Unmanned Systems, Vol. 1, No. 1 (2013) 143–162

DOI: 10.1142/S2301385013300011

UNMANNED SYSTEMS
http://www.worldscientific.com/worldscinet/us

Developments in Visual Servoing for Mobile Manipulation

Haoxiang Lang[*,§], Muhammad Tahir Khan[†,¶], Kok-Kiong Tan[‡,‖],
Clarence W. de Silva[*,**]

[*]Department of Mechanical Engineering, University of British Columbia, 2054-6250 Applied Science Lane, Vancouver, BC, V6T 1Z4, Canada

[†]Institute of Mechatronics University of Engineering and Technology, Peshawar, Pakistan

[‡]Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3 Singapore 117576, Singapore

A new trend in mobile robotics is to integrate visual information in feedback control for facilitating autonomous grasping and manipulation. The result is a visual servo system, which is quite beneficial in autonomous mobile manipulation. In view of mobility, it has wider application than the traditional visual servoing in manipulators with fixed base. In this paper, the state of art of vision-guided robotic applications is presented along with the associated hardware. Next, two classical approaches of visual servoing: image-based visual servoing (IBVS) and position-based visual servoing (PBVS) are reviewed; and their advantages and drawbacks in applying to a mobile manipulation system are discussed. A general concept of modeling a visual servo system is demonstrated. Some challenges in developing visual servo systems are discussed. Finally, a practical application of mobile manipulation system which is developed for applications of search and rescue and homecare robotics is introduced.

Keywords: Mobile robots; visual servoing; manipulation.

US

1. Introduction

In recent years, the emphasis of robotic research appears to have shifted from the development of robots for structured industrial environments to the development of autonomous mobile robots operating in unstructured and partially known natural environments, such as homes, planet surfaces and disaster scenes. These autonomous mobile robots are applicable in a number of challenging practical tasks such as cleaning of hazardous material, surveillance, rescue, and reconnaissance in unstructured environments which can be too hazardous for humans. It is expected that this new class of mobile robots will have extensive applications in activities where human capabilities are needed, yet not suitable or impractical for human presence.

An example for the necessity of autonomous search and rescue robots was highlighted in the 2001 attack and destruction of the World Trade Center (WTC) in New York City which resulted in the death of 343 firefighters and 65 other rescuers. A significant loss of life was caused in this situation due to search and rescue operations in an unsafe environment. Generally speaking, rescue workers have about 48 h to retrieve trapped humans subsequent to a disaster. However, much time is wasted because of the lack of necessary equipment and resources for accessing collapsed buildings or generally unsafe areas. Robotic rescuers will be able to carry out rescue operations more efficiently without further endangering human life.

Ten years after the WTC incident, an earthquake followed by a tsunami, in Sendai, Japan, resulted in a nuclear reactor crisis. It was another situation where robotic search and rescue would have been effective. Unfortunately, there was no report of the use of robots in that situation even though Japan is a leader in robotic applications. These facts highlight that there are still serious challenges and unsolved technical difficulties in the field of autonomous mobile robotic manipulation.

Received 22 January 2013; Revised 5 March 2013; Accepted 1 April 2013; Published 20 May 2013. This paper was recommended for publication in revised form by the Editors-in-Chief.
Email Addresses: [§]*hxlang@mech.ubc.ca*, [¶]*tahir@nwfpuet.edu.pk*, [‖]*eletankk@nus.edu.sg*, [**]*desilva@mech.ubc.ca*

Another good application of mobile manipulation systems is in homecare environments. Since the overall average age of the world population is growing, the percentage of nonworking elders is increasing with respect to the working population. In Canada, statistics show that the percentage of the senior population will reach about 25% of the overall population in 2050, while that percentage was a mere 8% in 1971. This dramatic change will cause financial problems (in particular, more public funds will be needed for healthcare and homecare) as well as social problems (in particular, a shortage of human caregivers and its effect on families). An effective solution to this problem is to employ autonomous mobile robots with adequate task capabilities and intelligence to carry out many of the needed tasks. Robots can be designed for taking care of elders (and the disabled) in their own home environments, providing assistance in daily activities, medical needs, surveillance, and so on. Not having to relocate the care receiver from the familiar home environment is a clear benefit in this regard. Also, round-the-clock care can be provided by a robotic system, in a consistent manner. Furthermore, the family members will have the freedom and peace of mind to pursue their own activities such as employment and education while being able to monitor the home scenario on line with the use of the system sensors and communication links.

As investigated in the present work, a mobile robotic manipulation system is rather complex and involves multi-domain technologies such as computer vision, artificial intelligence, mechanical and electrical design, signal processing, sensing and control. The objective of this paper is to review some important applications of vision-guided robotics systems, especially visual servoing. The state of art of vision-based mobile robotics and the associated hardware will be presented along with some targeted applications. Traditional visual servo methodologies and underlying challenges in applying them in autonomous mobile manipulation will be discussed. The integration of a vision system into a robot controller will be introduced along with the basic configurations of cameras in such systems. A general concept of controller design for visual servoing will be demonstrated. As an application, a mobile manipulation system, which has been developed in the Industrial Automation Laboratory of the University of British Columbia, for autonomous mobile rescue robotics will be demonstrated.

2. Vision-Guided Robotics Systems

Artificial vision is a powerful sensor as it mimics human sense of vision and allows noncontact measurement of a work environment. Moreover, it can generate a variety of sensory information such as position, distance, and features needed for object identification. Because of its advantages, researchers have paid attention to computer vision and practical applications. Since the invention of the camera, much effort has gone to applying vision as a feedback sensor in industrial control systems. Vision guided robotics, which integrates a vision system into a robot system, has been a major research area for decades. A common application in this field is the inspection of manufactured goods such as semiconductor chips, automobiles, food products, and pharmaceuticals. Other recent applications in this area involve vision-based object detection, monitoring of UGVs (Unmanned Ground Vehicles), tracking of mobile robots, and vision-based feedback control of robotic manipulator movements (visual servoing).

2.1. *State of the art of mobile robotic systems*

Among the projects of vision guided robotics applications, well known is the "DARPA Urban Challenge". This project promoted the development of autonomous mobile robots for a specific purpose which is autonomous driving of a vehicle in an urban environment. These robots were designed and developed to adapt to drive in urban traffic while performing complex maneuvers such as merging, passing, parking and negotiating intersections. Figure 1 shows some of the examples in the DARPA Urban Challenge. In these vehicles, camera is a primary sensor that provides feedback from the environment to the vehicle control system. In Refs. 1 and 2, the details of design and development of hardware and software for such vehicles were introduced and discussed.

Current applications of vision guided robotics and associated hardware are mainly intended for mobile navigation. Ma in 1999 developed a vision-guided navigation system where a nonholonomic mobile robot tracked an arbitrarily shaped continuous curve on ground.[3] They formulated this problem as one of controlling the shape of a curve on the image plane, and presented the corresponding control laws. Fang *et al.* in 2005 presented an adaptive tracking controller of a wheeled mobile robot, which used an un-calibrated camera system.[4] In their paper, the parameter uncertainty of the mechanical dynamics and the camera system was considered, and an adaptive controller was proposed to cope with this uncertainty. In order to implement robust vision-based autonomous navigation systems, various approaches have been explored such as planning of image-trajectory or image-memory,[5,6] embedded velocity fields,[7] specific geometry features (vanishing points and line orientations[8]), stereo cameras, and omni-directional or catadioptric cameras.[9–11] The possible reason for this limited activity is that vision-based mobile manipulation requires accurate and robust positioning performance

Fig. 1. Some example vehicles in DARPA Urban Challenge.

and hence it is more challenging than vision-based mobile robot navigation.

Vision-based automated manipulation has been playing a significant role in industrial and service applications for years. Many of the related applications have been reviewed in Refs. 12 and 13. However, they considered fixed base robotic manipulators only, which are typically applied in structured and known work environments. The new trend is to integrate visual servoing into a mobile robot for carrying out the activities of grasping or manipulation autonomously, resulting in a vision-based autonomous mobile manipulation system. Compared to traditional visual-servo control that is employed in fixed base robotic manipulators, a vision-based mobile manipulation system has many advantages. Obvious among them is that the robotic system becomes more flexible and has wider applications than a traditional fixed base manipulator system. Since the robotic arm (manipulator) and the cameras are usually mounted on a mobile base, a mobile manipulation system possesses better maneuverability and terrain coverage capability than a fixed base manipulator. The historical robotic systems that have been developed since the 1980s are reviewed in Ref. 14. In general, these robotic systems have the conventional architecture of four subsystems: mobile platform, robotic manipulator, tooling and sensors.

Figure 2 shows some of the newly developed mobile manipulation systems for different applications. They are Willow Garage PR2[15,16] and DLR Rollin' Justin[17] of homecare and service robots; NASA/GM Robonaut[18] and NASA Mars Rover[19] for space robotics; TALON IV engineer[20] and iRobot-PackBot EOD[21] which are applied in military; and KUKA OmniRob[22] and Neobotix[23] in applications of industrial robotics.

2.2. *Typical sensors*

Many off-the-shelf sensors (e.g., GPS, compasses, gyroscopes, and ultrasonic sensors) that are available for mobile robots are introduced in Ref. 24, giving their operating principles and performance limitations. Camera and laser range finder are the most commonly used sensors in mobile manipulation systems. A vision system is passive and has high resolution. With the help of advanced machine vision technologies, it can provide the location information of objects in the workspace using the vision information acquired from camera scenes. It appears to be the most promising sensor category for future generations of mobile robots.

Two general camera configurations are used in visual servo systems: eye-to-hand configuration and camera-in-hand configuration, which are shown in Fig. 3. In the eye-to-hand configuration (Fig. 3(a)), the camera is fixed in the workspace and observes the robotic "hand." Therefore, the field of view of the camera does not change. The camera keeps tracking the positions of both the robot end-effector and the target object in the image plane during the visual servo operation. In this configuration, the geometric relationship among the camera, robot and target object is fixed, the advantage of which is straightforward environmental modeling and 3D reconstruction. However, the end-effector can easily occlude the camera field of view and block the target object in the image plane. Moreover, since the camera is fixed in the environment, its field of view cannot be changed, which is a limitation in mobile manipulation applications.

