

# Robust visual tracking control system of a mobile robot based on a dual-Jacobian visual interaction model

Chi-Yi Tsai<sup>a</sup>, Kai-Tai Song<sup>a,\*</sup>, Xavier Dutoit<sup>b</sup>, Hendrik Van Brussel<sup>b</sup>, Marnix Nuttin<sup>b</sup>

<sup>a</sup> Department of Electrical and Control Engineering, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan, ROC

<sup>b</sup> Department of Mechanical Engineering, Division PMA, K. U. Leuven, Celestijnenlaan 300B, B-3001 Leuven, Belgium

## ARTICLE INFO

### Article history:

Received 26 March 2007

Received in revised form

6 January 2009

Accepted 7 January 2009

Available online 21 January 2009

### Keywords:

Visual tracking control

Visual estimation

Visual interaction model

Kalman filter

## ABSTRACT

This paper presents a novel design of a robust visual tracking control system, which consists of a visual tracking controller and a visual state estimator. This system facilitates human–robot interaction of a unicycle-modeled mobile robot equipped with a tilt camera. Based on a novel dual-Jacobian visual interaction model, a robust visual tracking controller is proposed to track a dynamic moving target. The proposed controller not only possesses some degree of robustness against the system model uncertainties, but also tracks the target without its 3D velocity information. The visual state estimator aims to estimate the optimal system state and target image velocity, which is used by the visual tracking controller. To achieve this, a self-tuning Kalman filter is proposed to estimate interesting parameters and to overcome the temporary occlusion problem. Furthermore, because the proposed method is fully working in the image space, the computational complexity and the sensor/camera modeling errors can be reduced. Experimental results validate the effectiveness of the proposed method, in terms of tracking performance, system convergence, and robustness.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

An intelligent robot uses its on-board sensors to collect information from the surroundings and react to the changes of its immediate environment. In recent years, vision systems have been widely used for various intelligent robots, and the research on visual tracking control has gained increasing attention in the area of robotic research [1–25]. In robotics, visual tracking control means vision-based robot motion control to track an interesting target. Based on the motion constraints of the robot, visual tracking control can be classified into two categories: visual servoing for holonomic manipulators and visual tracking for nonholonomic mobile robots. Although visual servoing of holonomic manipulators has been discussed extensively and many results can be found in the literature [1–3], mobile robots are commonly nonholonomic and the visual servoing results for holonomic manipulators are unsuitable for a mobile platform [4].

This paper addresses the problem of visual tracking control of unicycle-modeled or usually termed as wheeled mobile robots equipped with an *on-board* monocular vision system. Due to large number of mobile robot visual tracking control methods, we

classify the reported methods into four groups based on the type of the target to be tracked. Many efforts focus on the first group which aims to track a static target, such as a ground line, landmark, or reference image, for the purpose of mobile robot navigation or regulation (so-called homing) [5–17]. For a mobile robot to track a ground line, Ma et al. formulated the visual tracking control problem as controlling the shape of a ground curve in the image plane and proposed a closed-loop vision-guided control system for a nonholonomic mobile robot [5]. Coulaud et al. proposed a simple and stable feedback controller design, which avoids sophisticated image processing and control algorithms, for a mobile robot equipped with a fixed camera to track a line on the ground [6]. In the case of tracking the landmark, the reported controllers usually modify the visual servoing technique to satisfy the nonholonomic constraint for the motion control of the mobile robot [7–10]. In [11], Zhang and Ostrowski utilized an optimal control method to solve the visual motion-planning problem by generating a virtual trajectory in the image plane and the corresponding optimal control signals for the robot to follow. Nierobisch et al. proposed a visual tracking control method for a mobile robot with a pan-tilt camera to track visual reference landmarks in the acquired views during autonomous navigation [12]. Recently, the homography-based [13,14] and epipolar-based [4,15–17] visual tracking control approaches were proposed for a mobile robot equipped with a pinhole or an omni-directional (so-called central catadioptric) camera to track a reference image toward a desired configuration. These two approaches consider the mobile robot visual tracking

\* Corresponding author. Tel.: +886 3 5731865; fax: +886 3 5715998.

E-mail addresses: [chiyi.ece91g@nctu.edu.tw](mailto:chiyi.ece91g@nctu.edu.tw) (C.-Y. Tsai), [ksong@mail.nctu.edu.tw](mailto:ksong@mail.nctu.edu.tw) (K.-T. Song), [xavier.dutoit@mech.kuleuven.be](mailto:xavier.dutoit@mech.kuleuven.be) (X. Dutoit), [marnix.nuttin@mech.kuleuven.be](mailto:marnix.nuttin@mech.kuleuven.be) (M. Nuttin).

control problem as a visual servoing regulation or visual homing problem. In [13], Chen et al. developed a visual tracking controller based on the Euclidean homography to track a desired time-varying trajectory defined by a prerecorded image sequence of a stationary target viewed by the on-board camera as the mobile robot moves. However, the stability of their result is restricted by the non-zero reference velocity condition of the desired trajectory. To overcome this drawback, Fang et al. exploited Lyapunov-based techniques to construct a homography-based visual servoing regulation controller for proving asymptotic regulation of the mobile robot [14]. In [15], Mariottini et al. exploited the epipolar geometry defined by the current and desired camera views to develop a two-step visual servoing regulation controller. They also extend this design to the visual servoing regulation control of a mobile robot with a central catadioptric camera [16]. In [17], Goedemé et al. developed a vision-only navigation and homing system for mobile robots with an omni-directional camera. Their method divides the visual homing operation into two phases and computes visual homing vector based on epipolar geometry estimation. Although these approaches of the first group provide appropriate solutions for static target visual tracking control problem, they cannot guarantee to solve the moving (non-static) target visual tracking control problem.

The second group aims to track other robot teammates in a robot team for the formation control purpose [18,19]. The proposed approaches in this group usually are based on a central catadioptric camera model in order to detect all robot teammates at the same instant. The subject of the third group is to track a predictable moving target, such as a projectile or straight moving ball, for mobile robot interception purpose [20,21]. In [20], Borgstadt et al. utilized a human vision-based strategy to guide a mobile robot to intercept a projectile ball. Similarly, Capparella et al. extended the concept of human-like strategy to develop a vision-based two-level interception approach, which contains a lower level controller to control the on-board pan-tilt camera and a higher lever controller to operate the mobile robot platform, for intercepting a straight moving ball [21]. A common point of the second and third group is that the motion of the interesting target is known and predictable. However, in many robotic applications, a mobile robot requires to track a dynamic and unpredictable motion target, such as a human's face, for the purpose of pursuit or interaction. Thus, the existent methods of the aforementioned two groups are not suitable to solve the dynamic moving target visual tracking control problem.

The purpose of the fourth group is to solve the problem of tracking a dynamic moving target [22–25]. In [22], Wang et al. proposed an adaptive backstepping control law based on an image-based camera-target visual interaction model to track a dynamic moving target with unknown height parameter. Although the approach in [22] guarantees the asymptotic stability of a closed-loop visual tracking control system in tracking a dynamic moving target, the case of tracking a static target cannot be guaranteed due to the non-zero restrictions on the reference velocity of the mobile robot. In [23], Malis et al. integrated template-based visual tracking algorithms and model-free vision-based control techniques to build a flexible and robust visual tracking control system for various robotic applications. Because their visual tracking result is based on the homography estimation, which requires two images of the target pattern to estimate the optimal homography, the reported system only overcomes the partial occlusion problem but fails in the fully occlusion problem. In [24], Han et al. proposed an image-based visual tracking control scheme for a mobile robot to estimate the position of the target in the next image and track the target to the central area of the image. Since their method utilized the differential approximation method to estimate the velocity of the target in the image plane, the estimation result is very

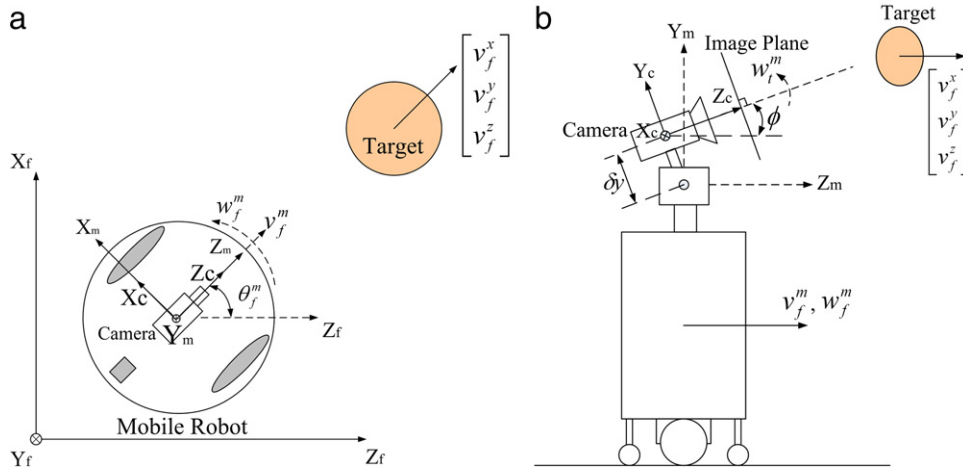
sensitive to image noise. Recently, a visual interaction controller had been proposed for a unicycle-modeled mobile robot to track a dynamic moving target such as a human's face [25]. The drawback of this method is that the controller requires the target's 3D motion velocity, which is difficult to estimate when only a monocular camera is used.

From the literature survey, we note that challenges in the mobile robot visual tracking control design is to develop a robust tracking control system to estimate the motion of the moving target and to track the target based on a stability criterion. This problem motivates us to derive a new model for developing a robust visual tracking control system to solve the tracking problem of dynamic moving target in the image plane directly and efficiently. To do so, the visual interaction model described in [25] is extended to derive a novel dual-Jacobian visual interaction model for designing a robust mobile robot visual tracking control system, which encompasses a visual state estimator and a visual tracking controller. The visual state estimator is constructed by a real-time self-tuning Kalman filter and aims to estimate the optimal system state and target motion in the image plane directly for later use by the visual tracking controller. The visual tracking controller then calculates the robot's control velocities in the image plane directly. The main differences between the proposed method and other existent approaches are summarized as follows:

- (1) The proposed dual-Jacobian visual interaction model considers not only the effect of mobile robot motion, but also the effect of target motion. Thus, based on the proposed model, the visual tracking control problem of a unicycle-modeled mobile robot for tracking a dynamic moving target can be solved with asymptotic convergence using a single controller. Moreover, the proposed model also considers the kinematics of a tilt camera platform mounted on the mobile robot. Therefore, the applicability of the proposed method is greatly increased.
- (2) The proposed visual tracking control system not only possesses some degree of robustness against the system model uncertainties, but also overcomes the unmodeled quantization effect in the velocity commands and the occlusion effect during visual tracking process. This advantage enhances the reliability of the proposed method in practical applications.
- (3) The proposed visual tracking control system works fully in the image space. Therefore, compared with position-based [23], homography-based [13,14], and epipole-based [4, 15–17] visual tracking control approaches, the computational complexity and the sensor/camera modeling errors can be much reduced due to the advantages of image-based visual servo control [2].
- (4) The proposed self-tuning Kalman filter can automatically choose a suitable observation covariance matrix in varying environmental conditions. This helps to improve the estimation performance when there is noise in the system observation.

To validate the performance and robustness of the proposed control system, computer simulation and experimental studies of tracking a moving target have been conducted. Simulation and experimental results will be presented and discussed to verify the effectiveness of the proposed control system, in terms of tracking performance, system convergence, and robustness. Note that a brief version of the research results has been published in [26]. This paper will present the complete design of the proposed tracking control system, including robustness analysis, computer simulation and experimental validation.

The rest of this paper is organized as follows. Section 2 describes the proposed dual-Jacobian visual interaction model. Section 3 presents the results of visual tracking controller design. Section 4 develops the visual state estimator using Kalman filter with self-tuning algorithm. Simulation and experimental results are reported in Section 5. Extended discussion of several interesting observations will be presented. Section 6 concludes the contributions of this paper.



