\mathcal{F}^*$, respectively. The a priori motion of $\mathcal{F}_d$ represents the desired trajectory of the coordinate system $\mathcal{F}$, where $\mathcal{F}$ and $\mathcal{F}_d$ are attached to identical objects, but at different points in time. The camera attached to $\mathcal{I}_R$ can be a different camera (with different calibration parameters) as the camera attached to $\mathcal{I}$. Based on these coordinate frame definitions, the problem considered in this section is to develop a kinematic controller for the object attached to $\mathcal{F}$ so that the time-varying rotation and translation of $\mathcal{F}$ converges to the desired time-varying rotation and translation of $\mathcal{F}_d$, where the motion of $\mathcal{F}$ is determined from the time-varying overhead camera attached to $\mathcal{I}$.

2.2 Geometric Relationships

Relationships between the various coordinate frames are summarized in Table I. In Table I, $R(t)$, $R^*(t)$, $R_r(t)$, $R'(t)$, $R_{rd}(t)$, $R_r^* \in SO(3)$ denote rotation matrices, and $x_{fr}(t)$, $x'_{fr}(t)$, $x_{frd}(t)$, $x_{fr}^* \in \mathbb{R}^3$ denote translation vectors. From Fig. 1, the translation $x'_{fr}(t)$ and the rotation $R'(t)$ can be expressed as

$$\begin{aligned} x'_{fr} &= x_{fr}^* + R_r^* R^{*T}(x_f - x_f^*) \\ R' &= R_r^* R^{*T} R. \end{aligned} \tag{1}$$

As illustrated in Fig. 1, π, π_d and π^* denote the planes of feature points associated with $\mathcal{F}$, $\mathcal{F}_d$, and $\mathcal{F}^*$, respectively. The constant Euclidean coordinates of the i-th feature point in $\mathcal{F}$ (and also $\mathcal{F}_d$) are denoted by $s_{1i} \in \mathbb{R}^3$ $\forall i = 1, 2, \cdots, n$ $(n \geq 4)$, and $s_{2i} \in \mathbb{R}^3$ $\forall i = 1, 2, \cdots, n$ denotes the constant Euclidean coordinates of the i-th feature point in $\mathcal{F}^*$. From the geometry between the coordinate frames depicted in Fig. 1, the following relationships

[1] Image processing techniques can often be used to select coplanar and non-collinear feature points within an image. However, if four coplanar target points are not available then the subsequent development can also exploit the virtual parallax method (Boufama & Mohr, 1995; Malis & Chaumette, 2000) where the non-coplanar points are projected onto a virtual plane.

Motion	Frames
$R(t), x_f(t)$	$\mathcal{F}$ to $\mathcal{I}$ in $\mathcal{I}$
$R^*(t), x_f^*(t)$	$\mathcal{F}^*$ to $\mathcal{I}$ in $\mathcal{I}$
$R_r(t), x_{fr}(t)$	$\mathcal{I}$ to $\mathcal{I}_R$
$R^{'}(t), x_{fr}^{'}(t)$	$\mathcal{F}$ to $\mathcal{I}_R$ in $\mathcal{I}_R$
R_r^*, x_{fr}^*	$\mathcal{F}^*$ to $\mathcal{I}_R$ in $\mathcal{I}_R$
$R_{rd}(t), x_{frd}(t)$	$\mathcal{F}_d$ to $\mathcal{I}_R$ in $\mathcal{I}_R$

Table 1. Coordinate frames relationships.

can be developed:

$$\bar{m}_i = x_f + Rs_{1i} \qquad \bar{m}_{rdi} = x_{frd} + R_{rd}s_{1i} \tag{2}$$

$$\bar{m}_{ri}^* = x_{fr}^* + R_r^* s_{2i} \qquad \bar{m}_i^{'} = x_{fr}^{\prime} + R^{'} s_{1i} \tag{3}$$

$$\bar{m}_i^* = x_f^* + R^* s_{2i}. \tag{4}$$

In (2)-(4), $\bar{m}_i(t)$, $\bar{m}_i^*(t) \in \mathbb{R}^3$ denote the Euclidean coordinates of the feature points on π and π^*, respectively, expressed in $\mathcal{I}$ as

$$\bar{m}_i(t) \triangleq \left[\begin{array}{ccc} x_i(t) & y_i(t) & z_i(t) \end{array} \right]^T \tag{5}$$

$$\bar{m}_i^*(t) \triangleq \left[\begin{array}{ccc} x_i^*(t) & y_i^*(t) & z_i^*(t) \end{array} \right]^T, \tag{6}$$

$\bar{m}_i^{'}(t)$, $\bar{m}_{rdi}(t) \in \mathbb{R}^3$ denote the actual and desired time-varying Euclidean coordinates, respectively, of the feature points on π expressed in $\mathcal{I}_R$ as

$$\bar{m}_i^{'}(t) \triangleq \left[\begin{array}{ccc} x_i^{'}(t) & y_i^{'}(t) & z_i^{'}(t) \end{array} \right]^T \tag{7}$$

$$\bar{m}_{rdi}(t) \triangleq \left[\begin{array}{ccc} x_{rdi}(t) & y_{rdi}(t) & z_{rdi}(t) \end{array} \right]^T, \tag{8}$$

and $\bar{m}_{ri}^* \in \mathbb{R}^3$ denotes the constant Euclidean coordinates of the feature points on the plane π^* expressed in $\mathcal{I}_R$ as

$$\bar{m}_{ri}^* \triangleq \left[\begin{array}{ccc} x_{ri}^* & y_{ri}^* & z_{ri}^* \end{array} \right]^T. \tag{9}$$

After some algebraic manipulation, the expressions in (2)-(4) can be rewritten as

$$\bar{m}_i^* = \bar{x}_n + R_n \bar{m}_i \tag{10}$$

$$\bar{m}_i = \bar{x}_f + \bar{R}\bar{m}_i^* \qquad \bar{m}_{rdi} = \bar{x}_{frd} + \bar{R}_{rd}\bar{m}_{ri}^* \tag{11}$$

$$\bar{m}_{ri}^* = x_{fr} + R_r \bar{m}_i^* \qquad \bar{m}_i^{'} = x_{fr} + R_r \bar{m}_i, \tag{12}$$

where $R_n(t)$, $\bar{R}(t)$, $\bar{R}_{rd}(t)$, $R_r(t) \in SO(3)$ and $\bar{x}_n(t)$, $\bar{x}_f(t)$, $\bar{x}_{frd}(t)$, $x_{fr}(t) \in \mathbb{R}^3$ are new rotation and translation variables, respectively, defined as[2]

$$\begin{aligned} R_n &= R^* R^T \qquad \bar{R} = R R^{*T} \\ \bar{R}_{rd} &= R_{rd} R_r^{*T} \qquad R_r = R_r^* R^{*T} \end{aligned} \tag{13}$$

$$\bar{x}_n = x_f^* - R_n \left(x_f - R \left(s_{2i} - s_{1i} \right) \right) \tag{14}$$

$$\bar{x}_f = x_f - \bar{R} \left(x_f^* + R^* \left(s_{2i} - s_{1i} \right) \right) \tag{15}$$

$$\bar{x}_{frd} = x_{frd} - \bar{R}_{rd} \left(x_{fr}^* + R_r^* \left(s_{2i} - s_{1i} \right) \right) \tag{16}$$

$$x_{fr} = x_{fr}^* - R_r x_f^* = x_{fr}' - R_r x_f. \tag{17}$$

To facilitate the development of a relationship between the actual Euclidean translation of $\mathcal{F}$ to the Euclidean translation that is reconstructed from the image information, projective relationships are developed from Fig. 1 as

$$d(t) = n^T \bar{m}_i \qquad d^*(t) = n^{*T} \bar{m}_i^* \qquad d_r^* = n_r^{*T} \bar{m}_{ri}^*, \tag{18}$$

where $d(t) \in \mathbb{R}$ represents the distance from the origin of $\mathcal{I}$ to π along the unit normal (expressed in $\mathcal{I}$) to π denoted as $n(t) \in \mathbb{R}^3$, $d^*(t) \in \mathbb{R}$ represents the distance from the origin of $\mathcal{I}$ to π^* along the unit normal (expressed in $\mathcal{I}$) to π^* denoted as $n^*(t) \in \mathbb{R}^3$, and $d_r^* \in \mathbb{R}$ represents the distance from the origin of $\mathcal{I}_R$ to π^* along the unit normal (expressed in $\mathcal{I}_R$) to π^* denoted as $n_r^* \in \mathbb{R}^3$ where $n^*(t) = R_r^T(t) n_r^*$. In (18), $d(t)$, $d^*(t)$, $d_r^* > \varepsilon$ for some positive constant $\varepsilon \in \mathbb{R}$. Based on (18), the relationships in (10)-(12) can be expressed as

$$\bar{m}_i^* = \left(R_n + \frac{\bar{x}_n}{d} n^T \right) \bar{m}_i \qquad \bar{m}_i = \left(\bar{R} + \frac{\bar{x}_f}{d^*} n^{*T} \right) \bar{m}_i^* \tag{19}$$

$$\bar{m}_{rdi} = \left(\bar{R}_{rd} + \frac{\bar{x}_{frd}}{d_r^*} n_r^{*T} \right) \bar{m}_{ri}^* \qquad \bar{m}_{ri}^* = \left(R_r + \frac{x_{fr} n^{*T}}{d^*} \right) \bar{m}_i^* \tag{20}$$

$$\bar{m}_i' = \left(R_r + \frac{x_{fr} n^T}{d} \right) \bar{m}_i. \tag{21}$$

As in Chen et al. (2005), the subsequent development requires that the constant rotation matrix R_r^* be known. The constant rotation matrix R_r^* can be obtained a priori using various methods (e.g., a second camera, additional on-board sensors, off-line calibration, Euclidean measurements). The subsequent development also assumes that the difference between the Euclidean distances $(s_{2i} - s_{1i})$ is a constant $\forall i = 1, ..., n$. While there are many practical applications that satisfy this assumption (e.g., a simple scenario is that the objects attached to $\mathcal{F}$ and $\mathcal{F}^*$ are the identical objects), the assumption is generally restrictive and is the focus of future research. As described in our preliminary work in Hu, Gans, Mehta & Dixon (2007),

[2] Note that $R_n(t)$, $\bar{R}(t)$ and $\bar{R}_{rd}(t)$ in (13) are the rotation matrices between $\mathcal{F}$ and $\mathcal{F}^*$, $\mathcal{F}^*$ and $\mathcal{F}$, and $\mathcal{F}^*$ and $\mathcal{F}_d$, respectively, but $\bar{x}_n(t)$, $\bar{x}_f(t)$ and $\bar{x}_{frd}(t)$ in (14)-(16) are not the translation vectors between the corresponding coordinate frames. Only the rotation matrices will be used in the controller development.

each of these assumptions can be avoided by using the geometric reconstruction approach in Dupree et al. (2007); Gans et al. (2008); Mackunis et al. (2007) under an alternative assumption that the Euclidean distance between two feature points is precisely known.