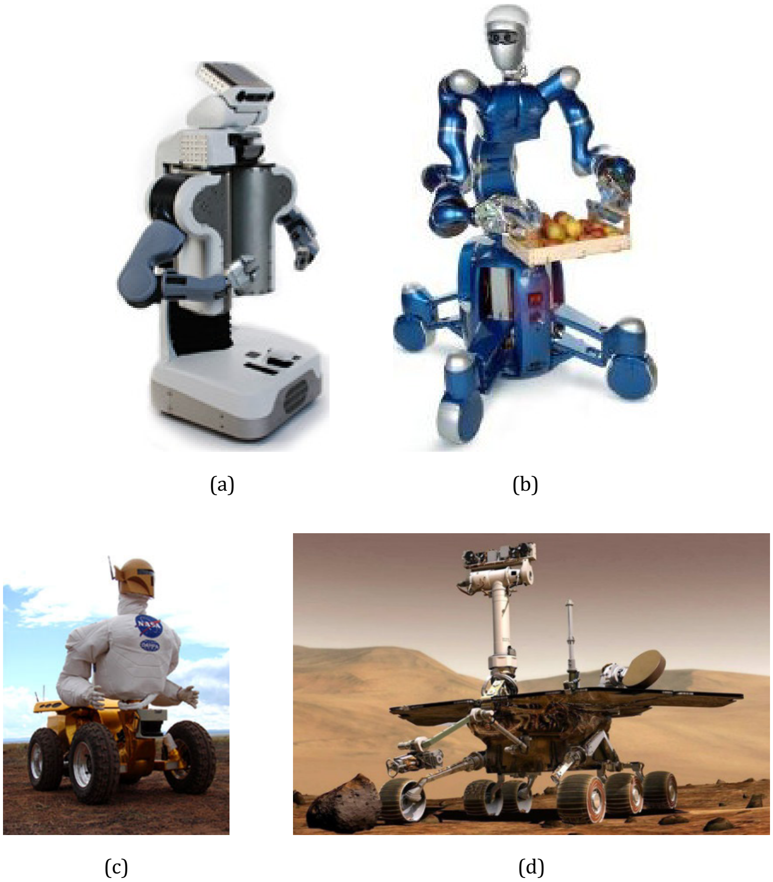


Fig. 2. Examples of mobile robot manipulation systems in different application areas. Service Robots: (a) Willow Garage PR2; (b) DLR Rollin Justin. Space Robots: (c) NASA/GM Robonaut; (d) NASA Mars Rover. Military Robots: (e) TALON IV engineer; (f) iRobot-PackBot EOD. Industrial Robots: (g) KUKA OmniRob; (h) Neobotix.

(e)

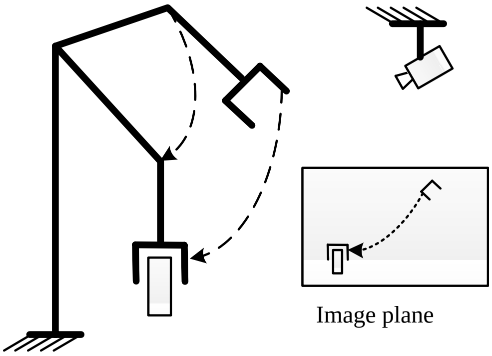
(f)

(g)

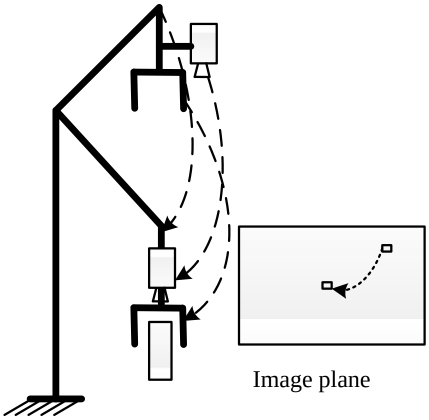
(h)

Fig. 2. (*Continued*)

In the eye-in-hand configuration, the camera is mounted at the end-effector of the robot. In this configuration, the camera will move with the robot. Therefore, the field of view of the camera changes constantly and depends on the position and orientation of the end-effector, which is an advantage. It is suitable in applications of visual servoing of mobile robots. Moreover, it can avoid the problem of occlusion since the camera is mounted on the end-effector



(a)



(b)

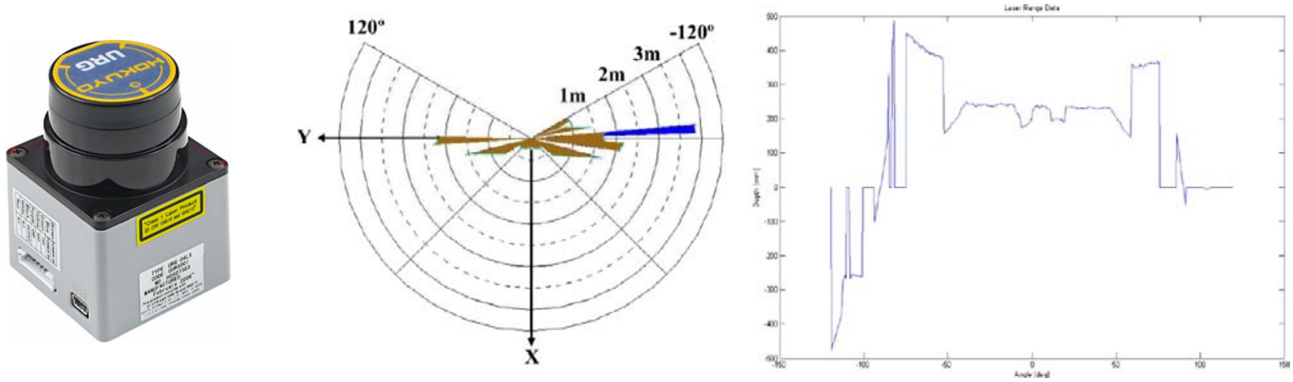
Fig. 3. (a) Eye-to-hand configuration; (b) Eye-in-hand configuration.

of the robot. However, it has disadvantages as well. For example, the geometry of the camera, robot and the target object changes as the end-effector moves, and the camera view may dramatically change even for small movements of the robot.

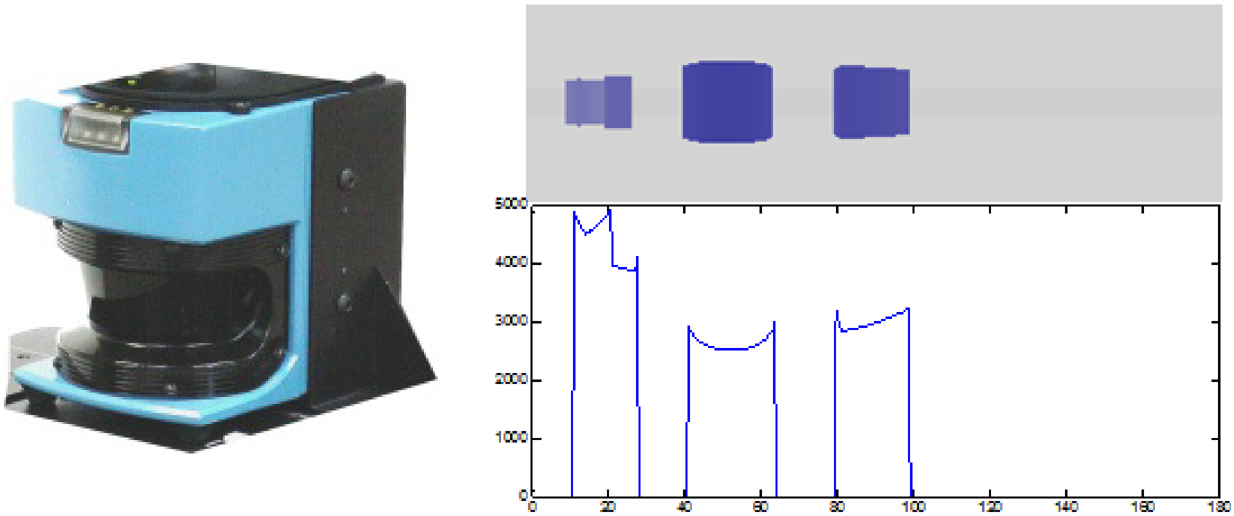
In the case of a mobile robot, the base frame is not fixed, and the base coordinate frame changes with time as the robot navigates. A pre-defined approach to manipulation, as typically used in industry, will not work in this case. Therefore, a camera fixed in the environment, which works as an eye-to-hand configuration, is not suitable because it is typically not possible to cover the entire workspace of the mobile robot by using a single camera. Therefore, the eye-in-hand configuration is more suitable for mobile manipulation applications.

In addition to the vision information, the distance information is also crucial in mobile manipulation. Sonar and laser distance finders are two typical sensors for sensing distance. Sonar is fast and inexpensive but is usually rather crude, whereas laser scanning is active, accurate and widely applied in mobile robotics. Figure 4 shows two type of laser distance finders and their sensing results. Hokuyo URG-04LX sensor can be thought of as miniature sonars that use light instead of sound to create 2-D maps of the neighborhood of objects close to the robot. The scanner has an angular scanning range of 240°, and the angular resolution is approximately 0.36° degrees with a scanning refresh rate of up to 10 Hz. Distances ranging from 20 mm to 4 m may be measured. The SICK LMS 200 2D scanner has a horizontal range of 180° with a maximum resolution of 0.5°.

Another way of measuring distance is to utilize multiple cameras. This approach is termed stereo vision. Figure 5 shows two commercial stereo vision sensors: BumbleBee 2 and Microsoft Kinect.

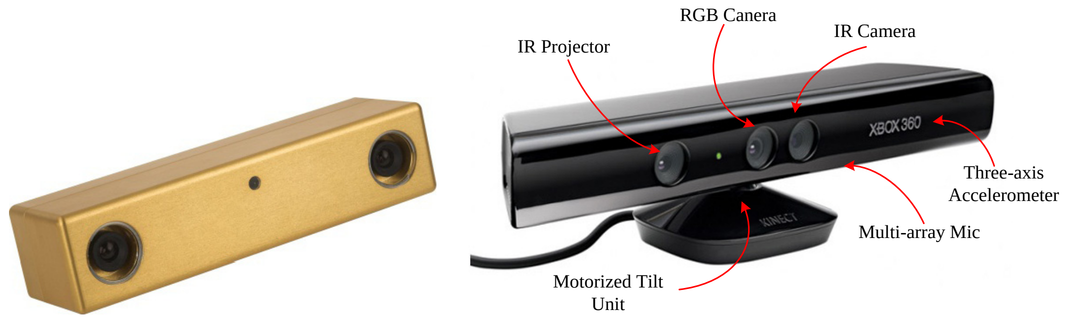


(a)



(b)

Fig. 4. Typical laser distance finders: (a) The Hokuyo URG-04LX sensor; (b) SICK LMS 200 2D scanner.



