# Research on Visual Servo Grasping of Household Objects for Nonholonomic Mobile Manipulator 

Huangsheng Xie, ${ }^{1}$ Guodong Li, ${ }^{1}$ Yuexin Wang, ${ }^{1}$ Zhihe Fu, ${ }^{1}$ and Fengyu Zhou ${ }^{2}$<br>${ }^{1}$ School of Physics and Mechanical and Electrical Engineering, Longyan College, Longyan, Fujian 364012, China<br>${ }^{2}$ School of Control Science and Engineering, Shandong University, Jinan 250061, China<br>Correspondence should be addressed to Guodong Li; 37729149@qq.com and Fengyu Zhou; zhoufengyu@sdu.edu.cn

Received 23 June 2014; Accepted 4 September 2014; Published 24 September 2014
Academic Editor: Wuneng Zhou
Copyright © 2014 Huangsheng Xie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.


#### Abstract

This paper focuses on the problem of visual servo grasping of household objects for nonholonomic mobile manipulator. Firstly, a new kind of artificial object mark based on QR (Quick Response) Code is designed, which can be affixed to the surface of household objects. Secondly, after summarizing the vision-based autonomous mobile manipulation system as a generalized manipulator, the generalized manipulator's kinematic model is established, the analytical inverse kinematic solutions of the generalized manipulator are acquired, and a novel active vision based camera calibration method is proposed to determine the hand-eye relationship. Finally, a visual servo switching control law is designed to control the service robot to finish object grasping operation. Experimental results show that QR Code-based artificial object mark can overcome the difficulties brought by household objects' variety and operation complexity, and the proposed visual servo scheme makes it possible for service robot to grasp and deliver objects efficiently.


## 1. Introduction

A classical mobile manipulator system (MMS) consists of a manipulator which is mounted on a nonholonomic mobile platform. This type of arrangement extends manipulator's workspace apparently and is widely used in service robot applications [1,2]. The development of MMS mainly involves two classical items, namely, motion planning [3-8] and coordinating control [9-13], which are used to overcome the mobile platform's nonholonomic constraint and make the MMS move quickly and efficiently.

When robots operate in unstructured environments, it is essential to include exteroceptive sensory information in the control loop. In particular, visual information provided by vision sensor such as charge-coupled device (CCD) cameras guarantees accurate positioning, robustness of calibration uncertainties, and reactivity of environmental changes. Much of the work related to CCD cameras and manipulators has focused on the applications about the manipulator's visual servo control, which specifies robotic tasks (such as object
grasping, assembling) in terms of desired image features extracted from a target object. The overview of visual servo can be seen in literature [14-16]. In general, visual servo approaches can be divided into three different kinds, namely, position-based visual servoing (PBVS) [17, 18], image-based visual sevoing (IBVS) [19, 20], and hybrid visual servoing (HYBVS) [21-23]. In PBVS, the feedback signals in vision loop are the intuitive relative 3D pose between current and desired cameras estimated by current and desired image features using homography matrix or fundamental matrix estimation and decomposition. In IBVS, the feedback signals are image features whose changing velocities are related to the velocity twist of the camera via the image jacobian matrix (also called interaction matrix). Compared with the PBVS, IBVS is robust to perturbations of the robot/camera models and can maintain the image features in the field of view (FOV) of the camera through path planning [24]. But the drawbacks are also obvious; the interaction matrix's depth information needs to be estimated and only local stability can be guaranteed for most IBVS schemes. To combine
the advantages of PBVS and IBVS, HYBVS is proposed. In HYBVS, the feedback signals consist of relative 3D pose and image features, the former is used to control a subset of the camera configuration vector while the latter are used to regulate the remaining camera configuration vector.

Relating CCD cameras and the mobile robots lead to the applications of vision-based autonomous navigation control. Ma et al. [25] have developed a vision-guided navigation system, where a nonholonomic mobile robot tracks an arbitrarily shaped continuous ground curve. Dixon et al. [26] present an adaptive tracking controller of a wheeled mobile robot via an uncalibrated camera system that the controller copes with the parameter uncertainty of the mechanical dynamics and the camera system. Amarasinghe et al. [27] have developed a vision-based hybrid control scheme for autonomous parking of a mobile robot that its controller consists of a discrete-event controller and a pixel-errordriven proportional controller. Vassallo et al. [28] present a similar project where a vision-based mobile robot attempts to autonomously navigate in a building.

The newest trend is to integrate CCD cameras into mobile manipulator to form a vision-based mobile manipulation system (VBMMS). Thanks to the capabilities of the vision subsystem, the VBMMS can work in an unstructured environment and has wider applications than a fixed-base manipulator and a mobile platform. Due to the lack of accurate and robust positioning performance of VBMMS, very few physical implementations have been reported. de Luca et al. [29] have considered the task-oriented modeling of the differential kinematics of nonholonomic mobile manipulators and has developed an image-based controller for VBMMS, but their approach is illustrated through simulation, not physical implementation. Mansard et al. [30] have attempted to control a humanoid robot to grasp an object while walking. Wang et al. [31] have developed a robust vision-based mobile manipulation system for wheeled mobile robots. In their research, an innovative controller with machine learning using Q-learning is proposed to guarantee visibility of visual features in servo process.

This paper presents a physical implementation of VBMMS in service robot intelligent space. It consists of two basic contributions. First, after summarizing the VBMMS as a generalized manipulator, the kinematics is analyzed analytically and an active vision-based camera calibration method is proposed to determine the hand-eye relationship consequently. Second, a novel switching control strategy is proposed which switches between eye-fixed approximation and position-based static look-and-move grasping. The remainder of the paper is organized as follows. Section 1 will introduce the design of the QR Code-based artificial object mark. In Section 2, the VBMMS is summarized as a generalized manipulator, then the kinematics, inverse kinematics, and hand-eye relationship determination are discussed. In Section 3, the switching control strategy which switches between eye-fixed approximation and static look-and-move grasping is designed. Two experiments will be presented in Section 4 to validate the designed switching controller. Conclusions will be drawn in Section 5.


Figure 1: QR Code-based artificial object mark.

## 2. Design of QR Code-Based Artificial Object Mark

As shown in Figure 1, the QR Code-based artificial object mark is composed of two parts: the internal information representation part which includes the object's property and operation information and the blue concentric ring region which is called external identification part. Due to the facilitated detection of the external identification part, the mark can be recognized from complex home environment rapidly using vision sensor.

The coding of information stored in internal information representation part of the mark is shown in Table 1.

It can be seen from Table 1 that there are two different kinds of information, namely, object's property information and object's operation information. The object's property information includes name, serial number, sizes, and material, in which the properties of name and serial number are used as the unique identification of an object and the property of sizes is to let robot know the gripper's opening degree. The object's operation information includes operation force, position, and orientation, which can assist the robot in finishing grasping operation in an appropriate way.

## 3. Kinematics, Inverse Kinematics, and Hand-Eye Relationship Determination of VBMMS

As shown in Figure 2, the eye-in-hand type VBMMS is combined with a nonholonomic tracked mobile robot (TMR), a 4-DOF (Degree of Freedom) manipulator, and a CMOS camera. The Grandar AS-RF type TMR uses differential drive structure and is easier to carry out steering control. The Schunk Powercube 4-DOF manipulator is mounted on the TMR, and the Gsou V80 CMOS camera is mounted on the manipulator's end-effector. In addition, the VBMMS is equipped with an onboard computer, whose computational capability can support real-time performance of the system.

TABLE 1: Coding of information stored in QR Code-based artificial object mark.

| Info type | Description | Bytes |
| :--- | :--- | ---: |
| Name | Object's name's first two words | 2 |
| Serial number | Such as $1,2, \ldots$ | 1 |
| Sizes | Long $*$ width $*$ height, 2 bytes for each items | 8 |
| Material | p (plastic), g (glass), z (paper), w (wood), s (metal), t (textile) | 1 |
| Operation force | H (huge), L (large), M (middle), S (small), T (tiny) | 1 |
| Operation position | u (upper), m (middle), b (bottom) | 1 |
| Operation orientation | a (above), f(front), b (back), l (left), r (right) | 1 |

Table 2: D-H parameters of the generalized manipulator.

| Link $i$ | $\alpha_{i}(\mathrm{rad})$ | $a_{i}(\mathrm{~cm})$ | $\theta_{i}(\mathrm{rad})$ | $d_{i}(\mathrm{~cm})$ | Type |
| :--- | :---: | :---: | :---: | :---: | :---: |
| 1 | 0 | 0 | $\pi / 2+\theta_{1}$ | 0 | R |
| 2 | $\pi / 2$ | 0 | 0 | $0+d_{2}$ | T |
| 3 | $-\pi / 2$ | 0 | $-\pi / 2+\theta_{3}$ | 25 | R |
| 4 | $-\pi / 2$ | -6.5 | $0+\theta_{4}$ | 0 | R |
| 5 | $\pi / 2$ | 6.5 | $0+\theta_{5}$ | 28.5 | R |
| 6 | $\pi / 2$ | 0 | $0+\theta_{6}$ | 0 | R |
| Tool | $-\pi / 2$ | 0 | 0 | 28 |  |

3.1. Kinematics. Due to the nonholonomic constraint of the TMR and the nonredundancy of the 4-DOF manipulator, using VBMMS to complete grasping task is very difficult, and so far, barely no related works can be found. Take the difficulties of controlling the TMR and manipulator separately into account, the VBMMS is summarized as a generalized manipulator shown in Figure 2. For the generalized manipulator, the TMR is considered as a 3-DOF (rotation-translation-rotation) manipulator and the manipulator's first degree is aborted because of its coincidence with the TMR's third degree. In that case, the generalized manipulator has six degrees, which guarantees that the end-effector can approach an arbitrary position at any pose.