2.3 Euclidean Reconstruction

The relationships given by (19)-(21) provide a means to quantify a translation and rotation error between the different coordinate systems. Since the pose of $\mathcal{F}$, $\mathcal{F}_d$, and $\mathcal{F}^*$ cannot be directly measured, a Euclidean reconstruction is developed to obtain the pose error by comparing multiple images acquired from the hovering monocular vision system. To facilitate the subsequent development, the normalized Euclidean coordinates of the feature points in π and π^* can be expressed in terms of $\mathcal{I}$ as $m_i(t), m_i^*(t) \in \mathbb{R}^3$, respectively, as

$$m_i \triangleq \frac{\bar{m}_i}{z_i} \qquad m_i^* \triangleq \frac{\bar{m}_i^*}{z_i^*}. \tag{22}$$

Similarly, the normalized Euclidean coordinates of the feature points in π, π_d and π^* can be expressed in terms of $\mathcal{I}_R$ as $m_i'(t), m_{rdi}(t), m_{ri}^* \in \mathbb{R}^3$, respectively, as

$$m_i'(t) \triangleq \frac{\bar{m}_i'(t)}{z_i'(t)} \quad m_{rdi}(t) \triangleq \frac{\bar{m}_{rdi}(t)}{z_{rdi}(t)} \quad m_{ri}^* \triangleq \frac{\bar{m}_{ri}^*}{z_{ri}^*}. \tag{23}$$

From the expressions given in (19) and (22), the rotation and translation between the coordinate systems $\mathcal{F}$ and $\mathcal{F}^*$, between $\mathcal{F}^*$ and $\mathcal{F}_d$, and between $\mathcal{I}$ and $\mathcal{I}_R$ can now be related in terms of the normalized Euclidean coordinates as

$$m_i = \alpha_i \left(\bar{R} + x_h n^{*T} \right) m_i^* \qquad m_i^* = \frac{1}{\alpha_i} \left(R_n + x_{nh} n^T \right) m_i \tag{24}$$

$$m_{rdi} = \alpha_{rdi} \left(\bar{R}_{rd} + x_{hrd} n_r^{*T} \right) m_{ri}^* \qquad m_{ri}^* = \alpha_{ri} \left(R_r + x_{hr} n^{*T} \right) m_i^* \tag{25}$$

where $\alpha_i(t), \alpha_{rdi}(t), \alpha_{ri}(t) \in \mathbb{R}$ denote depth ratios defined as

$$\alpha_i = \frac{z_i^*}{z_i} \qquad \alpha_{rdi} = \frac{z_{ri}^*}{z_{rdi}} \qquad \alpha_{ri} = \frac{z_i^*}{z_{ri}^*},$$

and $x_h(t), x_{nh}(t), x_{hrd}(t), x_{hr}(t) \in \mathbb{R}^3$ denote scaled translation vectors that are defined as

$$x_h = \frac{\bar{x}_f}{d^*} \quad x_{nh} = \frac{\bar{x}_n}{d} \quad x_{hrd} = \frac{\bar{x}_{frd}}{d_r^*} \quad x_{hr} = \frac{x_{fr}}{d^*}. \tag{26}$$

Since the normalized Euclidean coordinates in (24)-(25) can not be directly measured, the following relationships (i.e., the pin-hole camera model) are used to determine the normalized Euclidean coordinates from pixel information:

$$p_i = A_1 m_i \qquad p_i^* = A_1 m_i^* \qquad p_{rdi} = A_2 m_{rdi} \qquad p_{ri}^* = A_2 m_{ri}^*, \tag{27}$$

where $A_1, A_2 \in \mathbb{R}^{3\times3}$ are known, constant, and invertible intrinsic camera calibration matrices of the current camera and the reference camera, respectively. In (27), $p_i(t), p_i^*(t) \in \mathbb{R}^3$ represent the image-space coordinates of the Euclidean feature points on π and π^* expressed in terms of $\mathcal{I}$ as

$$p_i \triangleq \begin{bmatrix} u_i & v_i & 1 \end{bmatrix}^T \qquad p_i^* \triangleq \begin{bmatrix} u_i^* & v_i^* & 1 \end{bmatrix}^T, \tag{28}$$

respectively, where $u_i(t)$, $v_i(t)$, $u_i^*(t)$, $v_i^*(t) \in \mathbb{R}$. Similarly, $p_{rdi}(t)$, $p_{ri}^* \in \mathbb{R}^3$ represent the image-space coordinates of the Euclidean features on π_d and π^* expressed in terms of $\mathcal{I}_R$ as

$$p_{rdi} \triangleq \begin{bmatrix} u_{rdi} & v_{rdi} & 1 \end{bmatrix}^T \qquad p_{ri}^* \triangleq \begin{bmatrix} u_{ri}^* & v_{ri}^* & 1 \end{bmatrix}^T \tag{29}$$

respectively, where $u_{rdi}(t)$, $v_{rdi}(t)$, u_{ri}^*, $v_{ri}^* \in \mathbb{R}$. By using (24)-(25) and (29), the following relationships can be developed:

$$p_i = \alpha_i \underbrace{\left(A_1 \left(\bar{R} + x_h n^{*T}\right) A_1^{-1}\right)}_{G} p_i^* \qquad p_i^* = \frac{1}{\alpha_i} \underbrace{\left(A_1 \left(R_n + x_{nh} n^T\right) A_1^{-1}\right)}_{G_n} p_i \tag{30}$$

$$p_{rdi} = \alpha_{rdi} \underbrace{\left(A_2 \left(\bar{R}_{rd} + x_{hrd} n_r^{*T}\right) A_2^{-1}\right)}_{G_{rd}} p_{ri}^* \qquad p_{ri}^* = \alpha_{ri} \underbrace{\left(A_2 \left(R_r + x_{hr} n^{*T}\right) A_1^{-1}\right)}_{G_r} p_i^*, \tag{31}$$

where $G(t)$, $G_n(t)$, $G_{rd}(t)$, $G_r(t) \in \mathbb{R}^{3\times 3}$ denote projective homographies. Sets of linear equations can be developed from (30)-(31) to determine the projective homographies up to a scalar multiple. Various techniques can be used (e.g., see Faugeras & Lustman (1988); Zhang & Hanson (1995)) to decompose the Euclidean homographies, to obtain $\alpha_i(t)$, $\alpha_{rdi}(t)$, $\alpha_{ri}(t)$, $x_h(t)$, $x_{nh}(t)$, $x_{hrd}(t)$, $x_{hr}(t)$, $\bar{R}(t)$, $R_n(t)$, $\bar{R}_{rd}(t)$, $R_r(t)$, $n^*(t)$, n_r^*, $n(t)$. Given that the constant rotation matrix R_r^* is assumed to be known, the expressions for $\bar{R}_{rd}(t)$ and $R_r(t)$ in (13) can be used to determine $R_{rd}(t)$ and $R^*(t)$. Once $R^*(t)$ is determined, the expression for $\bar{R}(t)$ in (13) can be used to determine $R(t)$. Also, once R_r^*, $R^{*T}(t)$, and $R(t)$ have been determined, (1) can be used to determine $R'(t)$. Since $R_r(t)$, $x_{hr}(t)$, $\alpha_i(t)$, $n^*(t)$, n_r^*, $n(t)$, $m_i^*(t)$, and $m_i(t)$ can be determined, the following relationship can be used to determine $m_i'(t)$:

$$m_i' = \frac{z_i}{z_i'} \left(R_r + x_{hr} \alpha_i \frac{n^{*T} m_i^*}{n^T m_i} n^T\right) m_i, \tag{32}$$

where the inverse of the ratio $\dfrac{z_i(t)}{z_i'(t)}$ can be determined as

$$\frac{z_i'}{z_i} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \left(R_r + x_{hr} \alpha_i \frac{n^{*T} m_i^*}{n^T m_i} n^T\right) m_i. \tag{33}$$

2.4 Control Objective

The control objective is for a controlled agent (e.g., an UGV or an UAV) to track a desired trajectory that is determined by a sequence of images. This objective is based on the assumption that the control agent is physically able to follow the desired image trajectory, that the linear and angular velocities of the camera are control inputs that can be independently controlled (i.e., unconstrained motion), and that the reference and desired cameras are calibrated (i.e., A_1 and A_2 are known). The control objective can be stated as the desire for the Euclidean feature points on π to track the corresponding feature points on π_d, which can be mathematically stated as the desire for $\bar{m}_i'(t) \to \bar{m}_{rdi}(t)$. Equivalently, the control objective can also be stated in terms of the rotation and translation of the agent as the desire for $x_{fr}'(t) \to x_{frd}(t)$ and $R'(t) \to R_{rd}(t)$.

As stated previously, $R'(t)$ and $R_{rd}(t)$ can be computed by decomposing the projective homographies in (30)-(31) and using (1). Once these rotation matrices have been determined, a variety of parameterizations can be used to describe the rotation. The unit quaternion parameterization is used to describe the rotation in the subsequent problem formulation, control development, and stability analysis since the unit quaternion provides a globally nonsingular parameterization of the corresponding rotation matrices.
The unit quaternion is a four dimensional vector, which can be defined as

$$q \triangleq \begin{bmatrix} q_0 & q_v^T \end{bmatrix}^T \qquad q_v \triangleq \begin{bmatrix} q_{v1} & q_{v2} & q_{v3} \end{bmatrix}^T, \tag{34}$$

where $q_0(t)$, $q_{vi}(t) \in \mathbb{R}\ \forall i = 1,2,3$ satisfy the following nonlinear constraint

$$q^T q = 1. \tag{35}$$

Given the rotation matrices $R'(t)$ and $R_{rd}(t)$, the corresponding unit quaternions $q(t)$ and $q_d(t)$ can be calculated by using the numerically robust method presented in Hu et al. (2006) and Shuster (1993) based on the corresponding relationships

$$R' = \left(q_0^2 - q_v^T q_v\right) I_3 + 2q_v q_v^T + 2q_0 q_v^\times \tag{36}$$

$$R_{rd} = \left(q_{0d}^2 - q_{vd}^T q_{vd}\right) I_3 + 2q_{vd} q_{vd}^T + 2q_{0d} q_{vd}^\times \tag{37}$$

where I_3 is the 3×3 identity matrix, and the notation $q_v^\times(t)$ denotes a skew-symmetric form of the vector $q_v(t)$ as

$$q_v^\times = \begin{bmatrix} 0 & -q_{v3} & q_{v2} \\ q_{v3} & 0 & -q_{v1} \\ -q_{v2} & q_{v1} & 0 \end{bmatrix}, \forall q_v = \begin{bmatrix} q_{v1} \\ q_{v2} \\ q_{v3} \end{bmatrix}. \tag{38}$$

To quantify the rotation error between the feature points on π and π_d, the multiplicative error between rotation matrices $R'(t)$ and $R_{rd}(t)$ is defined as

$$\tilde{R} = R'^T R_{rd} = \left(\tilde{q}_0^2 - \tilde{q}_v^T \tilde{q}_v\right) I_3 + 2\tilde{q}_v \tilde{q}_v^T - 2\tilde{q}_0 \tilde{q}_v^\times, \tag{39}$$

where the error quaternion $\tilde{q}(t) = (\tilde{q}_0(t), \tilde{q}_v^T(t))^T$ is defined as

$$\tilde{q} = \begin{bmatrix} \tilde{q}_0 \\ \tilde{q}_v \end{bmatrix} = \begin{bmatrix} q_0 q_{0d} + q_v^T q_{vd} \\ q_{0d} q_v - q_0 q_{vd} + q_v^\times q_{vd} \end{bmatrix}. \tag{40}$$