(a) (b)

Fig. 5. Stereo vision sensors: (a) BumbleBee®2 stereo camera; (b) Microsoft Kinect.

An important benefit of a stereo camera is that it can provide object identification and distance data simultaneously. A stereo camera has two sets of cameras as shown in Fig. 6. Suppose that a point P is observed by two cameras simultaneously. Its projections to the image planes of the two cameras are at p and p', as shown. Note that O and O' are the optical centers of these two image planes. The line l', which is in image II, is called the epipolar line associated with point p of image I. Also, eL and eR are called epipoles of the two cameras. The epipolar e_R is the projection of the optical center O_L in the right camera frame.

Epipolar constraint[25] states that if p and p' are the projections of the same point P in different cameras, then p' must lie on the epipolar line associated with p. This result plays a fundamental role in stereo vision. The epipolar geometry can be simplified if the two camera image planes coincide, as shown in Fig. 7. In this case, the epipolar lines also coincide. Furthermore, the epipolar lines are parallel to the line between the focal points, and can in practice be aligned with the horizontal axes of the two images. This means that for each point in one image, its corresponding point in the other image can be found by looking only along a horizontal line.

The distance can be calculated by measuring the disparity, which is defined as

$$d = u_L - u_R = f\frac{B}{Z} \rightarrow Z = f\frac{B}{d}. \tag{1}$$

Here, Disparity $= k/\text{depth}$, and k is a system parameter, which can be acquired by calibration. By introducing the epipolar geometry and its constraints, the stereo vision problem is converted into a feature identification problem. In other words, the distance between the camera and the object in terms of the camera coordinates can be found by finding the position difference of the object in the right and left images (Fig. 7). Some applications of stereo vision are given in Refs. 26–28.

3. Basic Categories of Visual Servoing

A vision-based mobile manipulation system typically uses computer vision data as sensory feedback for controlling the motion of the robot. This is also called a visual servo system or a visual servo control system. It involves a fusion of many related technologies including machine vision, robot modeling, control theory, and real-time computing.

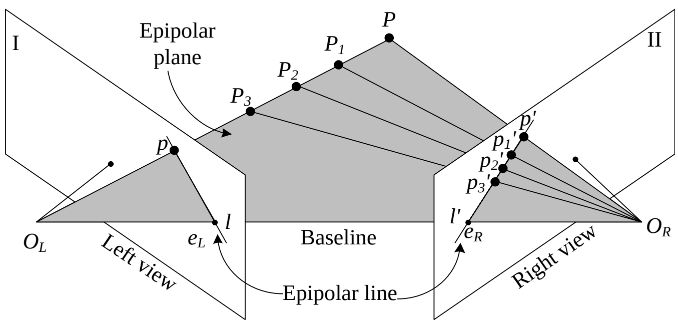


Fig. 6. Model of a stereo camera.

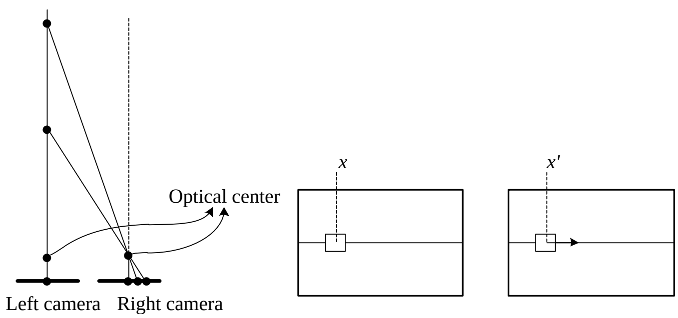


Fig. 7. Simplified case of epipolar geometry.

Specifically, in visual servoing, the vision system provides feedback information about the current state of the environment to the controller.[29]

The earliest research of visual servoing was reported in 1979.[30] However, the image processing procedure took seconds due to the limitation of the available computers and image sensing devices at that time, making real-time control virtually infeasible. Subsequently, in 1980, Sanderson and Weiss introduced a taxonomy of visual servo systems through a control structure.[31] It was called Dynamic Look-and-Move Structure where the control architecture was hierarchical and used vision to provide set-point inputs to the joint-level controller. The sub-control system utilized the joint feedback to internally stabilize the robot (Fig. 8).

With the rapid advancement of computer technologies and image sensing hardware (CCD and CMOS), computer vision is much faster now than it was in the 1980s. Direct visual servoing came to the attention of researchers in the 1990s. Direct visual servoing utilizes a visual servo controller which directly relies on vision information to compute joint inputs, thereby stabilizing the robot (Fig. 9). Hutchinson has reviewed many of the related work.[29] Since then the term "visual servoing" has come to be accepted as a generic description of any type of visual "feedback control" of a robotic system. The subject has been under study in various forms for more than 20 years, in contexts ranging from simple pick-and-place tasks to today's real-time, complex tasks involving multiple robots and objects, and autonomous cooperation, in dynamic, unstructured and unknown environments.

3.1. *Image-based visual servoing (IBVS)*

The classical visual servoing can be classified into the image-based approach and the position-based approach.[29] The two approaches share a common control block diagram with a slight difference in the control feedback loop and the reference input (Fig. 10).

The goal of IBVS is to eliminate the position errors of the feature points in the image plane by controlling the velocities of the robot joints. The underlying concept of IBVS is that the controller will continuously adjust the speeds of the robot joints so that the coordinates (u_i, v_i) of the feature points on the image are moved toward the desired position (u_{d_i}, v_{d_i}) on the image. In particular, the error vector of the image feature point is given by

$$e = \begin{bmatrix} u - u_d \\ v - v_d \end{bmatrix} = \begin{bmatrix} -s_x(r - r_d) \\ -s_y(c - c_d) \end{bmatrix}, \tag{2}$$

where s_x and s_y are the dimensions of the pixels of the image sensor, which is a charge-coupled device (CCD) in the present case. Also, r and c are the vision measurement of feature points in the image, and r_d and c_d are the desired positions of the feature points in the image.

Figure 11 shows the physical representation of the feature points in the IBVS system where the "*" are the desired positions and "o" are the current positions of the feature points in the camera scene (image). The operating principle of the system is to control the motion of the joints so that the feature points navigate from the current positions to the goal location. Some IBVS applications are reported in Refs. 32–34

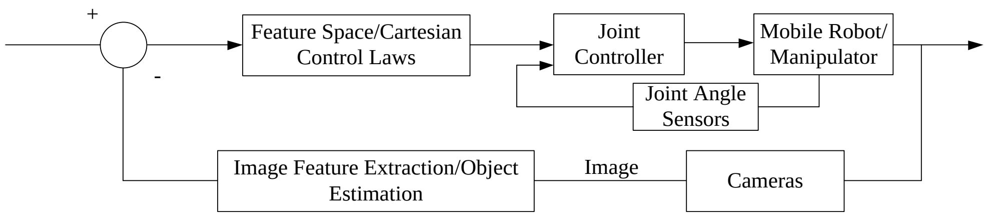


Fig. 8. Block diagram of a dynamic look-and-move system.

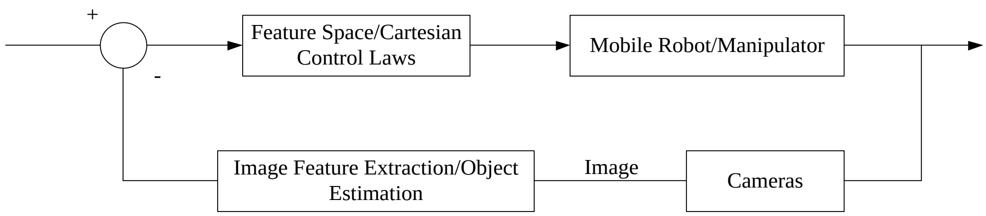


Fig. 9. Block diagram of a typical (direct) visual servo system.

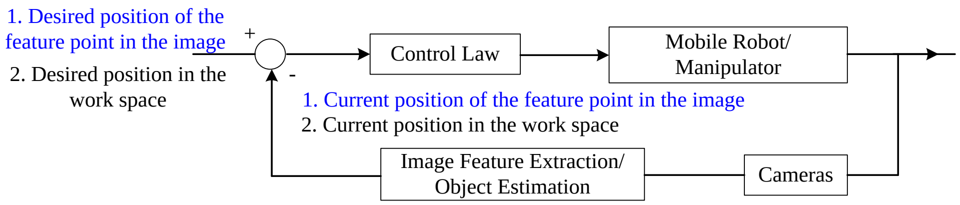


Fig. 10. Block diagram of a typical visual servo system (1. Position-based; 2. Image-based).

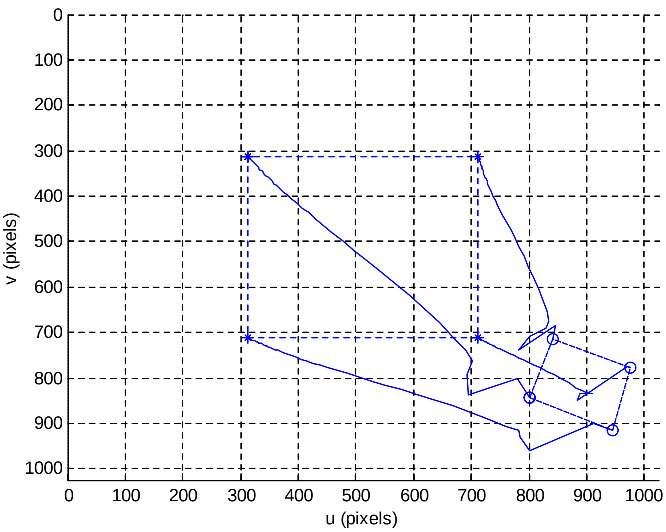


Fig. 11. The concept of IBVS.

