2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Adaptive Image-based Visual Servoing of Wheeled Mobile Robots with Fixed Camera Configuration

Xinwu Liang, Hesheng Wang and Weidong Chen

Abstract— **In this paper, we will study the uncalibrated vision-based positioning problem of wheeled mobile robots by using a ceiling-mounted camera. A new image-based visual servoing scheme will be proposed, which can cope with the unknown intrinsic and extrinsic parameters of the camera and the uncertain distance parameter of the feature point from geometric center of the mobile robot. The presented approach is developed via extending the depth-independent interaction matrix framework for robot manipulators to mobile robots such that the nonlinear dependence on unknown parameters can be removed from the image-Jacobian matrix and then, we can linearly parameterize uncertain parameters in the closed-loop system. In this way, an estimation scheme for the online updating of uncertain parameters can be developed very efficiently. To show that image errors can be guaranteed to be asymptotically convergent, stability analysis will be carried out by using Lyapunov theory. To validate the performance of the presented approach, simulation and experimental results will also be provided.**

Index Terms— **Adaptive visual servoing, depth-independent image Jacobian, fixed-camera, wheeled mobile robot.**

I. Introduction

Visual servoing is a key approach that adopts vision feedback from camera systems for the efficient control of robotic systems, which can be mainly classified into image-based, position-based and hybrid visual servoing approaches. Visual servoing approaches were originally proposed to solve the motion control of 6-DOFs robotic manipulators, where the motion constraints were usually not taken into account. In contrast, a wheeled mobile robot is usually suffered from nonholonomic constraints. It is well known that no continuous time-invariant state feedback control scheme can be designed to stabilize a nonholonomic system to a single equilibrium. Hence, the motion control problem of wheeled mobile robots can be more challenging compared with that of robot manipulators. To achieve high performance of nonholonomic systems, various strategies have been presented in recent years. It is noted that these methods were all developed by assuming that the robot states must be exactly available. Due to the uncertain kinematics and the slippage of the wheels, however, such an assumption is usually not satisfied in most applications.

To handle uncertainties in real environments and improve the control performance of mobile robot systems, applying the sensor information directly into controller designs is an alternative important strategy. Such a strategy is known as the sensor-based control of mobile robots. Up to now, lots of works have been done in the research area of vision-based control of mobile robots. Similar to robot manipulators, for visual servoing of mobile robots, there are also two basic configurations to perform a visual servoing task, i.e., eye-in-hand and fixed-camera configurations.

So far, various visual servoing strategies have been presented for wheeled mobile robots. In [1], a visual servoing scheme was presented for the nonholonomic wheeled mobile robots, where a pan and tilt unit was added onboard to increase the camera DOFs. To guarantee that the landmark is always kept in the camera field of view, a novel visual servoing strategy was developed to efficiently control nonholonomic vehicles in [2]. It is noted that all these methods are belong to position-based visual servoing approaches, and hence, metrical information is required to obtain all the states for feedback control such that the mobile robot is guaranteed to be asymptotically convergent to its desired pose.

To eliminate dependence on prior 3D information of the scene, a new vision-based control scheme was proposed in [3] to solve the positioning problem of wheeled mobile robots. Considering its good properties, epipolar geometry was adopted as a useful tool to effectively solve the vision-based control issue of wheeled mobile robots. A new two-step vision-based control method was developed in [4] without any prior information about the 3D scene geometry. Since we cannot directly control the distance of the current position from the desired position, the short baseline degeneracy problem may occur in strategies based on epipolar geometry. To solve this degeneracy problem, a direct vision-based control approach was proposed in [5] by introducing a virtual target and using essential matrix. However, the epipolar geometry will become ill-conditioned when all feature points are coplanar.

The planar homography strategies can overcome problems of strategies based on epipolar geometry. These strategies can be called 2.5D visual servoing. In [6], a homography-based vision-based control scheme was proposed to guarantee asymptotical convergence of the pose of nonholonomic

This work was supported in part by Specialized Research Fund for the Doctoral Program of Higher Education of China (20100073120020, 20100073110018), in part by Shanghai Municipal Natural Science Foundation (11ZR1418400), in part by China Postdoctoral Science Foundation (2012M511095), in part by Shanghai Postdoctoral Science Foundation (12R21414200), in part by the Special Financial Grant from the China Postdoctoral Science Foundation (2013T60448), in part by the China Domestic Research Project for the International Thermonuclear Experimental Reactor (ITER) (2012GB102008), in part by International Cooperation Project of Science and Technology Department, China (2011DFA11780), in part by the Natural Science Foundation of China (61105095, 61203361, 61175088, 61221003), and in part by the National High Technology Research and Development of China (2012AA041403).