Since $\tilde{q}(t)$ is a unit quaternion, (39) can be used to quantify the rotation tracking objective as

$$\|\tilde{q}_v(t)\| \to 0 \Longrightarrow \tilde{R}(t) \to I_3 \quad \text{as} \quad t \to \infty. \tag{41}$$

The translation error, denoted by $e(t) \in \mathbb{R}^3$, is defined as

$$e = m_e - m_{ed} \tag{42}$$

where $m_e(t)$, $m_{ed}(t) \in \mathbb{R}^3$ are defined as[3]

$$m_e = \begin{bmatrix} \frac{x_1'}{z_1'} & \frac{y_1'}{z_1'} & \ln(\frac{z_1'}{z_{r1}^*}) \end{bmatrix}^T \qquad m_{ed} = \begin{bmatrix} \frac{x_{rd1}}{z_{rd1}} & \frac{y_{rd1}}{z_{rd1}} & \ln(\frac{z_{rd1}}{z_{r1}^*}) \end{bmatrix}^T. \tag{43}$$

In (43), $\frac{z_1'(t)}{z_{r1}^*(t)}$ and $\frac{z_{rd1}(t)}{z_{r1}^*(t)}$ can be expressed in terms of known signals as

$$\frac{z_1'}{z_{r1}^*} = \frac{z_1'}{z_1}\frac{z_1}{z_1^*}\frac{z_1^*}{z_{r1}^*} = \frac{z_1'}{z_1}\frac{1}{\alpha_1}\alpha_{r1} \qquad \frac{z_{rd1}}{z_{r1}^*} = \frac{1}{\alpha_{rd1}}.$$

Based on (41) and (42), the subsequent control development targets the following objectives:

$$\|\tilde{q}_v(t)\| \to 0 \quad \text{and} \quad \|e(t)\| \to 0 \quad \text{as} \quad t \to \infty. \tag{44}$$

2.5 Control Development

2.5.1 Open-Loop Error System

Based on (39) and (40), the open-loop rotation error system can be developed as Dixon et al. (2003)

$$\dot{\tilde{q}} = \frac{1}{2}\begin{bmatrix} -\tilde{q}_v^T \\ \tilde{q}_0 I_3 + \tilde{q}_v^\times \end{bmatrix}\left(\omega_c - \tilde{R}\omega_{cd}\right), \tag{45}$$

where $\omega_{cd}(t)$ denotes the angular velocity of π_d expressed in $\mathcal{F}_d$ that can be calculated as Dixon et al. (2003)

$$\omega_{cd} = 2(q_{0d}\dot{q}_{vd} - q_{vd}\dot{q}_{0d}) - 2q_{vd}^\times \dot{q}_{vd}, \tag{46}$$

where $\left(q_{0d}(t),\, q_{vd}^T(t)\right)^T$, $\left(\dot{q}_{0d}(t),\, \dot{q}_{vd}^T(t)\right)^T$ are assumed to be bounded; hence, $\omega_{cd}(t)$ is also bounded. The open-loop translation error system can be derived as

$$z_{r1}^*\dot{e} = \frac{z_{r1}^*}{z_1'}L_v'R'\left(v_c + \omega_c^\times s_1\right) - z_{r1}^*\dot{m}_{ed}, \tag{47}$$

where $v_c(t), \omega_c(t) \in \mathbb{R}^3$ denote the linear and angular velocity of π expressed in $\mathcal{F}$, respectively, and the auxiliary measurable term $L_v'(t) \in \mathbb{R}^{3\times 3}$ is defined as

$$L_v' = \begin{bmatrix} 1 & 0 & -\frac{x_1'}{z_1'} \\ 0 & 1 & -\frac{y_1}{z_1'} \\ 0 & 0 & 1 \end{bmatrix}.$$

[3] Any point O_i can be utilized in the subsequent development; however, to reduce the notational complexity, we have elected to select the image point O_1, and hence, the subscript 1 is utilized in lieu of i in the subsequent development.

2.5.2 Closed-Loop Error System

Based on the open-loop rotation error system in (45) and the subsequent Lyapunov-based stability analysis, the angular velocity controller is designed as

$$\omega_c = -K_\omega \tilde{q}_v + \tilde{R}\omega_{cd}, \tag{48}$$

where $K_\omega \in \mathbb{R}^{3\times 3}$ denotes a diagonal matrix of positive constant control gains. From (45) and (48), the rotation closed-loop error system can be determined as

$$\dot{\tilde{q}}_0 = \frac{1}{2}\tilde{q}_v^T K_\omega \tilde{q}_v \qquad \dot{\tilde{q}}_v = -\frac{1}{2}\left(\tilde{q}_0 I_3 + \tilde{q}_v^\times\right) K_\omega \tilde{q}_v = -\frac{1}{2}K_\omega \tilde{q}_0 \tilde{q}_v. \tag{49}$$

Based on (47), the translation control input $v_c(t)$ is designed as

$$v_c = -\frac{z_1^{'}}{z_{r1}^*} R^{'T} L_v^{'-1}\left(K_v e - \hat{z}_{r1}^* \dot{m}_{ed}\right) - \omega_c^\times s_1, \tag{50}$$

where $K_v \in \mathbb{R}^{3\times 3}$ denotes a diagonal matrix of positive constant control gains. In (50), the parameter estimate $\hat{z}_{r1}^*(t) \in \mathbb{R}$ for the unknown constant z_{r1}^* is designed as

$$\dot{\hat{z}}_{r1}^* = -\gamma e^T \dot{m}_{ed}, \tag{51}$$

where $\gamma \in \mathbb{R}$ denotes a positive constant adaptation gain. By using (47) and (50), the translation closed-loop error system is

$$z_{r1}^* \dot{e} = -K_v e - \tilde{z}_{r1}^* \dot{m}_{ed}, \tag{52}$$

where $\tilde{z}_{r1}^*(t) \in \mathbb{R}$ denotes the parameter estimation error

$$\tilde{z}_{r1}^* = z_{r1}^* - \hat{z}_{r1}^*. \tag{53}$$

2.5.3 Stability Analysis

Theorem 1. *The controller given in (48) and (50), along with the adaptive update law in (51) ensures asymptotic tracking in the sense that*

$$\|\tilde{q}_v(t)\| \to 0, \ \|e(t)\| \to 0, \quad as \quad t \to \infty. \tag{54}$$

Proof. Let $V(t) \in \mathbb{R}$ denote the following differentiable non-negative function (i.e., a Lyapunov candidate):

$$V = \tilde{q}_v^T \tilde{q}_v + (1-\tilde{q}_0)^2 + \frac{z_{r1}^*}{2} e^T e + \frac{1}{2\gamma}\tilde{z}_{r1}^{*2}. \tag{55}$$

The time-derivative of $V(t)$ can be determined as

$$\begin{aligned}
\dot{V} &= -\tilde{q}_v^T K_\omega \tilde{q}_0 \tilde{q}_v - (1-\tilde{q}_0)\tilde{q}_v^T K_\omega \tilde{q}_v - e^T K_v e \\
&\quad + e^T\left(-K_v e - \tilde{z}_{r1}^* \dot{m}_{ed}\right) + \tilde{z}_{r1}^* e^T \dot{m}_{ed} \\
&= -\tilde{q}_v^T\left(\tilde{q}_0 I_3 + (1-\tilde{q}_0) I_3\right) K_\omega \tilde{q}_v - e^T K_v e \\
&= -\tilde{q}_v^T K_\omega \tilde{q}_v - e^T K_v e,
\end{aligned} \tag{56}$$

where (49) and (51)-(53) were utilized. Based on (55) and (56), $e(t)$, $\tilde{q}_v(t)$, $\tilde{q}_0(t)$, $\tilde{z}^*_{r1}(t) \in \mathcal{L}_\infty$ and $e(t)$, $\tilde{q}_v(t) \in \mathcal{L}_2$. Since $\tilde{z}^*_{r1}(t) \in \mathcal{L}_\infty$, it is clear from (53) that $\hat{z}^*_{r1}(t) \in \mathcal{L}_\infty$. Based on the fact that $e(t) \in \mathcal{L}_\infty$, (42) and (43) can be used to prove that $m'_1(t) \in \mathcal{L}_\infty$, and then $L'_v(t)$, $L'^{-1}_v(t) \in \mathcal{L}_\infty$. Based on the fact that $\tilde{q}_v(t) \in \mathcal{L}_\infty$ and $\omega_{cd}(t)$ is a bounded function, (48) can be used to conclude that $\omega_c(t) \in \mathcal{L}_\infty$. Since $\hat{z}^*_{r1}(t)$, $e(t)$, $m'_1(t)$, $L'_v(t)$, $L'^{-1}_v(t) \in \mathcal{L}_\infty$ and $\dot{m}_{ed}(t)$ is bounded, (50) can be utilized to prove that $v_c(t) \in \mathcal{L}_\infty$. From the previous results, (45)-(47) can be used to prove that $\dot{e}(t)$, $\dot{\tilde{q}}_v(t) \in \mathcal{L}_\infty$. Since $e(t)$, $\tilde{q}_v(t) \in \mathcal{L}_\infty \cap \mathcal{L}_2$, and $\dot{e}(t)$, $\dot{\tilde{q}}_v(t) \in \mathcal{L}_\infty$, we can utilize a corollary to Barbalat's Lemma Slotine & Li (1991) to conclude the result given in (54). □

3. Cooperative Tracking Control of A Nonholonomic Unmanned Ground Vehicle

In the previous section, a visual servo tracking controller is developed for a moving six-DOF agent based on daisy-chained image feedback from a moving camera where a stationary reference camera was used to encode a desired video. The development in this section and our preliminary work in Mehta, Hu, Gans & Dixon (2006) extends the previous section by allowing the reference camera to also move. The example of a reference camera in the previous section was a "stationary" satellite that was used to encode the desired trajectory. In this section, the desired trajectory could be encoded by a moving camera (e.g., attached to a moving satellite, a dirigible, or another UAV). In addition, instead of considering the general six-DOF control agent, the control agent in this section is a nonholonomic constrained UGV. The control objective is for the UGV to track a desired trajectory determined by a sequence of prerecorded images from some moving overhead camera. An additional technical issue resolved in this section is the challenge of comparing the relative velocity between a moving camera and a moving UGV to the relative desired trajectory recorded by a moving camera.

3.1 Problem Scenario

Previous results, such as Chen et al. (2006); Dixon et al. (2001); Fang et al. (2005), are typically focused on the results in which the targets are static with respect to the moving camera or the camera is stationary and recording images of the moving UGV. In contrast to these methods, the development in this section is motivated by the problem when a moving camera is recording images of a moving UGV so a second UGV can track a desired image trajectory.

The scenario examined in this section is depicted in Fig. 2, where various coordinate frames are defined again as a means to develop the subsequent Euclidean reconstruction and control methods. In Fig. 2, a single camera is navigating above the motion plane of an UGV. The moving coordinate frame $\mathcal{I}$ is attached to an overhead camera, which records an images for real-time tracking control. The moving coordinate frame $\mathcal{I}_M$ is attached to the overhead camera that recorded the desired image sequence, and the fixed coordinate frame $\mathcal{I}_R$ is some single snapshot of $\mathcal{I}_M$.