**Fig. 1.** (a) A model of the unicycle-modeled mobile robot and target in the world coordinate frame. (b) Side view of the mobile robot with a tilt camera mounted on top of its head to track a dynamic target.

## 2. Dual-Jacobian visual interaction model

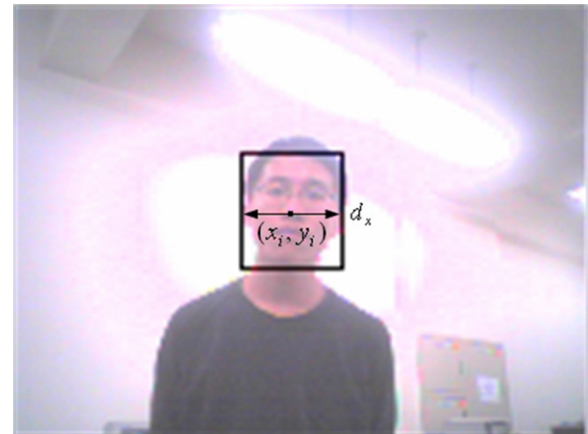
As shown in Fig. 1, the considered system is a unicycle-modeled mobile robot equipped with a tilt camera mounted on top of it to track a moving target, such as a human face, in the image plane. The robot can keep the target in the camera's field of view such that the optical axis of this camera faces the interested target. Fig. 1(a) illustrates the model of the unicycle-modeled mobile robot and target in the world coordinate frame  $F_f$ , in which the motion of the target is supposed to be holonomic. Fig. 1(b) is the side view of the scenario under consideration, in which the tilt angle  $\phi$  gives the relationship between camera coordinate frame  $F_c$  and the mobile coordinate frame  $F_m$ . The kinematics of the unicycle-modeled mobile robot and the target can be described [27], respectively, by

$$\begin{bmatrix} \dot{z}_f^m \\ \dot{x}_f^m \\ \dot{y}_f^m \\ \dot{\theta}_f^m \\ \dot{\phi}_f^m \end{bmatrix} = \begin{bmatrix} \cos \theta_f^m & 0 & 0 \\ \sin \theta_f^m & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_f^m \\ w_f^m \\ w_t^m \end{bmatrix}, \quad (1)$$

$$\dot{y}_f^m = 0 \quad \text{and} \quad \begin{bmatrix} \dot{z}_f^t \\ \dot{x}_f^t \\ \dot{y}_f^t \end{bmatrix} = \begin{bmatrix} v_f^z \\ v_f^x \\ v_f^y \end{bmatrix},$$

where  $(x_f^m, y_f^m, z_f^m)$  and  $(x_f^t, y_f^t, z_f^t)$  are, respectively, the positions of the mobile robot and the target in Cartesian coordinates.  $(\theta_f^m, \phi)$  are the orientation angle of the mobile robot and the tilt angle of the onboard camera.  $(v_f^m, w_f^m)$  and  $w_t^m$  are, respectively, the linear and angular velocity of the mobile robot and the tilt velocity of the camera.  $(v_f^x, v_f^y, v_f^z)$  are the target velocity in Cartesian coordinates.

In order for the mobile robot to interact with the target in the image coordinate frame, a visual interaction model was proposed in authors' previous work [25]. Fig. 2 shows the definition of the observed system state in the image plane used for the visual interaction model. In Fig. 2,  $x_i$  and  $y_i$  are, respectively, the horizontal and vertical position of the centroid of target in the image plane, and  $d_x$  is the width of target in the image plane. Let  $X_i = [x_i \ y_i \ d_x]^T$  denote the system state in the image plane,  $u = [v_f^m \ w_f^m \ w_t^m]^T$  is the control velocity of the mobile robot and on-board tilt camera,  $(f_x, f_y)$  represent fixed focal length along the image  $x$ -axis and  $y$ -axis, respectively [28], and  $W$  stands for the actual width of the target. The visual interaction model between robot and target in the image coordinate frame can be derived as in Box I, where  $\text{diag}(\cdot)$  denotes a 3-by-3 diagonal matrix,  $k_x = d_x/W$  and  $k_y = k_x f_y/f_x$  are two scalars, and  $\delta y$  is the distance between the



**Fig. 2.** Definition of the observed system state in the image plane.

center of robot tilt platform and the onboard camera. A detailed derivation of the visual interaction model (2) given in Box I is presented in Appendix.

The visual interaction model (2) given in Box I indicates that the elements of system matrix  $\mathbf{A}_i$  and vector  $C_i$  are functions of the target velocity. Thus, expression (2) given in Box I can be rewritten as in Box II, and  $V_t = [v_f^x \ v_f^y \ v_f^z]^T$  is the vector of target velocities in Cartesian coordinates. Expression (3) given in Box II shows that the visual interaction model consists of two parts: first, the effect of target motion  $\dot{X}_i^t \equiv [\dot{x}_i^t \ \dot{y}_i^t \ \dot{d}_x^t]^T = \mathbf{J}_t V_t$ , and second, the effect of mobile robot motion  $\dot{X}_i^m \equiv [\dot{x}_i^m \ \dot{y}_i^m \ \dot{d}_x^m]^T = \mathbf{B}_i u$ . Thus, the Eq. (3) given in Box II can be rewritten as a dual-Jacobian equation such that

$$\dot{X}_i = \dot{X}_i^t + \dot{X}_i^m = \mathbf{J}_t V_t + \mathbf{B}_i u, \quad (4)$$

where matrix  $\mathbf{J}_t$ , termed as *target image Jacobian*, transfers the target velocity  $V_t$  into target image velocity  $\dot{X}_i^t$ ; matrix  $\mathbf{B}_i$ , which denotes *robot image Jacobian*, transfers the mobile robot control velocity  $u$  into robot image velocity  $\dot{X}_i^m$ . In other words, the image velocity  $\dot{X}_i$  is caused by the combination of target image velocity  $\dot{X}_i^t$  and robot image velocity  $\dot{X}_i^m$ . Fig. 3 shows the concept of dual-Jacobian equation (4). Therefore, the visual interaction between robot and target in the image coordinate frame can be modeled as a dual-Jacobian visual interaction model (4).

$$\dot{X}_i = \mathbf{A}_i X_i + \mathbf{B}_i u + C_i \quad (2)$$

where  $\mathbf{A}_i = \text{diag}(A_1, A_2, A_1)$ ,  $A_1 = -f_x^{-1} k_x (v_f^x \cos \phi \sin \theta_f^m + v_f^y \sin \phi + v_f^z \cos \phi \cos \theta_f^m)$

$$A_2 = -f_y^{-1} k_y (v_f^x \cos \phi \sin \theta_f^m + v_f^y \sin \phi + v_f^z \cos \phi \cos \theta_f^m)$$

$$\mathbf{B}_i = \begin{bmatrix} f_x^{-1} k_x x_i \cos \phi & f_x^{-1} (x_i^2 + f_x^2) \cos \phi - f_y^{-1} f_x (k_y \delta y + y_i) \sin \phi & -f_y^{-1} x_i (k_y \delta y + y_i) \\ k_y (\sin \phi + f_y^{-1} y_i \cos \phi) & f_x^{-1} f_y x_i (\sin \phi + f_y^{-1} y_i \cos \phi) & -f_y^{-1} (y_i^2 + f_y^2 + k_y y_i \delta y) \\ f_x^{-1} k_x d_x \cos \phi & f_x^{-1} x_i d_x \cos \phi & -f_y^{-1} d_x (k_y \delta y + y_i) \end{bmatrix}$$

$$C_i = \begin{bmatrix} k_x (v_f^z \sin \theta_f^m - v_f^x \cos \theta_f^m) \\ k_y (v_f^y \cos \phi - v_f^x \sin \phi \sin \theta_f^m - v_f^z \sin \phi \cos \theta_f^m) \\ 0 \end{bmatrix}$$

Box I.

$$\dot{X}_i = (\mathbf{A}_i X_i + C_i) + \mathbf{B}_i u = \mathbf{J}_t V_t + \mathbf{B}_i u, \quad (3)$$

where

$$\mathbf{J}_t = \begin{bmatrix} -k_x (f_x^{-1} x_i \cos \phi \sin \theta_f^m + \cos \theta_f^m) & -k_x f_x^{-1} x_i \sin \phi & -k_x (f_x^{-1} x_i \cos \phi \cos \theta_f^m - \sin \theta_f^m) \\ -k_y (f_y^{-1} y_i \cos \phi \sin \theta_f^m + \sin \phi \sin \theta_f^m) & -k_y (f_y^{-1} y_i \sin \phi - \cos \phi) & -k_y (f_y^{-1} y_i \cos \phi \cos \theta_f^m + \sin \phi \cos \theta_f^m) \\ -k_x f_x^{-1} d_x \cos \phi \sin \theta_f^m & -k_x f_x^{-1} d_x \sin \phi & -k_x f_x^{-1} d_x \cos \phi \cos \theta_f^m \end{bmatrix}$$

Box II.

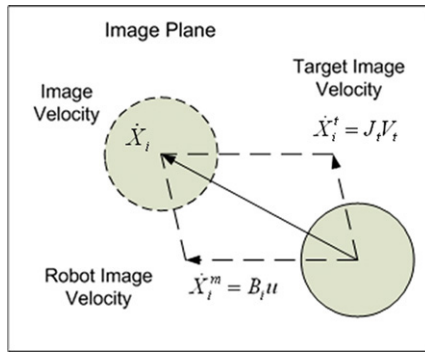


Fig. 3. Depicts the concept of dual-Jacobian visual interaction model (4).

Similar to the human's visual tracking behavior, the purpose of visual tracking control design aims to control the centroid position and width of target from an initial state to the desired state in the image plane. To achieve this, an error state model is introduced to design the tracking controller. Define the error coordinates in the image plane such that

$$X_e = [x_e \quad y_e \quad d_e]^T = \bar{X}_i - X_i^* \quad (5)$$

$$= [\bar{x}_i - x_i^* \quad \bar{y}_i - y_i^* \quad \bar{d}_x - d_x^*]^T,$$

where  $\bar{X}_i = [\bar{x}_i \quad \bar{y}_i \quad \bar{d}_x]^T$  is the desired state in the image plane;  $X_i^* = [x_i^* \quad y_i^* \quad d_x^*]^T$  is the estimated system state from a visual state estimator (VSE, see Section 4). If the system state converges to the desired state, the visual tracking control problem is solved. Furthermore, the dynamic error state model in the image plane can be derived directly by taking the derivative of (5). The result is given by

$$\dot{X}_e = -\dot{X}_i^t - \dot{X}_i^m = -\mathbf{J}_t V_t - \mathbf{B}_i u. \quad (6)$$

With the new coordinate  $X_e$ , the visual tracking control problem is transformed into a stability problem. In the rest of this paper, the new coordinate  $X_e$  will be used to solve the visual tracking control problem. If  $X_e$  converges to zero, then the visual tracking control problem is solved.

### 3. Visual tracking controller (VTC) design

In this section, a visual tracking control law is derived based on the error state model (6) to interact with an interesting target in the image plane by exploiting feedback linearization and pole placement techniques. The robustness analysis of the proposed controller against parametric uncertainties is also presented.