The authors are with the Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China. xinwu113@163.com;wanghesheng@sjtu.edu.cn;wdchen@sjtu.edu.cn

mobile robots. To compensate for the constant unmeasurable depth parameter, an adaptive vision-based controller was developed in [7] to handle the pose control problem of mobile robots. A homography-based visual tracking scheme was presented in [8] to make a nonholonomic mobile robot track a desired trajectory. As we know, in most homography-based strategies, homography estimation and decomposition are two necessary processes, and in the general case, two possible solutions will be found through the decomposition process. To avoid ambiguity issues of homography decomposition, the input-output linearization technique was used to develop a vision-based controller [9]. To overcome the oversensitivity problem of the strategy in [9] to various disturbances in the vision system, a two-level vision-based control method was developed in [10]. To completely remove both the homography estimation and decomposition processes, a 2.5D visual servoing approach was presented in [11] for the stabilization of mobile robots.

The previously mentioned approaches are all designed for the eye-in-hand configuration. Lots of studies have also been focused on the visual servoing problem of wheeled mobile robots for the fixed-camera configuration. A robust two-step technique was proposed in [12], where the lack of depth information and precise visual parameters can be compensated for efficiently. To handle both the unknown camera orientation and scaling parameters, a robust stabilization controller was presented in [13]. To cope with the vision-based tracking issue of nonholonomic mobile robots, a robust visual tracking controller was designed in [14]. A Jacobian estimation scheme based on the dynamic quasi-Newton strategy was presented in [15] for uncalibrated visual servoing of mobile robots, where the simultaneous navigation and obstacle avoidance problem was solved through optimization of potential field-based objective function. All these methods are developed only by considering the mobile robot kinematics. Considering the importance of the nonholonomic mobile robot dynamics to its control performance, a sliding-mode scheme was proposed in [16] to visually stabilize mobile robots with uncertain dynamics. In [17], a dynamic controller was developed for the position/orientation tracking of nonholonomic mobile robots. To apply these approaches, the depth of the imaging sensor from the planar ground is required to be invariant during the visual servoing process, i.e., the camera plane must be parallel to the plane of motion of mobile robots. Hence, their applicability can be very limited.

In the current work, an adaptive image-based visual servoing method is presented for wheeled mobile robots with fixed-camera configuration. The main novelty and one of the main contributions in this paper is that, the controller design in our scheme does not impose any conditions on the depth information, i.e., time-varying depth information is allowed and the camera image plane and the plane of motion of mobile robots are not required to be parallel. Furthermore, in the proposed scheme, exact knowledge of the camera parameters is not necessary. Via compensating for the uncertain time-varying depth parameter, the proposed approach can regulate the wheeled mobile robot toward its desired position. Asymptotical convergence of the closed-loop dynamics is proved by using Lyapunov-theory. Simulation and experimental results are also provided to validate our scheme.

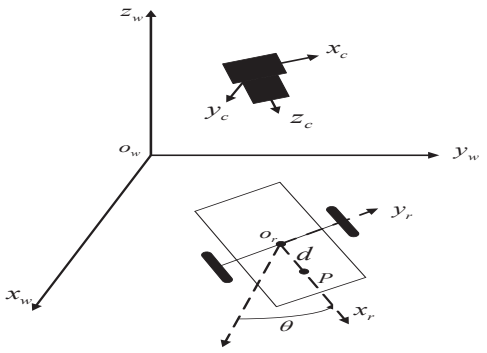

Fig. 1. Wheeled mobile robots with ceiling-mounted camera

II. Problem Formulation

The mobile robot system for the case of fixed-camera configuration is illustrated in Fig. 1. In this setup, we suppose that a pinhole camera is attached overhead to observe the robot motion but the motion plane and the camera image plane are not necessary to be parallel to each other. To execute the control task, we assume that in the period of visual servoing, the camera can always observe the feature point marked on board of the robot, which is tracked by the image processing system. In addition, we suppose that the location of the mass center of the mobile robot is the same as that of its geometric center. In this system, we define three related coordinate frames, i.e., the inertial frame W, the camera frame C and the body-fixed frame R, where the relationship of these three coordinate frames can be found in Fig. 1. As we know, the inertial frame W will be taken as the body-fixed frame R when the mobile robot is at its initial position. The homogeneous transformation from C to W is denoted by a 4×4 matrix $\mathbf{T}_w^c$, which represents the constant camera extrinsic parameters. The homogeneous transformation between W and R is denoted by a 4×4 matrix $\mathbf{T}_r^w$, which depends on the pose of the mobile robot. In the case of planar motion, the mobile robot position can be represented by $\mathbf{x}(t) = (x(t), y(t))^\top$, and its orientation is denoted as $\theta(t)$, which represents the angle between x-axis positive direction of W and the robot forward direction.

The nonholonomic constraint for mobile robots is given by

$$\dot{x}(t)\sin\theta(t) - \dot{y}(t)\cos\theta(t) = 0. \tag{1}$$

The nonholonomic kinematics of mobile robots is written as

$$\begin{aligned} \dot{x}(t) &= \nu\cos\theta(t) \\ \dot{y}(t) &= \nu\sin\theta(t) \\ \dot{\theta}(t) &= \omega \end{aligned} \tag{2}$$

where ν and ω respectively represent the forward and rotational velocities of the mobile robot.