Table 2 illustrates the modified D-H parameters of each link for serial-link manipulator, where $\alpha_{i}, a_{i}, \theta_{i}, d_{i}$ mean twist, length, angle, and offset of the $i$ th link, respectively. In addition to the last column of the table, R is for revolute while T is for prismatic. For the prismatic joint, $d_{i}$ is the variable whose value range is $d_{i}>0$, while for the revolute joint, $\theta_{i}$ is the variable whose value range is $-\pi \sim \pi$.

Table 2 results in a $4 \times 4$ homogeneous transformation matrix ${ }^{i-1} \mathbf{T}_{i}$ representing each link's coordinate frame with respect to the previous link's coordinate frame. Thereby, the representation of end-effector coordinate frame $\{E\}$ with respect to base coordinate frame $\{0\}$ can be shown in the following:

$$
{ }^{0} \mathbf{T}_{E}=\left[\begin{array}{cccc}
n_{x} & o_{x} & a_{x} & p_{x}  \tag{1}\\
n_{y} & o_{y} & a_{y} & p_{y} \\
n_{z} & o_{z} & a_{z} & p_{z} \\
0 & 0 & 0 & 1
\end{array}\right]
$$

where

$$
\begin{align*}
n_{x}= & c_{13} c_{4} c_{5} c_{6}+c_{13} s_{4} s_{6}-s_{13} s_{5} c_{6}  \tag{2}\\
n_{y}= & s_{13} c_{4} c_{5} c_{6}+s_{13} s_{4} s_{6}+c_{13} s_{5} c_{6}  \tag{3}\\
n_{z}= & -s_{4} c_{5} c_{6}+c_{4} s_{6}  \tag{4}\\
o_{x}= & -c_{13} c_{4} s_{5}-s_{13} c_{5}  \tag{5}\\
o_{y}= & -s_{13} c_{4} s_{5}+c_{13} c_{5}  \tag{6}\\
o_{z}= & s_{4} s_{5}  \tag{7}\\
a_{x}= & -c_{13} c_{4} c_{5} s_{6}+c_{13} s_{4} c_{6}+s_{13} s_{5} s_{6}  \tag{8}\\
a_{y}= & -s_{13} c_{4} c_{5} s_{6}+s_{13} s_{4} c_{6}-c_{13} s_{5} s_{6}  \tag{9}\\
a_{z}= & s_{4} c_{5} s_{6}+c_{4} c_{6}  \tag{10}\\
p_{x}= & -28 c_{13} c_{4} c_{5} s_{6}+28 c_{13} s_{4} c_{6}+6.5 c_{13} c_{4}  \tag{11}\\
& +28.5 c_{13} s_{4}-6.5 c_{13}+28 s_{13} s_{5} s_{6}+d_{2} c_{1} \\
p_{y}= & -28 s_{13} c_{4} c_{5} s_{6}+28 s_{13} s_{4} c_{6}+6.5 s_{13} c_{4} \\
& +28.5 s_{13} s_{4}-6.5 s_{13}-28 c_{13} s_{5} s_{6}+d_{2} s_{1}  \tag{12}\\
p_{z}= & 28 s_{4} c_{5} s_{6}+28 c_{4} c_{6}-6.5 s_{4}+28.5 c_{4}+25 \tag{13}
\end{align*}
$$

in (2)-(13), $c_{i}=\cos \theta_{i}, s_{i}=\sin \theta_{i}, c_{i j}=\cos \left(\theta_{i}+\theta_{j}\right), s_{i j}=$ $\sin \left(\theta_{i}+\theta_{j}\right), i, j=1,3,4,5,6$.

It is well known that the symbol ${ }^{E} \mathbf{J}$ can be used to note the generalized manipulator Jacobian matrix expressed in endeffector coordinate frame $\{E\}$ that transforms velocities in joint space to velocities of the end-effector in Cartesian space. For the 6-DOF generalized manipulator, the end-effector Cartesian velocity is

$$
\begin{equation*}
{ }^{E} \dot{\mathbf{X}}={ }^{E} \mathbf{J}(\mathbf{q}) \dot{\mathbf{q}}, \tag{14}
\end{equation*}
$$

where the 6-vector represented by ${ }^{E} \dot{\mathbf{X}}=\left[{ }^{E} \mathbf{v}^{T}{ }^{E} \boldsymbol{\omega}^{T}\right]^{T}$ is the end-effector's Cartesian velocity with respect to its own frame $\{E\}$. Based upon the resolved transformation matrix ${ }^{i-1} \mathbf{T}_{i}$ and ${ }^{6} \mathbf{T}_{E},{ }^{i} \mathbf{T}_{E}$ can be determined as

$$
{ }^{i} \mathbf{T}_{E}={ }^{i} \mathbf{T}_{i+1} \cdot{ }^{i+1} \mathbf{T}_{i+2} \cdots{ }^{6} \mathbf{T}_{E}=\left[\begin{array}{cccc}
{ }^{i} n_{x} & { }^{i} o_{x} & { }^{i} a_{x} & { }^{i} p_{x}  \tag{15}\\
{ }^{i} n_{y} & { }_{i} & o_{y} & { }_{i} \\
i_{y} & { }^{i} & p_{y} \\
{ }^{i} n_{z} & { }_{i} & o_{z} & { }^{i} a_{z} \\
{ }^{i} p_{z} \\
0 & 0 & 0 & 1
\end{array}\right],
$$



Figure 2: Vision-based mobile manipulation system and its link coordinate frames.
for prismatic joint corresponding $i=2$; the $i$ th column of ${ }^{E} \mathbf{J}$ can be constructed as follows:

$$
{ }^{E} \mathbf{J}_{i}=\left[\begin{array}{llllll}
{ }^{i} n_{z} & { }^{i} o_{z}{ }^{i} a_{z} & 0 & 0 & 0 \tag{16}
\end{array}\right]^{T}
$$

for revolute joint corresponding $i=1,3,4,5,6$; the $i$ th column of ${ }^{E} \mathbf{J}$ :

$$
\begin{gather*}
{ }^{E} \mathbf{J}_{i}=\left[-{ }^{i} n_{x}{ }^{i} p_{y}+{ }^{i} n_{y}{ }^{i} p_{x},-{ }^{i} o_{x}{ }^{i} p_{y}+{ }^{i} o_{y}{ }^{i} p_{x}\right. \\
\left.-{ }^{i} a_{x}{ }^{i} p_{y}+{ }^{i} a_{y}{ }^{i} p_{x},{ }^{i} n_{z},{ }^{i} o_{z},{ }^{i} a_{z}\right]^{T} \tag{17}
\end{gather*}
$$

The ${ }^{E} \mathbf{J}$ can finally be constructed as follows:

$$
{ }^{E} \mathbf{J}=\left[\begin{array}{llllll}
{ }^{E} \mathbf{J}_{1} & { }^{E} \mathbf{J}_{2} & { }^{E} \mathbf{J}_{3} & { }^{E} \mathbf{J}_{4} & { }^{E} \mathbf{J}_{5} & { }^{E} \mathbf{J}_{6} \tag{18}
\end{array}\right]
$$

3.2. Inverse Kinematics. From (2)-(13) we can achieve that the manipulator has 6 joints to allow arbitrary end-effector pose. In order to reach the specified end-effector position, the inverse kinematic solutions can be acquired through separating the unknown variable $\theta_{1}, d_{2}, \theta_{3}$ in (2)-(13).