3.2. *Position-based visual servoing (PBVS)*

In PBVS,[35,36] features are extracted from the images obtained by one or more cameras, and used in conjunction with the camera model and a geometric model of the target object to estimate the pose of the camera with respect to the object. Three coordinate frames are involved in the control system: current frame f_c, desired frame f_d, and reference frame f_r. The reference frame is usually attached to the object in the workspace. The controller seeks to reduce the error between the current pose and the desired pose in a 3D (three-dimensional) workspace by adjusting the speed of the robotic joints. Since the control scheme is defined in the Cartesian space (3D workspace) for PBVS, it is usually referred to as 3D visual servoing, while IBVS is called 2D visual servoing since it only considers the feature points in the image (2D workspace). A feature in PBVS can be defined as:

$$s = (T, \theta U), \tag{3}$$

where T is the translation vector and θU is the orientation vector. The error dynamic equation can be derived as:

$$e = (T_r^c - T_d, \theta U). \tag{4}$$

The interaction matrix may be written as:

$$L = \begin{bmatrix} -I_3 & [T_r^c]_\times \\ 0 & L_{\theta U} \end{bmatrix}, \tag{5}$$

where $L_{\theta U}$ is given by Ref. 35:

$$L_{\theta U} = I_3 - \frac{\theta}{2}[U]_\times + \left(1 - \frac{\operatorname{sinc}\theta}{\operatorname{sinc}^2\frac{\theta}{2}}\right)[U]_\times^2, \tag{6}$$

where $\sin cx$ is the sine cardinal defined such that $x \sin cx = \sin x$ and $\sin c0 = 1$. The velocity control scheme may be written as:

$$v_c = -kL^{-1}e. \tag{7}$$

Figure 12 shows a physical experimentation of PBVS. There are four feature points in the workspace which represent the object of interested. The system controls the motions of the robot joints in order to orient the camera toward a goal pose from the current pose in 3D workspace.

Both IBVS and PBVS have their advantages and disadvantages. In applications of autonomous mobile rescue robotics, the workspace is usually unknown; the base frame of the robot is always moving and changing; and the objects are not predefined well. Therefore, the 3D models of the targets are unavailable. In this situation, IBVS has the advantages of reducing the computational burden, omitting unnecessary image interpretation, and eliminating the calibration errors in sensors and cameras; and the eye-in-hand camera configuration is more suitable than the eye-to-hand configuration. However, the IBVS system introduces nonlinear and time varying terms into the system model which makes it a more complex control problem. Moreover, obtaining the depth information is a challenging problem.

4. Modeling of Visual Servo System

Modeling of mobile manipulation system usually consists of kinematic modeling of robots and modeling of cameras. The former incorporates extensions of the principles of rigid motions and homogeneous transformation,[37] and the latter incorporates modeling of perspective projection of a pinhole camera.[38]

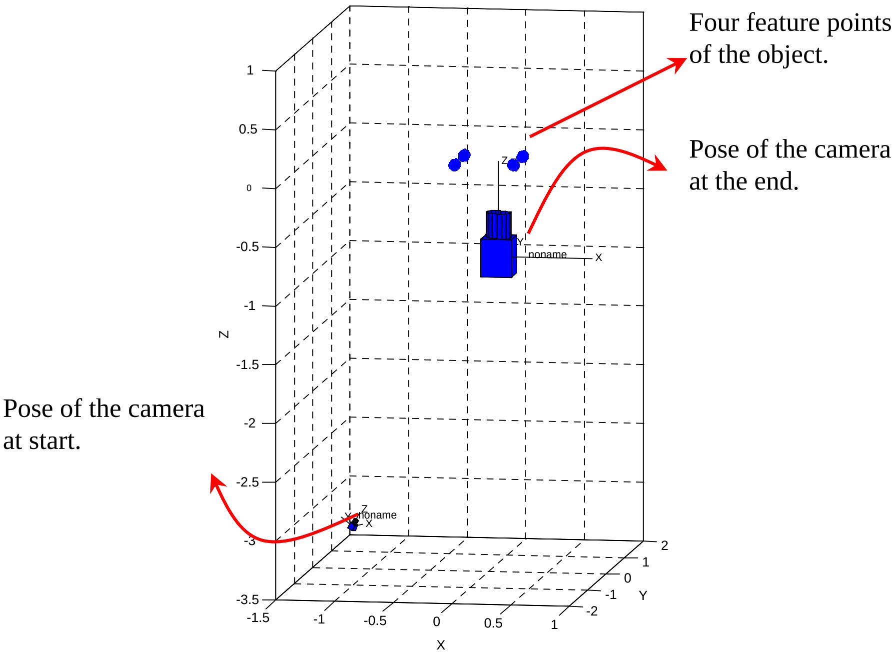


Fig. 12. The concept of PBVS.

4.1. *Kinematic modeling of robots*

Traditionally, the rigid motions of a robot can be represented by the matrix

$$H = \begin{bmatrix} R & d \\ 0 & 1 \end{bmatrix}, \quad R \in SO(3), \quad d \in \mathbb{R}^3, \tag{8}$$

where R represents a 3×3 rotational matrix; and d is a distance vector. The matrix in Eq. (8) is called a homogeneous transformation, which represents the rotational and displacement relationships between two coordinate frames. It is a 4×4 matrix. The last row of the matrix consists of three zeros and a one; and remaining elements are composed of a rotation matrix and a position vector.

In robotic applications, a commonly used convention for defining the coordinate frames of reference and generating the homogeneous transformation matrices is the Denavit–Hartenberg (DH) convention. It was introduced by Denavit and Hartenberg[13] to simplify the modeling procedure in generating forward kinematics (i.e., expressing the end-effector movement in terms of the joint movements) of a robot. There are four parameters in this representation: θ_i, d_i, a_i and α_i. The detailed definitions of these four parameters are given in Table 1.
Suppose that the coordinates are assigned based on the two DH rules:

(DH1): The axis x_i is perpendicular to the axis z_{i-1} ($x_i \perp z_{i-1}$).

(DH2): The axis x_i intersects the axis z_{i-1} ($x_i \cap z_{i-1} \neq \emptyset$).

Then, there exists a unique homogeneous transformation matrix that takes the coordinates from one frame to the base frame, as given by:

$$\begin{aligned} H &= \mathrm{Rot}_{z,\theta}\mathrm{Trans}_{z,d}\mathrm{Trans}_{x,a}\mathrm{Rot}_{x,\alpha} \\ &= \begin{bmatrix} c\theta_i & -s\theta_i & 0 & 0 \\ s\theta_i & c\theta_i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &\quad \times \begin{bmatrix} 1 & 0 & 0 & a_i \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c\alpha_i & -s\alpha_i & 0 \\ 0 & s\alpha_i & c\alpha_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c\theta_i & -s\theta_i c\alpha_i & s\theta_i s\alpha_i & a_i c\theta_i \\ s\theta_i & c\theta_i c\alpha_i & -c\theta_i s\alpha_i & a_i s\theta_i \\ 0 & s\alpha_i & c\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{9}$$

where $\mathrm{Rot}_{z,\theta}$ is the rotational matrix about the z-axis, $\mathrm{Trans}_{z,d}$ is the translational matrix along the z-axis, $\mathrm{Trans}_{x,a}$ is the translational matrix along the x-axis, and $\mathrm{Rot}_{x,a}$ is the rotational matrix along the x-axis. Figure 13 shows an example of coordinate assignment according to the DH rules where $x_i \perp z_{i-1}$ and $x_i \cap z_{i-1} \neq 0$.

Table 1. Denavit–Hartenberg convention.

Parameter	Axis	Description
θ_i	z_{i-1}	Joint Angle: The angle from x_{i-1} to x_i measured about z_{i-1} — Variable for revolute joints.
d_i	z_{i-1}	Link Offset: Distance along z_{i-1} from o_{i-1} to the intersection of the axes x_i and z_{i-1} — Variable for prismatic joints.
a_i	x_i	Link Length: Distance along x_i from the intersection of the axes x_i and z_{i-1} axes to o_i — Constant perpendicular distance between z_{i-1} and z_{i-1}.
α_i	x_i	Link Twist: The angle from z_{i-1} to z_i measured about x_i — Constant angle between z_{i-1} and z_i.

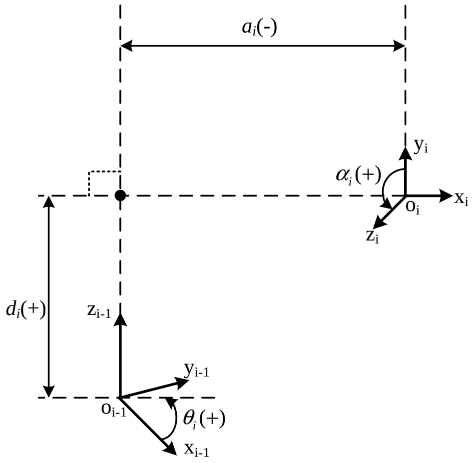


Fig. 13. An example of coordinate frames satisfying DH convention.

If the coordinates are defined according to the DH convention, the forward kinematic model of the robot can be represented in the form given in Ref. 9.

4.2. *Camera modeling*

The simplest model of a camera is the pinhole camera model, which is shown in Fig. 14(a). In this model, light from a point on the object (P) in the world space passes through a pinhole and projects onto the camera image surface called the image plane. For an ideal pinhole camera, the camera plane is located at a distance f (focal length) behind the pinhole. The origin of the coordinate frame is called the center of projection in the image plane. The intersection of the z-axis and the image plane is known as the principal point.

Figure 14(b) shows a pinhole camera model with the reflected image plane model.[24] This model has the advantage that the projected object (x') in the image plane is reversed when compared with the object in the world space (X). Therefore, the modified pinhole camera model simply shifts the image plane to the front of the principal point by

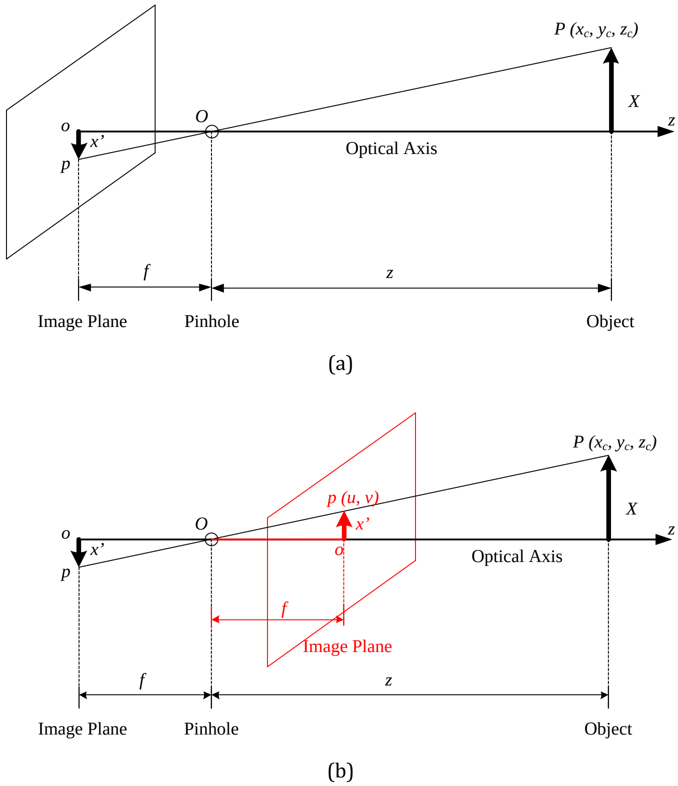


Fig. 14. (a) Pinhole camera model; (b) Pinhole camera model with reflected image plane.