#### 3.1. Feedback linearization and pole placement

According to the dynamic error state model (6), a feedback control law can be obtained by feedback linearization such that

$$u = \mathbf{B}_i^{-1} (\mathbf{K}_g X_e - \mathbf{J}_t V_t) = \mathbf{B}_i^{-1} (\mathbf{K}_g X_e - \dot{X}_i^t), \quad (7)$$

where  $\mathbf{K}_g$  is a 3-by-3 gain matrix. Substituting (7) into (6) yields

$$\dot{X}_e = -\mathbf{K}_g X_e. \quad (8)$$

It is clear that if all eigen values of matrix  $-\mathbf{K}_g$  are constant and lie inside the left-half complex plane, then the system error state  $X_e(t)$  will decay exponentially to zero. Thus, we choose the gain matrix by pole placement such that

$$\mathbf{K}_g = \text{diag}(\alpha_1, \alpha_2, \alpha_3), \quad (9)$$

in which  $(\alpha_1, \alpha_2, \alpha_3)$  are three positive constants. Substituting (9) into (8) yields

$$\dot{X}_e = -\mathbf{K}_g X_e = -\text{diag}(\alpha_1, \alpha_2, \alpha_3) X_e. \quad (10)$$

Suppose that the initial error state  $X_e(t_0)$  is within the image plane. Then expression (10) indicates that

$$X_e(t)|_{t_0, X_e(t_0)} \equiv X_e(t; t_0, X_e(t_0)) \\ = \text{diag}(e^{-\alpha_1(t-t_0)}, e^{-\alpha_2(t-t_0)}, e^{-\alpha_3(t-t_0)}) X_e(t_0) \quad (11)$$

for some  $t_0 \geq 0$ .

Because  $(\alpha_1, \alpha_2, \alpha_3)$  are three positive constants, expression (11) leads to the following inequality:

$$\|X_e(t; t_0, X_e(t_0))\| \leq e^{-\lambda_{\min}(\mathbf{K}_g)(t-t_0)} \|X_e(t_0)\| \quad \text{for all } t \geq t_0, \quad (12)$$

where  $\lambda_{\min}(\mathbf{A})$  denotes the minimum eigenvalue of matrix  $\mathbf{A}$ , and  $\|B\|$  denotes the 2-norm value of vector  $B$ . From (12), it is clear that the system error state satisfies  $\|X_e(t; t_0, X_e(t_0))\| \leq \|X_e(t_0)\|$



and  $\lim_{t \rightarrow \infty} \|X_e(t; t_0, X_e(t_0))\| = 0$ , and thus the visual tracking control problem is solved. Summarizing the above discussions, we obtain the following proposition.

**Proposition 1.** Suppose that the initial system state  $X_i$  is within the image plane. Let  $(\alpha_1, \alpha_2, \alpha_3) > 0$  be three positive constants. Consider the dual-Jacobian visual interaction system (4). If the matrix  $\mathbf{B}_i$  is nonsingular, then the closed-loop visual tracking system (6) is asymptotically stable by using the control law

$$u = \mathbf{B}_i^{-1}(\mathbf{K}_g X_e - \dot{X}_i^t), \quad (13)$$

where  $\dot{X}_i^t = \mathbf{J}_i V_t$  is the target image velocity defined in (4), matrices  $\mathbf{B}_i$  and  $\mathbf{J}_i$  are the robot and target image Jacobian defined in (2) and (3), respectively,  $X_e$  is the system error state defined in (5), and  $\mathbf{K}_g$  is a 3-by-3 diagonal gain matrix defined in (9).

**Proof.** Consider the closed-loop visual tracking system (6). We first define a positive-definite Lyapunov function associated with the system error state

$$V(x_e, y_e, d_e) = \frac{1}{2}(x_e^2 + y_e^2 + d_e^2). \quad (14)$$

Taking the derivative of (14) yields

$$\dot{V} = X_e^T \dot{X}_e = -(X_e^T \mathbf{J}_i V_t + X_e^T \mathbf{B}_i u) = -X_e^T (\dot{X}_i^t + \mathbf{B}_i u) \equiv -f(u), \quad (15)$$

where  $f(u) = X_e^T (\dot{X}_i^t + \mathbf{B}_i u)$ . In view of Lyapunov theory [29], expression (15) tells us that if  $f(u) > 0$  then the equilibrium point of (6) is asymptotically stable. Substituting the control law (13) into  $f(u)$ , we then have

$$f(u) = X_e^T \mathbf{K}_g X_e, \quad (16)$$

where  $\mathbf{K}_g = \text{diag}(\alpha_1, \alpha_2, \alpha_3)$ , and  $(\alpha_1, \alpha_2, \alpha_3)$  are three positive constants. Since  $\mathbf{K}_g$  is a symmetric positive definite (SPD) matrix, the following inequality holds:

$$f(u) \geq \lambda_{\min}(\mathbf{K}_g) \|X_e\|^2 = \min(\alpha_1, \alpha_2, \alpha_3) \|X_e\|^2 > 0, \quad (17)$$

where  $\lambda_{\min}(\mathbf{A})$  denotes the minimum eigenvalue of matrix  $\mathbf{A}$ . Expression (17) concludes that the closed-loop visual tracking system (6) is asymptotically stable and hence completes the proof. ■

### 3.2. Singularity analysis

The feedback linearization control law (13) poses a singularity problem of matrix  $\mathbf{B}_i$ . In [25], the singularity condition of matrix  $\mathbf{B}_i$  is derived such that

$$f_y = (y_i + S d_x) \tan \phi, \quad (18)$$

where  $S = (f_y \delta y) / (f_x W)$  is a fixed scalar factor. Let  $(x_c, y_c, z_c)$  represents the related position between robot and target in the camera coordinate frame. Because of  $f_y = k_y z_c$ ,  $y_i = k_y y_c$ , and  $d_x = k_x W$  (see (A.2) in Appendix), singularity condition (18) can be rewritten such that

$$z_c = (y_c + \delta y) \tan \phi, \quad (19)$$

where  $z_c$  is the distance from the camera to the target. Fig. 4 shows the particular configuration of the singular condition (19). In Fig. 4, the distance  $z_\phi$  equals:

$$z_\phi = (y_c + \delta y) \tan \phi. \quad (20)$$

From (19) and (20), it is clear that the physical meaning of the singularity condition (19) is that the distance between the camera and the target equals to the distance  $z_\phi$ . In general, this situation usually will not happen unless the target directly locates on  $Y_m$ -axis (the  $Y$ -axis of the mobile coordinate frame  $F_m$ ), which means that the target is directly above or directly below the robot. Under such circumstances, the robot will be unable to approach the target in any way due to deficient degree-of-freedom, and the robot should stop tracking temporarily.

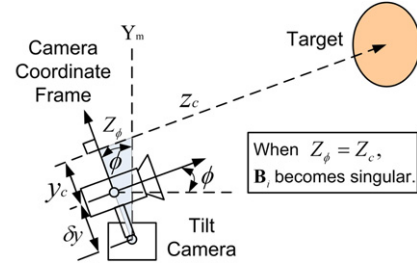


Fig. 4. Particular configuration of the singularity condition (10).

**Remark 1.** The proposed control law (13) is associated with the target's velocity in the working space or target's image velocity in the image plane. If the 3D velocity of the target  $V_t$  is known, the target image velocity  $\dot{X}_i^t$  can be calculated by using target image Jacobian  $\mathbf{J}_i$  without the estimation process. However, in practical applications, it is difficult to measure the 3D velocity of the target when using only one camera in real-time operations. In this situation, an estimation process is required to estimate the target image velocity  $\dot{X}_i^t$  in the image plane directly. In Section 4, a VSE will be proposed to accomplish this task. This design will facilitate more general performance of the proposed tracking control scheme in the image plane.

### 3.3. Robustness analysis

In this subsection, we investigate the robustness of the proposed VTC (13) against model uncertainties on camera parameters  $(f_x, f_y)$ , robot parameters  $(\theta_f^m, \phi)$ , and target parameters  $(W, \dot{x}_i^t, \dot{y}_i^t, \dot{d}_x^t)$ . Consider the following closed-loop visual tracking system with parametric uncertainties:

$$\dot{X}_e = -\dot{X}_i^t - \bar{\mathbf{B}}_i u = -(\dot{X}_i^t + \delta \dot{X}_i^t) - (\mathbf{B}_i + \delta \mathbf{B}_i) u, \quad (21)$$

where  $\delta \dot{X}_i^t$  and  $\delta \mathbf{B}_i$  are unknown bounded disturbances. Recall the positive-definite Lyapunov function defined in (14), the derivative of (14) with parametric uncertainties becomes

$$\dot{V} = -X_e^T (\dot{X}_i^t + \bar{\mathbf{B}}_i u) = -[f(u) + \delta f(u)] \equiv -\bar{f}(u), \quad (22)$$

where  $\delta f(u) = X_e^T (\delta \dot{X}_i^t + \delta \mathbf{B}_i u)$  is unknown. Assume that  $\delta f(u)$  is bounded and there exists a SPD matrix  $\bar{\mathbf{Q}}$  such that

$$\delta f(u) < \|X_e\|_{\bar{\mathbf{Q}}}^2, \quad (23)$$

where  $\|X\|_{\mathbf{A}} = (X^T \mathbf{A} X)^{1/2}$  denotes a weighted vector norm with a SPD matrix  $\mathbf{A}$ . Now, the main result is presented as follows.

**Theorem 1.** Consider the dual-Jacobian visual interaction system (4) with unknown bounded parametric uncertainties  $\delta \dot{X}_i^t$  and  $\delta \mathbf{B}_i$  defined in (21). Let  $\bar{\mathbf{Q}} > 0$  be a SPD matrix defined in (23). Choose the controller  $u$  as given in the expression (13) with a constant SPD matrix  $\mathbf{K}_g = \text{diag}(\alpha_1, \alpha_2, \alpha_3) > 0$ . Then, the closed-loop visual tracking system (6) is asymptotically stable for all  $\alpha_i > \lambda_{\max}(\bar{\mathbf{Q}})$ ,  $i = 1, 2, 3$ .

**Proof.** From (22) and (23), it follows that

$$f(u) = \bar{f}(u) - \delta f(u) > \bar{f}(u) - \|X_e\|_{\bar{\mathbf{Q}}}^2. \quad (24)$$

Expression (24) implies that  $\bar{f}(u) - \|X_e\|_{\bar{\mathbf{Q}}}^2$  is a lower-bound of  $f(u)$ . If  $\bar{f}(u) - \|X_e\|_{\bar{\mathbf{Q}}}^2 > 0$  can be guaranteed, then  $f(u) > 0$  is satisfied and thus the system has the robust property w.r.t. the parametric uncertainties.

Choose the controller  $u$  as in (13) with parametric uncertainties such that

$$u = \bar{\mathbf{B}}_i^{-1}(\mathbf{K}_g X_e - \dot{X}_i^t), \quad (25)$$

where  $\dot{X}_i^t = \bar{J}_i V_i$ . Substituting (25) into (22) yields

$$\dot{V} = -\bar{f}(u) = -\|X_e\|_{\mathbf{K}_g}^2, \quad (26)$$

where  $\mathbf{K}_g = \text{diag}(\alpha_1, \alpha_2, \alpha_3) > 0$  is a constant SPD matrix. From (24) and (26), it is clear that

$$f(u) > \|X_e\|_{\mathbf{K}_g}^2 - \|X_e\|_{\mathbf{Q}}^2 \geq [\lambda_{\min}(\mathbf{K}_g) - \lambda_{\max}(\mathbf{Q})] \|X_e\|^2, \quad (27)$$

where  $\lambda_{\min}(\mathbf{A})$  is defined in (17), and  $\lambda_{\max}(\mathbf{A})$  denotes the maximum eigenvalue of matrix  $\mathbf{A}$ . Expression (27) tells us that if  $\lambda_{\min}(\mathbf{K}_g) - \lambda_{\max}(\mathbf{Q})$  is positive, then  $f(u) > 0$  is satisfied and thus the equilibrium point of (6) is asymptotically stable. This means that the proposed VTC (13) is robust against the unknown parametric uncertainties and hence completes the proof. ■

**Remark 2.** In realization of the control schemes, it is worth noting that the quantization error in velocity commands degrade the performance of the controller and might make the system unstable. In order to overcome this problem, the proposed VTC (13) will be combined with a robust control law presented in [25] for improving the robustness of the visual tracking control system. Interested readers refer to [25] for more technical details.