Note that ${ }^{0} \mathbf{T}_{3}={ }^{0} \mathbf{T}_{1}{ }^{1} \mathbf{T}_{2}{ }^{2} \mathbf{T}_{3}$, by premultiplicating ${ }^{0} \mathbf{T}_{3}^{-1}$ in (1), we can acquire ${ }^{3} \mathbf{T}_{E}$. On the other hand, ${ }^{3} \mathbf{T}_{E}$ can be computed by ${ }^{3} \mathbf{T}_{E}={ }^{3} \mathbf{T}_{4}{ }^{4} \mathbf{T}_{5}{ }^{5} \mathbf{T}_{6}{ }^{6} \mathbf{T}_{E}$. Let the two different
forms of ${ }^{3} \mathbf{T}_{E}$ be equal; four sets of analytical inverse kinematic solutions can be achieved:

$$
\begin{aligned}
& \theta_{4} \\
&= A \tan 2(28.5,6.5) \\
&+A \tan 2\left(\frac{28 a_{z}-p_{z}+25}{\sqrt{6.5^{2}+28.5^{2}}}, \pm \sqrt{1-\frac{\left(28 a_{z}-p_{z}+25\right)^{2}}{6.5^{2}+28.5^{2}}}\right) \\
& \theta_{1}+ \theta_{3} \\
&= A \tan 2\left(\frac{-c_{4} o_{z}}{s_{4} \sqrt{o_{x}^{2}+o_{y}^{2}}}, \pm \sqrt{\left.1-\frac{\left(c_{4} o_{z}\right)^{2}}{s_{4}^{2}\left(o_{x}^{2}+o_{y}^{2}\right)}\right)}\right. \\
&-A \tan 2\left(o_{x}, o_{y}\right) \\
& \theta_{5}= A \tan 2\left(\frac{o_{z}}{s_{4}},-s_{13} o_{x}+c_{13} o_{y}\right) \\
& \theta_{6}= A \tan 2\left(\frac{s_{13} a_{x}-c_{13} a_{y}}{s_{5}}, \frac{-s_{13} n_{x}+c_{13} n_{y}}{s_{5}}\right) \\
& d_{2}=\left\{\left(c_{13} p_{x}+s_{13} p_{y}+28 c_{4} c_{5} s_{6}-28 s_{4} c_{6}\right.\right. \\
&\left.\quad-6.5 c_{4}-28.5 s_{4}+6.5\right)^{2} \\
& \theta_{3}= A \tan 2\left(-28 s_{5} s_{6}+s_{13} p_{x}-c_{13} p_{y}\right. \\
&\left.+\left(-28 s_{5} s_{6}+s_{13} p_{x}-c_{13} p_{y}\right)^{2}\right\}{ }^{1 / 2} \\
&=
\end{aligned}
$$



Figure 3: Scheme of hand-eye relationship determination.

$$
\begin{align*}
& c_{13} p_{x}+s_{13} p_{y}+28 c_{4} c_{5} s_{6}-28 s_{4} c_{6} \\
& \left.-6.5 c_{4}-28.5 s_{4}+6.5\right) \tag{19}
\end{align*}
$$

Given an arbitrary end-effector pose, whose orientation is expressed by RPY description method, the four sets of inverse kinematic solutions can be solved by (19), as shown in Table 3.

After solving all the solutions, they can be processed so that they verify the joints' value ranges, and then the optimal set of solution can be acquired using a certain optimality criteria such as the shortest path.
3.3. Hand-Eye Relationship Determination. Hand-eye relationship determination is a key issue in visual servo control of a robot hand-eye system, and much work has been done. In this paper, we propose a novel method which is based on Zhang's camera calibration method and tensor theory. As we know, Zhang's algorithm proposed in 1999 is very representative because of its easy use, flexibility, and high accuracy [32]. The algorithm only requires the camera to observe a planar pattern (such as planar checkerboard) shown at a few different orientations. In consequence, a closed-form solution of camera intrinsic parameter matrix $\mathbf{K}$ and the pose of planar object's coordinate frame $\{o\}$ with respect to the camera frame $\{c\}$ noted as ${ }^{c} \mathbf{T}_{o}$ can be computed, which are followed by a nonlinear refinement based on the maximum likelihood criterion so as to improve the accuracy.

Figure 3 shows the scheme of hand-eye relationship determination, where $\{E\},\{c\}$ are the current end-effector frame and current camera frame, respectively.

When the manipulator executes the $i$ th $(i=1,2, \ldots, n)$ known movement, they move to a new position noted as $\left\{E^{i}\right\}$ and $\left\{c^{i}\right\}$.

From Figure 3, there is

$$
\begin{equation*}
{ }^{E} \mathbf{T}_{c} \cdot{ }^{c} \mathbf{T}_{o}={ }^{E} \mathbf{T}_{E^{i}} \cdot{ }^{E^{i}} \mathbf{T}_{c^{i}} \cdot{ }^{c^{i}} \mathbf{T}_{o} . \tag{20}
\end{equation*}
$$

In (20), ${ }^{E} \mathbf{T}_{c}={ }^{E^{i}} \mathbf{T}_{c^{i}}$ is the unknown hand-eye relationship, ${ }^{E} \mathbf{T}_{E^{i}}$ can be known from the known motion of the manipulator, and ${ }^{c} \mathbf{T}_{o}$ and ${ }^{c^{i}} \mathbf{T}_{o}$ can be known from Zhang's method.

Note that

$$
\begin{align*}
& { }^{E^{i}} \mathbf{T}_{E}={ }^{0} \mathbf{T}_{E^{i}}^{-1} \cdot{ }^{0} \mathbf{T}_{E}=\left[\begin{array}{ll}
\mathbf{R}_{A}^{i} & \mathbf{t}_{A}^{i} \\
\mathbf{0}^{T} & 1
\end{array}\right] \\
& { }^{E} \mathbf{T}_{c}=\left[\begin{array}{ll}
\mathbf{R} & \mathbf{t} \\
\mathbf{0}^{T} & 1
\end{array}\right]  \tag{21}\\
& { }^{c} \mathbf{T}_{c^{i}}={ }^{c} \mathbf{T}_{o} \cdot{ }^{c^{i}} \mathbf{T}_{o}^{-1}=\left[\begin{array}{ll}
\mathbf{R}_{B}^{i} & \mathbf{t}_{B}^{i} \\
\mathbf{0}^{T} & 1
\end{array}\right] .
\end{align*}
$$

By substituting (21) into (20), one obtains

$$
\begin{gather*}
\mathbf{R}_{A}^{i} \cdot \mathbf{R} \cdot \mathbf{R}_{B}^{i}=\mathbf{R}  \tag{22}\\
\mathbf{R}_{A}^{i} \cdot \mathbf{R} \cdot \mathbf{t}_{B}^{i}+\mathbf{R}_{A}^{i} \cdot \mathbf{t}+\mathbf{t}_{A}^{i}=\mathbf{t} . \tag{23}
\end{gather*}
$$

It can be seen from (22) that once we solve the unknown $\mathbf{R}$, then $\mathbf{t}$ will be solved easily with (23). On the other hand, (22) can be summarized as the form of $\mathbf{A X B}=\mathbf{X}$, considering a $M \times N$ matrix $\mathbf{X}$ to be a second-order mixed tensor whose element of $p$ th row and $q$ th column is noted as $x_{q}^{p}$; the equation $\mathbf{A X B}=\mathbf{X}$ can be written as follows:

$$
\begin{equation*}
a_{k}^{p} x_{l}^{k} b_{q}^{l}=x_{q}^{p} \tag{24}
\end{equation*}
$$

where $p, q$ are free indexes, and $k, l$ are dummy indexes. Take Kronecker dinto account:

$$
\delta_{\beta}^{\alpha}= \begin{cases}1, & \text { if } \alpha=\beta  \tag{25}\\ 0, & \text { if } \alpha \neq \beta\end{cases}
$$

Equation (24) can be written as

$$
\begin{equation*}
\left(a_{k}^{p} b_{q}^{l}-\delta_{k}^{p} \delta_{q}^{l}\right) x_{l}^{k}=0 \tag{26}
\end{equation*}
$$

According to the different values of the free indexes $p$ and $q$, we can acquire $M \times N$ sets of equations from (26). Let $\mathbf{x}=$ $\left[\mathbf{x}^{1 T}, \mathbf{x}^{2 T}, \ldots, \mathbf{x}^{m T}\right]^{T}$, where $\mathbf{x}^{i T}$ is the $i$ th row vector of Matrix $\mathbf{X}$. Equation (26) can finally be written as

$$
\begin{equation*}
\mathbf{H x}=\mathbf{0}, \tag{27}
\end{equation*}
$$

where $\mathbf{H} \in R^{(M \times N) \times(M \times N)}$, and its element

$$
\begin{equation*}
h_{(k-1) \cdot N+l}^{(p-1) \cdot N+q}=a_{k}^{p} b_{q}^{l}-\delta_{k}^{p} \delta_{q}^{l} . \tag{28}
\end{equation*}
$$