$2f$. In this model, the image reversion problem is resolved. Meanwhile, the size of the projected object does not change, in view of the property of similar triangles. If camera focal length f is known, and the coordinates of the point object in terms of the camera frame is known as well, which is denoted by (x_c, y_c, z_c), we can determine the coordinates of the point in the real world in terms of the camera frame by sensing the position of the projected point in the image plane (u, v):

$$k\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} u \\ v \\ f \end{bmatrix}, \tag{10}$$

where u, v are the coordinates of the projected point in the image plane. It is seen that these equations can help identify the 3D position of a point in the camera frame if u and v are measureable. However, it is not easy to measure these variables. The sensor of a digital camera is a two-dimensional array whose elements are called pixels, as shown in Fig. 15. These pixels are not continuous, and there exist gaps among the pixels. Moreover, a pixel element is usually rectangular and not square, especially in inexpensive cameras. The dimensions of the pixel element are descried by s_x and s_y in Fig. 15. Also, the origin of the image coordinate frame, which is the principal point, is usually not the center of the image/pixel plane. There exist offsets in real situations.

In Fig. 14(b), P is a point in the work environment with coordinates (x, y, z^c) relative to the camera frame, and p is the projection of P on the image plane with coordinates (u, v) relative to the image plane frame and with coordinates (r, c) relative to the pixel coordinate frame. The distance between the origin of the camera frame and the image plane is denoted by λ, and the coordinates of the principal point are (o_r, o_c) with respect to the pixel coordinate frame. The coordinate transformation between the frames is given by:

$$\begin{cases} u = -s_x(r - o_r) \\ v = -s_y(c - o_c) \end{cases}, \tag{11}$$

$$\begin{cases} r = -f_x \dfrac{x^c}{z^c} + o_r \\[2ex] c = -f_y \dfrac{y^c}{z^c} + o_c \end{cases}, \tag{12}$$

where s_x and s_y are the horizontal and vertical dimensions, respectively, of a pixel as given by $f_x = \frac{\lambda}{s_x}$ and $f_y = \frac{\lambda}{s_y}$.

Let p be a feature point on the image plane with coordinates (u, v). The moving velocity of p can be expressed in terms of the camera velocity using the interaction matrix L[39] as:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = L\xi_c^c$$

$$= \begin{bmatrix} -\dfrac{\lambda}{z^c} & 0 & \dfrac{u}{z^c} & \dfrac{uv}{\lambda} & -\dfrac{\lambda^2 + u^2}{\lambda} & v \\[2ex] 0 & -\dfrac{\lambda}{z^c} & \dfrac{v}{z^c} & \dfrac{\lambda^2 + v^2}{\lambda} & -\dfrac{uv}{\lambda} & -u \end{bmatrix} \xi_c^c. \tag{13}$$

In Eqs. (11)–(13), there are several unknowns: the dimensions of the sensor (s_x and s_y), focal length of the camera (f_x and f_y), the location of the principal point (o_r and o_c), and the velocity of the camera with respect to the camera frame (ξ_c^c). These parameters can be determined through camera calibration, which involves finding two groups of parameters:

(1) Intrinsic camera parameters
(2) Extrinsic camera Parameters.

Camera parameters

Since cameras are usually rigidly attached to the robots or mounted on tripods, the rotational matrix R_c^w and the translational vector T_c^w can be easily determined. Therefore, if we know the 3D position of a point P in the camera frame, the coordinates of this point in the world frame can be determined by using

$$P^w = R_c^w P^c + T_c^w \tag{14}$$

or conversely,

$$P^c = R_w^c(P^w - T_c^w) = R_w^c P^w - R_w^c T_c^w, \tag{15}$$

where R_w^c and $R_w^c T_c^w$ are called the extrinsic camera parameters. We will denote them by R and T, respectively, for brevity. Consider the rotational matrix:

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \tag{16}$$

and the translational vector:

$$T = [\,T_x \quad T_Y \quad T_z\,]^T \tag{17}$$

with respect to the world coordinate frame. The coordinates of the point in terms of the camera frame can be found by

$$x^c = r_{11}x^w + r_{12}y^w + r_{13}z^w + T_x, \tag{18}$$

$$y^c = r_{21}x^w + r_{22}y^w + r_{23}z^w + T_y, \tag{19}$$

$$z^c = r_{31}x^w + r_{32}y^w + r_{33}z^w + T_z. \tag{20}$$

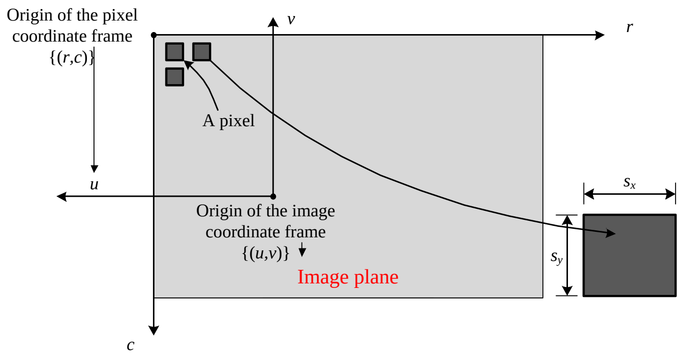


Fig. 15. Image plane and pixel plane.

Substituting (18)–(20) into (12), we obtain

$$r - o_r = -f_x \frac{x^c}{z^c} = -f_x \frac{r_{11}x^w + r_{12}y^w + r_{13}z^w + T_x}{r_{31}x^w + r_{32}y^w + r_{33}z^w + T_z}, \qquad (21)$$

$$c - o_c = -f_y \frac{x^c}{z^c} = -f_y \frac{r_{21}x^w + r_{22}y^w + r_{23}z^w + T_y}{r_{31}x^w + r_{32}y^w + r_{33}z^w + T_z}. \qquad (22)$$

Combining (21) and (22), we obtain:

$$\begin{aligned} &f_x(c - o_c)(r_{11}x^w + r_{12}\,y^w + r_{13}z^w + T_x) \\ &\quad = f_y(r - o_r)(r_{21}x^w + r_{22}\,y^w + r_{23}z^w + T_y). \end{aligned} \qquad (23)$$

This relation has nine unknowns $(f_x, f_y, r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, T_x, T_y)$. Here R and T are the extrinsic parameters, which denote the coordinate transformations from the 3D world coordinates to the 3D camera coordinates. Equivalently, the extrinsic parameters define the position of the camera center and the camera heading in world coordinates, which may be determined by the homogeneous transformation as discussed previously. The intrinsic parameters are determined by a camera calibration tool, as presented in Ref. 40.

An example of intrinsic parameter determination

The intrinsic parameters determine the characteristics of a camera sensor, and they will remain constant for a camera. A chess board (Fig. 16(a)) and the Matlab camera calibration toolbox are utilized to determine the parameters of a Logitech web camera (Fig. 19). The results show that the Logitech camera has the focal length: $f = [562.17889 \quad 522.49795] \pm [1.38737 \quad 1.29051]$; and the principal point: $cc = [332.45990 \quad 236.88157] \pm [2.13199 \ 1.73026]$.

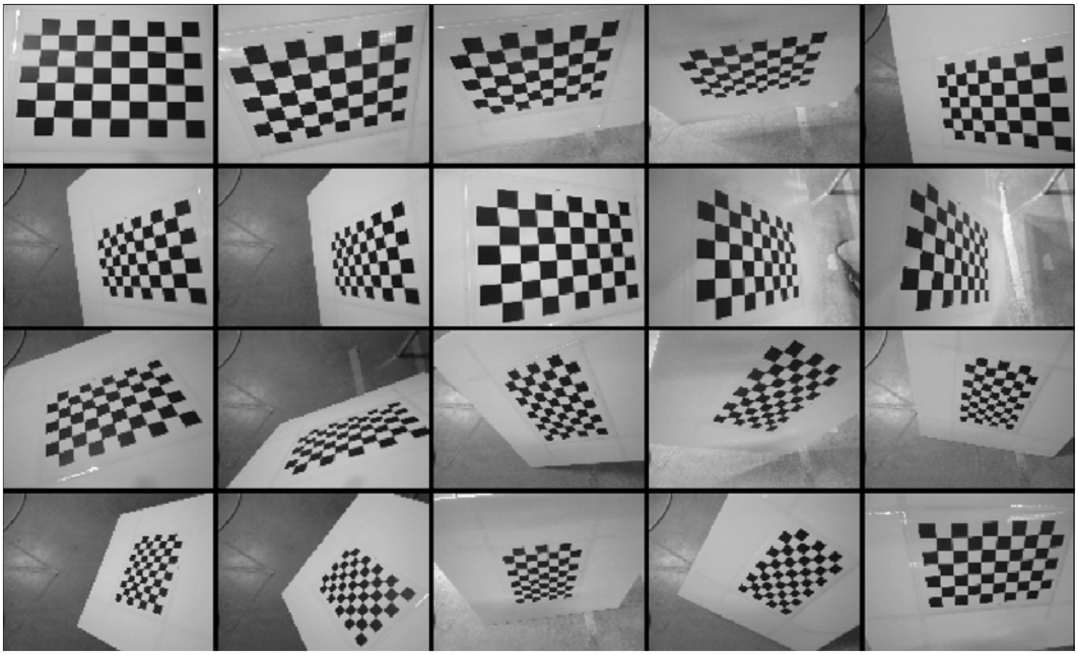
5. Challenges in Visual Servoing

An overview of the process of visual servoing is shown in Fig. 17. In general, the visual servo system acquires and estimates features from the image plane and matches them to the Cartesian space in the camera coordinate frame by utilizing the image Jacobian. Then robot Jacobian is utilized to relate the differential changes between the robot joints and the camera. Finally, the relationship between the joint velocities and the velocities of the feature is generated.

There are several challenges in developing visual servoing for a robot manipulation system. Next, three main challenges are indicated in developing a visual servo controller for mobile manipulation.