#### 4. Visual state estimator (VSE) design

As implied by Proposition 1, the VTC (13) requires the information of target image velocity  $\dot{X}_i^t$ . This requirement poses two questions: first, how the target image velocity can be estimated in the image plane directly; second, what estimation methods can be used. In this section, we first formulate the problem of target image velocity estimation in the image space. A VSE will then be proposed in order to compute the optimal estimates of target status  $X_i$  and target image velocity  $\dot{X}_i^t$  in the weighted least squared error sense for later used by the VTC.

##### 4.1. Problem formulation

Since actual image processing is discrete, the first step of VSE design is to discretize the system model (4) into a corresponding discrete form. By the definition  $\dot{x}(t) = \lim_{T \rightarrow 0} [x(t) - x(t - T)]/T$ ,  $T$  denotes the sampling time of the digital system, one can approximate the system model (4) as

$$X_i[n] = X_i[n-1] + T\dot{X}_i^t[n-1] + T\mathbf{B}_i u_{n-1}, \quad (28)$$

for  $n = 1, 2, \dots$

where  $u_n = [v_f^m \quad w_f^m \quad w_t^m]^T$  is the discrete-time control signal at sample instant  $n$ . Suppose that the target motion can be approximated by a smooth motion during a sampling period. Then the target image velocity has the following relationship between two consecutive sampling instants

$$\dot{X}_i^t[n] = \dot{X}_i^t[n-1]. \quad (29)$$

Based on (28) and (29), the propagation model of the visual state estimator is given by

$$X_n = \begin{bmatrix} \mathbf{I}_3 & T\mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} X_{n-1} + \begin{bmatrix} T\mathbf{B}_i \\ \mathbf{0}_3 \end{bmatrix} u_{n-1} \equiv \mathbf{A}_{\text{est}} X_{n-1} + \mathbf{B}_{\text{est}} u_{n-1}, \quad (30)$$

where  $X_n^T = [(X_i[n])^T \quad (\dot{X}_i^t[n])^T]^T$  is the vector of system estimates at instant  $n$ ,  $\mathbf{I}_3$  is a 3-by-3 identity matrix, and  $\mathbf{0}_3$  is a 3-by-3 zero matrix. Next, because the observed image only contains the information of target status  $X_i$  in each instant, the observation model of VSE is given by

$$Z_n = [\mathbf{I}_3 \quad \mathbf{0}_3] X_n \equiv \mathbf{H}_{\text{est}} X_n. \quad (31)$$

Based on the propagation model (30) and observation model (31), the estimation problem in the image plane becomes to find the state estimate  $X_n^*$  that minimizes the weighted least square

distance:

$$X_n^* = \arg \min_X [(X_n - X)^T \mathbf{P}_n^{-1} (X_n - X) + (Z_n - \mathbf{H}_{\text{est}} X)^T \mathbf{R}_n^{-1} (Z_n - \mathbf{H}_{\text{est}} X)], \quad (32)$$

where  $\mathbf{P}_n = \mathbf{A}_{\text{est}} \mathbf{P}_{n-1} \mathbf{A}_{\text{est}}^T$  is the covariance matrix of propagation model at instant  $n$ , and  $\mathbf{R}_n$  is the covariance matrix of observation model at instant  $n$ .

##### 4.2. Self-tuning Kalman filter

Because the propagation model (30) and observation model (31) are both linear equations, a Kalman filter will provide the optimal estimate  $X_n^*$  based on the performance criterion (32) when (30) and (31) have Gaussian uncertainties [30]:

$$\text{Propagation: } X_n = \mathbf{A}_{\text{est}} X_{n-1}^* + \mathbf{B}_{\text{est}} u_{n-1} + \delta X_{n-1} \quad (33a)$$

$$\text{Propagation Covariance Matrix: } \mathbf{P}_n = \mathbf{A}_{\text{est}} \mathbf{P}_{n-1}^* \mathbf{A}_{\text{est}}^T + \mathbf{Q}_{n-1} \quad (33b)$$

$$\text{Observation: } Z_n = \mathbf{H}_{\text{est}} X_n + \delta Z_n \quad (33c)$$

where  $(X_n, \mathbf{P}_n)$  are the propagation state and the corresponding covariance matrix at instant  $n$ ;  $(X_{n-1}^*, \mathbf{P}_{n-1}^*)$  are the optimal estimate and the corresponding covariance matrix at instant  $n-1$ ;  $\delta X_n^T = [(\delta X_i[n])^T \quad (\delta \dot{X}_i^t[n])^T]^T \sim N(0, \mathbf{Q}_n)$  represents Gaussian propagation uncertainty with zero mean and covariance matrix  $\mathbf{Q}_n$  at instant  $n$ ; and  $\delta Z_n \sim N(0, \mathbf{R}_n)$  denotes Gaussian observation uncertainty with zero mean and covariance matrix  $\mathbf{R}_n$  at instant  $n$ . Based on Eqs. (33a)–(33c), the local minimum solution of performance criterion (32) and the corresponding covariance matrix at instant  $n$  are given by

$$X_n^* = X_n^p + \mathbf{K}_n (Z_n - \mathbf{H}_{\text{est}} X_n^p) \quad \text{and} \quad \mathbf{P}_n^* = (\mathbf{I}_6 - \mathbf{K}_n \mathbf{H}_{\text{est}}) \mathbf{P}_n, \quad (34)$$

where  $X_n^p = \mathbf{A}_{\text{est}} X_{n-1}^* + \mathbf{B}_{\text{est}} u_{n-1}$  is the ideal propagation state,  $\mathbf{K}_n = \mathbf{P}_n \mathbf{H}_{\text{est}}^T (\mathbf{H}_{\text{est}} \mathbf{P}_n \mathbf{H}_{\text{est}}^T + \mathbf{R}_n)^{-1}$  is the Kalman gain matrix, and  $\mathbf{I}_6$  is a 6-by-6 identity matrix.

Although expression (34) provides the best linear estimates at each instant, the filter performance still depends on the covariance matrices  $\mathbf{Q}_n$  and  $\mathbf{R}_n$ . Thus, a difficult problem in Kalman filter applications is to determine the values of matrices  $\mathbf{Q}_n$  and  $\mathbf{R}_n$  for computing Kalman gain matrix  $\mathbf{K}_n$  [31]. Moreover, the observation uncertainty usually varies with the conditions of target motion (such as orientation and rotation of the human face) and working environment (such as light variation and occlusion); the corresponding covariance matrix  $\mathbf{R}_n$  would be time-varying for various operating conditions. These problems motivate us to combine a self-tuning algorithm with a Kalman filter to choose a suitable observation covariance matrix  $\mathbf{R}_n$  in varying environmental conditions. On the other hand, because the propagation uncertainty and the corresponding covariance matrix  $\mathbf{Q}_n$  are difficult to estimate online, the propagation covariance matrix  $\mathbf{Q}_n$  will be fixed at initialization without updating in this design.

The proposed self-tuning algorithm attempts to estimate the minimum variance of a set of observation data recorded over time. To do so, a linear-least-squares regression method is adopted to analyze the observed time series data [32]. The typical linear regression model for a discrete time series is given by

$$y_n = an + b + \varepsilon_n, \quad (35)$$

where the residual  $\varepsilon_n$  is a random variable with zero mean,  $(a, b)$  are the parameters to be determined by minimizing the variance of residuals. Fig. 5 shows the concept of the linear-least-squares regression, in which the solid line is the observed time series, and the dotted line indicates the best linear fitting  $\hat{y}_n = an + b$  with minimum residual variance  $\sigma^2$ .

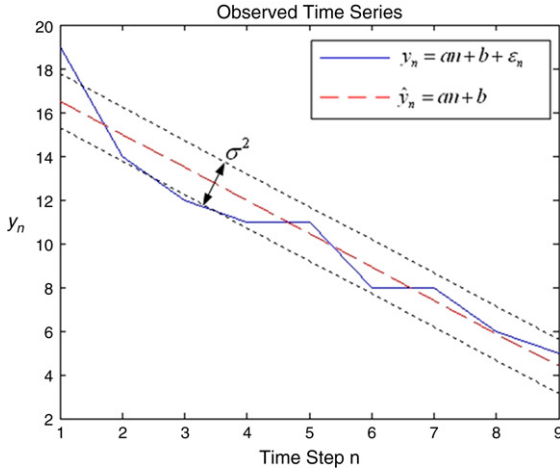


Fig. 5. Concept of time series linear-least-squares regression.

Let  $k$  denote the length of the observed time series. Based on the linear regression model (35), the observed time series can be modeled as

$$Y = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ k & 1 \end{bmatrix} \theta + \varepsilon \equiv \mathbf{A}_{st} \theta + \varepsilon, \quad (36)$$

where  $Y = [y_1 \ y_2 \ \dots \ y_k]^T$  is the vector of observed data over time, and  $\varepsilon = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_k]^T$  represents the corresponding residuals.  $\theta = [a \ b]^T$  is the parameter vector to be detected such that

$$\theta^* = \min_{\theta} \text{var}(\varepsilon) = \min_{\theta} \|\varepsilon\| = \min_{\theta} \|Y - \mathbf{A}_{st} \theta\|, \quad (37)$$

where  $\text{var}(x)$  is the variance of vector  $x$ , and  $\|x\|$  is the norm of vector  $x$ . The optimal solution of (37) will be the least-squares solution such that

$$\theta^* = \mathbf{A}_{st}^+ Y, \quad (38)$$

where  $\mathbf{A}_{st}^+ = (\mathbf{A}_{st}^T \mathbf{A}_{st})^{-1} \mathbf{A}_{st}^T$  denotes the pseudo-inverse matrix of  $\mathbf{A}_{st}$ . Substituting (38) into (36), the residual vector with minimum variance can be obtained by

$$\varepsilon^* = Y - \mathbf{A}_{st} \mathbf{A}_{st}^+ Y = (\mathbf{I}_k - \mathbf{A}_{st} (\mathbf{A}_{st}^T \mathbf{A}_{st})^{-1} \mathbf{A}_{st}^T) Y \equiv \mathbf{T}_{st} Y, \quad (39)$$

where  $\mathbf{T}_{st} = \mathbf{I}_k - \mathbf{A}_{st} (\mathbf{A}_{st}^T \mathbf{A}_{st})^{-1} \mathbf{A}_{st}^T$  is a fixed  $k$ -by- $k$  coefficient matrix, and  $\mathbf{I}_k$  is a  $k$ -by- $k$  identity matrix. Expression (39) tells us that the minimum variance residual vector  $\varepsilon^*$  is the linear transformation of observed data vector  $Y$  through a fixed transformation matrix  $\mathbf{T}_{st}$ . This observation provides us an efficient method for detecting the minimum variance of an observed data sequence in real-time. For instance, let  $X_1^k$ ,  $Y_1^k$ , and  $D_1^k$  denote, respectively, the observed data sequence of  $x_i$ ,  $y_i$ , and  $d_x$  over time steps 1 to  $k$ . Using (39), the minimum variances of  $X_1^k$ ,  $Y_1^k$ , and  $D_1^k$  are given by

$$\sigma_x^2 = \text{var}(\mathbf{T}_{st} X_1^k), \quad \sigma_y^2 = \text{var}(\mathbf{T}_{st} Y_1^k), \quad \text{and} \quad (40)$$

$$\sigma_d^2 = \text{var}(\mathbf{T}_{st} D_1^k).$$

Based on (40), the covariance matrix  $\mathbf{R}_n$  is updated as

$$\mathbf{R}_n = \mathbf{R}_0 + \text{diag}((\sigma_x^2)^2, (\sigma_y^2)^2, (\sigma_d^2)^2), \quad (41)$$

where  $\mathbf{R}_0$  is the initial covariance matrix of  $\mathbf{R}_n$ . Combining the self-tuning equations (40)–(41) with Kalman filter equations (33)–(34), the implemented self-tuning Kalman filter is summarized in Fig. 6. The processing steps are listed as follows:

- (1) Choose two initial covariance matrices  $\mathbf{Q}_0$  and  $\mathbf{R}_0$ , usually by a trial-and-error procedure.
- (2) Assume that the initial position of target locates in the field-of-view of the camera, then initialize the estimated system state  $X_0^*$  and propagation covariance matrix  $\mathbf{P}_0$  by the first observation such that  $X_0^* = [Z_0^T \ 0 \ 0 \ 0]^T$  and  $\mathbf{P}_0 = \mathbf{I}_6$ .
- (3) Store current observed measurement in a shift register with length  $k$ . If the length of storage data is equal to  $k$ , then compute the variance of the observed data sequences by (40) and update covariance matrix  $\mathbf{R}_n$  by (41); else set  $\mathbf{R}_n = \mathbf{R}_0$ ; go to step (4).
- (4) Compute the ideal propagated state  $X_n^p$  defined in (32) and the corresponding propagation covariance matrix  $\mathbf{P}_n$  using (33b).
- (5) If the target is detected in the observed image, then compute the Kalman gain matrix  $\mathbf{K}_n$  and update the estimated state vector  $X_n^*$  with the corresponding covariance matrix  $\mathbf{P}_n^*$  using (34); else set  $X_n^* = X_n^p$  and  $\mathbf{P}_n^* = \mathbf{P}_n$ ; go to step (6).
- (6) Let  $X_{n-1}^* = X_n^*$ ,  $\mathbf{P}_{n-1}^* = \mathbf{P}_n^*$  and  $\mathbf{Q}_{n-1} = \mathbf{Q}_0$ ; go to step (3). ■

## 5. Simulation and experimental results

Computer simulations and several interesting experiments have been performed to validate the robustness and tracking performance of the proposed control system. In the computer

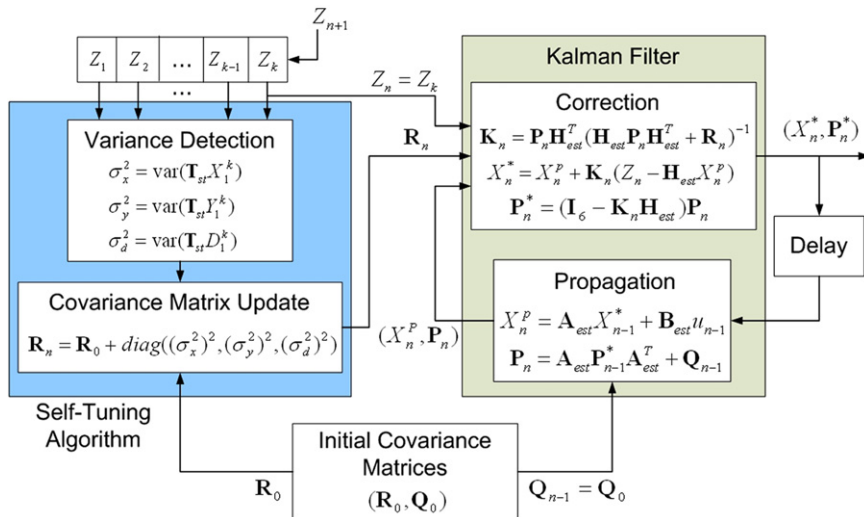
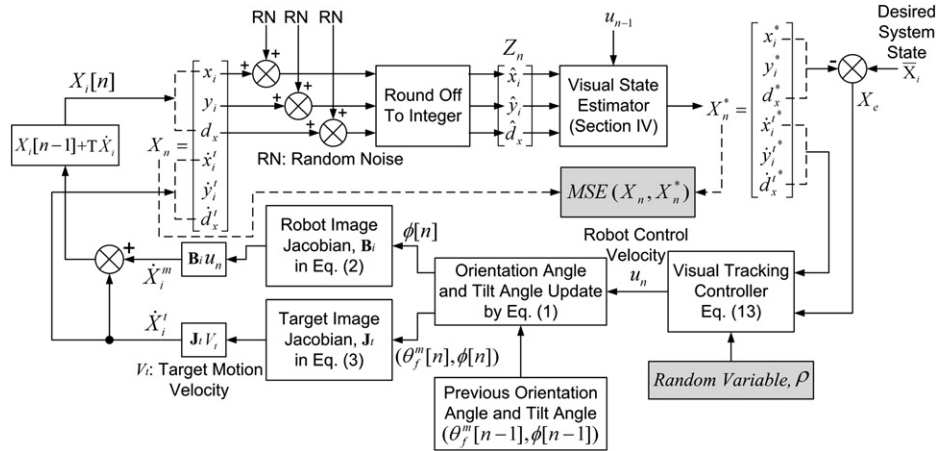


Fig. 6. Architecture of the proposed self-tuning Kalman filter.

**Table 1**  
Parameters used in the simulations and experiments.

Symbol	Quantity	Description
$(f_x, f_y)$	(393.4, 391.8) pixels	Camera focal length in retinal coordinates [28].
$W$	12 cm	Width of the target.
$D$	40 cm	Distance between two drive wheels.
$\delta y$	10 cm	Distance between the center of robot tilt platform and the onboard camera
$T$	100 ms	Sampling period of the control system.
$\mathbf{Q}_0$	diag(5, 5, 5, 20, 20, 20)	Initial propagation covariance matrix
$\mathbf{R}_0$	diag(5, 5, 5)	Initial observation covariance matrix
$(\bar{x}_i, \bar{y}_i, \bar{d}_x)$	(0, 0, 35)	Desired system state in the image plane.
$(\alpha_1, \alpha_2, \alpha_3)_s$	(1/2, 3/4, 1/3)	Positive control gains of simulation.
$(\alpha_1, \alpha_2, \alpha_3)_1$	(5/4, 3, 1/2)	Positive control gains of experiment.
$(z_f^m, x_f^m, \theta_f^m, \phi) _{t=0}$	(0, 0, 0, 0)	Initial pose of tracking robot.



**Fig. 7.** Simulation setup for the robustness and performance evaluation of the visual tracking control system.

simulation, MATLAB was used to verify the robustness of the proposed visual tracking control system against the parametric uncertainties. An experiment was performed on an experimental mobile robot to validate the tracking performance and robustness against the occlusion uncertainty. The experiment adopts the proposed control system to control a mobile platform with the tilt function of the camera platform. Table 1 shows the parameters used in the simulations and the experiment. Note that the processing time of the proposed visual tracking control system is less than 50 ms including face detection, estimation and control computations. This means the overall tracking system is of low computation load and can track the user's face in real time. However, the sampling period of the control system  $T$  was set to 100 ms in the experiments due to other image processing computations such as image compression and storage.

### 5.1. Simulation setup

Fig. 7 shows the simulation setup for the evaluation of system robustness and tracking performance. In Fig. 7,  $X_n$  denotes the ideal state needed to be estimated by the VSE presented in Section 4.  $X_i[n]$  is the ideal system state at time instance  $n$ , and  $\dot{X}_i^t$  is the ideal target image velocity at time instance  $n$ . The observation signal  $Z_n$  is obtained by the rounding off the value of  $X_i[n]$  with random noise (RN) to an integer. The random noise used in this paper is given by  $RN = \sigma_n(2\omega - 1)$ , (42)

where  $\sigma_n \in [0, 10]$  is a constant noise gain, and  $\omega \in [0, 1]$  is a random noise signal with uniform distribution. Next, the VSE aims to filter RN and provide the optimal estimates. The performance of the VSE is then validated by mean-squared-error (MSE) criterion between the ideal signal  $X_n$  and the estimated signal  $X_n^*$ .

In order to validate the robustness of the VTC (13) against the parametric uncertainty, a random variable is utilized to control the

variation of the system parameters  $(f_x, f_y, W, \delta y)$  such that

$$\begin{aligned} \bar{f}_x &= (1 + \rho)f_x, & \bar{f}_y &= (1 + \rho)f_y, & \bar{W} &= (1 + \rho)W, \\ \text{and } \bar{\delta y} &= (1 + \rho)\delta y, \end{aligned} \quad (43)$$

where  $\rho \in [-0.1, 0.1]$  is a random variable with uniform distribution introduced in the practical system parameters  $(\bar{f}_x, \bar{f}_y, \bar{W}, \bar{\delta y})$  that will be used to calculate the practical robot image Jacobian  $\bar{\mathbf{B}}_i$ . In the simulation, the motion of the target is set as a dynamic and smooth motion with velocity

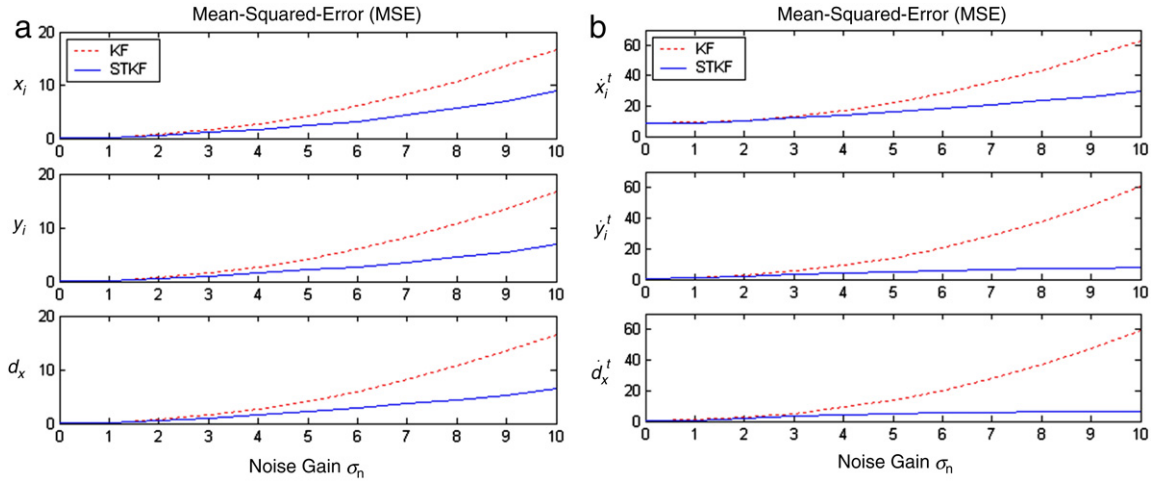
$$(v_f^x, v_f^y, v_f^z) = (v_f^t \sin \theta_f^t, 0, v_f^t \cos \theta_f^t), \quad (44)$$

where  $v_f^t(n) = v_f^t(0) \cos(n\pi T/20)$  (cm/s) for  $v_f^t(0) = 15$  and  $\theta_f^t(n) = \theta_f^t(n-1) + T(5\pi/72)$  rad for  $\theta_f^t(0) = 0$ . Expression (44) will be used to calculate the ideal target image velocity  $\dot{X}_i^t$ .

### 5.2. Simulation results of the proposed self-tuning Kalman filter

Two visual state estimators are used to compare the performance: the conventional Kalman filter (KF) and the proposed self-tuning Kalman filter (STKF). Fig. 8 shows the evolution of the average MSE measurements as the noise gain  $\sigma_n$  increased in the simulations. Note that each average MSE measurement is out of 20 simulations for each  $\sigma_n$ , and the value of the random variable  $\rho$  is randomly chosen in the beginning of each simulation. In Fig. 8, it is clear that the estimation results of the KF are very sensitive to the intensity of the observation noise. As the noise gain  $\sigma_n$  increased from 0 to 10 with interval 1, the MSE measurements of the KF increase much faster than those of STKF. Further, when the noise gain  $\sigma_n = 10$  (the observation signal has the largest noise intensity), the proposed STKF provides improved estimation results compared with the KF.





**Fig. 8.** MSE measurements of the simulation results using Kalman filter and self-tuning Kalman filter. (a) Average MSE measurements of system state ( $x_i, y_i, d_x$ ). (b) Average MSE measurements of target image velocity ( $\dot{x}_i^t, \dot{y}_i^t, \dot{d}_x^t$ ).