Generally speaking, the mobile robot position $\mathbf{x}(t)$ can be derived via encoders of the motors or other sensors. In complex environments, however, obtaining this information can be very difficult. To solve this problem, visual information can be considered as an effective solution. In our case, a black circular mark on board of the robot, which is denoted by P, is used to facilitate the detection and task specification of the robot through the ceiling-mounted camera system. In this way, image coordinates of P can be used for the localization of the mobile robot and to define a desired robot position. In contrast, the orientation angle $\theta(t)$ of the mobile robot can be easily obtained via angle sensors and will be used directly in the vision-based controller. P is located at the x-axis of frame R, and let d denote the distance of P from the origin of frame R. Then, $\mathbf{x}_p(t) = (x_p(t), y_p(t))^\top$, the position of P on the motion plane with respect to frame W, is given by

$$\mathbf{x}_p(t) = \begin{bmatrix} x_p(t) \\ y_p(t) \end{bmatrix} = \begin{bmatrix} x(t) + d\cos\theta(t) \\ y(t) + d\sin\theta(t) \end{bmatrix}. \quad (3)$$

Differentiating (3) and then substituting (2), we have

$$\dot{\mathbf{x}}_p(t) = \begin{bmatrix} \cos\theta(t) & -d\sin\theta(t) \\ \sin\theta(t) & d\cos\theta(t) \end{bmatrix} \begin{bmatrix} \nu \\ \omega \end{bmatrix}. \quad (4)$$

The homogeneous coordinates of P with respect to W can be given by $\mathbf{x}_p^w(t) = (\mathbf{x}_p^\top(t), 0, 1)^\top$, whose homogeneous coordinates with respect to frame C can be expressed as

$$\mathbf{x}_p^c(t) = \mathbf{T}_w^c \mathbf{x}_p^w(t) = \begin{bmatrix} \mathbf{R}_w^c & \mathbf{t}_w^c \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \mathbf{x}_p^w(t) \quad (5)$$

where $\mathbf{R}_w^c$ and $\mathbf{t}_w^c$ are respectively the rotational and translational components of $\mathbf{T}_w^c$.

By applying the pinhole camera model, image coordinates of P $\mathbf{y}_p(t) = (u_p(t), v_p(t))^\top$ can be written as

$$\begin{bmatrix} \mathbf{y}_p(t) \\ 1 \end{bmatrix} = \frac{1}{z_p^c(t)} \mathbf{\Omega} \mathbf{x}_p^c(t) \quad (6)$$

where $\mathbf{\Omega} \in \Re^{3\times 4}$ only depends on intrinsic parameters of the camera $\mathbf{\Omega} = [\mathbf{A} \ \mathbf{0}_{3\times 1}]$, with $\mathbf{A} \in \Re^{3\times 3}$ given in [18], and $z_p^c(t)$ represents the depth of P expressed in the camera frame. Substituting (5) into (6) and using $\mathbf{x}_p^w(t)$ yields

$$\begin{bmatrix} \mathbf{y}_p(t) \\ 1 \end{bmatrix} = \frac{1}{z_p^c(t)} \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t}_w^c \end{bmatrix} \begin{bmatrix} \mathbf{x}_p(t) \\ 1 \end{bmatrix} \quad (7)$$

where $\mathbf{R}_w^c = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$. Let $\mathbf{H} = \mathbf{A}[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{t}_w^c]$, which is actually the homography between the camera image plane and the motion plane and is determined by uncertain camera parameters, (7) can be simplified as

$$\begin{bmatrix} \mathbf{y}_p(t) \\ 1 \end{bmatrix} = \frac{1}{z_p^c(t)} \mathbf{H} \begin{bmatrix} \mathbf{x}_p(t) \\ 1 \end{bmatrix}. \quad (8)$$

Based on the notation of $\mathbf{H}$, (8) can be further rewritten as

$$\mathbf{y}_p(t) = \frac{1}{z_p^c(t)} \begin{bmatrix} \mathbf{h}_1^\top \\ \mathbf{h}_2^\top \end{bmatrix} \begin{bmatrix} \mathbf{x}_p(t) \\ 1 \end{bmatrix} \quad (9)$$

where $\mathbf{h}_i^\top$ represents the ith row of $\mathbf{H}$. Similarly, the depth $z_p^c(t)$ can be expressed as

$$z_p^c(t) = \mathbf{h}_3^\top \begin{bmatrix} \mathbf{x}_p(t) \\ 1 \end{bmatrix}. \quad (10)$$