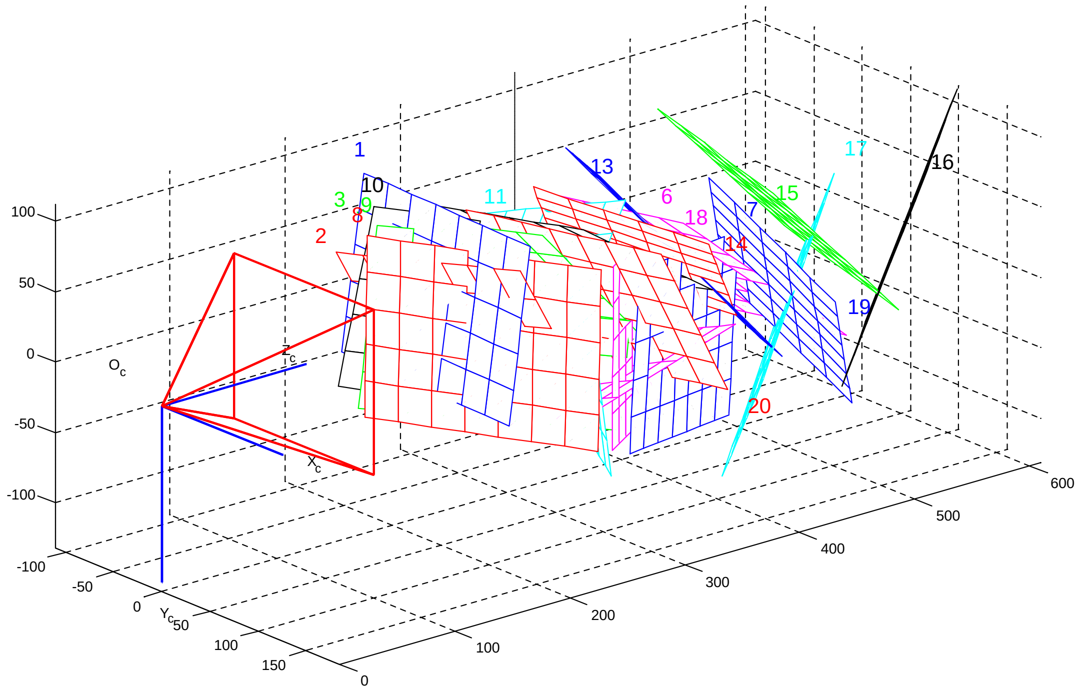
Visual measurements

The manipulation operation heavily relies on the results of identification and tracking because the approach of visual servoing utilizes camera information in the feedback loop of the robot control system. The failure to correctly identify



(a)

Fig. 16. (a) Images for calibration; (b) Camera reference and extrinsic parameters.



(b)

Fig. 16. (*Continued*)

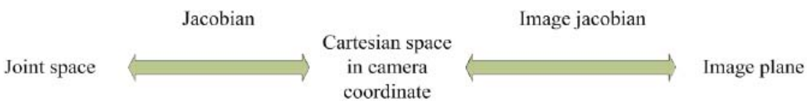


Fig. 17. The process of visual servoing.

and track during manipulation will result in failure of the entire system.

Color is a very useful feature in object detection. Color tracking algorithms are widely used for object identification and tracking in visual servoing because of its simplicity and efficiency. In Ref. 41, a color-based object identification system is developed for object pose estimation in mobile robot application. The system shows promising results when the color of the object is unique in the workspace and the background of the camera scene is relatively clean. However, when the number of objects becomes large and the camera scene becomes complex, it is difficult to decide the threshold of the color tracking system, which leads to poor identification results.

Template-based method utilizes unique shapes of objects for identification and tracking. Corners,[42] edges,[43] and some predefined patterns[44] on objects are the candidates for template-based object identification and tracking. Because these features may not be natural and may not be applicable to a general application, it can limit the implementation of template-based method.

Feature-based methods learn the characteristics of object appearance from a set of training images by capturing the variability in the object class. It is more accurate and reliable than the previously discussed approaches. In particular, Lowe proposed an algorithm for object recognition and tracking, called the Scale Invariant Feature Transform (SIFT),[45,46] which uses a class of local image features. In his algorithm, the detected features are invariant to changes in illumination, noise, rotation and scaling; and it has been proven that this approach has high robustness and reliability. However, this method consumes considerable computing time, which leads to slow response of the control system. This drawback could result in instability of the overall system.

Visibility constraint

Research of visual servoing has been challenged by an important requirement: how to keep the visual features within

(a)

(b)

Fig. 18. (a) Robot search and rescue scenario; (b) Homecare robotics scenario.

the field of view of the camera. To date, both IBVS and PBVS have used measurements of visual features to compute the controller outputs. Due to motion of the robot or some unknown disturbance, if these visual features move outside the field of view of the camera, the controller will completely fail. This problem becomes even more severe when visual servoing is applied to mobile robotic tasks, because mobile robots move through long distances than fixed base robotic manipulators.

It is common for visual features to move out of the field of view due to camera/robot calibration errors or controller design in a visual-servo system. Since the visibility of the visual features directly affects the robustness of the system, a significant effort has gone into solving this problem. A popular solution is to employ the potential field approach to push visual features toward the center of the field of view when the features approach the edge of the image. A representative work in this area has been completed in Ref. 47, where a potential function has been incorporated into an IBVS controller to repel visual feature points from the boundary of the image plane. Chesi and Hung[48] employed a similar approach to solve the visibility constraint problem in their global path planner for optimal visual servoing. The approach of navigation function, which guarantees a unique minimum, has been introduced into image-space path planners of visual servoing to generate a desired image trajectory while keeping the visual features in the field of view.[49,50]

Another popular method of keeping visual features in the field of view is to employ a path-planning technique to plan a camera trajectory or a "virtual path" on the image plane while meeting the visibility constraint.[51–53] It is common as well to employ a pan/tilt camera to enlarge the field of view so that the visual features will remain within it Refs. 54 and 55. Further approaches are available in the literature to solve the visibility problem. For example, Remazeilles and Chaumette[56] have employed several specific visual features to ensure that the robot navigates within the visibility path. Cowan and Chang[57] have developed a diffeomorphism-based IBVS controller to keep the features within the field of view and avoid self-occlusion.

There exists a common shortcoming in the approaches mentioned above, which are based on the techniques of potential-field, navigation-function, or planning: they represent static solutions designed in advance by a human. These approaches cannot improve their performance online and also cannot autonomously adapt to changing control tasks.

Controller design

Most visual-servo projects primarily concern object modeling and the quality of the vision feature feedback while paying less attention to controller design. Notably, a simple P (Proportional) control law or a PID (Proportional, Integral, Derivative) control law is commonly used in the literature. However, a PID controller may not be adequate to handle the issues of robustness and stability in real-life mobile robot applications. Spong *et al.*[13] discussed proportional control with Lyapunov stability, which is the controller that is most commonly used by researchers. Although this control law[12,13] can guarantee system stability, its controller output is not optimal and it is unable to consider various constraints (robot location constraints, visibility constraints, velocity constraints, and so on) which are common in a mobile visual-servo system.

In order to obtain optimal controller outputs, Ginhoux *et al.*[58] proposed to use GPC (generalized predictive controller) in visual servoing of a robotized surgical task. In particular, a 6 DOF (Degree of Freedom) robotic arm was developed to track the motion of a pig's heart based on the visual feedback. Since they employed the basic GPC scheme, no constraints were considered. In addition, they used the same idea in teleoperated laparoscopic surgery[59] and 3D profile.[60] Since the constraint issue is quite important in visual-servo tasks, Sauvee *et al.*[61] implemented a nonlinear model predictive controller for visual servoing of a robotic arm using vision feedback from ultrasound images. In their project, they considered various constraints and linearized

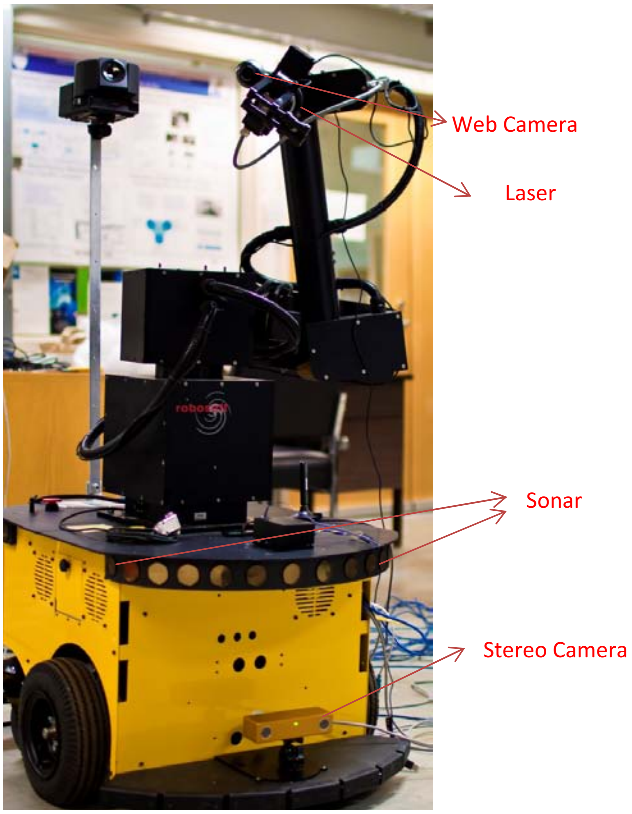


Fig. 19. Physical configuration of the test bed.

the nonlinear model into a constant linear model at the equilibrium point.

Applying MPC (model predictive control) to visual servoing is still in its preliminary stages, and the reported work is limited, as presented above. Although some encouraging results have been obtained, they shared a common shortcoming: in applying the existing (basic) MPC technique to visual-servo tasks, a constant linear model is assumed for the nonlinear plant. As a result, the implemented system can only work in a small neighborhood of the equilibrium point about which the model is linearized. If the current operating point of the system is farther from the equilibrium point (which is common in a mobile manipulation task), the controller performance will deteriorate quickly due to the mismatch between the model and the plant.

6. An Example of Visual Servo System

Two research applications of mobile manipulation are carried out in the Industrial Automation Laboratory (IAL) of the University of British Columbia: (1) Robotic search and rescue project; (2) Homecare robotics project. Figure 18 gives graphic representations of these two scenarios. Figure 18(a) demonstrates a future city where there are general-purpose robots that perform surveillance. If a natural disaster or some emergency situation occurs in the city, these robots will arrive at the scene for search and rescue work. Figure 18(b) presents a home environment where a mobile robot is implemented as a care-giver robot. It is capable of monitoring the inhabitants of the home while simultaneously doing some housework in an autonomous manner. If an abnormal behavior of an inhabitant is detected, the robot will reach the person to check the situation, provide assistance, and if necessary, may report it to a monitoring center.

A mobile manipulation system is being developed at IAL in order to carry out research and development activities related to these two applications. Figure 19 shows the physical configuration of the mobile manipulation system. It contains a Powerbot which is the mobile base, a RubuArm (the manipulator), and different types of sensors (web camera, laser, sonar, stereo camera, etc.).

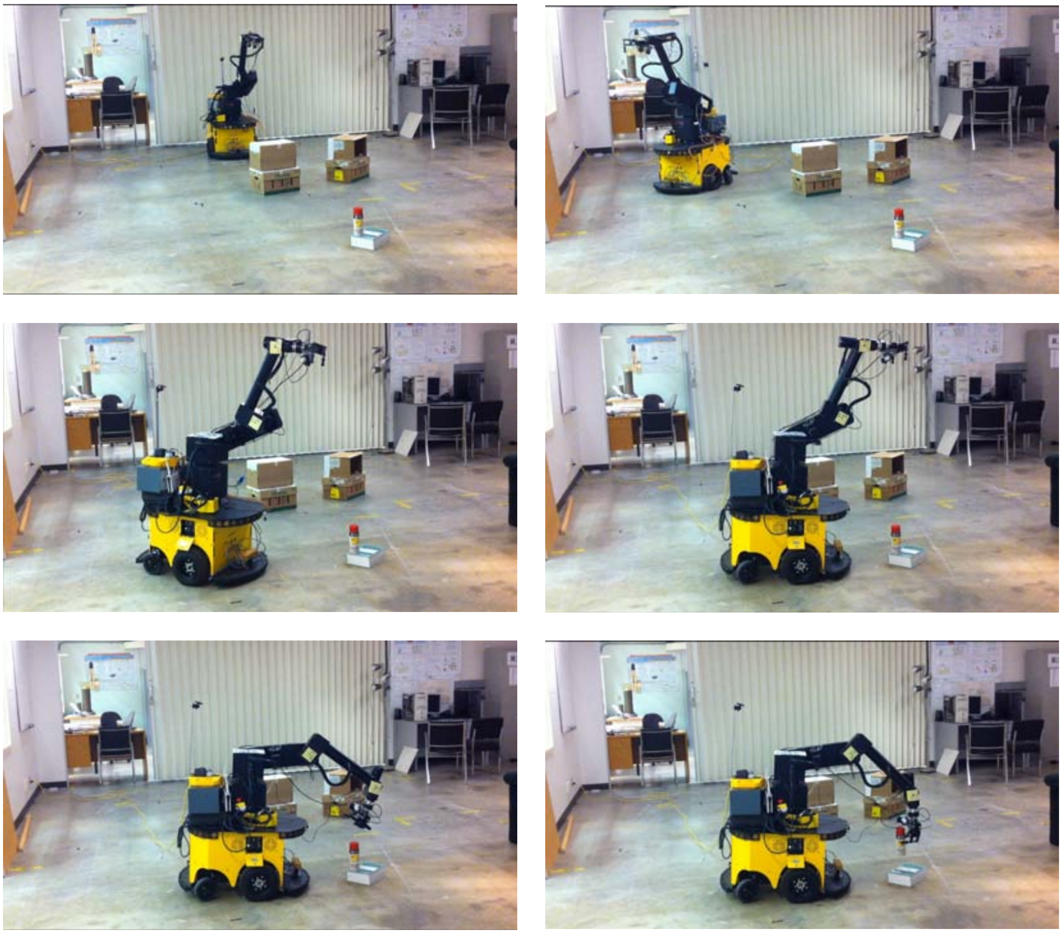


Fig. 20. Snapshots of the video clip of mobile manipulation experiment.

The Powerbot is a high-payload differential-drive robot developed by Adept Mobile Robots Company. It is able to move at a speed of up to 1.6 m/s with a payload of up to 100 kg. Hence it is an acceptable platform for research tasks involving delivery, navigation, and manipulation. The Powerbot is equipped with a gyro sensor and two groups of sonars. The readings from the gyro and the encoders are able to provide the position and orientation of the mobile robot. There are 32 sonar cones embedded in the mobile base and they can measure the distance between the robot and the neighborhood obstacles.

RobuArm is a six-DOF robotic manipulator which is developed by Robosoft Company in France. It has six revolute joints with a payload of 10 kg, and is rather powerful. The main advantage of this arm is the controllability of the acceleration, velocity, and position.

There are other sensors in the developed robotic system; for example, a Logitech QuickCam® Communicate STX™ CCD webcam is used as the monocular eye-in-hand camera; a BumbleBee 2 Stereo camera is configured as both eye-in-hand for the mobile based and eye-to-hand for the robotics manipulator; and a Hokuyo URG-04LX laser distance sensor which has an angular scanning range of 240° with distance measurement capability from 20 mm to 4 m.

A physical experiment that has been carried out using our test bed is shown in Fig. 20. The purpose of the experiment was to find an object of interest in the workspace. In this experiment, the robot approaches the object using machine learning capabilities and visual servoing.[62] Then, the robot grasps the object using the manipulator guided by its visual servo controller.[63] The performance of the overall system was found to be quite satisfactory in a series of experiments that were conducted.

7. Conclusion

In this paper, the state of the art of vision-guided robotic applications was introduced and the available mobile manipulation systems were discussed. Categorization of the traditional visual servo approaches was reviewed, and their advantages and drawbacks in applying in a mobile manipulation system were discussed. The general concept of modeling a visual servo system was demonstrated with examples. Some challenges in developing visual servo systems for practical applications were indicated. Finally, a practical application of mobile manipulation system which has been developed for applications of search-and-rescue and home-care robotics, in the Industrial Automation Laboratory in University of the British Columbia was introduced.

Acknowledgment

This work has been supported by research grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Canada Foundation for Innovation (CFI), the British Columbia Knowledge Development Fund (BCKDF), and the Senior Canada Research Chair in Mechatronics and Industrial Automation held by C.W. de Silva.

References

[1] B. J. Patz, Y. Papelis, R. Pillat, G. Stein and D. Harper, A practical approach to robotic design for the DARPA Urban Challenge, *J. Field Robot.* **25**(8) (2008) 528–566.

[2] M. Motemerlo, J. Becker, S. Shat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt and S. Thrun, Junior: The Stanford entry in the urban challenge Montemerlo, *J. Field Robot.* **25**(9) (2008) 569–597.

[3] Y. Ma, J. Kosecka and S. S. Sastry, Vision guided navigation for a nonholonomic mobile robot, *IEEE Trans. Robot. Autom.* **15**(3) (1999) 521–536.

[4] Y. Fang, W. E. Dixon, D. M. Dawson and P. Chawda, Homography-based visual servo regulation of mobile robots, *IEEE Trans. Syst., Man, Cybern. B, Cybern.* **35**(5) (2005) 1041–1050.

[5] 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 Trans. Syst., Man, Cybern. B, Cybern.* **31**(3) (2001) 341–352.

[6] A. Remazeilles and F. Chaumette, Image-based robot navigation from an image memory, *Robot. Autonom. Syst.* **55**(4) (2007) 345–356.

[7] R. Kelly, E. Bugarin and V. Sanchez, Image-based visual control of nonholonomic mobile robots via velocity fields: case of partially calibrated inclined camera, in *Proc. 45th IEEE Conf. Decision & Control*, San Diego, CA, December 2006.

[8] Z. Zhang, R. Weiss and A. R. Hanson, Visual servoing control of autonomous robot calibration and navigation, *J. Robot. Syst.* **16**(6) (1999) 313–328.

[9] G. L. Mariottini and D. Prattichizzo, Image-based visual servoing with central servoing with central catadioptric cameras, *Int. J. Robot. Res.* **27**(1) (2008) 41–56.

[10] P. Zhang and M. Hebert, First results in omni-directional visual servoing, *Adv. Robot.* **14**(3) (2000) 205–220.

[11] J. Gaspar, N. Winters and J. Santos-Victor, Vision-based navigation and environmental representations with an omnidirectional camera, *IEEE Trans. Robot. Autom.* **16**(6) (2000) 890–898.

[12] F. Chaumette and S. Hutchinson, Visual servo control part 1: Basic approaches, *IEEE Robot. Autom. Mag.* **13**(4) (2006) 82–90.

[13] M. W. Spong, S. Hutchinson and M. Vidyasagar, *Robot Modeling and Control* (Wiley, 2005).

[14] M. Hvilshj, S. Bgh, O. Nielsen and O. Madsen, Autonomous industrial mobile manipulation (AIMM): Past, present and future, *Ind. Robot* **39**(2) (2012) 120–135.

[15] M. Joseph, K. Hsiao, G. Niemeyer, S. Chitta and K. J. Kuchenbecker, Human-inspired robotic grasp control with tactile sensing, *IEEE Trans. Robot.* **27**(6) (2011) 1067–1079.

[16] J. Bohren, R. B. Rusu, E. G. Jones, E. Marder-Eppstein, C. Pantofaru, M. Wise, L. Mösenlechner, W. Meeussen and S. Holzer, Towards autonomous robotic butlers: Lessons learned with the PR2, in *Proc. IEEE Int. Conf. Robotics and Automation, ICRA 2011* (2011), pp. 5568–5575.

[17] M. Fuchs, C. Borst, P. R. Giordano, A. Baumann, E. Kraemer, J. Langwald, R. Gruber, N. Seitz, G. Plank, K. Kunze, R. Burger, F. Schmidt, T. Wimboeck and G. Hirzinger, Rollin' Justin — Design considerations and realization of a mobile platform for a humanoid upper body, *IEEE Int. Conf. Robotics and Automation ICRA 09* (2009), pp. 4131–4137.

[18] W. Bluethmann, R. Ambrose, M. Diftler, S. Askew, E. Huber, M. Goza, F. Rehnmark, C. Lovchik and D. Magruder, Robonaut: A robot designed to work with humans in space, *Auton. Robots* **14**(2–3) (2003) 179–197.

[19] J. Uhlmann, S. Julier, B. Kamgar-Parsi, M. Lanzagorta and H. Shyu, NASA Mars rover: A testbed for evaluating applications of covariance intersection, in *Proc. SPIE — The International Society for Optical Engineering*, Vol. 3693 (1999), pp. 140–149.

[20] J. Kucera and M. Sirak, Bomb-clearing robots to enter full-rate production, *Jane's Defence Weekly*, July 2005.

[21] J. Kucera, US Army seeks to hone bomb-clearing robots, *Jane's Defence Weekly*, September 2005.

[22] R. Bischoff, U. Huggenberger and E. Prassler, KUKA youBot — A mobile manipulator for research and education, *IEEE Int. Conf. Robotics and Automation, ICRA* (2011).

[23] F. Vohr and T. May, Robot hardware design in the DESIRE project, Springer Tracts in Advanced Robotics, Vol. 76, STAR (Springer, 2012), pp. 11–18.

[24] R. Siegwart and L. R. Nourbakhs, *Introduction to Autonomous Mobile Robots* (The MIT Press, Cambridge, Massachusetts, London, England, 2004).

[25] R. Szeliski, *Computer Vision: Algorithms and Applications* (Springer, New York, 2011).

[26] D. C. Mutto, Z. Pietro and C. M. Guido, Fusion of geometry and color information for scene segmentation, *IEEE J. Sel. Topics Signal Process.* **6**(5) (2012) 505–521.

[27] T. Huntsberger, H. Aghazarian, A. Howard and D. C. Trotz, Stereo vision-based navigation for autonomous surface vessels, *J. Field Robot.* **28**(1) (2011) 3–18.

[28] C. Lee, Y. Lim, S. Kwon and J. Lee, Stereo vision-based vehicle detection using a road feature and disparity histogram, *Opt. Eng.* **50** (2) (2011) 1–23.

[29] S. Hutchinson, G. D. Hager and P. I. Corke, A tutorial on visual servo control, *IEEE Trans. Robot. Autom.* **12**(5) (1996) 651–670.

[30] G. J. Agin, Real time control of a robot with a mobile camera, Technical Note 179, SRI International, Menlo Park (1979).

[31] A. C. Sanderson and L. E. Weiss, Image-based visual servo control using relational graph error signals, in *Proc. IEEE* (1980), pp. 1074–1077.

[32] F. Conticelli, B. Allotta and P. K. Khosla, Image-based visual servoing of nonholonomic mobile robots, in *Proc. 38th IEEE Conf. Decision and Control*, Vol. 4, Phoenix, AZ (1999), pp. 3496–3501.

[33] Y. Wang, H. Lang, H. Lin and C. W. de Silva, Vision-based grasping using mobile robots and nonlinear model predictive control, *Int. J. Contr. Intell. Syst.* **40**(1) (2012) 3–10.

[34] Y. Wang, H. Lang and C. W. de Silva, A hybrid visual servoing controller for robust manipulation using mobile robots, *IEEE/ASME Trans. Mechatron.* **15**(5) (2010) 757–769.

[35] E. Malis, F. Chaumette and S. Boudet, 1-1/2-D visual servoing, *IEEE Trans. Robot. Autom.* **15**(2) (1999) 238–250.

[36] W. J. Wilson, C. C. W. Hulls and G. S. Bell, Relative end-effector control using Cartesian position-based visual servoing, *IEEE Trans. Robot. Autom.* **12**(5) (1996) 684–696.

[37] J. J. Crag, *Introduction to Robotics: Mechanics and Control*, 3rd edn. (Prentice Hall, Upper Saddle River, NJ, 2004).

[38] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd edn. (Prentice Hall, Upper Saddle River, NJ, 2002).

[39] E. C. Dean-Le'on, V. Parra-Vega and A. Espinosa-Romero, Visual servoing for constrained planar robots subject to complex friction, *IEEE/ASME Trans. Mechatron.* **11**(4) (2006) 389–400.

[40] Z. Y. Zhang, Flexible camera calibration by viewing a plane from unknown orientations, in *Proc. Seventh IEEE Int. Conf. Computer Vision*, Vol. 1, September 1999, pp. 666–673.

[41] H. Lang, Y. Wang and C. W. de Silva, Mobile robot localization and object pose estimation using optical encoder, vision and laser sensors, *IEEE Int. Conf. Automation and Logistics*, Qingdao, China (2008).

[42] M. Bertozzi, A. Broggi and S. Castelluccio, A real-time oriented system for vehicle detection, *J. Syst. Architect.* (1997) 317–325.

[43] N. Matthews, P. An, D. Charnley and C. Harris, Vehicle detection and recognition in greyscale imagery, *Contr. Eng. Pract.* **4** (1996) 473–479.

[44] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner and W. V. Seelan, An image processing system for driver assistance, *Image Vis. Comput.* **18**, (2000) 367–376.

[45] D. G. Lowe, Object recognition from local scale-invariant features, *Int. Conf. Computer Vision*, Corfu, Greece, September 1999, pp. 1150–1157.

[46] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* **60**(2) (2004) 91–110.

[47] P. I. Corke and S. A. Hutchinson, A new partitioned approach to image-based visual servo control, *IEEE Trans. Robot. Autom.* **17**(4) (2001) 507–515.

[48] G. Chesi and Y. S. Hung, Global path-planning for constrained and optimal visual servoing, *IEEE Trans. Robot.* **23**(5) (2007) 1050–1060.

[49] J. Chen, D. M. Dawson, W. E. Dixon and V. K. Chitrakaran, Navigation function-based visual servo control, *Automatica* **43**(7) (2007) 1165–1177.

[50] N. J. Cowan, J. D. Weingarten and D. E. Koditschek, Visual servoing via navigation functions, *IEEE Trans. Robot. Autom.* **18**(4) (2002) 521–532.

[51] F. Schramm and G. Morel, Ensuring visibility in calibration-free path planning for image-based visual servoing, *IEEE Trans. Robot.* **22**(4) (2006) 848–854.

[52] H. Zhang and J. P. Ostrowski, Visual motion planning for mobile robots, *IEEE Trans. Robot. Autom.* **18**(2) (2002) 199–208.

[53] G. Chesi and Y. S. Hung, Global path-planning for constrained and optimal visual servoing, *IEEE Trans. Robot.* **23**(5) (2007) 1050–1060.

[54] F. Capparella, L. Freda, M. Malagnino and G. Oriolo, Visual servoing of a wheeled mobile robot for intercepting a moving object, in *Proc. 2005 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Edmonton, Canada, August 2005.

[55] T. Nierobisch, W. Fischer and F. Hoffman, Large view visual servoing of a mobile robot with a pan-tilt camera, in *Proc. 2006 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Beijing, China, October 2006.

[56] A. Remazeilles and F. Chaumette, Image-based robot navigation from an image memory, *Robot. Autonom. Syst.* **55**(4) (2007) 345–356.

[57] N. J. Cowan and D. E. Chang, Geometric visual servoing, *IEEE Trans. Robot.* **21**(6) (2005) 1128–1138.

[58] R. Ginhoux, J. Gangloff and M. de Mathelin, Active filtering of physiological motion in robotized surgery using predictive control, *IEEE Trans. Robot.* **21**(1) (2005) 67–79.

[59] J. Gangloff, R. Ginhoux and M. de Mathelin, Model predictive control for compensation of cyclic organ motions in teleoperated laparoscopic surgery, *IEEE Trans. Control Syst. Technol.* **14**(2) (2006) 235–245.

[60] J. Gangloff and M. de Mathelin, Visual servoing of a 6-DOF manipulator for unknown 3-D profile following, *IEEE Trans. Robot. Autom.* **18**(4) (2002) 511–519.

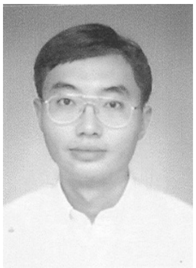
[61] M. Sauvee, P. Poignet and E. Dombre, Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control, *Int. J. Robot. Res.* **27**(1) (2008) 25–39.

[62] H. Lang and C. W. de Silva, An Application of Mobile Manipulation, *CSME International Congress*, Winnipeg, Manitoba, Canada (2012).

[63] H. Lang and C. W. de Silva, Predictive Control for Image-Based Visual Servoing of a Mobile Manipulation System, *CSME International Congress*, Winnipeg, Manitoba, Canada (2012).

Haoxiang Lang received his Ph.D. degree from the University of British Columbia in 2012; M.A.Sc. degree in Mechanical Engineering from the University of British Columbia in 2008; and Bachelor degree from Ningbo University in 2003. He worked in the Motorola Cellular Equipment Company as a software engineering from January 2004 to September 2005. He is now a Post-Doctoral Fellow with the Industrial Automation Laboratory at the University of British Columbia. His research and development areas are Robotics, Computer Vision, Controls and Sensor Fusion.

M. Tahir Khan received his Ph.D. degree from the University of British Columbia, Vancouver, BC, Canada, in 2010; Master's degree in Mechatronics from University of New South Wales, Sydney, Australia, in 1999; and Bachelor's degree in mechanical engineering from NWFP University of Engineering & Technology, Peshawar, Pakistan, in 1997. He was a Post-Doctoral Fellow with the Industrial Automation Laboratory at the University of British Columbia, Vancouver, BC, Canada for two years until January 2012. He currently holds a faculty position with the Institute of Mechatronics at the University of Engineering and Technology, Peshawar, Pakistan. His research interests include Robotics, Intelligent Control Systems, Neurofuzzy and Artificial Immune Systems.



Kok Kiong, Tan received his Ph.D. in 1995 from the National University of Singapore (NUS). Prior to joining the university, he was a research fellow at SIMTech, a national R&D institute spearheading the promotion of R&D in the local manufacturing industry, where he has been involved in managing industrial projects in IT and automation. He is currently an associate professor at NUS, and his current research interests are in Internet Applications, Remote Control and Monitoring, Precision Motion Control and Instrumentation, Advanced Process Control and Autotuning, and General Industrial Automation.

Clarence W. de Silva is a Fellow of: IEEE, ASME, the Canadian Academy of Engineering, and the Royal Society of Canada. He received Ph.D. degrees from the Massachusetts Institute of Technology (1978); and the University of Cambridge, U.K. (1998); and an honorary D.Eng. degree from the University of Waterloo, Canada (2008). He has been a Professor of Mechanical Engineering and NSERC-BC Packers Chair in Industrial Automation, at the University of British Columbia, Vancouver, Canada since 1988, and currently occupies the Tier 1 Canada Research Chair in Mechatronics & Industrial Automation. He has authored 20 books and over 400 papers, approximately half of which are in journals. His recent books published by Taylor & Francis/CRC are: *Mechatronics — A Foundation Course* (2010); *Modeling and Control of Engineering Systems* (2009); *Sensors and Actuators — Control System Instrumentation* (2007); *VIBRATION — Fundamentals and Practice, 2nd Ed.* (2007); *Mechatronics — An Integrated Approach* (2005); and by Addison Wesley: *Soft Computing and Intelligent Systems Design — Theory, Tools, and Applications* (with F. Karray, 2004).