**Table 2**

Average MSE measurements of computer simulations.

MSE	Value	KF	STKF	MSE	Value	KF	STKF
$x_i$	$\sigma_n = 10$	16.7670	<b>8.8669</b>	$\dot{x}_i^t$	$\sigma_n = 10$	62.6014	<b>29.8525</b>
	$\sigma_n = 0$	0.0442	<b>0.0441</b>		$\sigma_n = 0$	<b>8.3137</b>	8.5533
	MSE gap	16.7228	<b>8.8228</b>		MSE gap	54.2878	<b>21.2992</b>
$y_i$	$\sigma_n = 10$	16.6215	<b>6.9174</b>	$\dot{y}_i^t$	$\sigma_n = 10$	60.9062	<b>8.0422</b>
	$\sigma_n = 0$	<b>0.0441</b>	0.0442		$\sigma_n = 0$	<b>0.4667</b>	0.5341
	MSE gap	16.5774	<b>6.8732</b>		MSE gap	60.4395	<b>7.5081</b>
$d_x$	$\sigma_n = 10$	16.5441	<b>6.3912</b>	$\dot{d}_x^t$	$\sigma_n = 10$	59.0913	<b>6.6815</b>
	$\sigma_n = 0$	0.0981	<b>0.0964</b>		$\sigma_n = 0$	<b>0.2977</b>	0.3132
	MSE gap	16.4460	<b>6.2948</b>		MSE gap	58.7935	<b>6.3683</b>

Table 2 records the MSE gap between  $\sigma_n = 10$  and  $\sigma_n = 0$ . A small MSE gap implies a large robustness against the intensity of observation noise. In Table 2, the bold font denotes the smallest value of MSE measurement across each row. Table 2 shows that the MSE gaps of the KF for all estimates are larger than that of STKF. In other words, the proposed STKF provides high robustness against the observation uncertainty compared with the KF. Therefore, the simulation results validate the robustness of the proposed STKF.

**Remark 3.** Because the effect of estimation error is not considered in the current control design, a large estimation error will degrade the tracking performance of the visual tracking control system. In other words, a robust VSE that provides a small estimation error is helpful in improving the tracking performance. The results in Table 2 show that the proposed STKF provides better estimation results than the conventional KF does. Therefore, the proposed STKF is more helpful to the proposed VTC when the observation is perturbed with random noise.

### 5.3. Simulation results of robustness to the system parametric uncertainty

Fig. 9 presents the computer simulation results of the visual tracking control system with random noise defined in (42) and parametric uncertainty defined in (43). The results shown in Fig. 9 are obtained from the average of 220 simulations (20 simulations for each  $\sigma_n$  and one random value  $\rho$  for each simulation). Fig. 9(a) and (b) show, respectively, the tracking errors and the estimates of target image velocity. In Fig. 9(a) and (b), the dotted lines illustrate the ideal values while the solid lines show the estimation results. It is clear that all tracking errors converge to zero, and each estimate converges to the corresponding ideal value. Fig. 9(c) shows the control velocities of the mobile robot and the tilt camera.

Fig. 9(d) shows the transition of  $f(u)$  defined in (15) in order to clarify the stability of the closed-loop visual tracking control system. In the simulation, the uncertainty function  $\delta f(u)$  defined in (22) is computable since  $\delta \dot{X}_i^t$  and  $\delta \mathbf{B}_i$  are known. Thus, according to (24), a low-bound (LB) of  $f(u)$  can be obtained by the following equation:

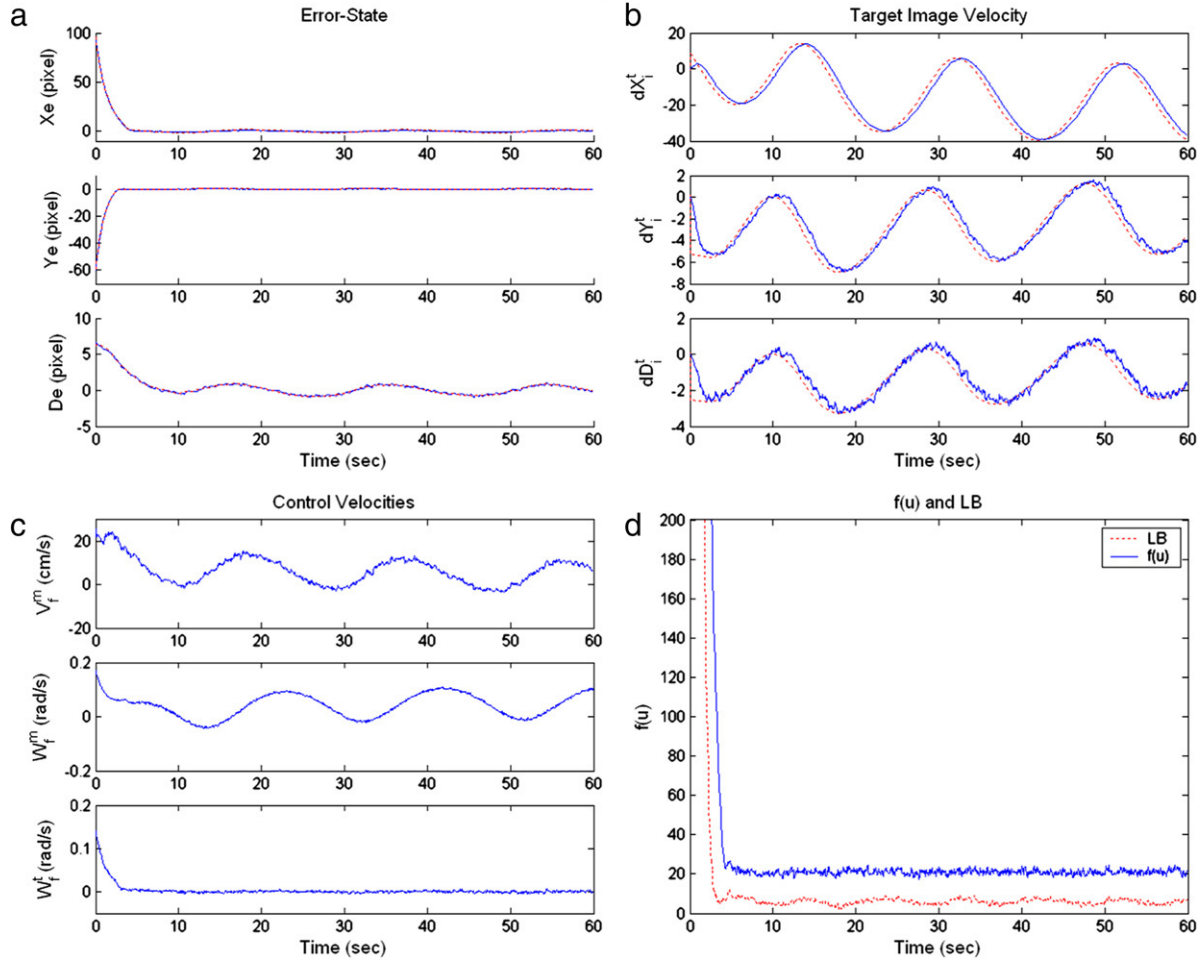
$$LB = \lambda_{\min}(\mathbf{K}_g) \|X_e\|^2 - \delta f(u) \leq \bar{f}(u) - \delta f(u) = f(u). \quad (45)$$

Expression (45) tells us that if LB is positive, then  $f(u)$  is guaranteed to be positive and the system is stable. In Fig. 9(d), the dotted line shows the transition of LB defined in (45) while the solid line indicates the transition of  $f(u)$ . It is clear that LB is positive during the visual tracking task, and hence the closed-loop visual tracking control system with parametric uncertainty is stable. Therefore, these simulation results validate that the proposed visual tracking control system not only overcomes the random noise in the observation, but also overcomes the parametric uncertainty in the system model.

### 5.4. Experiment setup

Fig. 10 shows the experimental mobile robot equipped with an on-board 1.6G industrial personal computer (IPC), USB 2.0 camera and a pan-tilt camera platform for the study of human–robot interaction through visual tracking control. Fig. 11 illustrates the complete robust visual tracking control system constructed using the proposed VSE and VTC. The function of each block shown in Fig. 11 is listed below:

- (1) Feature detection and tracking: perform face detecting and tracking algorithms proposed in [33] to extract the state of the user's face  $[x_i \ y_i \ d_x]^T$  in the image captured from the camera.



**Fig. 9.** The computer simulation results of the visual tracking control system with random noise and parametric uncertainty. (a) Tracking errors. (b) Estimated target image velocity. (c) Control velocities of the mobile robot and the tilt camera. (d) Transition of  $f(u)$  defined in (15) and its corresponding low-bound defined in (45).



**Fig. 10.** Experimental mobile robot interacts with a user using a real-time face tracking algorithm and the proposed robust visual tracking control system.

- (2) Visual state estimator: estimate the optimal system state  $[x_i^* \ y_i^* \ d_x^*]^T$  and the target image velocity  $[\dot{x}_i^t \ \dot{y}_i^t \ \dot{d}_x^t]^T$  by using the proposed self-tuning Kalman filter described in Section 4.
- (3) Visual tracking control law: compute desired robot control velocity  $[v_f^m \ w_f^m \ w_t^m]^T$  using (13).
- (4) Velocity transformation: transform the desired linear and angular control velocities into desired left- and right-wheel

- (5) Scaling processing: scale the desired control velocity  $[v_l \ v_r \ w_t^m]^T$  to satisfy the maximum velocity and acceleration limitations of the specific robot system. Note that this processing will degrade the tracking performance, but increase the smoothness and safety, of the practical robot system.
- (6) Quantization processing: quantize the scaled control velocity  $[\tilde{v}_l \ \tilde{v}_r \ \tilde{w}_t^m]^T$  dependent on the resolution of motion control module. The resolution of the motion control card used in the experiments is 8-bit, which means it can command the linear wheel velocity from  $-128$  to  $127$  cm/s in integer. For example, suppose that the scaled left-wheel velocity command  $\tilde{v}_l$  is  $2.9925$  cm/s. After quantization processing, the quantized velocity command  $\tilde{v}_l$  is  $[2.9925] = 2$  cm/s, where  $[x]$  is the largest integer smaller than  $x$ , and the corresponding quantization error  $\delta v_l$  is  $\tilde{v}_l - \tilde{v}_l = -0.9925$  cm/s.
- (7) Robust control law: compute robust control velocity  $[\bar{v}_l^* \ \bar{v}_r^* \ \bar{w}_t^{m*}]^T$  to overcome the velocity quantization problem by using the robust control law presented in [25]. The interested reader is referred to [25] for more technical details.
- (8) Velocity inverse transformation: transfer the robot's current left- and right-wheel velocities into linear and angular velocities using  $v_c = (v_{lc} + v_{rc})/2$  and  $w_c = (v_{rc} - v_{lc})/D$ . Let  $u_{n-1} = [v_c \ w_c \ w_{tc}]^T$  denote the previous robot control velocity for the visual state estimator to calculate the current propagation system state  $X_n^p$ .