The moving coordinate frame $\mathcal{F}$ is attached to the UGV at the center of the rear wheel axis (for simplicity and without loss of generality). The UGV is represented in the camera image by four feature points that are coplanar and not collinear. The Euclidean distance (i.e., $s_{1i} \in \mathbb{R}^3$ $\forall i = 1, 2, 3, 4$) from the origin of $\mathcal{F}$ to one of the feature points is assumed to be known. A priori information (such as a known target in the initial FOV Davison et al. (2007)) is sometimes used in VSLAM methods to establish scale. The plane defined by the UGV motion (i.e., the plane defined by the xy-axis of $\mathcal{F}$) and the UGV feature points is denoted by π. The linear

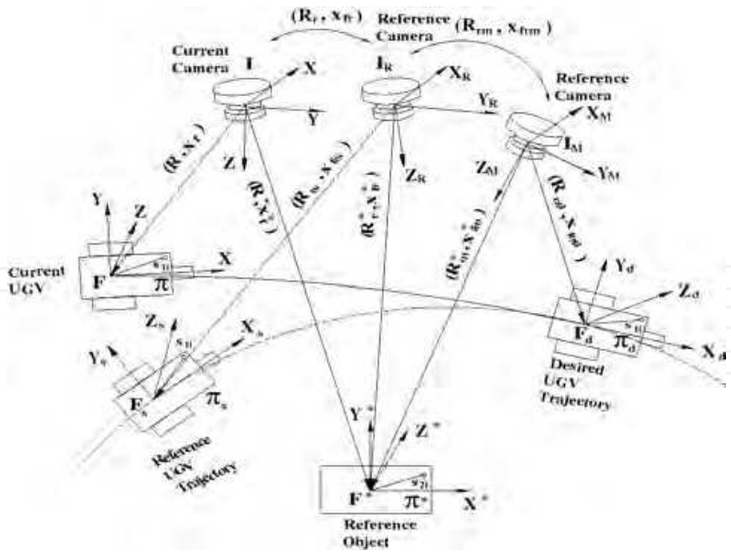


Fig. 2. Geometric model for a moving camera, moving target and moving reference camera.

velocity of the UGV along the x-axis of $\mathcal{F}$ is denoted by $v_c(t) \in \mathbb{R}$, and the angular velocity $\omega_c(t) \in \mathbb{R}$ is about the z-axis of $\mathcal{F}$.

While viewing the feature points of the UGV, the camera is assumed to also view four additional coplanar and noncollinear feature points of a stationary reference object. The four additional feature points define the plane π^* in Fig. 2. The stationary coordinate frame $\mathcal{F}^*$ is attached to the object where a distance from the origin of the coordinate frame to one of the feature points (i.e., $s_{2i} \in \mathbb{R}^3$) is assumed to be known. The plane π^* is assumed to be parallel to the plane π. When the camera is coincident with $\mathcal{I}_R$, a fixed (i.e., a single snapshot) reference pose of the UGV, denoted by $\mathcal{F}_s$, is assumed to be in the camera's FOV. A desired trajectory is defined by a prerecorded time-varying trajectory of $\mathcal{F}_d$ that is assumed to be second-order differentiable where $v_{cd}(t), \omega_{cd}(t) \in \mathbb{R}$ denote the desired linear and angular velocity of $\mathcal{F}_d$, respectively. The feature points that define π^* are also assumed to be visible when the camera is a priori located coincident with the pose of the stationary coordinate frame $\mathcal{I}_R$ and the time-varying coordinate frame $\mathcal{I}_M$. Based on these coordinate frame definitions, the problem considered in this section is to develop a kinematic controller for the object attached to $\mathcal{F}$ so that the time-varying rotation and translation of $\mathcal{F}$ converges to the desired time-varying rotation and translation of $\mathcal{F}_d$, where the motion of $\mathcal{F}$ is determined from the time-varying overhead camera attached to $\mathcal{I}$.

3.2 Geometric Relationships

The rotation matrices and translation vectors in Table I (except the last line) are also valid for this section. Additional relationships between the various coordinate frames are summarized in Table II. In Table II, R_{rs}, $R_{md}(t)$, $R_m^*(t)$, $R_{rm}(t)$, $R'_{md}(t) \in SO(3)$ denote rotation matrices, and x_{frs}, $x_{fmd}(t)$, $x_{fm}^*(t)$, $x_{frm}(t)$, $x'_{frm}(t) \in \mathbb{R}^3$ denote translation vectors.

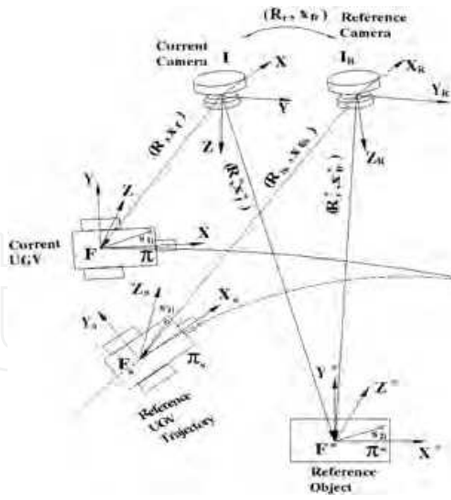


Fig. 3. Geometric model for a moving camera, moving target and a stationary coordinate frame attached to a snapshot of the moving reference camera.

Motion	Frames
R_{rs}, x_{frs}	$\mathcal{F}_s$ to $\mathcal{I}_R$
$R_{md}(t)$, $x_{fmd}(t)$	$\mathcal{F}_d$ to $\mathcal{I}_M$
$R^*_m(t)$, $x^*_{fm}(t)$	$\mathcal{F}^*$ to $\mathcal{I}_M$
$R_{rm}(t)$, $x_{frm}(t)$	$\mathcal{I}_M$ to $\mathcal{I}_R$
$R'_{md}(t)$, $x'_{frm}(t)$	$\mathcal{F}$ to $\mathcal{I}_R$ in $\mathcal{I}_M$

Table 2. Coordinate frame relationships.

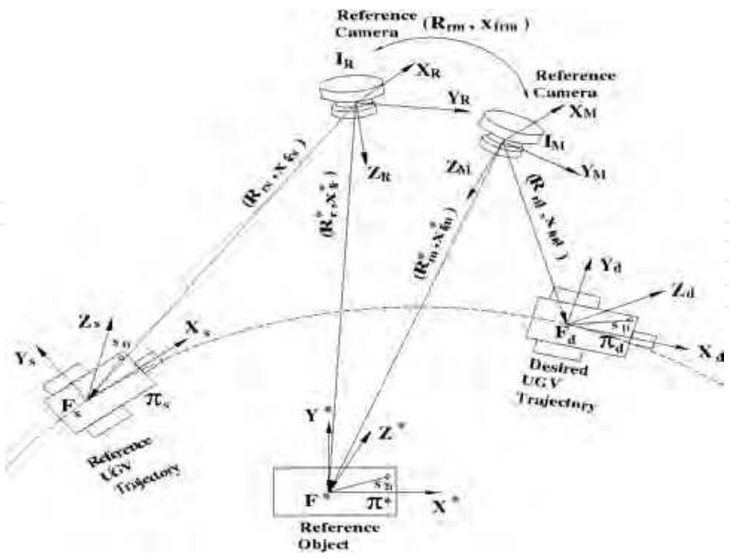


Fig. 4. Geometric model for a moving target, moving reference camera, desired trajectory, and a stationary coordinate frame attached to a snapshot of the moving reference camera.

3.3 Euclidean Reconstruction

The coordinate frame representation in Fig. 2 can be separated into Figs. 3 and 4 to relate $\mathcal{I}$ to $\mathcal{I}_R$ and $\mathcal{I}_R$ to $\mathcal{I}_M$, respectively. The coordinate frames in each figure have the same relationships as in Fig. 1. Therefore, the same Euclidean reconstruction process as presented in Section 2.1-2.3 can be used twice to build the Euclidean relationships for this example.

To reconstruct the Euclidean relationship for the geometric model as shown in Fig. 3, let $\bar{m}_{rsi} \in \mathbb{R}^3$ denote the constant reference Euclidean coordinates of the feature points on π_s expressed in $\mathcal{I}_R$ as

$$\bar{m}_{rsi} \triangleq \begin{bmatrix} x_{rsi} & y_{rsi} & z_{rsi} \end{bmatrix}^T,$$

and let $p_{rsi} \in \mathbb{R}^3$ represent the constant image-space coordinates of the feature points on π_s taken by the camera attached to $\mathcal{I}_M$ when $\mathcal{I}_M$ is coincident with $\mathcal{I}_R$

$$p_{rsi} \triangleq \begin{bmatrix} u_{rsi} & v_{rsi} & 1 \end{bmatrix}^T.$$

Following the development in Section 2.2 and 2.3, relationships can be obtained to determine the homographies and depth ratios as[4]

$$p_i = \alpha_i \underbrace{\left(A\left(\bar{R} + x_h n^{*T}\right)A^{-1}\right)}_{G} p_i^* \qquad p_{rsi} = \alpha_{rsi} \underbrace{\left(A\left(\bar{R}_{rs} + x_{hrs} n_r^{*T}\right)A^{-1}\right)}_{G_{rs}} p_{ri}^* \tag{57}$$

$$p_{ri}^* = \alpha_{ri} \underbrace{\left(A\left(R_r + x_{hr} n^{*T}\right)A^{-1}\right)}_{G_r} p_i^* \tag{58}$$

where

$$\begin{aligned} \alpha_i &= \frac{z_i^*}{z_i} \qquad \alpha_{rsi} = \frac{z_{ri}^*}{z_{rsi}} \qquad \alpha_{ri} = \frac{z_i^*}{z_{ri}^*} \\ \bar{R} &= R R^{*T} \quad \bar{R}_{rs} = R_{rs} R_r^{*T} \quad R_r = R_r^* R^{*T}. \end{aligned} \tag{59}$$

Furthermore, the normalized Euclidean coordinates $m_i(t)$ can be related to $m_i'(t)$ as

$$m_i' = \frac{z_i}{z_i'}\left(R_r + x_{hr}\alpha_i \frac{n^{*T} m_i^*}{n^{*T} m_i} n^{*T}\right) m_i \tag{60}$$

$$\frac{z_i'}{z_i} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\left(R_r + x_{hr}\alpha_i \frac{n^{*T} m_i^*}{n^T m_i} n^{*T}\right) m_i. \tag{61}$$

To reconstruct the Euclidean relationship for the geometric model as shown in Fig. 4, let $\bar{m}_{mdi}(t)$, $\bar{m}_{mi}^*(t) \in \mathbb{R}^3$ denote the Euclidean coordinates of the feature points on π_d and π^* expressed in $\mathcal{I}_M$ as

$$\begin{aligned} \bar{m}_{mdi}(t) &\triangleq \begin{bmatrix} x_{mdi}(t) & y_{mdi}(t) & z_{mdi}(t) \end{bmatrix}^T \\ \bar{m}_{mi}^*(t) &\triangleq \begin{bmatrix} x_{mi}^*(t) & y_{mi}^*(t) & z_{mi}^*(t) \end{bmatrix}^T, \end{aligned}$$