Consider that there are $n$ sets of known manipulator movement, $\mathbf{H}_{i}$, corresponding to the $i$ th movement, therefore we obtain

$$
\left[\begin{array}{c}
\mathbf{H}_{1}  \tag{29}\\
\cdots \\
\mathbf{H}_{n}
\end{array}\right] \mathbf{x}=\mathbf{B} \mathbf{x}=\mathbf{0} .
$$

Based upon least-squares solution of a homogeneous system of linear equations, $\mathbf{x}$ is the last column of $\mathbf{V}$, where $\mathbf{B}=\mathbf{U D V}^{T}$. Note

$$
\mathbf{R}^{\prime}=\left[\begin{array}{rrr}
x_{1}^{1} & x_{2}^{1} & x_{3}^{1}  \tag{30}\\
x_{1}^{2} & x_{2}^{2} & x_{3}^{2} \\
x_{1}^{3} & x_{2}^{3} & x_{3}^{3}
\end{array}\right]
$$

TAbLE 3: Inverse kinematic results of the generalized manipulator.

| Roll/(rad) | Pitch/(rad) | Yaw/(rad) | $p_{x} /(\mathrm{cm})$ | $p_{y} /(\mathrm{cm})$ | $p_{z} /(\mathrm{cm})$ | $\theta_{1} /(\mathrm{rad})$ | $d_{2} /(\mathrm{cm})$ | $\theta_{3} /(\mathrm{rad})$ | $\theta_{4} /(\mathrm{rad})$ | $\theta_{5} /(\mathrm{rad})$ | $\theta_{6} /(\mathrm{rad})$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.4449 | 0.2566 | -1.2466 | 129.4143 | 15.7087 | 76.7968 | 0.2417 | 122.6 | -0.7854 | 0.1366 | -0.6283 | -0.3927 |
|  |  |  |  |  |  | 0.1922 | 123.7487 | 1.1579 | 0.1366 | -2.5133 | -0.6142 |
|  |  |  |  |  |  | 0.0880 | 127.4757 | -1.3771 | -0.5851 | 0.1454 | -1.0836 |
|  |  |  |  |  |  | -5.9305 | 121.3515 | 1.7428 | -0.5851 | 2.9962 | 0.0767 |
| -0.2411 | -0.4617 | -0.2266 | 12.5929 | $22.4547$ | $69.2228$ | 1.0203 | 19.3710 | -1.0229 | 0.5987 | -0.2428 | 1.0851 |
|  |  |  |  |  |  | 0.2661 | 40.1966 | 2.4693 | 0.5987 | -2.8988 | -0.0848 |
|  |  |  |  |  |  | 0.1653 | 52.60 | -0.4488 | -1.0472 | 0.1571 | -0.5417 |
|  |  |  |  |  |  | -4.5806 | 20.1375 | 1.3136 | -1.0472 | 2.9845 | 1.5420 |

because $\mathbf{R}$ is a rotation matrix whose Frobenius norm is $\sqrt{3}$ and $\operatorname{det}(\mathbf{R})>0$, we can determine $\mathbf{R}$ as

$$
\begin{equation*}
\mathbf{R}= \pm \sqrt{3} \mathbf{R}^{\prime} \tag{31}
\end{equation*}
$$

Till then, the unique $\mathbf{R}$ verifying constraint $\operatorname{det}(\mathbf{R})>0$ can be distinguished from (31). Furthermore, we choose $\mathbf{R}_{F}=$ $\mathbf{U} \mathbf{V}^{T}$, where $\mathbf{R}=\mathbf{U D V}{ }^{T}$ as the final solution of (22).

After solving $\mathbf{R}$, the unknown $\mathbf{t}$ can also be solved using the next equation:

$$
\mathbf{t}=\left[\begin{array}{c}
\mathbf{I}-\mathbf{R}_{A}^{1}  \tag{32}\\
\cdots \\
\mathbf{I}-\mathbf{R}_{A}^{n}
\end{array}\right]^{+} \cdot\left[\begin{array}{c}
\mathbf{R}_{A}^{1} \cdot \mathbf{R} \cdot \mathbf{t}_{B}^{1}+\mathbf{t}_{A}^{1} \\
\cdots \\
\mathbf{R}_{A}^{n} \cdot \mathbf{R} \cdot \mathbf{t}_{B}^{n}+\mathbf{t}_{A}^{n}
\end{array}\right] .
$$

## 4. Design of Visual Servo Switch Control Scheme for Grasping

The visual servo control scheme designed in this section consists of two steps: eye-fixed approximation of the household object and static look-and-move grasping. Once the VBMMS is commanded to grasp a household object that is in its camera's FOV, the VBMMS starts to approximate the object with its eye gazing it. When the distance between VBMMS's camera and household object reaches a certain value, the process of approximation switches to the static look-andmove grasping.
4.1. Eye-Fixed Approximation. Figure 4 illustrates the scheme of eye-fixed approximation. In Figure $4, m_{00}$ and $m_{00}^{*}$ correspond to current image and desired image, respectively, and both are called zero-order moment of the blue concentric ring region to the object mark. ${ }^{c} \mathbf{m}$ and ${ }^{c^{*}} \mathbf{m}$ are called nonhomogeneous projective coordinate of the center of blue concentric ring region to the object mark, which also correspond to current image and desired image, respectively.

It is well known that for nonhomogeneous projective coordinate ${ }^{c} \mathbf{m}=[x, y]^{T}=\left[{ }^{c} X /{ }^{c} Z,{ }^{c} Y /{ }^{c} Z\right]^{T}$, its


Figure 4: Scheme of eye-fixed approximation.
time derivative ${ }^{c} \dot{\mathbf{m}}$ is linearly related to the joints' velocity $\dot{\mathbf{q}}$ through the interaction matrix $\mathbf{L}_{c_{\mathrm{m}}}$ :

$$
\begin{align*}
\mathbf{L}_{c_{\mathrm{m}}}= & {\left[\begin{array}{cccccc}
-\frac{1}{{ }^{c} Z} & 0 & \frac{x}{{ }^{c} Z} & x y & -\left(1+y^{2}\right) & y \\
0 & -\frac{1}{{ }^{c} Z} & \frac{y}{{ }^{c} Z} & 1+y^{2} & -x y & -x
\end{array}\right] }  \tag{33}\\
& \times\left[\begin{array}{cc}
{ }^{E} \mathbf{R}_{c}^{T} & -{ }^{E} \mathbf{R}_{c}^{T}\left[{ }_{c}^{E} \mathbf{t}_{c}\right]_{\times} \\
0 & { }^{E} \mathbf{R}_{c}^{T}
\end{array}\right]{ }^{E} \mathbf{J}(\mathbf{q}),
\end{align*}
$$

where $\mathbf{P}=\left[{ }^{c} X,{ }^{c} Y,{ }^{c} Z\right]^{T}$ is the coordinates of a 3 D point relative to current camera frame, ${ }^{E} \mathbf{R}_{c},{ }^{E} \mathbf{t}_{c}$ are the results of hand-eye relationship, and ${ }^{E} \mathbf{J}(\mathbf{q})$ is the generalized manipulator's Jacobian in end-effector coordinate frame $\{E\}$. For the unknown variable ${ }^{c} Z$ in $\mathbf{L}_{c_{\mathrm{m}}}$, its estimation ${ }^{c} \widehat{Z}$ can be chosen as follows:

$$
\begin{equation*}
{ }^{c} \widehat{Z}={ }^{c^{*}} Z \sqrt{\frac{m_{00}^{*}}{m_{00}}} \tag{34}
\end{equation*}
$$

where ${ }^{c^{*}} Z$ is the constant depth of the 3 D point relative to the desired camera frame.

We partition the interaction matrix so as to isolate the second degree of freedom of the generalized manipulator. Note the following:

$$
\begin{align*}
\mathbf{L}_{d} & =\mathbf{L}_{c_{\mathrm{m}}} \operatorname{diag}\left(\begin{array}{llllll}
0 & 1 & 0 & 0 & 0 & 0
\end{array}\right)  \tag{35}\\
\mathbf{L}_{d^{\perp}} & =\mathbf{L}_{c_{\mathrm{m}}} \operatorname{diag}\left(\begin{array}{lllllll}
1 & 0 & 0 & 1 & 1 & 1
\end{array}\right)
\end{align*}
$$

in which $\operatorname{diag}(\cdot)$ is a diagonal matrix, $\mathbf{L}_{d}$ only includes the second column of $\mathbf{L}_{c_{\mathrm{m}}}$, and $\mathbf{L}_{d^{\perp}}$ includes the first, fourth, fifth, and sixth columns of $\mathbf{L}_{c_{m}}$. We get

$$
\begin{equation*}
{ }^{c} \dot{\mathbf{m}}=\mathbf{L}_{d} \dot{\mathbf{q}}_{d}+\mathbf{L}_{d^{\perp}} \dot{\mathbf{q}}_{d^{\perp}} . \tag{36}
\end{equation*}
$$



Figure 5: Scheme of static look-and-move grasping.

Note $e_{2}=m_{00}-\alpha m_{00}^{*}$, where $\alpha$ is a constant coefficient; a simple approximation control law can be acquired by using feature $e_{2}$, as

$$
\dot{\mathbf{q}}_{d}=\left[\begin{array}{llllll}
0 & -\lambda_{1} \operatorname{sgn}\left(e_{2}\right) & 0 & 0 & 0 & 0 \tag{37}
\end{array}\right]^{T},
$$

where $\operatorname{sgn}(\cdot)$ is sign function and $\lambda_{1}$ is a time-invariant coefficient. Note $\mathbf{e}_{1}={ }^{c} \mathbf{m}-{ }^{c^{*}} \mathbf{m}$; let $\mathbf{e}_{1}$ be exponential convergent $\dot{\mathbf{e}}_{1}=-\lambda \mathbf{e}_{1}$. There is

$$
\begin{equation*}
\dot{\mathbf{q}}_{d^{\perp}}=-\mathbf{L}_{d^{\perp}}^{+}\left(\lambda \mathbf{e}_{1}+\mathbf{L}_{d} \dot{\mathbf{q}}_{d}\right) . \tag{38}
\end{equation*}
$$

The vector $\lambda \mathbf{e}_{1}+\mathbf{L}_{d} \dot{\mathbf{q}}_{d}$ can be considered as a modified error that incorporates the original error $\mathbf{e}_{1}$ while taking into account the error that will be induced by robot motion $\dot{\mathbf{q}}_{d}$.

Under the influence of control input shown in (37) and (38), the VBMMS approaches to object with a constant speed untile $e_{2}=0$, and during the approximation process, the object is maintained in the FOV from beginning to end.
4.2. Static Look-and-Move Grasping. Figure 5 illustrates the scheme of static look-and-move grasping.

In Figure 5, 3D object point $\mathbf{P}$ is on the plane $\pi$ whose normal is $\mathbf{n}$. Meanwhile, there are three different camera frames: the desired camera frame $\left\{c^{*}\right\}$, the initial camera frame $\{c\}$, and the current camera frame $\left\{c^{\prime}\right\}$, that is, after a manipulator's known movement from $\mathbf{q}_{i}$ to $\mathbf{q}_{c}$. The homogeneous transformation ${ }^{c^{\prime}} \mathbf{T}_{c}$ can be computed as

$$
{ }^{c^{\prime}} \mathbf{T}_{c}=\left[\begin{array}{cc}
\mathbf{R}^{\prime} & \mathbf{t}^{\prime}  \tag{39}\\
0^{T} & 1
\end{array}\right]={ }^{E} \mathbf{T}_{c}^{-1} \cdot{ }^{0} \mathbf{T}_{E^{\prime}}^{-1} \cdot{ }^{0} \mathbf{T}_{E} \cdot{ }^{E} \mathbf{T}_{c}
$$

where ${ }^{E} \mathbf{T}_{c}$ is the hand-eye relationship, ${ }^{0} \mathbf{T}_{E}$ is the initial end-effector's pose, and ${ }^{0} \mathbf{T}_{E^{\prime}}$ is the current end-effector's pose. Our ultimate objective is to determine the pose of $\left\{c^{\prime}\right\}$
relative to $\left\{c^{*}\right\}$, namely, $\mathbf{R}, \mathbf{t}$. In order to solve them, two steps are needed.

Step 1. Structure reconstruction of plane $\pi$ with respect to camera frame $\left\{c^{\prime}\right\}$.

Take current camera frame $\{c\}$ as the reference frame, the plane $\pi$ can be noted as ${ }^{c} \pi=\left[{ }^{c} \mathbf{n}^{T},{ }^{c} d\right]^{T}$ and point $\left[{ }^{c} \mathbf{m}^{T}, 1\right]^{T}$ in frame $\{c\}$ is related to the corresponding point $\left[{ }^{\left[c^{\prime}\right.} \mathbf{m}^{T}, 1\right]^{T}$ in frame $\left\{c^{\prime}\right\}$ by Euclidean homography $\mathbf{H}^{\prime}$ :

$$
\left[\begin{array}{c}
{ }^{c^{\prime}} \mathbf{m}  \tag{40}\\
1
\end{array}\right] \cong\left(\mathbf{R}^{\prime}-\frac{\mathbf{t}^{\prime} \cdot{ }^{c} \mathbf{n}^{T}}{{ }^{c} d}\right)\left[\begin{array}{c}
{ }^{c} \mathbf{m} \\
1
\end{array}\right] \cong \mathbf{H}^{\prime}\left[\begin{array}{c}
{ }^{c} \mathbf{m} \\
1
\end{array}\right]
$$

where $\cong$ means that the $\mathbf{H}^{\prime}$ transforms one vector into the other, up to a scale factor. Given at least four sets of image point correspondences ${ }^{c^{\prime}} \mathbf{m}_{i} \leftrightarrow{ }^{c} \mathbf{m}_{i}, i=1, \ldots, n$, the $\mathbf{H}^{\prime}$ satisfying $\left\|\mathbf{H}^{\prime}\right\|_{F}=1$ can be derived using direct linear transformation (DLT) method.

After solving $\mathbf{H}^{\prime}$, written as $\mathbf{H}^{\prime}=\left[\begin{array}{lll}\mathbf{h}_{1}^{\prime} & \mathbf{h}_{2}^{\prime} & \mathbf{h}_{3}^{\prime}\end{array}\right]$ and $\mathbf{R}^{\prime}=$ $\left[\begin{array}{lll}\mathbf{r}_{1} & \mathbf{r}_{2} & \mathbf{r}_{3}\end{array}\right],{ }^{c} \mathbf{n}=\left[\begin{array}{lll}n_{x} & n_{y} & n_{z}\end{array}\right]^{T}$, there are

$$
\begin{align*}
& \lambda \mathbf{h}_{1}^{\prime}=\mathbf{r}_{1}-\frac{n_{x}}{c} \mathbf{t}^{\prime} \\
& \lambda \mathbf{h}_{2}^{\prime}=\mathbf{r}_{2}-\frac{n_{y}}{c} \mathbf{t}^{\prime}  \tag{41}\\
& \lambda \mathbf{h}_{3}^{\prime}=\mathbf{r}_{3}-\frac{n_{z}}{{ }^{c} d} \mathbf{t}^{\prime} .
\end{align*}
$$

Furthermore

$$
\left[\begin{array}{cccc}
\mathbf{r}_{1}^{T} \mathbf{h}_{1}^{\prime} & \mathbf{r}_{1}^{T} \mathbf{t}^{\prime} & 0 & 0  \tag{42}\\
\mathbf{r}_{2}^{T} \mathbf{h}_{2}^{\prime} & 0 & \mathbf{r}_{2}^{T} \mathbf{t}^{\prime} & 0 \\
\mathbf{r}_{3}^{T} \mathbf{h}_{3}^{\prime} & 0 & 0 & \mathbf{r}_{3}^{T} \mathbf{t}^{\prime} \\
\mathbf{r}_{2}^{T} \mathbf{h}_{1}^{\prime} & \mathbf{r}_{2}^{T} \mathbf{t}^{\prime} & 0 & 0 \\
\mathbf{r}_{3}^{T} \mathbf{h}_{1}^{\prime} & \mathbf{r}_{3}^{T} \mathbf{t}^{\prime} & 0 & 0 \\
\mathbf{r}_{1}^{T} \mathbf{h}_{2}^{\prime} & 0 & \mathbf{r}_{1}^{T} \mathbf{t}^{\prime} & 0 \\
\mathbf{r}_{3}^{T} \mathbf{h}_{2}^{\prime} & 0 & \mathbf{r}_{3}^{T} \mathbf{t}^{\prime} & 0 \\
\mathbf{r}_{1}^{T} \mathbf{h}_{3}^{\prime} & 0 & 0 & \mathbf{r}_{1}^{T} \mathbf{t}^{\prime} \\
\mathbf{r}_{2}^{T} \mathbf{h}_{3}^{\prime} & 0 & 0 & \mathbf{r}_{2}^{T} \mathbf{t}^{\prime}
\end{array}\right]\left[\begin{array}{l}
\lambda \\
\frac{n_{x}}{c_{d}} \\
\frac{n_{y}}{c} d \\
\frac{n_{z}}{c d}
\end{array}\right]=\left[\begin{array}{l}
1 \\
1 \\
1 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{array}\right] .
$$

By solving (42), ${ }^{c} \pi$ can be determined. Finally, using

$$
\begin{equation*}
{ }^{c^{\prime}} \mathbf{n}=\mathbf{R}^{\prime} \cdot{ }^{c} \mathbf{n} \quad c^{\prime} d={ }^{c} d-{ }^{c^{\prime}} \mathbf{n}^{T} \cdot \mathbf{t}^{\prime} \tag{43}
\end{equation*}
$$

we can acquire the structure of plane ${ }^{c^{\prime}} \pi$ with respect to camera frame $\left\{c^{\prime}\right\}$.

Step 2. Computation of $\mathbf{R}, \mathbf{t}$ of camera frame $\left\{c^{\prime}\right\}$ with respect to desired camera frame $\left\{c^{*}\right\}$.

Similar to Step 1, point $\left[{ }^{c^{\prime}} \mathbf{m}^{T}, 1\right]^{T}$ in frame $\left\{c^{\prime}\right\}$ is related to the corresponding point $\left[{ }^{c^{*}} \mathbf{m}^{T}, 1\right]^{T}$ in frame $\left\{c^{*}\right\}$ by Euclidean homography $\mathbf{H}$ :

$$
\left[\begin{array}{c}
{ }^{d} \mathbf{m}  \tag{44}\\
1
\end{array}\right] \cong\left(\mathbf{R}-\frac{\mathbf{t} \cdot{ }^{c^{\prime}} \mathbf{n}^{T}}{c^{\prime} d}\right)\left[\begin{array}{c}
c^{\prime} \mathbf{m} \\
1
\end{array}\right] \cong \mathbf{H}\left[\begin{array}{c}
c^{\prime} \mathbf{m} \\
1
\end{array}\right]
$$

and the $\mathbf{H}$ satisfying $\|\mathbf{H}\|_{F}=1$ can also be derived using DLT method if there are $n>4$ sets of image point correspondences.

Adjust the solved $\mathbf{H}$ so that it satisfies $\operatorname{det}(\mathbf{H})>0$; there are

$$
\begin{equation*}
\alpha \mathbf{H}=\mathbf{R}-\frac{\mathbf{t} \cdot{ }^{c^{\prime}} \mathbf{n}^{T}}{c^{\prime} d} \tag{45}
\end{equation*}
$$

where the scale factor $\alpha>0$. Furthermore,

$$
\begin{equation*}
\alpha\left[c^{c^{\prime}} \mathbf{n}\right]_{\times} \mathbf{H}^{T}=\left[{ }^{c^{\prime}} \mathbf{n}\right]_{X} \mathbf{R}^{T} . \tag{46}
\end{equation*}
$$

As we all know, the rotation matrix $\mathbf{R}_{k, \theta}$ whose axis and angle are $\mathbf{k}$ and $\theta$ can be described as

$$
\begin{equation*}
\mathbf{R}_{k, \theta}=\cos \theta \mathbf{I}_{3 \times 3}+\sin \theta[\mathbf{k}]_{\times}+(1-\cos \theta) \mathbf{k} \mathbf{k}^{T} \tag{47}
\end{equation*}
$$

Note $\mathbf{e}_{3}=\left[\begin{array}{lll}0 & 0 & 1\end{array}\right]^{T}$. Choose

$$
\begin{align*}
& \mathbf{k}_{1}=\frac{c^{\prime} \mathbf{n} \times \mathbf{e}_{3}}{\left\|c^{c^{\prime}} \mathbf{n} \times \mathbf{e}_{3}\right\|_{2}}  \tag{48}\\
& \theta_{1}=\arccos ^{c^{\prime}} \mathbf{n}^{T} \mathbf{e}_{3}
\end{align*}
$$

then

$$
\begin{equation*}
\mathbf{R}_{k_{1}, \theta_{1}}{ }^{c^{\prime}} \mathbf{n}=\mathbf{e}_{3} . \tag{49}
\end{equation*}
$$

Substitute ${ }^{c^{\prime}} \mathbf{n}=\mathbf{R}_{k_{1}, \theta_{1}}^{T} \mathbf{e}_{3}$ into (46); according to Nanson's equation $\mathbf{A}^{T}[\mathbf{A t}]_{\times} \mathbf{A}=[\mathbf{t}]_{\times}$, one obtains

$$
\alpha\left[\begin{array}{ccc}
0 & -1 & 0  \tag{50}\\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right] \mathbf{R}_{k_{1}, \theta_{1}} \mathbf{H}^{T}=\left[\begin{array}{ccc}
0 & -1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{array}\right] \mathbf{R}_{k_{1}, \theta_{1}} \mathbf{R}^{T} .
$$

Equation (50) means that the first two rows of matrix $\alpha \mathbf{R}_{k_{1}, \theta_{1}} \mathbf{H}^{T}$ and $\mathbf{R}_{k_{1}, \theta_{1}} \mathbf{R}^{T}$ are equal, respectively. Note

$$
\mathbf{R}_{k_{1}, \theta_{1}} \mathbf{H}^{T}=\left[\begin{array}{c}
\mathbf{r}_{1}^{T}  \tag{51}\\
\mathbf{r}_{2}^{T} \\
\mathbf{c}^{T}
\end{array}\right]
$$

then the positive scale factor $\alpha$ and the rotation matrix $\mathbf{R}$ are

$$
\begin{gather*}
\alpha=\frac{2}{\left\|\mathbf{r}_{1}\right\|_{2}+\left\|\mathbf{r}_{2}\right\|_{2}}  \tag{52}\\
\mathbf{R}=\left[\alpha \mathbf{r}_{1}, \alpha \mathbf{r}_{2}, \alpha^{2}\left(\mathbf{r}_{1} \times \mathbf{r}_{2}\right)\right] \mathbf{R}_{k_{1}, \theta_{1}} . \tag{53}
\end{gather*}
$$

To refine the solved $\mathbf{R}$, apply singular value decomposition to it. As $\mathbf{R}=\mathbf{U D V}^{T}$, choose $\mathbf{R}=\mathbf{U V}^{T}$ as the ultimate result. By substituting (52) and refined $\mathbf{R}$ into (45), the unknown $\mathbf{t}$ can finally be determined:

$$
\begin{equation*}
\mathbf{t}={ }^{c^{\prime}} d(\mathbf{R}-\alpha \mathbf{H})^{c^{\prime}} \mathbf{n} . \tag{54}
\end{equation*}
$$

Till then, the homogeneous transformation ${ }^{c^{*}} \mathrm{~T}_{c^{\prime}}$ can be acquired:

$$
{ }^{c^{*}} \mathbf{T}_{c^{\prime}}=\left[\begin{array}{cc}
\mathbf{R} & \mathbf{t}  \tag{55}\\
0^{T} & 1
\end{array}\right] .
$$

Furthermore, the transformation ${ }^{0} \mathbf{T}_{E^{*}}$, which represents the pose of the desired end-effector frame $\left\{E^{*}\right\}$ relative to the base frame $\{0\}$, can be computed as

$$
\begin{equation*}
{ }^{0} \mathbf{T}_{E^{*}}={ }^{0} \mathbf{T}_{E^{\prime}} \cdot{ }^{E} \mathbf{T}_{c} \cdot{ }^{c^{*}} \mathbf{T}_{c^{\prime}}^{-1} \cdot{ }^{E} \mathbf{T}_{c}^{-1} . \tag{56}
\end{equation*}
$$

In (56), ${ }^{0} \mathbf{T}_{E^{\prime}}$ can be known from the VBMMS's kinematic model established in Section 3.1, and the hand-eye relationship ${ }^{E} \mathbf{T}_{c}$ has already been known from Section 3.3. Finally, we can acquire the VBMMS's control input $\dot{\mathbf{q}}$ from the analytical inverse kinematics proposed in Section 3.2.

## 5. Experiments and Analysis

In this section, we will discuss two experiments: hand-eye relationship determination and the VBMMS's switch control scheme separately.
5.1. Experiment of Hand-Eye Relationship Determination. In order to simplify the complexity of grid corner extraction, choose a model plane which contains a pattern of $4 \times 4$ squares as the calibration object; the size of each square is $22 \mathrm{~mm} \times 22 \mathrm{~mm}$. Attach an object frame $\{o\}$ to it; the upper left corner of the first square is selected to be associated to the origin point of $\{o\}$, the $Z$ axis is perpendicular to the model plane pointing outward, the $X Y$-plane of $\{o\}$ is aligned with the plane, so that points on the plane have zero $Z$-coordinate, and $X, Y$ axes are paralleled to the sides of the checkerboard's square, respectively.

Let the TMR be static; take six images of the plane under different orientations caused by several known manipulator movements. The images are shown in Figure 6, whose resolutions are all $640 \times 480$, and the corners are detected as the intersection of straight lines fitted to each square.

Using Zhang's calibration method, the refined camera intrinsic parameter matrix $\mathbf{K}$, together with the refined extrinsic parameters ${ }^{c^{i}} \mathbf{T}_{o}(i=1, \ldots, 6)$ corresponding to the six images respectively, can be acquired:

$$
\mathbf{K}=\left[\begin{array}{ccc}
816.09924 & 0 & 391.05409  \tag{57}\\
0 & 800.56461 & 221.07601 \\
0 & 0 & 1
\end{array}\right]
$$

The known movements of the manipulator corresponding to six images are shown in Table 4.


Figure 6: Six images of a model plane under different orientation, together with the corners (indicated by red cross and blue square).


Figure 7: Images of the target for the desired/initial camera position and object mark recognition using Gaussian model and Hough transformation.

(a)


$$
\begin{array}{ll}
-q_{1} & \cdots \cdots \cdot q_{5} \\
--q_{4} & -\cdots \cdot q_{6}
\end{array}
$$

(b)

Figure 8: The computed eye-fixed approximation control input.

From Table 4, the homogeneous transformation ${ }^{0} \mathbf{T}_{E^{i}}$ can be computed using generalized manipulator's kinematic model. The computed ${ }^{0} \mathbf{T}_{E^{i}}$ and corresponding ${ }^{c^{i}} \mathbf{T}_{o}$ are shown in Table 5.

Combine image 1 and image 2 (noted as \{12\}). Using calibration method mentioned in Section 3.3, matrix $\mathbf{H}_{1} \in R^{9 \times 9}$ can be determined. Similarly, combinations of $\{13\},\{14\},\{15\}$, and $\{16\}$ can determine $\mathbf{H}_{2}, \mathbf{H}_{3}, \mathbf{H}_{4}, \mathbf{H}_{5}$, respectively. The hand-eye relationship ${ }^{E} \mathbf{T}_{c}$ can finally be determined:

$$
{ }^{E} \mathbf{T}_{c}=\left[\begin{array}{cccc}
0.0644 & 0.909 & 0.4117 & -10.4233  \tag{58}\\
-0.9879 & 0.1164 & -0.1023 & -6.1634 \\
-0.1409 & -0.4002 & 0.9055 & -17.2431 \\
0 & 0 & 0 & 1
\end{array}\right]
$$

Table 4: The known movements of the generalized manipulator.

| Image | $\theta_{1}, d_{2}, \theta_{3}$ | $\theta_{4}$ | $\theta_{5}$ | $\theta_{6}$ |
| :--- | :---: | :---: | :---: | :---: |
| 1 | 0 | 0.5756 | -0.0245 | -1.1241 |
| 2 | 0 | 0.5710 | -0.2224 | -1.0983 |
| 3 | 0 | 0.4560 | -0.2256 | -1.2930 |
| 4 | 0 | 0.2143 | 0.0360 | -1.5370 |
| 5 | 0 | -0.0012 | 0.1302 | -1.6249 |
| 6 | 0 | -0.3974 | -0.1872 | -1.9276 |

5.2. Experiment of VBMMS's Switch Control Scheme. The images corresponding to the desired and initial camera position and the object mark recognition using Gaussian model and Hough transformation are illustrated in Figure 7.

(a) Image corresponding to $\{c\}$

(b) Image corresponding to $\left\{c^{\prime}\right\}$

(c) The extracted corner correspondences matched by RANSAC

Figure 9: Two images of the object mark corresponding to $\{c\}$ and $\left\{c^{\prime}\right\}$, together with the extracted corner correspondences.

(a) Image corresponding to $\left\{c^{\prime}\right\}$

(b) Image corresponding to $\left\{c^{*}\right\}$

(c) The extracted corner correspondences matched by RANSAC

Figure 10: Two images of the object mark corresponding to $\left\{c^{\prime}\right\}$ and $\left\{c^{*}\right\}$, together with the extracted corner correspondences.

Table 5: The pose information corresponding to known movements of the generalized manipulator.

|  |  | Roll | Pitch | Yaw | $p_{x}$ | $p_{y}$ | $p_{z}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathrm{I}_{1}$ | ${ }^{0} \mathbf{T}_{E 1}$ | 2.9713 | 1.4401 | -2.9833 | 42.2281 | -0.6184 | 41.7720 |
|  | ${ }^{c 1} \mathbf{T}_{o}$ | -2.5304 | -0.1911 | -1.6459 | -6.2760 | -7.2463 | 45.1793 |
| $\mathrm{I}_{2}$ | ${ }^{0} \mathbf{T}_{E 2}$ | 1.9856 | 1.3546 | -2.0969 | 41.7209 | -5.4980 | 43.0438 |
|  | ${ }^{c 2} \mathbf{T}_{o}$ | -2.6322 | -0.0152 | -1.7485 | -17.8379 | -1.7928 | 44.5287 |
| $\mathrm{I}_{3}$ | ${ }^{0} \mathbf{T}_{E 3}$ | 2.2297 | 1.2953 | -2.3112 | 38.8304 | -6.0230 | 43.0616 |
|  | ${ }^{c 3} \mathbf{T}_{o}$ | -2.5396 | 0.0100 | -1.7451 | -20.4513 | -6.5954 | 45.0891 |
| $\mathrm{I}_{4}$ | ${ }^{0} \mathbf{T}_{E 4}$ | -2.9439 | 1.3868 | 2.9484 | 33.4385 | 1.0060 | 46.4443 |
|  | ${ }^{c 4} \mathbf{T}_{o}$ | -2.4884 | -0.2583 | -1.6244 | -2.9982 | -11.5786 | 54.2764 |
|  | ${ }^{0} \mathbf{T}_{E 5}$ | -1.9581 | 1.4303 | 1.9545 | 27.6891 | 3.6303 | 52.0272 |
| $\mathrm{I}_{5}$ | ${ }^{c 5}$ | $\mathbf{T}_{o}$ | -2.6579 | -0.3423 | -1.6116 | 2.7396 | -2.0687 |
|  | ${ }^{0} \mathbf{T}_{E 6}$ | 1.3769 | 1.3922 | -1.3078 | 16.0177 | -4.8820 | 63.8751 |
| $\mathrm{I}_{6}$ | ${ }^{c 6} \mathbf{T}_{o}$ | -2.7377 | -0.0032 | -1.5332 | -23.9188 | 54.7535 |  |
|  |  |  |  |  |  | 1.5681 | 71.5180 |

As can be seen in Figure 7, the blue concentric ring region of the object mark can be detected robustly and conveniently using Gaussian model for color-based segmentation and Hough transformation for segmented region's ellipse fitting.

Choose $\alpha=0.5$, and the gain involved in the approximation control law (37) $\lambda_{1}=0.01$. The gain involved in the eyefixed control law (38) $\lambda=0.15, c^{*} Z$ is set to be 25 cm while its true value is 31 cm ; the computed eye-fixed approximation control input is given in Figure 8.

Figure 9 shows the images corresponding to the camera position $\{c\}$, where the eye-fixed approximation process terminated, and the position $\left\{c^{\prime}\right\}$, where the VBMMS does a known movement from $\mathbf{q}=[0,0,0,1.2820,-0.0240,-0.2175]^{T}$ to $\mathbf{q}^{\prime}=$ $[0,0,0,1.3894,-0.1964,-0.0416]^{T}$, together with the extracted corner corre-spondences, matched by RANSAC (random sample consensus) method. In addition to the corner extraction, the ROI (region of interest) is selected to be within the concentric ring.

The corner correspondences given in Figure 9 is leading to the homography $\mathbf{H}^{\prime}$ satisfying $\left\|\mathbf{H}^{\prime}\right\|_{F}=1$. Using Step 1 mentioned in Section 3.2, the 3D structure of the object mark plane with respect to the camera frame $\left\{c^{\prime}\right\}$ can be computed:

$$
c^{\prime} \pi=\left[\begin{array}{llll}
-0.2428 & 0.2620 & -0.9340 & 15.7297 \tag{59}
\end{array}\right]^{T}
$$

Figure 10 shows the images corresponding to the camera position $\left\{c^{\prime}\right\}$ and the desired camera position $\left\{c^{*}\right\}$, together with the extracted corner correspondences matched by RANSAC.

Applying DLT to the corner correspondences, the homography $\mathbf{H}$ satisfying $\operatorname{det}(\mathbf{H})>0$ can be acquired; then using Step 2 mentioned in Section 3.2, the transformation ${ }^{\prime} \mathbf{T}_{c^{*}}$ can be computed. Finally the ${ }^{0} \mathbf{T}_{E^{*}}$ and the control input of the VBMMS can be determined from the inverse kinematics discussed in Section 3.2:

$$
\dot{\mathbf{q}}=\left[\begin{array}{llllll}
-0.1847 & 5.3829 & 0.1237 & 1.3305 & -0.0058 & -0.1220 \tag{60}
\end{array}\right]^{T} .
$$

Affected by the computed control input, the VBMMS moved to a position corresponding to the desired camera
frame $\left\{c^{*}\right\}$. Thereby, the visual servo grasping task is executed successfully.

## 6. Conclusion

It is nearly impossible for VBMMS to finish household objects' grasping and delivering operation without some prior knowledge (such as objects' color, texture, sizes, and localization) provided by people, not only for the household objects' variety and operation complexity, but also for the difficulties of the VBMMS's kinematic modeling and its nonholonomic constraint's handling. On the one hand, a new QR Code-based artificial object mark is designed, which can store object's property information and operation information and can be easily distinguished from complex family environment. On the other hand, in order to model the VBMMS, we summarize it as a generalized manipulator, followed by acquiring its analytical inverse kinematic solutions and determining the hand-eye relationship based upon active vision. Meanwhile, in order to deal with the VBMMS's nonholonomic constraint, a visual servo switching control law which is composed of eye-fixed approximation part and static look-and-move grasping part is designed. The proposed scheme can solve the household objects' grasping and delivering problem well, and makes it possible to let VBMMS-type service robot provide better housekeeping service.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This project was supported by the National Natural Science Foundation of China (no. 61375084) and Advanced Mechanical Design and Manufacturing Public Service Platform Building, Longyan Science and Technology Plan Project. (no. 2012ly01).

## References

[1] S. Ekvall, D. Kragic, and P. Jensfelt, "Object detection and mapping for service robot tasks," Robotica, vol. 25, no. 2, Article ID 00323, pp. 175-187, 2007.
[2] K. Severinson-Eklundh, A. Green, and H. Hüttenrauch, "Social and collaborative aspects of interaction with a service robot," Robotics and Autonomous Systems, vol. 42, no. 3-4, pp. 223-234, 2003.
[3] W. F. Carriker, P. K. Khosla, and B. H. Krogh, "Path planning for mobile manipulators for multiple task execution," IEEE Transactions on Robotics and Automation, vol. 7, no. 3, pp. 403408, 1991.
[4] Q. Huang, K. Tanie, and S. Sugano, "Coordinated motion planning for a mobile manipulator considering stability and manipulation," International Journal of Robotics Research, vol. 19, no. 8, pp. 732-742, 2001.
[5] A. Mohri, S. Furuno, and M. Yamamoto, "Trajectory planning of mobile manipulator with end-effector's specified path," in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 4, pp. 2264-2269, Maui, Hawaii, USA, November 2001.
[6] K. Tchon, J. Jakubiak, and R. Muszynski, "Regular Jacobian motion planning algorithms for mobile manipulators," in Proceedings of the 15th IFAC World Congress, vol. 15, Barcelona, Spain, 2002.
[7] J. Vannoy and J. Xiao, "Real-time adaptive motion planning (RAMP) of mobile manipulators in dynamic environments with unforeseen changes," IEEE Transactions on Robotics, vol. 24, no. 5, pp. 1199-1212, 2008.
[8] S. Ide, T. Takubo, K. Ohara, Y. Mae, and T. Arai, "Realtime trajectory planning for mobile manipulator using model predictive control with constraints," in Proceedings of the 8th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI '11), pp. 244-249, Incheon, Republic of Korea, November 2011.
[9] J. H. Chung, S. A. Velinsky, and R. A. Hess, "Interaction control of a redundant mobile manipulator," International Journal of Robotics Research, vol. 17, no. 12, pp. 1302-1309, 1998.
[10] A. Mazur, "Hybrid adaptive control laws solving a path following problem for non-holonomic mobile manipulators," International Journal of Control, vol. 77, no. 15, pp. 1297-1306, 2004.
[11] S. Lin and A. A. Goldenberg, "Robust damping control of mobile manipulators," IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics, vol. 32, no. 1, pp. 126-132, 2002.
[12] M. Mailah, E. Pitowarno, and H. Jamaluddin, "Robust motion control for mobile manipulator using resolved acceleration and proportional-integral active force control," International Journal of Advanced Robotic Systems, vol. 2, no. 2, pp. 125-134, 2005.
[13] M. Galicki, "Control of mobile manipulators in a task space," IEEE Transactions on Automatic Control, vol. 57, no. 11, pp. 2962-2967, 2012.
[14] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," IEEE Transactions on Robotics and Automation, vol. 12, no. 5, pp. 651-670, 1996.
[15] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," IEEE Robotics and Automation Magazine, vol. 13, no. 4, pp. 82-90, 2006.
[16] F. Chaumette and S. Hutchinson, "Visual servo control. II. Advanced approaches," IEEE Robotics and Automation Magazine, vol. 14, no. 1, pp. 109-118, 2007.
[17] W. J. Wilson, C. C. W. Hulls, and G. S. Bell, "Relative endeffector control using cartesian position based visual servoing," IEEE Transactions on Robotics and Automation, vol. 12, no. 5, pp. 684-696, 1996.
[18] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, "Position based visual servoing: keeping the object in the field of vision," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02), vol. 2, pp. 1624-1629, Washington, DC, USA, May 2002.
[19] P. I. Corke and S. A. Hutchinson, "A new partitioned approach to image-based visual servo control," IEEE Transactions on Robotics and Automation, vol. 17, no. 4, pp. 507-515, 2001.
[20] R. Mahony, P. Corke, and T. Hamel, "Dynamic image-based visual servo control using centroid and optic flow features," Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, vol. 130, no. 1, Article ID 011005, 2008.
[21] E. Malis, F. Chaumette, and S. Boudet, "2-1/2-D visual servoing," IEEE Transactions on Robotics and Automation, vol. 15, no. 2, pp. 238-250, 1999.
[22] E. Malis and F. Chaumette, "21/2D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement," International Journal of Computer Vision, vol. 37, no. 1, pp. 79-97, 2000.
[23] E. Malis and F. Chaumette, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," IEEE Transactions on Robotics and Automation, vol. 18, no. 2, pp. 176-186, 2002.
[24] Y. Mezouar and F. Chaumette, "Path planning in image space for robust visual servoing," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '00), vol. 3, pp. 2759-2764, April 2000.
[25] Y. Ma, J. Kosěcká, and S. S. Sastry, "Vision guided navigation for a nonholonomic mobile robot," IEEE Transactions on Robotics and Automation, vol. 15, no. 3, pp. 521-536, 1999.
[26] 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 B: Cybernetics, vol. 31, no. 3, pp. 341-352, 2001.
[27] D. Amarasinghe, G. K. I. Mann, and R. G. Gosine, "Visionbased hybrid control scheme for autonomous parking of a mobile robot," Advanced Robotics, vol. 21, no. 8, pp. 905-930, 2007.
[28] R. F. Vassallo, H. J. Schneebeli, and J. Santos-Victor, "Visual servoing and appearance for navigation," Robotics and Autonomous Systems, vol. 31, no. 1, pp. 87-97, 2000.
[29] A. de Luca, G. Oriolo, and P. R. Giordano, "Image-based visual servoing schemes for nonholonomic mobile manipulators," Robotica, vol. 25, no. 2, Article ID 00326, pp. 131-145, 2007.
[30] N. Mansard, O. Stasse, F. Chaumette, and K. Yokoi, "Visuallyguided grasping while walking on a humanoid robot," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '07), pp. 3041-3047, April 2007.
[31] Y. Wang, H. Lang, and C. W. de Silva, "A hybrid visual servo controller for robust grasping by wheeled mobile robots," IEEE/ASME Transactions on Mechatronics, vol. 15, no. 5, pp. 757769, 2010.
[32] Z. Zhang, "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, 2000.


International Journal of Distributed
Sensor Networks


## Hindawi

Submit your manuscripts at
http://www.hindawi.com