To further simplify the developments, the sub-matrix consisting of the first two columns of $\mathbf{H}$ is denoted by $\bar{\mathbf{H}}$. Differentiating (10) and then substituting (4), we have

$$\dot{z}_p^c(t) = \bar{\mathbf{h}}_3^\top \begin{bmatrix} \cos\theta(t) & -d\sin\theta(t) \\ \sin\theta(t) & d\cos\theta(t) \end{bmatrix} \begin{bmatrix} \nu \\ \omega \end{bmatrix} \triangleq \mathbf{b}_p^\top(t) \begin{bmatrix} \nu \\ \omega \end{bmatrix} \quad (11)$$

where $\bar{\mathbf{h}}_i^\top$ represents the ith row of $\bar{\mathbf{H}}$ and $\mathbf{b}_p^\top(t)$ is a 1×2 row vector which depends on the camera parameters and d. Differentiating (9) and then substituting (11) arrives at

$$\begin{aligned} \dot{\mathbf{y}}_p(t) &= \frac{1}{z_p^c(t)} \begin{bmatrix} \bar{\mathbf{h}}_1^\top - u_p(t)\bar{\mathbf{h}}_3^\top \\ \bar{\mathbf{h}}_2^\top - v_p(t)\bar{\mathbf{h}}_3^\top \end{bmatrix} \begin{bmatrix} \cos\theta(t) & -d\sin\theta(t) \\ \sin\theta(t) & d\cos\theta(t) \end{bmatrix} \begin{bmatrix} \nu \\ \omega \end{bmatrix} \\ &\triangleq \frac{1}{z_p^c(t)} \mathbf{Q}_p(t) \begin{bmatrix} \nu \\ \omega \end{bmatrix} \end{aligned} \quad (12)$$

where $\mathbf{Q}_p(t) \in \Re^{2\times 2}$ is known as the depth-independent image Jacobian, which depends on camera parameters and d.

Note that $\mathbf{Q}_p(t)$ and $\mathbf{b}_p^\top(t)$ have the following property:

Property 1: For any 2×1 vector ξ, the terms $\mathbf{Q}_p(t)\xi$ and $\mathbf{b}_p^\top(t)\xi$ can be linearly parameterized as

$$\mathbf{Q}_p(t)\xi = \mathbf{N}_q(\mathbf{y}_p(t), \theta(t), \xi)\rho \qquad \mathbf{b}_p^\top(t)\xi = \mathbf{n}(\theta(t), \xi)\rho \quad (13)$$

where $\mathbf{N}_q(\mathbf{y}_p(t), \theta(t), \xi) \in \Re^{2\times 14}$ and $\mathbf{n}(\theta(t), \xi) \in \Re^{1\times 14}$ are regressor matrices which can be calculated without the use of unknown parameters and $\rho = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{11}d, h_{12}d, h_{21}d, h_{22}d, h_{31}d, h_{32}d)^\top \in \Re^{14\times 1}$, with h_{ij} being the element in the ith row and jth column of $\mathbf{H}$.

The problem to be addressed can be described as follows.

Problem: Given the desired image coordinates of P, design a vision-based controller to provide kinematics-based control inputs $\mathbf{v} = (\nu, \omega)^\top$ for the mobile robot such that P can be driven asymptotically toward its desired image location without exact knowledge of camera parameters and d.

III. Adaptive Image-Based Visual Servoing

In this section, we propose an adaptive kinematics-based visual servoing scheme for wheeled mobile robots.

Let $\mathbf{y}_{pd}$ be the desired constant image coordinates of P. Subtracting the desired position from the current one, we have

$$\Delta\mathbf{y}_p(t) = \mathbf{y}_p(t) - \mathbf{y}_{pd}. \quad (14)$$

Assume that we have no exact knowledge of the camera parameters and the parameter d and let $\hat{\rho}(t)$ denote the estimated value of ρ. For convenience, we introduce

$$\mathbf{D}_p(t) = \mathbf{Q}_p(t) + \frac{1}{2}\Delta\mathbf{y}_p(t)\mathbf{b}_p^\top(t). \quad (15)$$

By substituting $\hat{\rho}(t)$, the estimated matrix $\hat{\mathbf{D}}_p(t)$ can be calculated from (15). Then, we design the following controller to drive the robot toward its desired position

$$\mathbf{v} = -\hat{\mathbf{D}}_p^{\top}(t)\mathbf{K}_p\Delta\mathbf{y}_p(t) \tag{16}$$

where $\mathbf{K}_p \in \Re^{2\times 2}$ is a symmetric positive-definite matrix.

Based on ***Property 1***, $\mathbf{D}_p(t)\mathbf{v}$ can be expressed as

$$\mathbf{D}_p(t)\mathbf{v} = \mathbf{N}\Big(\mathbf{y}_p(t), \theta(t), \mathbf{v}\Big)\rho \tag{17}$$

where $\mathbf{N}\Big(\mathbf{y}_p(t), \theta(t), \mathbf{v}\Big) \in \Re^{2\times 14}$ denotes the regressor matrix and can be calculated without the knowledge of ρ. It is not difficult to derive that

$$\mathbf{N}\Big(\ \cdot\ \Big) = \mathbf{N}_q(\mathbf{y}_p(t), \theta(t), \mathbf{v}) + \frac{1}{2}\Delta\mathbf{y}_p(t)\mathbf{n}(\theta(t), \mathbf{v}). \tag{18}$$

To update $\hat{\rho}(t)$, we propose the following adaptive law

$$\dot{\hat{\rho}}(t) = \mathbf{\Gamma}^{-1}\mathbf{N}^{\top}(\mathbf{y}_p(t), \theta(t), \mathbf{v})\mathbf{K}_p\Delta\mathbf{y}_p(t) \tag{19}$$

where $\mathbf{\Gamma}$ is a symmetric positive-definite matrix.