[4] To simplify the notations, the cameras are assumed to have the same calibration matrix A in the following development. The readers can refer to Section 2.1 for the deductions that the calibration matrices are different.

let $\bar{m}'_{mdi}(t) \in \mathbb{R}^3$ denote the desired Euclidean coordinates of the feature points on π_d expressed in $\mathcal{I}_R$ as

$$\bar{m}'_{mdi}(t) \triangleq \begin{bmatrix} x'_{mdi}(t) & y'_{mdi}(t) & z'_{mdi}(t) \end{bmatrix}^T,$$

and let $p_{mdi}(t)$, $p^*_{mi}(t) \in \mathbb{R}^3$ represent the image-space coordinates of the feature points on π_d and π^* captured by the camera attached to $\mathcal{I}_M$, respectively, as

$$p_{mdi} \triangleq \begin{bmatrix} u_{mdi} & v_{mdi} & 1 \end{bmatrix}^T \qquad p^*_{mi} \triangleq \begin{bmatrix} u^*_{mi} & v^*_{mi} & 1 \end{bmatrix}^T.$$

The normalized coordinates of $\bar{m}'_{mdi}(t)$ and $\bar{m}_{mdi}(t)$, denoted as $m'_{mdi}(t)$, $m_{mdi}(t) \in \mathbb{R}^3$, respectively, are defined as

$$m'_{mdi}(t) \triangleq \frac{\bar{m}'_{mdi}(t)}{z'_{mdi}(t)} \qquad m_{mdi}(t) \triangleq \frac{\bar{m}_{mdi}(t)}{z_{mdi}(t)}. \tag{62}$$

Following the development in Section 2.2 and 2.3, relationships can be developed to compute the homographies and depth ratios as

$$p_{mdi} = \alpha_{mdi} \underbrace{\left(A\left(\bar{R}_{md} + x_{hmd} n^{*T}_m\right) A^{-1}\right)}_{G_{md}} p^*_{mi} \qquad p^*_{ri} = \alpha_{rmi} \underbrace{\left(A\left(R_{rm} + x_{hrm} n^{*T}_m\right) A^{-1}\right)}_{G_{rm}} p^*_{mi}, \tag{63}$$

where

$$\begin{aligned} \alpha_{mdi} &= \frac{z^*_{mi}}{z_{mdi}} \qquad \alpha_{rmi} = \frac{z^*_{mi}}{z^*_{ri}} \\ \bar{R}_{md} &= R_{md} R^{*T}_m \qquad R_{rm} = R^*_r R^{*T}_m. \end{aligned} \tag{64}$$

Relationships between $m_{mdi}(t)$ and $m'_{mdi}(t)$ can be established as

$$m'_{mdi} = \frac{z_{mdi}}{z'_{mdi}} \left(R_{rm} + x_{hrm} \alpha_{mdi} \frac{n^{*T}_m m^*_{mi}}{n^{*T}_m m_{mdi}} n^{*T}_m\right) m_{mdi} \tag{65}$$

$$\frac{z'_{mdi}}{z_{mdi}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \left(R_{rm} + x_{hrm} \alpha_{mdi} \frac{n^{*T}_m m^*_{mi}}{n^{*T}_m m_{mdi}} n^{*T}_m\right) m_{mdi}. \tag{66}$$

In (57)-(63), $n^*(t)$, $n^*_m(t)$, and $n^*_r \in \mathbb{R}^3$ denote the constant unit normal to the planes π and π^* as expressed in $\mathcal{I}$, $\mathcal{I}_M$, and $\mathcal{I}_R$ respectively, $x_h(t)$, $x_{hrs}(t)$, $x_{hr}(t)$, $x_{hmd}(t)$, $x_{hrm}(t) \in \mathbb{R}^3$ denote the corresponding scaled translation vectors, and $G(t)$, G_{rs}, $G_r(t)$, $G_{md}(t)$, $G_{rm}(t) \in R^{3\times 3}$ denote projective homographies.

Sets of linear equations in (57)-(58) and (63) can be used to determine and decompose homographies to obtain $\alpha_i(t)$, α_{rsi}, $\alpha_{mdi}(t)$, $\alpha_{ri}(t)$, $\alpha_{rmi}(t)$, $x_h(t)$, x_{hrs}, $x_{hr}(t)$, $x_{hmd}(t)$, $x_{hrm}(t)$, $\bar{R}(t)$, $\bar{R}_{rs}$, $R_r(t)$, $\bar{R}_{md}(t)$, and $R_{rm}(t)$. Given that the rotation matrix $R^*_r(t)$ is assumed to be known, the expressions for $\bar{R}_{rs}(t)$ and $R_r(t)$ in (59) can be used to determine $R_{rs}(t)$ and $R^*(t)$. Once $R^*(t)$ is determined, the expression for $\bar{R}(t)$ and $R_{rm}(t)$ in (59) and (64) can be used to determine $R(t)$ and $R^*_m(t)$. The rotation $R^*_m(t)$ can then be used to calculate $R_{md}(t)$ from the relationship for $\bar{R}_{md}$ in (64). Based on the definitions for $R(t)$, $R^*(t)$, $R_{md}(t)$, $R^*_m(t)$, R^*_r,

and R_{rs} provided in the previous development, the rotation from $\mathcal{F}$ to $\mathcal{F}_s$ and from $\mathcal{F}_d$ to $\mathcal{F}_s$, denoted by $R_1(t)$, $R_{d1}(t) \in SO(3)$, respectively, are defined as

$$R_1(t) = R_{rs}^T R_r^* R^{*T}(t) R(t) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{67}$$

$$R_{d1}(t) = R_{rs}^T R_r^* R_m^{*T}(t) R_{md}(t) = \begin{bmatrix} \cos\theta_d & \sin\theta_d & 0 \\ -\sin\theta_d & \cos\theta_d & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{68}$$

where $\theta(t) \in \mathbb{R}$ denotes the right-handed rotation angle about the z-axis that aligns $\mathcal{F}$ with $\mathcal{F}_s$, and $\theta_d(t) \in \mathbb{R}$ denotes the right-handed rotation angle about the z-axis that aligns $\mathcal{F}_d$ with $\mathcal{F}_s$. From the definitions of $\theta(t)$ and $\theta_d(t)$, it is clear that

$$\dot{\theta} = \omega_c \qquad \dot{\theta}_d = \omega_{cd} \tag{69}$$

where $\omega_c(t), \omega_{cd}(t) \in \mathbb{R}$ denote the desired angular velocities of $\mathcal{F}$ and $\mathcal{F}_d$, respectively. Based on the fact that $R(t)$, $R^*(t)$, $R_{md}(t)$, $R_m^*(t)$, R_r^*, and R_{rs} are known, it is clear from (67)-(69) that $\theta(t)$ and $\theta_d(t)$ are known signals that can be used in the subsequent control development. To facilitate the subsequent development, $\theta(t)$ and $\theta_d(t)$ are assumed to be confined to the following regions

$$-\pi < \theta(t) \leqslant \pi \qquad -\pi < \theta_d(t) \leqslant \pi. \tag{70}$$

3.4 Control Objective

The objective is to develop a visual servo controller that ensures that the coordinate system $\mathcal{F}$ tracks the time-varying trajectory of $\mathcal{F}_d$ (i.e., $\bar{m}_i(t)$ measured in $\mathcal{I}$ tracks $\bar{m}_{mdi}(t)$ measured in $\mathcal{I}_M$). To ensure that $\bar{m}_i(t)$ tracks $\bar{m}_{mdi}(t)$, the control objective can be stated by using the Euclidean reconstruction given in (57)-(63) as the desire for $\bar{m}_1'(t) \rightarrow \bar{m}_{md1}'(t)$. To quantify the control objective, translation and rotation tracking error vector, denoted by $e(t) \triangleq [e_1(t), e_2(t), e_3(t)]^T \in \mathbb{R}^3$, is defined as Dixon et al. (2003)

$$e_1 \triangleq \eta_1 - \eta_{d1} \qquad e_2 \triangleq \eta_2 - \eta_{d2} \qquad e_3 \triangleq \theta - \theta_d \tag{71}$$

where $\theta(t)$ and $\theta_d(t)$ are introduced in (67) and (68), respectively, and the auxiliary signals $\eta(t) \triangleq [\eta_1(t), \eta_2(t), \eta_3(t)]^T$, $\eta_d(t) \triangleq [\eta_{d1}(t), \eta_{d2}(t), \eta_{d3}(t)]^T \in \mathbb{R}^3$ are defined as

$$\eta(t) \triangleq \frac{1}{z_{r1}^*} R^T(t) R^*(t) R_r^{*T} \bar{m}_1'(t) \tag{72}$$

$$\eta_d(t) \triangleq \frac{1}{z_{r1}^*} R_{md}^T(t) R_m^*(t) R_r^{*T} \bar{m}_{md1}'(t).$$

Also, the normal unit vector n_r^* is defined as Mehta, Hu, Gans & Dixon (2006)

$$\begin{aligned} n_r^* &= R_r^* R^{*T}(t) R(t) \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T \\ &= R_r^* R_m^{*T}(t) R_{md}(t) \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T. \end{aligned} \tag{73}$$

The expressions in (72) and (73) can be used to determine that

$$\eta_3 = \eta_{d3} = \frac{-d_r}{z_{r1}^*}. \tag{74}$$

The expressions in (57)-(66) can be used to rewrite $\eta\,(t)$ and $\eta_d(t)$ in terms of the measurable signals $\alpha_1(t)$, $\alpha_{r1}(t)$, $\alpha_{rm1}(t)$, $\alpha_{md1}(t)$, $R(t)$, $R^*(t)$, R_r^*, $R_{md}(t)$, $R_m^*(t)$, $p_1(t)$, and $p_{md1}(t)$ as

$$\begin{aligned} \eta\,(t) &= \frac{\alpha_{r1}}{\alpha_1} R^T(t) R^*(t) R_r^{*T} H_r' A^{-1} p_1 \\ \eta_d(t) &= \frac{\alpha_{rm1}}{\alpha_{md1}} R_{md}^T(t) R_m^*(t) R_r^{*T} H_{rm}' A^{-1} p_{md1}. \end{aligned} \tag{75}$$

Based on (71), (75), and the fact that $\theta(t)$ and $\theta_d(t)$ are measurable, it is clear that $e(t)$ is measurable. By examining (71)-(74), the control objective is achieved if $\|e(t)\| \to 0$. Specifically, if $e_3(t) \to 0$, then it is clear from (71) that $R_1(t) \to R_{d1}(t)$. If $e_1(t) \to 0$ and $e_2(t) \to 0$, then from (71) and (74) it is clear that $\eta\,(t) \to \eta_d(t)$. Given that $R_1(t) \to R_{d1}(t)$ and that $\eta\,(t) \to \eta_d(t)$, then (72) can be used to conclude that $m_1'(t) \to m_{md1}'(t)$. If $m_1'(t) \to m_{md1}'(t)$ and $R_1(t) \to R_{d1}(t)$, then the Euclidean relationships in the geometric model can be used to prove that $\bar{m}_i(t)$ measured in terms of $\mathcal{I} \to \bar{m}_{mdi}(t)$ measured in terms of $\mathcal{I}_M$.