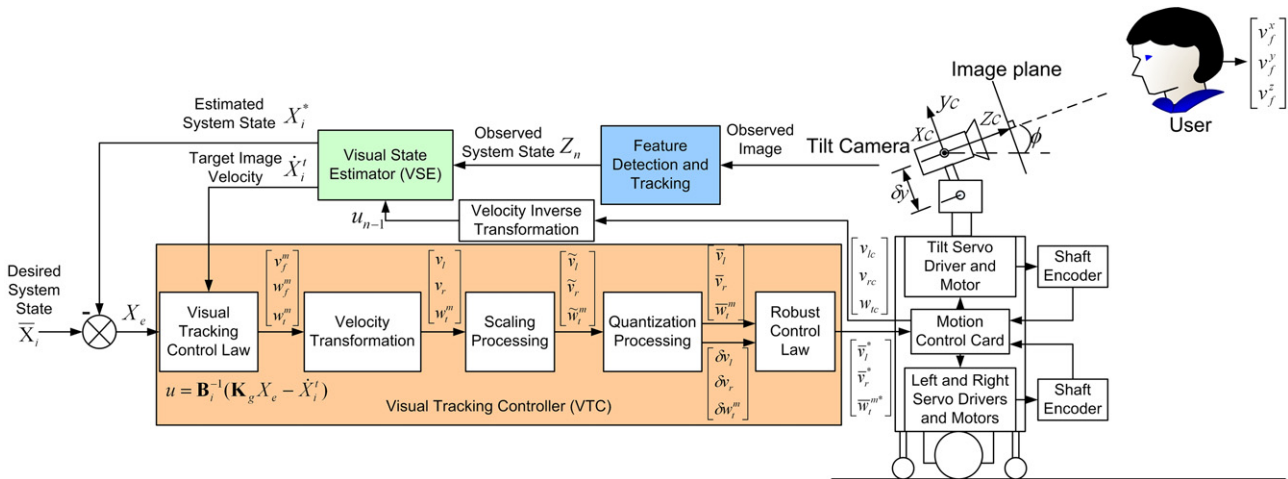


Fig. 11. Implemented robust visual tracking control system of a wheeled mobile robot with on-board tilt camera.

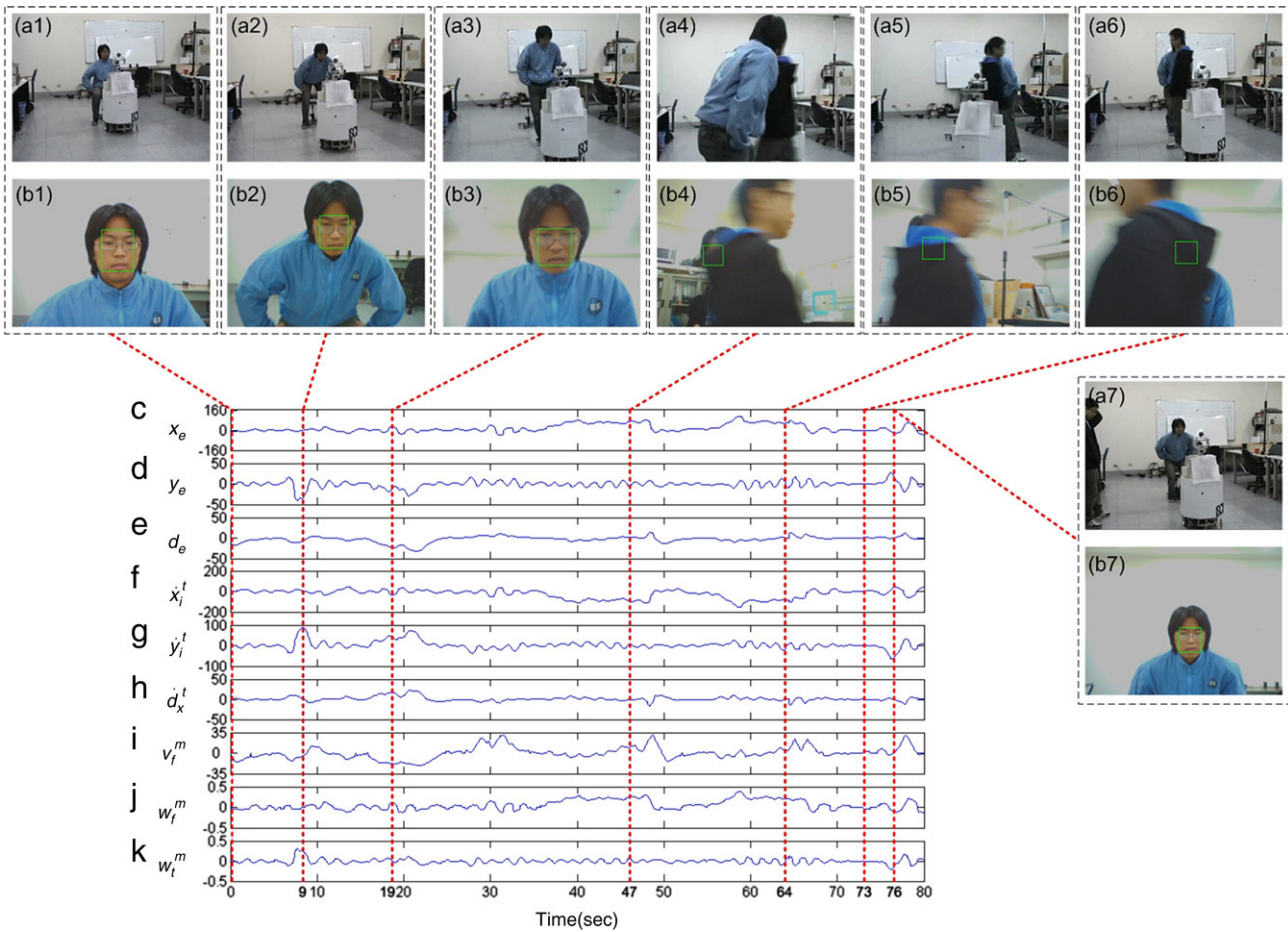


Fig. 12. Experimental results. (a1–a7): Image sequence recorded from a DV camera. (b1–b7): Corresponding image sequence recorded from on-board USB camera. (c–e): Recorded tracking errors in the image plane. (f–h): Target image velocity estimates. (i–j): Command linear and angular velocities of the mobile robot. (k): Command velocity of the tilt camera.

5.5. Experimental results of robustness in visual tracking

Fig. 12 presents the recorded images and responses of the mobile robot and tilt camera in the experiment that includes occlusions to validate the robustness of the proposed visual tracking control system. Fig. 12(a1–a7) show the recorded pictures from a digital video (DV) camera, and Fig. 12(b1–b7) are the corresponding pictures recorded by the on-board USB camera. Fig. 12(c–e) and Fig. 12(f–h) depict the response of the tracking

errors ( $x_e, y_e, d_e$ ) and target image velocity estimates ( $\dot{x}_i^t, \dot{y}_i^t, \dot{d}_x^t$ ), respectively. Fig. 12(i–k) illustrate the response of robot and tilt camera control velocities ( $v_f^m, w_f^m, w_t^m$ ).

In the beginning, the user statically sat on a stool, and the robot started to track his face using the proposed visual tracking control system. From Fig. 12(f–h), one can see that the target image velocity estimates all approach to zero when robot started working for about 5 s. Next, the user stood up (Fig. 12(a2)) and the tilt camera worked to keep tracking his face. From Fig. 12(g),



we observe that the target image velocity estimate  $\dot{y}_i^t$  increased when the user stood up. In the following, the user walked forward (Fig. 12(a3)) and the robot moved backward to keep tracking user's face. From Fig. 12(h), it is clear that the image velocity estimate  $\dot{d}_i^t$  increased when the user walked forward. These estimation results are consistent with the practical situation.

When the user walked around in the room, the robot kept following and tracking the user's face. While the user was walking, another person walked across between them temporarily (Fig. 12(a4–a6)). Thus, in Fig. 12(b4–b6), the user's face was temporarily fully blocked by the walking person. In this situation, the variation of observed time sequence data will rise, and the corresponding variance will become large. Based on the proposed self-tuning algorithm, the prediction information will dominate the estimation results of the Kalman filter even if the target is fully unobservable. Therefore, as shown in Fig. 12(c–h), the STKF still estimated the positions and velocities of the unobservable moving target in the image plane successfully even during full occlusion. This occlusion experiment validates the robust estimation performance of the proposed visual tracking control system. Finally, the user sat down on the stool, and the robot tracked the user continuously. The video clips of the experiments are available online [34].

**Remark 4.** It is interesting to know how long of the occlusion can be handled by using the proposed visual tracking control system. In the current design, the optimal covariance matrix  $\mathbf{P}_n^*$  of the self-tuning Kalman filter output is used to evaluate the reliability of the estimation result. If one of the diagonal values of the covariance matrix  $\mathbf{P}_n^*$  is larger than a preset threshold, then it implies that the estimation result is not reliable, and thus the visual tracking control system will stop and reinitialize. Therefore, the occlusion handling time is dependent on the value of preset threshold. In the experiments, the value of preset threshold is 4800, and the occlusion handling time is about 3 s, which is enough to overcome the temporary occlusion problem.

## 6. Conclusion and future work

A novel dual-Jacobian visual interaction model that represents the visual interaction between a mobile robot and a dynamic moving target in the image plane has been derived in this paper. Based on this visual interaction model, a robust visual tracking control system, which includes a VTC and a VSE, is proposed to solve dynamic moving target visual tracking control problem without the information of target's 3D motion velocity. In the robustness analysis, we have shown that the proposed VTC provides some degree of robustness against system model uncertainties. The proposed VSE is implemented by a real-time self-tuning Kalman filter in order to estimate the optimal system state and the target image velocity in the image plane. This design not only provides the target motion information for the later VTC to enhance the tracking performance, but also overcomes the temporary occlusion problem during visual tracking process. Experimental results validate the tracking performance and robustness of the proposed system. Therefore, the proposed visual tracking control scheme provides a useful solution for visual tracking control of a unicycle-modeled mobile robot with a pan-tilt camera to interact with a dynamic moving target such as a human's face.

Because Kalman filter is applied to the VSE design, there are some restrictions on the current VSE due to the conditions of Kalman filter such as Gaussian distribution uncertainty, smoothness motion, and uniform sampling rate. These disadvantages restrict the performance of the proposed VSE, which might fail in tracking a highly dynamic moving target. Therefore, the future work will focus on developing other types of VSE, such as neural-networks based VSE, to solve this problem and improve the accuracy of the visual estimation results.

## Acknowledgment

The authors would like to thank Fu-Sheng Huang, Chen-Yang Lin, and Chun-Wei Chen for their assistance in experiments. This work was supported by National Science Council of Taiwan, ROC under grant NSC 95-2218-E-009-024, 95-EC-17-A-04-S1-054 and the research foundation Flanders FWO, Belgium under grant G.0317.05.