Now, it is time to state the main result in the following.

Theorem 1: Assume that P is always in the field of view of the camera in the period of task execution, the control law (16) and the estimation law (19) can guarantee asymptotical convergence of image errors of P in the following way

$$\lim_{t\to\infty} \hat{\mathbf{D}}_p^{\top}(t)\mathbf{K}_p\Delta\mathbf{y}_p(t) = \mathbf{0}. \tag{20}$$

That is to say, the mobile robot can be ensured to arrive at its desired position despite the lack of exact camera parameters and exact position of P.

Proof: Introduce the following positive function:

$$V(t) = \frac{1}{2}\Big(z_p^c(t)\Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\Delta\mathbf{y}_p(t) + \Delta\rho^{\top}(t)\mathbf{\Gamma}\Delta\rho(t)\Big) \tag{21}$$

where $\Delta\rho(t) = \hat{\rho}(t) - \rho$. By noting that $\Delta\dot{\mathbf{y}}_p(t) = \dot{\mathbf{y}}_p(t)$ and $\Delta\dot{\rho}(t) = \dot{\hat{\rho}}(t)$, differentiating (21) and then substituting (11) and (12), we can obtain

$$\dot{V}(t) = \Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\mathbf{D}_p(t)\mathbf{v} + \Delta\rho^{\top}(t)\mathbf{\Gamma}\dot{\hat{\rho}}(t) \tag{22}$$

where the definition given in (15) has been used. Adding and subtracting the term $\Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\hat{\mathbf{D}}_p(t)\mathbf{v}$ into (22) leads to

$$\begin{aligned}\dot{V}(t) = {} & \Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\hat{\mathbf{D}}_p(t)\mathbf{v} + \Delta\rho^{\top}(t)\mathbf{\Gamma}\dot{\hat{\rho}}(t) \\ & - \Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\mathbf{N}\Big(\mathbf{y}_p(t), \theta(t), \mathbf{v}\Big)\Delta\rho(t),\end{aligned} \tag{23}$$

where (17) has been used. Substituting the controller (16) and the adaptive law (19) into (23), we can easily obtain

$$\dot{V}(t) = -\Delta\mathbf{y}_p^{\top}(t)\mathbf{K}_p\hat{\mathbf{D}}_p(t)\hat{\mathbf{D}}_p^{\top}(t)\mathbf{K}_p\Delta\mathbf{y}_p(t) \tag{24}$$

from which $\dot{V}(t) \leq 0$ can be easily concluded, and we know that the mobile robot system is guaranteed to be stable by applying the designed visual servoing controller and adaptive law. Then, from LaSalle's Invariance Principle, we can arrive at the conclusion of (20). □

Remark 1: Since Cartesian position $\mathbf{x}(t)$ is not required and hence, in real environments, the robot can be driven to its desired location more precisely. On the other hand, $\theta(t)$ can be measured by angle sensors with high precision and positioning performance will not be affected much.

Remark 2: In order to make the proposed scheme applicable to the environments without angle sensors and to further extend its applicability, it is a good idea to remove the dependence on $\theta(t)$ from the scheme to make its implementation easier, and this is one of our main objectives in the future.

Remark 3: It is noted that, (20) can only show that $(\nu, \omega) \to 0$. If $\hat{\mathbf{D}}_p(t)$ can be guaranteed to be always nonsingular, we can directly derive that $\Delta\mathbf{y}_p(t) \to 0$ from (20). As demonstrated by the simulation and experimental results, image errors are actually convergent to zeros, which confirms the conclusion that $\Delta\mathbf{y}_p(t) \to 0$. Though we have never encountered, it is possible that, in some cases, $\hat{\mathbf{D}}_p(t)$ happens to be singular during task execution and then, the conclusion that $\Delta\mathbf{y}_p(t) \to 0$ cannot be derived. To theoretically guarantee $\Delta\mathbf{y}_p(t) \to 0$ in a strict manner, in the future, we will investigate improved schemes to make $\hat{\mathbf{D}}_p(t)$ always nonsingular.

IV. Simulation Results

In this section, simulation results based on a two-wheeled mobile robot will be provided to show the control performance of the presented approach. In this simulation setup, we use a ceiling-mounted camera to observe the motion of the mobile robot. Intrinsic parameters of the camera are given by $k_u = k_v = 1800$ pixels/m, $u_0 = 280$ pixels, $v_0 = 250$ pixels and $f = 0.035$ m, where (k_u, k_v) represent the scaling factors, (u_0, v_0) is the principal point, and f is the focal length. We can directly obtain $\mathbf{A}$ from the intrinsic parameters. To accomplish the control task, a black circular mark is attached on board of the mobile robot. The distance parameter d between the geometric center of the mark and that of the mobile robot is given by $d = 0.2$ m.

We suppose that the constant homogeneous transformation matrix $\mathbf{T}_c^w$ between frames C and W (which is the same as the initial body-fixed frame of the mobile robot) is given by

$$\mathbf{T}_c^w = Rt(\mathbf{x}, \pi - \alpha_1)Rt(\mathbf{y}, \alpha_2)Rt(\mathbf{z}, \alpha_3)Tr(a_1, a_2, -a_3)$$

where $Rt(\iota, \phi)$ represents a rotation matrix about ι-axis by angle ϕ, $Tr(a_1, a_2, -a_3)$ is a translation matrix, and α_1, α_2, α_3, a_1, a_2 and a_3 are all positive constants. Then, extrinsic parameters of the camera can be derived as $\mathbf{T}_w^c = (\mathbf{T}_c^w)^{-1}$. In this setup, the ceiling-mounted camera is not parallel to the robot motion plane, which means that the depth parameter is time-varying. Hence, previous vision-based controllers cannot be applied in this case, but our scheme can handle this problem effectively. In this simulation, the values $\alpha_1 = \pi/18$ rad, $\alpha_2 = \alpha_3 = \pi/4$ rad, $a_1 = a_2 = 1$ m and $a_3 = 3$ m are used, from which we can easily calculate the camera extrinsic parameters $\mathbf{T}_w^c$. From $\mathbf{T}_w^c$, we can respectively obtain its rotational and translational components $\mathbf{R}_w^c = [\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{r}_3]$ and $\mathbf{t}_w^c$. Then, the planar homography $\mathbf{H}$ between the image plane and the robot motion plane can be computed by $\mathbf{H} = \mathbf{A}[\mathbf{r}_1\ \mathbf{r}_2\ \mathbf{t}_w^c]$.

In the simulation, we set $\hat{\alpha}_1(0) = \hat{\alpha}_2(0) = \hat{\alpha}_3(0) = 0$ rad, $\hat{a}_1(0) = \hat{a}_2(0) = 3$ m, $\hat{a}_3(0) = 8$ m, $\hat{f}(0) = 0.35$

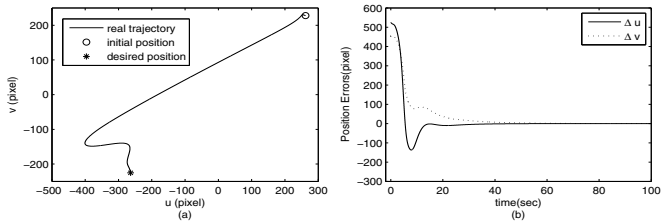

Fig. 2. Simulation results: Trajectory and position errors of the feature point on the image plane. (a) Image trajectory. (b) Image errors.

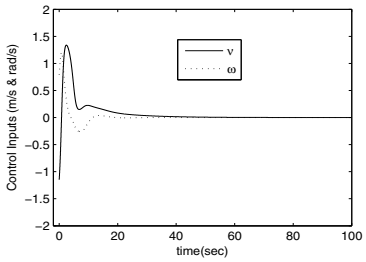

Fig. 3. Simulation results: Controller inputs (ν, ω).

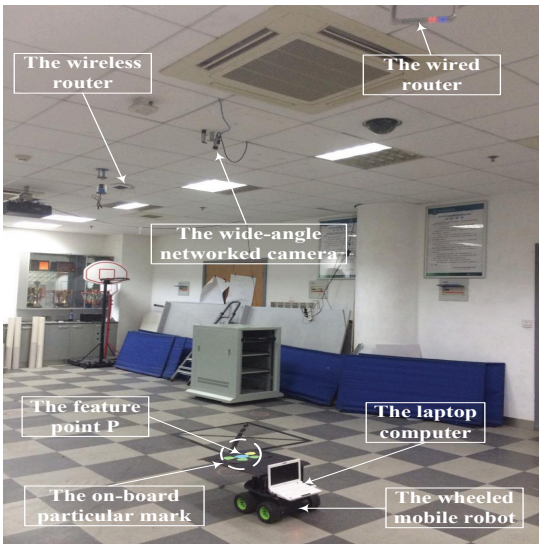

Fig. 4. Experimental system setup.

m, $\hat{u}_0(0) = \hat{v}_0(0) = 200$ pixels, $\hat{k}_u(0) = \hat{k}_v(0) = 1000$ pixels/m to calculate the initial estimation of the planar homography matrix $\hat{\mathbf{H}}(0)$. In addition, the initial estimation of the distance parameter is given by $\hat{d}(0) = 0.8$ m. Then, we can easily obtain the initial value $\hat{\rho}(0)$ by using both $\hat{\mathbf{H}}(0)$ and $\hat{d}(0)$.

Assumed that initial states of the mobile robot are given by $\mathbf{x}(0) = (0, 0)^\top$ and $\theta(0) = 0$. The desired image position of P is acquired when the mobile robot is at its desired position $\mathbf{x}_{pd} = (-2, 7)^\top$ m, from which the desired image position can be calculated as $\mathbf{y}_{pd} = (-263.6081, -225.5603)^\top$ pixels and will be used in the feedback law. The simulation results with the control gains $\mathbf{K}_p = 5 \times 10^{-5}\mathbf{I}_2$ and $\mathbf{\Gamma}^{-1} = 10^{-5}\mathbf{I}_{14}$ are given in Fig. 2 and Fig. 3. It can be clearly seen that convergence of the image errors to zeros is indeed guaranteed even without knowing the camera parameters and d.

Remark 4: Due to the page limitation, the simulation results only show evaluation of the proposed scheme for one test with noiseless data. To further test its robustness and show its superiority, much more simulations should be done by taking simulated camera noises into consideration, and this is one of our main objectives in the future.

V. Experimental Results

To further validate the performance of the proposed scheme, in this section, we will provide experimental results based on a wheeled mobile robot. The experimental system setup is given in Fig. 4. In this setup, a wide-angle networked camera is fixed to the ceiling to track the robot motion, where the angle between the optical-axis and the vertical line is about 30 degrees. The robot motion can be detected by observing a particular mark attached at the mobile robot, which consists of five colored circular blobs. This particular mark is designed in such a way that the robot motion can be robustly tracked even in light-varying conditions. The geometric center of the biggest circular blob is selected as the feature point P. The captured images are transmitted to the host computer via wireless network, where the feature point P is tracked continuously and then its image coordinates are provided to the laptop computer, where the proposed visual servoing algorithm is carried out. Once the controller inputs are generated by the proposed algorithm executed on the laptop computer, they will be sent to the embedded system on-board the mobile robot via a USB port. After receiving commands from the laptop computer, the low-level controller implemented in the embedded system will output the corresponding driving signals for the wheels to drive the mobile robot toward its destination. The previous processes form a feedback loop and execute repeatedly at a frequency of about 10 Hz until the mobile robot has arrived at its goal position with enough precision. The wide-angle camera has been calibrated without taking into account the nonlinear lens distortion effects. But in the calculation of $\hat{\mathbf{H}}(0)$, we only use the approximate intrinsic parameters, and directly use the same initial values $\hat{\alpha}_1(0)$, $\hat{\alpha}_2(0)$, $\hat{\alpha}_3(0)$, $\hat{a}_1(0)$, $\hat{a}_2(0)$ and $\hat{a}_3(0)$ given in the simulation without calibrating the camera extrinsic parameters. The distance of P from the geometric center of the mobile robot is approximately 0.4 m, while $\hat{d}(0)$ in the simulation is used in the experiment. By combining $\hat{\mathbf{H}}(0)$ and $\hat{d}(0)$, the initial estimation $\hat{\rho}(0)$ can be obtained and will be updated by the proposed adaptive law.

In the implementation, the control gains are given by $\mathbf{K}_p = diag(0.003, 0.003)$ and $\mathbf{\Gamma}^{-1} = 10^{-6}\mathbf{I}_{14}$. Satisfied experimental results are given in Fig. 5 and Fig. 6. From Fig. 5, we can see that, even in the presence of large camera lens distortion effects, the proposed scheme guarantees that the mobile robot can be driven to its target position. In other words, the image errors are asymptotically convergent to zeros, which further confirms the validity of our approach.

Remark 5: Though the mobile robot moves on a planar ground, the feature depth is still time varying since in the

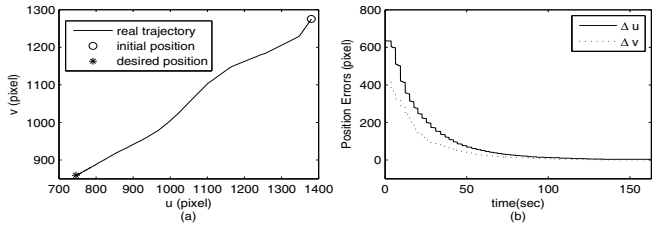

Fig. 5. Experimental results: Trajectory and position errors of the feature point on the image plane. (a) Image trajectory. (b) Image errors.

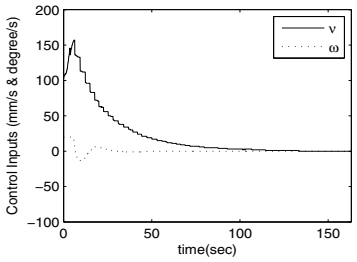

Fig. 6. Experimental results: Controller inputs (ν, ω).

experiment, the camera is in a general position with respect to the ground (the camera image plane is not parallel to the ground plane). The variation level of the depth will depend on the angle δ between the camera optical axis and the vertical line, and δ is about 30 degrees in the current experimental results. The variation of the depth will become larger when a bigger δ is used. To further test the performance of the proposed scheme, in the future, we will carry out more experiments for different δ in a more systematic way.

Remark 6: As we can see from the attached video of the experiment, we obtain the desired image position of the mobile robot by placing it at the target location. This is known as the "teach-by-showing" approach. Actually, the proposed scheme can work if a big black cross is selected as the target location while the mobile robot is never placed there before. In the current work, we just used another different Human Machine Interface (HMI) design methodology.

VI. Conclusion

In this paper, we presented an adaptive image-based kinematics-level visual servoing approach to visually guide the wheeled mobile robots with ceiling-mounted camera, which can efficiently deal with both the unknown camera parameters and the unknown position of feature points. The proposed scheme can handle the positioning control problem of mobile robots and can allow the depth parameters of feature points to be time varying. Image errors were shown to be asymptotically convergent to zeros by theoretical analysis so that the mobile robot can be driven to its desired position. Simulation results based on a two-wheeled mobile robot were also provided to show the satisfying performance of the proposed method. To further validate the performance of the proposed scheme, experimental tests on real robotic systems were also provided. It should be pointed out that we only consider the adaptive vision-based positioning problem without considering the orientation. In the future, we will extend the proposed scheme to solve the stabilization problem of simultaneous control of position and orientation for wheeled mobile robots.

References

[1] Y. Masuutani, M. Mikawa, N. Maru, and F. Miyazaki, "Visual servoing for nonholonomic mobile robots," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, Munich, Germany, Sep. 1994, pp. 1133–1140.

[2] N. R. Gans and S. A. Hutchinson, "A stable vision-based control scheme for nonholonomic vehicles to keep a landmark in the field of view," in *Proc. IEEE International Conference on Robotics and Automation*, Roma, Italy, Apr. 2007, pp. 2196–2201.

[3] F. Conticelli, B. Allotta, and P. K. Khosla, "Image-based visual servoing of nonholonomic mobile robots," in *Proc. IEEE International Conference on Decision and Control*, Phoenix, Arizona, USA, Dec. 1999, pp. 3496–3501.

[4] G. L. Mariottini, G. Oriolo, and D. Prattichizzo, "Image-based visual servoing for nonholonomic mobile robots using epipolar geometry," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 87–100, Feb. 2007.

[5] G. Lopez-Nicolas, C. Sagues, and J. J. Guerrero, "Parking with the essential matrix without short baseline degeneracies," in *Proc. IEEE International Conference on Robotics and Automation*, Kobe, Japan, May 2009, pp. 1098–1103.

[6] Y. Fang, D. M. Dawson, W. E. Dixon, and M. S. deQueiroz, "Homography-based visual servoing of wheeled mobile robots," in *Proc. IEEE International Conference on Decision and Control*, Las Vegas, Nevada, USA, Dec. 2002, pp. 2866–2871.

[7] Y. Fang, W. E. Dixon, D. M. Dawson, and P. Chawda, "Homography-based visual servo regulation of mobile robots," *IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, vol. 35, no. 5, pp. 1041–1050, Oct. 2005.

[8] J. Chen, W. E. Dixon, D. M. Dawson, and M. McIntyre, "Homography-based visual servo tracking control of a wheeled mobile robot," *IEEE Trans. Robot.*, vol. 22, no. 2, pp. 407–416, Apr. 2006.

[9] G. Lopez-Nicolas, S. Bhattacharya, J. J. Guerrero, C. Sagues, and S. Hutchinson, "Switched homography-based visual control of differential drive vehicles with field-of-view constraints," in *Proc. IEEE International Conference on Robotics and Automation*, Rome, Italy, Apr. 2007, pp. 4238–4244.

[10] Y. Fang, X. Liu, and X. Zhang, "Adaptive active visual servoing of nonholonomic mobile robots," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 1, pp. 486–497, Jan. 2012.

[11] X. Zhang, Y. Fang, and X. Liu, "Motion-estimation-based visual servoing of nonholonomic mobile robots," *IEEE Trans. Robot.*, vol. 27, no. 6, pp. 1167–1175, Dec. 2011.

[12] C. Wang, W. Niu, Q. Li, and J. Jia, "Visual servoing based regulation of nonholonomic mobile robots with uncalibrated monocular camera," in *Proc. IEEE International Conference on Control and Automation*, Guangzhou, China, May 2007, pp. 214–219.

[13] C. Wang, "Visual servoing feedback based robust regulation of nonholonomic wheeled mobile robots," in *Proc. IEEE International Conference on Robotics and Automation*, Shanghai, China, May 2011, pp. 6174–6179.

[14] Q. Li, C. Wang, and W. Niu, "Tracking of nonholonomic control systems based on visual servoing feedback," in *Proc. Chinese Control Conference*, Zhangjiajie, Hunan, China, Jul. 2007, pp. 459–463.

[15] J. A. Piepmeier, "Uncalibrated vision-based mobile robot obstacle avoidance," in *Proc. the 33rd Southeastern Symposium on System Theory*, Athens, OH, Mar. 2001, pp. 251–255.

[16] F. Yang and C. Wang, "Adaptive stabilization for uncertain nonholonomic dynamic mobile robots based on visual servoing feedback," *Acta Automatica Sinica*, vol. 37, no. 7, pp. 857–864, Jul. 2011.

[17] W. E. Dixon, D. M. Dawson, E. Zergeroglu, and A. Behal, "Adaptive tracking control of a wheeled mobile robot via an uncalibrated camera system," *IEEE Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, vol. 31, no. 3, pp. 341–352, Jun. 2001.

[18] E. Malis, F. Chaumette, and S. Boudet, "2d 1/2 visual servoing," *IEEE Trans. Robot. Autom.*, vol. 15, no. 2, pp. 234–246, Apr. 1999.