3.5 Control Development

The open-loop error system can be obtained by taking the time derivative of (72) as

$$\dot{\eta} = \frac{v}{z_{r1}^*} + \left[\eta - \frac{s_{11}}{z_{r1}^*}\right]^{\times} \omega \tag{76}$$

where $v(t), \omega(t) \in \mathbb{R}^3$ denote the respective linear and angular velocity of an UGV expressed in $\mathcal{F}$ as

$$v \triangleq \begin{bmatrix} v_c & 0 & 0 \end{bmatrix}^T \qquad \omega \triangleq \begin{bmatrix} 0 & 0 & \omega_c \end{bmatrix}^T. \tag{77}$$

Without loss of generality, the location of the feature point s_1 is taken as the origin of $\mathcal{F}$, so that $s_{11} = [0, 0, 0]^T$. Then, based on (76) and (77), the error system can be further written as

$$\dot{\eta}_1 = \frac{v_c}{z_{r1}^*} + \eta_2 \omega_c \qquad \dot{\eta}_2 = -\eta_1 \omega_c. \tag{78}$$

Since the desired trajectory is assumed to be generated in accordance with UGV motion constraints, a similar expression to (78) can be developed as

$$\dot{\eta}_{d1} = \frac{v_{cd}}{z_{r1}^*} + \eta_{2d} \omega_{cd} \qquad \dot{\eta}_{d2} = -\eta_{d1} \omega_{cd}\,. \tag{79}$$

where $v_{cd}(t) \in \mathbb{R}$ denotes the desired linear velocity of $\mathcal{F}_d$. From (69), (71), (76) and (78), the open-loop error system can be obtained as

$$\begin{aligned} z_{r1}^* \dot{e}_1 &= v_c + z_{r1}^* (\eta_2 \omega_c - \dot{\eta}_{d1}) \\ \dot{e}_2 &= -\eta_1 \omega_c + \eta_{d1} \dot{\theta}_d \\ \dot{e}_3 &= \omega_c - \dot{\theta}_d. \end{aligned} \tag{80}$$

To facilitate the subsequent development, the auxiliary variable $\bar{e}_2(t) \in \mathbb{R}$ is defined as

$$\bar{e}_2 \triangleq e_2 + \eta_{d1} e_3. \tag{81}$$

After taking the time derivative of (81) and utilizing (80), the following expression is obtained:

$$\dot{\bar{e}}_2 = -e_1 \omega_c + \dot{\eta}_{d1} e_3. \tag{82}$$

Based on (81), it is clear that if $\bar{e}_2(t)$, $e_3(t) \to 0$, then $e_2(t) \to 0$. Based on this observation and the open-loop dynamics given in (82), the following control development is based on the desire to show that $e_1(t), \bar{e}_2(t), e_3(t)$ are asymptotically driven to zero.
Based on the open-loop error systems in (80) and (82), the linear and angular velocity control inputs for an UGV are designed as

$$v_c \triangleq -k_v e_1 + \bar{e}_2 \omega_c - \hat{z}_{r1}^* (\eta_2 \omega_c - \dot{\eta}_{d1}) \tag{83}$$

$$\omega_c \triangleq -k_\omega e_3 + \dot{\theta}_d - \dot{\eta}_{d1} \bar{e}_2 \tag{84}$$

where k_v, $k_\omega \in \mathbb{R}$ denote positive, constant control gains. In (83), the parameter update law $\hat{z}_{r1}^*(t) \in \mathbb{R}$ is generated by the differential equation

$$\dot{\hat{z}}_{r1}^* = \gamma_1 e_1 (\eta_2 \omega_c - \dot{\eta}_{d1}) \tag{85}$$

where $\gamma_1 \in \mathbb{R}$ is a positive, constant adaptation gain. After substituting the kinematic control signals designed in (83) and (84) into (80), the following closed-loop error systems are obtained:

$$\begin{aligned} z_{r1}^* \dot{e}_1 &= -k_v e_1 + \bar{e}_2 \omega_c + \tilde{z}_{r1}^* (\eta_2 \omega_c - \dot{\eta}_{d1}) \\ \dot{\bar{e}}_2 &= -e_1 \omega_c + \dot{\eta}_{d1} e_3 \\ \dot{e}_3 &= -k_\omega e_3 - \dot{\eta}_{d1} \bar{e}_2 \end{aligned} \tag{86}$$

where (82) was utilized, and the depth-related parameter estimation error, denoted by $\tilde{z}_{r1}^*(t) \in \mathbb{R}$, is defined as

$$\tilde{z}_{r1}^* \triangleq z_{r1}^* - \hat{z}_{r1}^* . \tag{87}$$

Theorem 2. *The control input designed in (83) and (84) along with the adaptive update law defined in (85) ensure asymptotic tracking for UGV in the sense that*

$$\|e(t)\| \to 0 \tag{88}$$

provided the time derivative of the desired trajectory satisfies the following condition

$$\dot{\eta}_{d1} \nrightarrow 0. \tag{89}$$

Lyapunov-based analysis methods and Barbalat's lemma can be used to proved Theorem 2 based on a positive definite function $V(t) \in \mathbb{R}$ defined as Mehta, Hu, Gans & Dixon (2006)

$$V \triangleq \frac{1}{2} z_{r1}^* e_1^2 + \frac{1}{2} \bar{e}_2^2 + \frac{1}{2} e_3^2 + \frac{1}{2\gamma_1} \tilde{z}_{r1}^{*2} . \tag{90}$$

4. Simultaneous Tracking, Localization and Mapping

For vision-based autonomous systems applications (e.g., tracking, localization and mapping), the given reference object can leave the camera's FOV while another reference object enters the FOV. In comparison to the single reference object problem presented in Section 3, multiple reference objects are taken into consideration in this section. The daisy-chaining method is further developed to achieve asymptotic tracking of the UGV by mapping each reference object to a global coordinate system. Moreover, the time-varying Euclidean position of the UGV and the stationary position of the reference objects can be localized with respect to the global coordinate system. In addition to achieving the visual servo tracking and localization objectives, the developed method generates data for the SLAM problem.

4.1 Problem Scenario

The geometric model in this section is the same as in Section 3, except that multiple reference objects are taken into consideration. While viewing the feature points of the UGV, the camera is assumed to also view four additional coplanar and noncollinear feature points of a stationary reference object, such that at any instant of time along the camera motion trajectory at least one such reference object is in the FOV and two reference objects are required to be visible for the frames switching from one reference to the other. The four additional feature points define the plane π_j^* in Fig. 5. The stationary coordinate frame $\mathcal{F}_j^*$ $(j = 1, 2, ..., k)$ is attached to the object where distance from the origin of the coordinate frame to one of the feature points is assumed to be known, i.e., $s_{2ji} \in \mathbb{R}^3$ $\forall i = 1, 2, 3, 4$. The plane π_j^* is assumed to be parallel to the plane π. The feature points that define π_1^*, corresponding to a reference object $\mathcal{F}_1^*$ (i.e., $\mathcal{F}_j^*$ corresponding to $j = 1$), are also assumed to be visible when the camera is a priori located coincident with the pose of the stationary coordinate frame $\mathcal{I}_R$. The fixed coordinate frame $\mathcal{I}_R$ is a snapshot of I_M at the time instant that the first reference object π_1^* is visible to the reference camera. The reference object π_1^* is visible to $\mathcal{I}_R$, but the other reference objects $\pi_j^*(j > 1)$ are not.

4.2 Geometric Relationships

In addition to the notations in Tables I and II, more relationships between the various coordinate frames are summarized in Table III. In Table III, $R_j^*(t)$, $R_{rj}^*(t)$, $R_{mj}^*(t) \in SO(3)$ denote rotation matrices and $x_{fj}^*(t)$, x_{frj}^*, $x_{fmj}^*(t) \in \mathbb{R}^3$ denote translation vectors.

Motion	Frames
$R_j^*(t)$, $x_{fj}^*(t)$	$\mathcal{F}_j^*$ to $\mathcal{I}$ in $\mathcal{I}$
$R_{rj}^*(t)$, x_{frj}^*	$\mathcal{F}_j^*$ to $\mathcal{I}_R$ in $\mathcal{I}_R$
$R_{mj}^*(t)$, $x_{fmj}^*(t)$	$\mathcal{F}_j^*$ to $\mathcal{I}_M$ in $\mathcal{I}_M$

Table 3. Coordinate frame relationships.

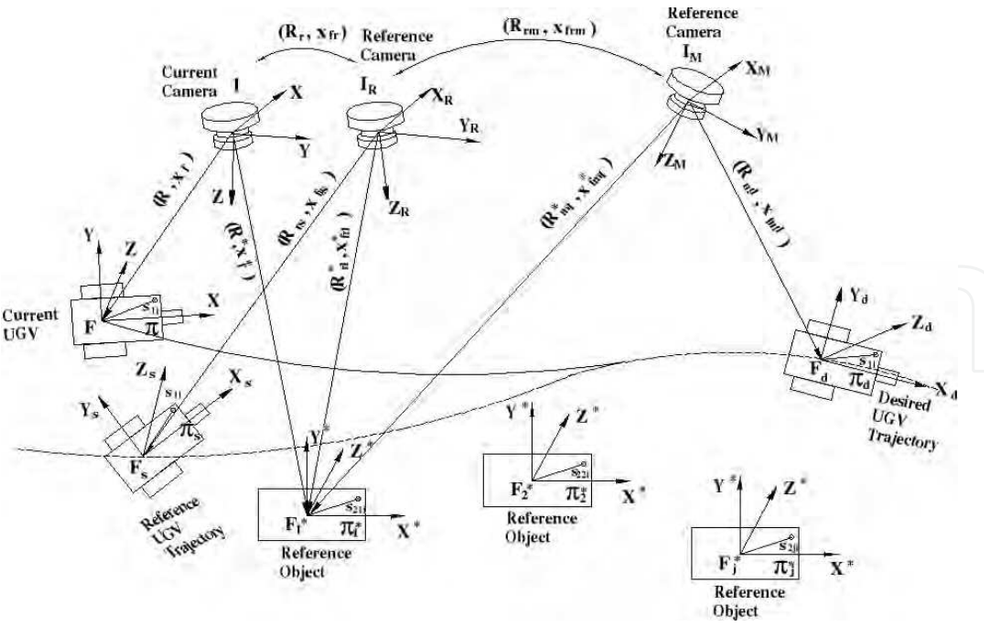


Fig. 5. Geometric model for a moving camera, moving target, moving reference camera and multiple reference objects for simultaneous tracking, localization, and mapping.

4.3 Euclidean Reconstruction

The Euclidean reconstruction for the geometric model in Fig. 5 can be separated into three cases. Case 1: a single reference object π_1^* is within the reference camera's FOV and therefore π_1^* is used as the reference object. Case 2: two reference objects (e.g., π_1^* and π_2^*) are within the camera's FOV, and the reference object in use is going to be switched from one to the other (e.g., from π_1^* to π_2^*). Case 3: π_j^* $(j \geq 2)$ is used as the reference object.