## Appendix

This appendix presents the derivation of visual interaction model (2). Let  $X_f^m = [x_f^m \ y_f^m \ z_f^m]^T$  and  $X_f^t = [x_f^t \ y_f^t \ z_f^t]^T$  denote, respectively, the position of the mobile robot and the target in the world coordinate frame, and  $X_c = [x_c \ y_c \ z_c]^T$  denotes the relative position between robot and target in the camera coordinate frame. Then  $X_c$  can be calculated by the coordinate transformation such that

$$X_c = \mathbf{R}(\phi, \theta_f^m)(X_f^t - X_f^m) - \delta Y, \quad (\text{A.1})$$

where

$$\mathbf{R}(\phi, \theta_f^m) = \begin{bmatrix} \cos \theta_f^m & 0 & -\sin \theta_f^m \\ -\sin \phi \sin \theta_f^m & \cos \phi & -\sin \phi \cos \theta_f^m \\ \cos \phi \sin \theta_f^m & \sin \phi & \cos \phi \cos \theta_f^m \end{bmatrix},$$

and  $\delta Y = [0 \ \delta y \ 0]^T$ . By taking the derivative of (A.1), the kinematics of the interaction between robot and target in the camera frame can be obtained such that

$$\dot{X}_c = \mathbf{A}_c X_c + \mathbf{B}_c u + \mathbf{R}(\phi, \theta_f^m) V_t, \quad (\text{A.2})$$

where

$$\mathbf{A}_c = \begin{bmatrix} 0 & w_f^m \sin \phi & -w_f^m \cos \phi \\ -w_f^m \sin \phi & 0 & -w_t^m \\ w_f^m \cos \phi & w_t^m & 0 \end{bmatrix},$$

$$\text{and } \mathbf{B}_c = \begin{bmatrix} 0 & \delta y \sin \phi & 0 \\ \sin \phi & 0 & 0 \\ -\cos \phi & 0 & \delta y \end{bmatrix}.$$

Next, the position  $X_c$  is transformed into the image coordinate frame for deriving the visual interaction model (2) by using the expression (A.2). According to the pinhole camera model, the coordinate transformation from the camera frame into the image frame can be defined by standard projection equations [28] such that:

$$X_i = \mathbf{A}_c^i X_c, \quad \text{and} \quad X_c = (\mathbf{A}_c^i)^{-1} X_i, \quad (\text{A.3})$$

where  $\mathbf{A}_c^i = \text{diag}(-k_x, k_y, k_x k_w)$ ,  $k_x = f_x/z_c$ ,  $k_y = f_y/z_c$  and  $k_w = W/z_c$ . By taking the derivative of (A.3), the kinematic of the interaction between robot and target in the image frame can be found such that

$$\dot{X}_i = \mathbf{P}_i \dot{X}_c, \quad (\text{A.4})$$

where  $\mathbf{P}_i = \begin{bmatrix} -k_x & 0 & -k_x f_x^{-1} x_i \\ 0 & k_y & -k_y f_y^{-1} y_i \\ 0 & 0 & -k_x f_x^{-1} d_x \end{bmatrix}$ . Substituting (A.2) and (A.3) into (A.4), the visual interaction model (2) then can be obtained. ■

## References

- [1] N.P. Papanikolopoulos, P.K. Khosla, T. Kanade, Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision, *IEEE Transactions on Robotics and Automation* 9 (1) (1993) 14–35.
- [2] S. Hutchinson, G.D. Hager, P.I. Corke, A tutorial on visual servo control, *IEEE Transactions Robotics and Automation* 12 (5) (1996) 651–670.
- [3] J. Su, R. Qiu, H. Ma, P.-Y. Woo, Calibration-free robotic eye–hand coordination based on an auto disturbance-rejection controller, *IEEE Transactions on Robotics and Automation* 20 (5) (2002) 899–907.



- [4] G. López-Nicolás, C. Sagüés, J.J. Guerrero, D. Kragic, P. Jensfelt, Nonholonomic epipolar visual servoing, in: Proceedings of IEEE International Conference on Robotics and Automation, Orlando, FL, 2006, pp. 2378–2384.
- [5] Y. Ma, J. Koščeká, S.S. Sastry, Vision guided navigation for a nonholonomic mobile robot, IEEE Transactions on Robotics and Automation 15 (3) (1999) 521–536.
- [6] J.-B. Coulaud, G. Campion, G. Bastin, M.D. Wan, Stability analysis of a vision-based control design for an autonomous mobile robot, IEEE Transactions on Robotics 22 (5) (2006) 1062–1069.
- [7] Y. Masutani, M. Mikawa, N. Maru, F. Miyazaki, Visual servoing for nonholonomic mobile robots, in: Proceedings IEEE/RSJ/CI International Conference on Intelligent Robots and Systems, Munich, Germany, 1994, pp. 1133–1140.
- [8] R. Pissard-Gibollet, P. Rives, Applying visual servoing techniques to control a mobile hand-eye system, in: Proceedings IEEE International Conference on Robotics and Automation, Nagoya, Aichi, Japan, 1995, pp. 166–171.
- [9] F. Conticelli, B. Allotta, P.K. Khosla, Image-based visual servoing of nonholonomic mobile robots, in: Proceedings IEEE 38th Conference Decision and Control, Phoenix, AZ, USA, 1999, pp. 3496–3501.
- [10] D. Burschka, J. Geiman, G. Hager, Optimal landmark configuration for vision-based control of mobile robot, in: Proceedings IEEE International Conference on Robotics and Automation, Taipei, Taiwan, 2003, pp. 3917–3922.
- [11] H. Zhang, J.P. Ostrowski, Visual motion planning for mobile robots, IEEE Transactions on Robotics and Automation 18 (2) (2002) 199–208.
- [12] T. Nierobisch, W. Fischer, F. Hoffmann, Large view visual servoing of a mobile robot with a pan-tilt camera, in: Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006, pp. 3307–3312.
- [13] J. Chen, W.E. Dixon, D.M. Dawson, M. McIntyre, Homography-based visual servo tracking control of a wheeled mobile robot, IEEE Transactions on Robotics 22 (2) (2006) 407–416.
- [14] Y. Fang, W.E. Dixon, D.M. Dawson, P. Chawda, Homography-based visual servo regulation of mobile robots, IEEE Transactions on System, Man, and Cybernetics—Part B: Cybernetics 35 (5) (2005) 1041–1049.
- [15] G.L. Mariottini, G. Oriolo, D. Prattichizzo, Image-based visual servoing for nonholonomic mobile robots using epipolar geometry, IEEE Transactions on Robotics 23 (1) (2007) 87–100.
- [16] G.L. Mariottini, D. Prattichizzo, G. Oriolo, Image-based visual servoing for nonholonomic mobile robots with central catadioptric camera, in: Proceedings IEEE International Conference on Robotics and Automation, Orlando, FL, USA, 2006, pp. 538–544.
- [17] T. Goedemé, M. Nuttin, T. Tuytelaars, L. Van Gool, Omnidirectional vision based topological navigation, International Journal of Computer Vision 74 (3) (2007) 219–236.
- [18] A.K. Das, R. Fierro, V. Kumar, J.P. Ostrowski, J. Spletzer, C.J. Taylor, A vision-based formation control framework, IEEE Transactions on Robotics and Automation 18 (5) (2002) 813–825.
- [19] R. Vidal, O. Shakernia, S. Sastry, Following the flock [formation control], IEEE Robotics and Automation Magazine 11 (4) (2004) 14–20.
- [20] J.A. Borstadt, N.J. Ferrier, Interception of a projectile using a human vision-based strategy, in: Proceedings IEEE International Conference on Robotics and Automation, San Francisco, USA, 2000, pp. 3189–3196.
- [21] L. Freda, G. Oriolo, Vision-based interception of a moving target with a nonholonomic mobile robot, Journal of Robotics and Autonomous Systems 55 (6) (2007) 419–432.
- [22] H.Y. Wang, S. Itani, T. Fukao, N. Adachi, Image-based visual adaptive tracking control of nonholonomic mobile robots, in: Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, HI, USA, 2001, pp. 1–6.
- [23] E. Malis, S. Benhimane, A unified approach to visual tracking and servoing, Journal of Robotics and Autonomous Systems 52 (1) (2005) 39–52.
- [24] Y. Han, H. Hahn, Visual tracking of a moving target using active contour based SSD algorithm, Journal of Robotics and Autonomous Systems 53 (3–4) (2005) 265–281.
- [25] C.-Y. Tsai, K.-T. Song, Face tracking interaction control of a nonholonomic mobile robot, in: Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006, pp. 3319–3324.
- [26] C.-Y. Tsai, K.-T. Song, X. Dutoit, H. Van Brussel, M. Nuttin, Robust mobile robot visual tracking control system using self-tuning Kalman filter, in: Proceedings of IEEE Symposium on Computational Intelligence in Robotics and Automation, Jacksonville, FL, 2007, pp. 161–166.
- [27] T.-C. Lee, C.-Y. Tsai, K.-T. Song, Fast parking control of mobile robots: A motion planning approach with experimental validation, IEEE Transactions on Control Systems Technology 12 (5) (2004) 661–676.
- [28] W.I. Grosky, L.A. Tamburino, A unified approach to the linear camera calibration problem, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7) (1990) 663–671.
- [29] J.-J.E. Slotine, W. Li, Applied Nonlinear Control, Prentice-Hall, Englewood Cliffs, NJ, 1991.
- [30] J.D. Schutter, J.D. Geeter, T. Lefebvre, H. Bruyninckx, Kalman filters: A tutorial, Journal A 40 (4) (1999) 52–59.
- [31] O.V. Korniyenko, M.S. Sharawi, D.N. Aloji, Neural network based approach for tuning Kalman filter, in: Proceedings of IEEE International Conference on Electro/Information Technology, Lincoln, Nebraska, 2005, pp. 1–5.

- [32] D.G. Luenberger, Linear and Nonlinear Programming, Addison-Wesley, Reading, MA, 1989.
- [33] K.-T. Song, J.-S. Hu, C.-Y. Tsai, C.-M. Chou, C.-C. Cheng, W.-H. Liu, C.-H. Yang, Speaker attention system for mobile robots using microphone array and face tracking, in: Proceedings IEEE International Conference on Robotics and Automation, Orlando, FL, 2006, pp. 3624–3629.
- [34] The experiment video website. [Online]. Available: <http://isci.cn.nctu.edu.tw/video/RVTC/>.



and computer vision.

**Chi-Yi Tsai** was born in Kaohsiung, Taiwan, Republic of China in 1978. He received the B.S. and M.S. degree in electrical engineering from National Yunlin University of Science and Technology, Yunlin, Taiwan, ROC, in 2000 and 2002, respectively. He received the Ph.D. degree in electrical and control engineering, at the National Chiao Tung University, Taiwan, ROC, in 2008. He is currently an engineer in software R&D department of EPC business unit, ASUSTek Computer Incorporation. His research interests include image processing, color enhancement processing, visual tracking control for mobile robots, visual servoing



**Kai-Tai Song** was born in Taipei, Taiwan, Republic of China in 1957. He received the B.S. degree in Power Mechanical Engineering from National Tsing Hua University in 1979 and the Ph.D. degree from the Katholieke Universiteit Leuven, Belgium in 1989. He was with Chung Shan Institute of Science and Technology from 1981 to 1984. Since 1989 he has been on the faculty and is currently a Professor in the Department of Electrical and Control Engineering, National Chiao Tung University, Taiwan. He has also served as the Associate Dean of Research & Development Office of National Chiao Tung University from 2007 till now. He served as the Chair of IEEE Robotics & Automation Chapter, Taipei Section in the term of 1999. His areas of research interest include mobile robotics, image processing, visual tracking, embedded systems, and mechatronics.



**Xavier Dutoit** was born in Switzerland in 1981. He received his M.S. degree in Communications Systems from the Ecole Polytechnique Federale de Lausanne, in Switzerland, in 2004. Since 2005, he is a Ph.D. student at the Katholieke Universiteit of Leuven, in Belgium, where he is conducting research in machine learning and mobile learning robots.



**Hendrik Van Brussel** (1944), is full professor of mechatronics and automation at the Faculty of Engineering, Katholieke Universiteit Leuven (K.U.Leuven), Belgium, and chairman of the Division of Production engineering, Machine design and Automation (PMA), Department of Mechanical Engineering.

He was a pioneer in robotics research in Europe and an active promoter of the mechatronics idea as a new paradigm in machine design. He has published extensively on different aspects of robotics, mechatronics and flexible automation. His present research interest is also shifting towards holonic manufacturing systems, behaviour based robots, and micro and precision engineering, including microrobotics.

He is Fellow of SME and IEEE. He was President of CIRP (International Academy for Production Engineering). He is a Member of the Royal Flemish Academy of Belgium for Sciences and Arts, and Foreign Member of the Royal Swedish Academy of Engineering Sciences (IVA). He is President of euspenn (European Society for Precision Engineering and Nanotechnology).



**Marnix Nuttin** was born in 1969 in Belgium. He obtained his engineering degree from the Katholieke Universiteit Leuven in 1992, and his Ph.D. in 1998. He has been conducting research on mobile learning robots from 1992 to 2008 in the department of Mechanical Engineering. He developed learning techniques for Programming by Demonstration and adaptive shared autonomy for wheelchair control. He participated to the ITEA AMBIENCE project (context-aware environments for ambient services). In FP6 he was principle investigator of the projects MOVEMENT (Modular Versatile Mobility Enhancement Technology), MAIA (Non Invasive Brain Interaction with Robots) and DYNAVIS (Dynamically Reconfigurable Quality Control). He is currently affiliated to the Faculty of Engineering,