Let $\bar{m}_{mji}^*(t), \bar{m}_{rji}^{'*} \in \mathbb{R}^3$ denote the Euclidean coordinates of the feature points on π_j^* expressed in $\mathcal{I}_M$ and $\mathcal{I}_R$, respectively. Also, let $p_{mji}^*(t), p_{rji}^{'*} \in \mathbb{R}^3$ represent the image-space coordinates of the feature points on π_j^* captured by the reference camera attached to $\mathcal{I}_M$ and $\mathcal{I}_R$, respectively. When $j = 1$, $p_{r1i}^{'*}$ are measurable and denoted by p_{r1i}^*, however, when $j > 1$, $p_{rji}^{'*}$ can not be measured directly and needs to be computed based on the corresponding normalized coordinates obtained from the daisy-chaining multi-view geometry. The normalized coordinates of $\bar{m}_{mji}^*(t)$ and $\bar{m}_{rji}^{'*}$ are denoted as $m_{mji}^*(t), m_{rji}^{'*} \in \mathbb{R}^3$, respectively.

For the first case, the Euclidean reconstruction is exactly the same as that in Section 3. For the second case, consider the feature point planes π_1^* and π_2^* as an example. Similar to the Euclidean reconstruction development in Section 3.3, relationships can be established between

$m_{r2i}^{'*}(t)$ and $m_{r1i}^{*}(t)$ as

$$m_{r2i}^{'*} = \frac{z_{r1i}^{*}}{z_{r2i}^{'*}} \left(R_{21} + x_{h21}\alpha_{rm1i} \frac{n_{m1}^{*T} m_{m1i}^{*}}{n_{r1}^{*T} m_{r1i}^{*}} n_{r1}^{*T} \right) m_{r1i}^{*} \tag{91}$$

$$\frac{z_{r2i}^{'*}}{z_{r1i}^{*}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \left(R_{21} + x_{h21}\alpha_{rm1i} \frac{n_{m1}^{*T} m_{m1i}^{*}}{n_{r1}^{*T} m_{r1i}^{*}} n_{r1}^{*T} \right) m_{r1i}^{*}. \tag{92}$$

From $m_{r2i}^{'*}(t)$ in (91), the virtual pixel coordinates $p_{r2i}^{'*}(t)$ can be computed. Based on (91) and (92), the Euclidean coordinates of the feature points on π_2^* can be related to fixed coordinate frame $\mathcal{I}_R$. Following the same idea used to relate π_2^* and π_1^*, π_j^* can be related to π_{j-1}^* ($j = 3, ..., k$) as

$$m_{rji}^{'*} = \frac{z_{r(j-1)i}^{*}}{z_{rji}^{'*}} \left(R_{j(j-1)} + x_{hj(j-1)}\alpha_{rm(j-1)i} \frac{n_{m(j-1)}^{*T} m_{m(j-1)i}^{*}}{n_{r(j-1)}^{*T} m_{r(j-1)i}^{*}} n_{r(j-1)}^{*T} \right) m_{r(j-1)i}^{*} \tag{93}$$

$$\frac{z_{rji}^{'*}}{z_{r(j-1)i}^{*}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \left(R_{j(j-1)} + x_{hj(j-1)}\alpha_{rm(j-1)i} \frac{n_{m(j-1)}^{*T} m_{m(j-1)i}^{*}}{n_{r(j-1)}^{*T} m_{r(j-1)i}^{*}} n_{r(j-1)}^{*T} \right) m_{r(j-1)i}^{*}. \tag{94}$$

Recursively, from (91)-(94), $m_{rji}^{'*}(t)$ can be related to the known normalized Euclidean coordinate m_{r1i}^{*}. Once $m_{rji}^{'*}(t)$ is computed based on Case 2, the geometric model in Fig. 5 is equivalent to that in Fig. 2. Therefore, the Euclidean reconstruction in Section 3 can be used to build the Euclidean relationships among different coordinate frames.

4.4 Tracking and Mapping

The tracking control design is the same as that in Section 3, once the Euclidean relationship between $\mathcal{F}$ and $\mathcal{F}_d$ is obtained based on the Euclidean reconstruction analysis as shown in Section 4.3. The time-varying Euclidean position of the UGV and the stationary position of the reference objects can be localized with respect to the global coordinate system $\mathcal{I}_R$. Using the known geometric length s_{21i} and a unit normal n_{r1}^{*} (i.e., the normal to π_1^* expressed in $\mathcal{I}_R$), the geometric reconstruction method in Dupree et al. (2007); Gans et al. (2008); Mackunis et al. (2007) can be utilized to obtain $\bar{m}_{r1i}^{*}(t)$. Based on the computed $\bar{m}_{r1i}^{*}(t)$, (92) can be used to find $z_{r2i}^{'*}$, and then (91) can be used to find $\bar{m}_{r2i}^{'*}(t)$. Recursively, based on (93) and (94), the Euclidean coordinates of the other reference objects denoted as $\bar{m}_{rji}^{'*}(t)$ ($j = 3, ..., k$) can be computed. Similarly, using the known geometric length s_{1i} and a unit normal $n(t)$ (i.e., the normal to π expressed in $\mathcal{I}$), the geometric reconstruction method in Gans et al. (2008) can also be utilized to obtain $\bar{m}_i(t)$. Based on the computed $\bar{m}_i(t)$, (60) and (61) can be used to find $\bar{m}_i^{'}(t)$.

4.5 Simulation Results

A numerical simulation was performed to illustrate the localization and mapping performance given the controller in (83), (84), and the adaptive update law in (85). The origins of the coordinate frames $\mathcal{F}$, $\mathcal{F}_1^*$, $\mathcal{F}_2^*$, and $\mathcal{F}_d$, and the four coplanar feature points on the planes π, π_1^* π_2^*, and π_d are chosen such that the Euclidean coordinates of the feature points in $\mathcal{F}$, $\mathcal{F}_1^*$,

$\mathcal{F}_2^*$, and $\mathcal{F}_d$ are given by s_i (where $i = 1, 2, 3, 4$) i.e., the feature points are located at the same distance from the origins of the coordinate frames $\mathcal{F}$, $\mathcal{F}_1^*$, $\mathcal{F}_2^*$, and $\mathcal{F}_d$. The Euclidean space trajectory of the time-varying feature points attached to the plane π_d and the moving reference camera $\mathcal{I}_M$ along with the performance of the visual servo tracking controller is shown in Fig. 6, which shows the Euclidean space trajectory of the feature points attached to the planes π and π_d, taken by $\mathcal{I}$ and $\mathcal{I}_M$, respectively and the time-varying trajectory of the current and reference camera, $\mathcal{I}$ and $\mathcal{I}_M$, respectively. From Fig. 6, it can be seen that the current trajectory corresponding to the time-varying UGV $\mathcal{F}(t)$ is indistinguishable from the desired trajectory corresponding to the time-varying UGV $\mathcal{F}_d(t)$. The tracking error is depicted in Fig. 7. Fig. 8 shows the results of localization of the current UGV attached to $\mathcal{F}(t)$ and mapping of reference targets attached to $\mathcal{F}_1^*$ and $\mathcal{F}_2^*$ expressed in constant reference frame $\mathcal{I}_R$.

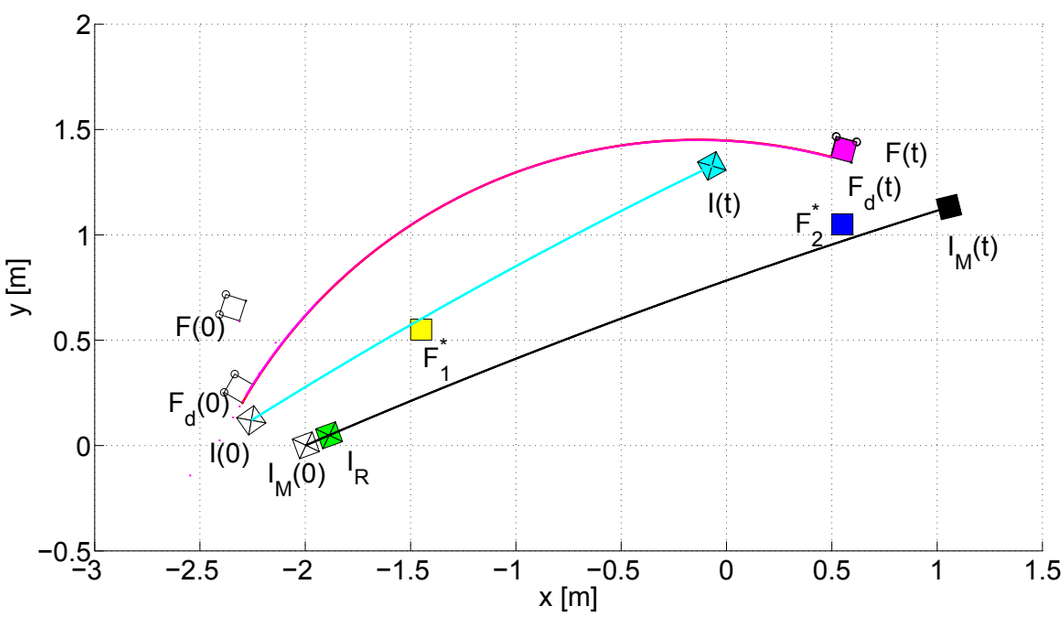


Fig. 6. Euclidean space trajectory of the feature points attached to the current (i.e. $\mathcal{F}(t)$) and desired (i.e. $\mathcal{F}_d(t)$) UGV taken by $\mathcal{I}$ and $\mathcal{I}_M$, respectively and the time-varying trajectory of the current and reference camera, $\mathcal{I}$ and $\mathcal{I}_M$, respectively. $\mathcal{F}(0)$ denotes the initial position of the current UGV, $\mathcal{F}(t)$ denotes the time-varying position of the current UGV, $\mathcal{F}_d(0)$ denotes the initial position of the desired UGV, $\mathcal{F}_d(t)$ denotes the time-varying position of the desired UGV, $\mathcal{I}(0)$ denotes the initial position of the current camera, $\mathcal{I}(t)$ denotes the time-varying position of the current camera, $\mathcal{I}_M(0)$ denotes the initial position of the time-varying reference camera, $\mathcal{I}_M(t)$ denotes the time-varying position of the time-varying reference camera, $\mathcal{I}_R$ denotes the position of the stationary reference camera, and $\mathcal{F}_1^*$ and $\mathcal{F}_2^*$ denote the position of the stationary reference objects.

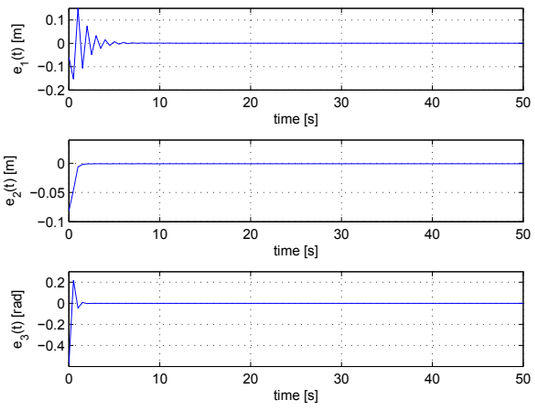


Fig. 7. Linear (i.e. $e_1(t)$ and $e_2(t)$) and angular (i.e. $e_3(t)$) tracking error.

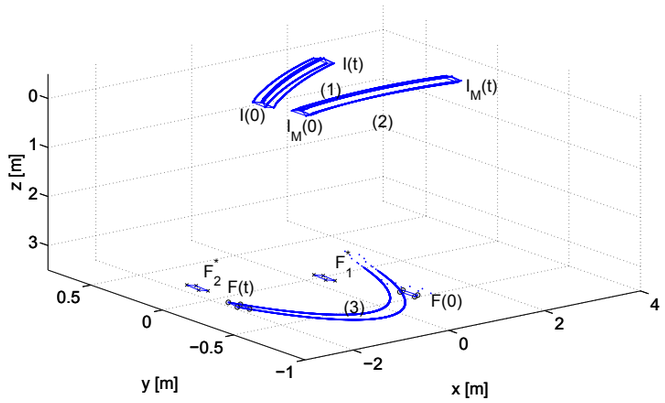


Fig. 8. Results of localization of the current UGV attached to $\mathcal{F}(t)$ and mapping of reference targets attached to $\mathcal{F}_1^*$ and $\mathcal{F}_2^*$ expressed in constant reference frame $\mathcal{I}_R$. Specifically, trajectory (1) shows the time-varying pose of the current camera attached to $\mathcal{I}(t)$, trajectory (2) shows the time-varying pose of the moving camera attached to $\mathcal{I}_M(t)$, and trajectory (3) shows the time-varying pose of the current UGV attached to $\mathcal{F}(t)$ measured in the stationary reference camera frame $\mathcal{I}_R$. $\mathcal{F}(0)$ denotes the initial position of the current UGV, $\mathcal{I}(0)$ denotes the initial position of the current camera, $\mathcal{I}_M(0)$ denotes the initial position of the time-varying reference camera, and $\mathcal{F}_1^*$ and $\mathcal{F}_2^*$ denote the position of the stationary reference objects.

5. References

Allen, P. K., Timcenko, A., Yoshimi, B. & Michelman, P. (1993). Automated tracking and grasping of a moving object with a robotic hand-eye system, *IEEE Trans. Robot. Automat.* **9**(2): 152–165.

Baker, S. & Nayar, S. (1999). A theory of single-viewpoint catadioptric image formation, *Int. J. Computer Vision* **35**(2): 175–196.

Bensalah, F. & Chaumette, F. (1995). Compensation of abrupt motion changes in target tracking by visual servoing, *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 181–187.

Boufama, B. & Mohr, R. (1995). Epipole and fundamental matrix estimation using virtual parallax, *Proc. IEEE Int. Conf. Computer Vision*, pp. 1030–1036.

Burschka, D. & Hager, G. (2001). Vision-based control of mobile robots, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 1707–1713.

Chen, J., Dawson, D. M., Dixon, W. E. & Behal, A. (2005). Adaptive homography-based visual servo tracking for a fixed camera configuration with a camera-in-hand extension, *IEEE Trans. Contr. Syst. Technol.* **13**(5): 814– 825.

Chen, J., Dixon, W. E., Dawson, D. M. & McIntyre, M. (2006). Homography-based visual servo tracking control of a wheeled mobile robot, *IEEE Trans. Robot.* **22**(2): 406–415.

Das, A., Fierro, R., Kumar, V., Southall, B., Spletzer, J. & Taylor, C. (2001). Real-time vision-based control of a nonholonomic mobile robot, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 1714–1719.

Davison, A. J., Reid, I. D., Molton, N. D. & Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM, *IEEE Trans. Pattern Anal. Machine Intell.* **29**(6): 1052–1067.

Dixon, W. E., Behal, A., Dawson, D. M. & Nagarkatti, S. (2003). *Nonlinear Control of Engineering Systems: A Lyapunov-Based Approach*, Birkhäuser Boston.

Dixon, W. E., Dawson, D. M., Zergeroglu, E. & Behal, A. (2001). Adaptive tracking control of a wheeled mobile robot via an uncalibrated camera system, *IEEE Trans. Syst., Man, Cybern. - Part B: Cybern.* **31**(3): 341–352.

Dupree, K., Gans, N., Mackunis, W. & Dixon, W. E. (2007). Euclidean feature tracking for a rotating satellite, *Proc. American Control Conf.*, pp. 3874–3879.

Eustice, R., Singh, H., Leonard, J., Walter, M. & Ballard, R. (2005). Visually navigating the RMS titanic with SLAM information filters, *Proc. Robotics: Science and Systems*.

Fang, Y., Dixon, W. E., Dawson, D. M. & Chawda, P. (2005). Homography-based visual servoing of wheeled mobile robots, *IEEE Trans. Syst., Man, Cybern. - Part B: Cybern.* **35**(5): 1041–1050.

Faugeras, O. & Lustman, F. (1988). Motion and structure from motion in a piecewise planar environment, *Int. J. Pattern Recognition and Artificial Intelligence* **2**(3): 485–508.

Gans, N. R., Dani, A. & Dixon, W. E. (2008). Visual servoing to an arbitrary pose with respect to an object given a single known length, *Proc. American Control Conf.*, pp. 1261–1267.

Goncalves, L., di Bernardo, E., Benson, D., Svedman, M., Ostrowski, J., Karlsson, N. & Pirjanian, P. (2005). A visual front-end for simultaneous localization and mapping, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 44–49.

Hagar, G., Kriegman, D., Georghiades, A. & Ben-Shahar, O. (1998). Toward domain-independent navigation: Dynamic vision and control,", *Proc. IEEE Conf. Decision Control*, pp. 3257–3262.

Hager, G. D., Chang, W.-C. & Morse, A. S. (1995). Robot hand-eye coordination based on stereo vision, *IEEE Contr. Syst. Mag.* **15**(1): 30–39.

Hu, G., Dixon, W. E., Gupta, S. & Fitz-coy, N. (2006). A quaternion formulation for homography-based visual servo control, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 2391–2396.

Hu, G., Gans, N., Mehta, S. & Dixon, W. E. (2007). Daisy chaining based visual servo control part II: Extensions, applications and open problems, *Proc. IEEE Multi-Conf. Syst. Control*, pp. 729–734.

Hu, G., Mehta, S., Gans, N. & Dixon, W. E. (2007). Daisy chaining based visual servo control part I: Adaptive quaternion-based tracking control, *Proc. IEEE Multi-Conf. Syst. Control*, pp. 1474–1479.

Hutchinson, S., Hager, G. & Corke, P. (1996). A tutorial on visual servo control, *IEEE Trans. Robot. Automat.* **12**(5): 651–670.

Jensfelt, P., Kragic, D., Folkesson, J. & Bjorkman, M. (2006). A framework for vision based bearing only 3D SLAM, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 1944–1950.

Jung, I.-K. & Lacroix, S. (2003). High resolution terrain mapping using low attitude aerial stereo imagery, *Proc. IEEE Int. Conf. Computer Vision*, pp. 946–951.

Kim, B., Roh, D., Lee, J., Lee, M., Son, K., Lee, M., Choi, J. & Han, S. (2001). Localization of a mobile robot using images of a moving target, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 253–258.

Kim, J.-H. & Sukkarieh, S. (2003). Airborne simultaneous localisation and map building, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 406–411.

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

Mackunis, W., Gans, N., Kaiser, K. & Dixon, W. E. (2007). Unified tracking and regulation visual servo control for wheeled mobile robots, *Proc. IEEE Multi-Conf. Syst. Control*, pp. 88–93.

Malis, E. & Chaumette, F. (2000). 2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement, *Int. J. Computer Vision* **37**(1): 79–97.

Mehta, S., Dixon, W. E., MacArthur, D. & Crane, C. D. (2006). Visual servo control of an unmanned ground vehicle via a moving airborne monocular camera, *Proc. American Control Conf.*, pp. 5276–5211.

Mehta, S., Hu, G., Gans, N. & Dixon, W. E. (2006). Adaptive vision-based collaborative tracking control of an ugv via a moving airborne camera: A daisy chaining approach, *Proc. IEEE Conf. Decision Control*, pp. 3867–3872.

Papanikolopoulos, N., Khosla, P. & Kanade, T. (1993). Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision, *IEEE Trans. Robot. Automat.* **9**(1): 14–35.

Se, S., Lowe, D. & Little, J. (2002). Global localization using distinctive visual features, *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, pp. 226–231.

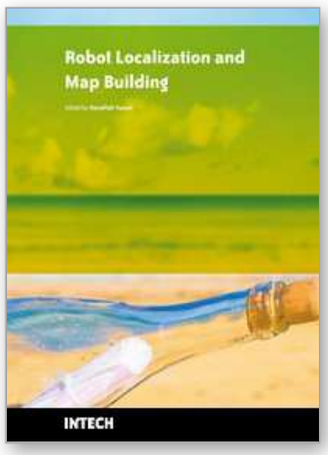
Shuster, M. (1993). A survey of attitude representations, *J. Astronautical Sciences* **41**(4): 439–518.

Slotine, J. J. & Li, W. (1991). *Applied Nonlinear Control*, Prentice Hall, Inc., Englewood Cliff, NJ.

Song, K.-T. & Huang, J.-H. (2001). Fast optical flow estimation and its application to real-time obstacle avoidance, *Proc. IEEE Int. Conf. Robot. Automat.*, pp. 2891–2896.

Wiiesoma, S., Wolfe, D. & Richards, R. (1993). Eye-to-hand coordination for vision-guided robot control applications, *Int. J. Robot. Res.* **12**(1): 65–78.

Zhang, Z. & Hanson, A. R. (1995). Scaled euclidean 3D reconstruction based on externally uncalibrated cameras, *IEEE Symp. Computer Vision*, pp. 37–42.



Robot Localization and Map Building

Edited by Hanafiah Yussof

ISBN 978-953-7619-83-1
Hard cover, 578 pages
Publisher InTech
Published online 01, March, 2010
Published in print edition March, 2010

Localization and mapping are the essence of successful navigation in mobile platform technology. Localization is a fundamental task in order to achieve high levels of autonomy in robot navigation and robustness in vehicle positioning. Robot localization and mapping is commonly related to cartography, combining science, technique and computation to build a trajectory map that reality can be modelled in ways that communicate spatial information effectively. This book describes comprehensive introduction, theories and applications related to localization, positioning and map building in mobile robot and autonomous vehicle platforms. It is organized in twenty seven chapters. Each chapter is rich with different degrees of details and approaches, supported by unique and actual resources that make it possible for readers to explore and learn the up to date knowledge in robot navigation technology. Understanding the theory and principles described in this book requires a multidisciplinary background of robotics, nonlinear system, sensor network, network engineering, computer science, physics, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

S. S. Mehta and W.E. Dixon (2010). A Daisy-Chaining Visual Servoing Approach with Applications in Tracking, Localization, and Mapping, Robot Localization and Map Building, Hanafiah Yussof (Ed.), ISBN: 978-953-7619-83-1, InTech, Available from: http://www.intechopen.com/books/robot-localization-and-map-building/a-daisy-chaining-visual-servoing-approach-with-applications-in-tracking-localization-and-mapping

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen