Model Predictive Control for Uncalibrated and Constrained Image-Based Visual Servoing Without Joint Velocity Measurements

ZHOUJINGZI QIU, SHIQIANG HU[iD], AND XINWU LIANG[iD]

School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Shiqiang Hu (sqhu@ sjtu.edu.cn)

ABSTRACT This paper presents a novel scheme for image-based visual servoing (IBVS) of a robot manipulator by considering robot dynamics without using joint velocity measurements in the presence of constraints, uncalibrated camera intrinsic and extrinsic parameters and unknown feature position parameters. An approach to design model predictive control (MPC) method based on identification algorithm and sliding mode observer has been proposed. Based on the MPC method, the IBVS tasks can be considered as a nonlinear optimization problem while the constraints due to the visibility constraint and the torque constraint can be explicitly taken into account. By using the depth-independent interaction matrix framework, the identification algorithm can be used to update the unknown parameters and the prediction model. In addition, many existing controllers require the joint velocity measurements which can be contaminated by noises, thus resulting in the IBVS performance degradation. To overcome the problem without joint velocity measurements, the sliding mode observer is designed to estimate the joint velocities of the IBVS system. The simulation results for both eye-in-hand and eye-to-hand camera configurations are presented to verify the effectiveness of the proposed control method.

INDEX TERMS Image-based visual servoing, model predictive control, constrained optimization control, depth-independent interaction matrix, sliding mode observer.

I. INTRODUCTION

Visual servoing has been used widely in robotics field to make the machines more flexible and intelligent. In general, the visual features are used as the feedback signals to control the motion of a robot manipulator. The visual information can be obtained from the vision system either mounted on the end-effector of the robot manipulator (eye-in-hand) or fixed at the position near the manipulator (eye-to-hand). The visual servoing schemes can be divided into three types which depend on the feedback information: position-based visual servoing (PBVS) [1]–[3], image-based visual servoing [4]–[8] and hybrid visual servoing [9]–[11]. In the position-based visual servoing, the control errors are the Cartesian errors of relative pose between the end-effector and the target. The relative pose is computed by using the pose estimation algorithm from the image measurements [12]. In the image-based visual servoing, the control errors are the image errors between the current features and the desired features. The hybrid visual servoing is also called 2 1/2-D visual servoing which combines the PBVS and IBVS, the control errors are chosen visual features as the feedback signals defined partly in 2-D, and partly in 3-D [5].

The classical image-based visual servoing uses the proportional control law which is designed based on the traditional image Jacobian matrix [4], where there may exist the problem of local minima and singularities in the image Jacobian matrix. In [13], an IBVS control scheme with eye-in-hand camera configuration is proposed to solve the problem of regulation of robot's end-effector with the remote radio unit, where the PD type controller is designed on the basis of the idea of sliding surface from variable structure control. In [14], the reinforcement learning scheme is used to adaptively adjust the proportional servoing gain of IBVS system,

instead of employing a constant gain, where a bagging method is used to compute the inverse kinematics. However, these methods [4], [13], [14] have not explicitly handled the constraints which are important for the IBVS controller designing. For instance, if the initial and desired positions of the features are far away, the features may get out of the camera's field of view and the actuator limitations may be violated in the visual servoing process, which will lead to the failure of the visual servoing tasks. Hence, the satisfaction of the constraints is an crucial issue. A great deal of approaches focus on the constraints handling in the design of visual servoing controllers.

One approach to cope with constraints is via path planning technique [15]–[17]. A collision-free trajectory can be generated by the potential field approach in [15], where the constraints such that the target stays into the camera's field of view or the robot avoids the joint limitations are considered. In [16], [17], the path planning in the camera space has been designed by using LMI optimization, where the camera intrinsic and extrinsic parameters and the model of the target should be known. Another way to handle the constraints in the vision-based control is through numerous advanced control schemes, such as optimal control [18]–[20], LMI [21], [22], sensor-based control [23], [24], switch control [25], [26], neural network [27], and predictive control [28]–[31]. In [22], an LMI optimization is proposed by minimizing the error norm to handle visibility and kinematic constraints. A sensor-based control method is proposed in [24] to deal with constraints, where a weighting matrix is chosen to make the contribution of each feature with regularity. In [25], the authors proposed a switching approach among position-based visual servoing schemes and backward motion to keep all features within the camera's field of view, where the extrinsic parameters should be given and the chattering phenomenon would be caused by the switches between control patterns. In [27], a recurrent neural network scheme with eye-in-hand camera configuration is proposed for the IBVS system to deal with joint angle and velocity limits of the robot manipulators and the pseudoinversion performing of the image Jacobian matrix are not needed, where the camera intrinsic and extrinsic parameters are assumed to be known and the nonlinear robot dynamics are not considered. Moreover, it is well known that one of the advantage of model predictive control is that constraints can be handled explicitly, and numerous MPC-based control approaches [28]–[31] are proposed to deal with constraints in image-based visual servoing system. In [29], a MPC-based IBVS method designed based on the traditional image Jacobian matrix is proposed, where the contribution of the image prediction is indicated. A quasi-min-max MPC method designed by using the polytopic model of IBVS with the fixed depth value is proposed in [30] for constrained IBVS. In [31], a robust MPC scheme is introduced to cope with system constraints and prevent the visual features from leaving the camera's field of view. However, the above mentioned model predictive control

methods [28]–[31] for constrained visual servoing systems require the knowledge of the camera intrinsic and extrinsic parameters, and the depth information should be given. Moreover, these model predictive control methods design the velocity control command by using visual feedback without taking into account the nonlinear robot dynamics, which assumes that the robot manipulator can perfectly perform the velocity control. When nonlinear forces have dominated influences in high-speed tasks, the kinematic-based schemes cannot ensure satisfactory performance.

In this paper, we propose a novel model predictive control approach to regulate a set of feature points on the image plane to desired positions by considering nonlinear robot dynamics without the use of joint velocity measurements in the presence of unknown camera intrinsic and extrinsic parameters, unknown feature position parameters, the visibility constraint and the torque constraint. The MPC scheme based on the identification algorithm and sliding mode observer for the IBVS system has been presented. By using the MPC strategy, the visual servoing task is written as a nonlinear optimization control problem on the image plane, where the visibility constraint and the torque constraint can be easily handled. Due to the prediction algorithm and the receding horizon strategy, the performance of the control system can achieve highly efficient. Moreover, since the camera calibration is usually a costly, tedious and error prone process, and it is difficult to measure the depth online for the monocular vision, then, we adopt the depth-independent interaction matrix framework, and the identification algorithm is incorporated in the MPC for obtaining the model parameters and the MPC controller determines the control input and updates the parameter estimates at every time instant. In addition, to address the problem of absence of the joint velocity measurements, the MPC-based IBVS controller using estimated joint velocities, which is generated by a sliding mode observer, is designed. Compared to the traditional sliding mode observer [32], the proposed sliding mode observer here is designed based on a sigmoid function, which can reduce the chattering problem effectively. The proposed control approach can be used for both fixed and eye-in-hand camera configurations. Simulation results are provided to demonstrate the performance of the proposed control approach.

This article is organized as follows: Section 2 discusses the visual servoing model. The proposed control method is presented in Section 3, which provides the MPC formulation for IBVS system, the uncertainty and learning for IBVS, and the sliding mode observer design. In Section 4, simulation results are presented to demonstrate the effectiveness of the proposed control method. Finally, the conclusion is provided in Section 5.

II. VISUAL SERVOING MODEL
The image formation models of the feature points for the eye-in-hand camera configuration and eye-to-hand camera

configuration have the same structure, here, we present a kinematics framework which can be applied to both camera configurations. The image projected coordinates $s_m = [u_m, v_m]^T$ of the feature point can be obtained by using the perspective projection model

$$s_m = \frac{1}{Z^c} \begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} T \begin{bmatrix} y \\ 1 \end{bmatrix}, \tag{1}$$

where Z^c denotes the depth of the feature point, p_i^T denotes the *i*th row vector of the unknown camera intrinsic and extrinsic matrix $P \in \Re^{3 \times 4}$, $T \in \Re^{4 \times 4}$ denotes the homogeneous transformation matrix defined by the forward kinematics of the robot manipulator, and $y \in \Re^{3 \times 1}$ denotes the unknown 3-D position coordinates of the feature point.

Differentiating (1) with respect to time, we can obtain the following relationship which depicts the visual variation with respect to the joint velocity

$$\dot{s}_m = \frac{1}{Z^c} D \dot{q}, \tag{2}$$

where $\dot{q} \in \Re^{n \times 1}$ denotes the joint velocity, *n* denotes the number of DOFs, and $D \in \Re^{2 \times n}$ denotes the depth-independent image Jacobian matrix, which is expressed as

$$D = \begin{pmatrix} p_1^T - u_m p_3^T \\ p_2^T - v_m p_3^T \end{pmatrix} \frac{\partial \left(T \begin{pmatrix} y \\ 1 \end{pmatrix} \right)}{\partial q}. \tag{3}$$

The depth Z^c can be written as

$$Z^c = p_3^T T \begin{pmatrix} y \\ 1 \end{pmatrix}. \tag{4}$$

The dynamic equation of the robot manipulator can be expressed as

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + K(q) = \tau, \tag{5}$$

where $M(q) \in \Re^{n \times n}$ denotes the inertia matrix, $C(q, \dot{q}) \in \Re^{n \times n}$ denotes the centripetal and Coriolis torque matrix, $K(q) \in \Re^{n \times 1}$ denotes the gravitational torques and $\tau \in \Re^{n \times 1}$ denotes the input torque vector. Consider the dynamics given in (5), the dynamics can be rewritten as

$$\ddot{q} = -M(q)^{-1}C(q, \dot{q})\dot{q} - M(q)^{-1}K(q) + M(q)^{-1}\tau, \tag{6}$$

for the system of (6), the feedback linearization control can be obtained as

$$\tau = C(q, \dot{q})\dot{q} + K(q) + M(q)v, \tag{7}$$

and defining *v* as

$$v = w + K_v \dot{q}, \tag{8}$$

where *w* is the new input, and an internal state feedback is implemented via matrix K_v on the velocity loop, then, from (6), (7) and (8), we can obtain the dynamic model

$$\ddot{q} = f(q, \dot{q}, w) = w + K_v \dot{q}, \tag{9}$$

which is considered as the model employed for the open loop prediction in the MPC method.

To use MPC to control the robotic IBVS system, the discrete time model is used. Defining a state vector $\begin{bmatrix} s_m^T & \dot{q}^T \end{bmatrix}^T$, (2) and (9) can be incorporated and the overall system dynamics can be expressed as follows:

$$\begin{pmatrix} s_m(k+1) \\ \dot{q}(k+1) \end{pmatrix} = \begin{pmatrix} s_m(k) \\ \dot{q}(k) \end{pmatrix} + \begin{pmatrix} \frac{1}{Z^c}D(k)\dot{q}(k)T_e \\ (w + K_v\dot{q})T_e \end{pmatrix}, \tag{10}$$

where $s_m(k)$ is both the system's states and outputs, and T_e denotes the sampling period.

III. CONTROLLER DESIGN

The control objective is to tackle the visual servoing problem in the presence of the constraints, parametric uncertainties associated with the robot and camera models, while the joint velocity measurements are unavailable.

A. THE MPC FORMULATION FOR IBVS SYSTEM
Given the overall system dynamics model (10), the visual predictive control can be formulated by finding the optimal control input by solving the constrained finite-time optimization control problem. The quadratic cost function at sampling time *k* is defined as:

$$J(k) = \sum_{i=1}^{N_p} \|s_d(k+i) - s_m(k+i|k)\|_F^2 + \sum_{i=1}^{N_c} \|w(k+i-1)\|_G^2, \tag{11}$$

where N_p is the prediction horizon and N_c is the control horizon, usually $N_p \geq N_c$, *F* and *G* are weighting matrices, $\|\cdot\|$ represents the Euclidean vector norm, $s_d(k+i)$ denotes the desired projection location of the feature point at the future sampling instant $k+i$, $s_m(k+i|k)$ denotes the predicted output at the future sampling instant $k + i$, which are predicted at the current instant *k* by using the prediction model involving the input $w(k)$, $w(k + i - 1)$ denotes the input at the future sampling instant $k + i - 1$.

The MPC approach for IBVS by considering the system dynamics, the visibility constraint and the torque constraint can be formulated as:

$$\min_{\{w(j)\}_{j=k}^{k+N_c-1}} J(k) \tag{12}$$

$$\text{subject to} \quad s_m(k+1) = s_m(k) + \frac{1}{\hat{Z}^c}\hat{D}(k)\dot{q}(k)T_e, \tag{13}$$

$$\dot{q}(k+1) = \dot{q}(k) + (w + K_v\dot{q})T_e, \tag{14}$$

$$u_m^{\min} \leq u_m(k) \leq u_m^{\max}, \tag{15}$$

$$v_m^{\min} \leq v_m(k) \leq v_m^{\max}, \tag{16}$$

$$\tau^{\min} \leq \tau(k) \leq \tau^{\max}, \tag{17}$$

where $u_m^{\max}$ and $v_m^{\max}$ are the maximum coordinates of the image plane, $u_m^{\min}$ and $v_m^{\min}$ are the minimum coordinates of the image plane, and (15) and (16) represent the visibility constraints, which can be used to ensure the feature point

within the camera's field of view. $\tau^{\max}$ and $\tau^{\min}$ are the lower and upper bounds of the force torque, and (17) represents the torque constraint. The constraints can be written as the functions of the optimal sequence of the input. The MPC method is implemented based on a receding horizon strategy. At every sampling instant, the optimal input sequence $\{w(j)\}_{j=k}^{k+N_c-1}$ is calculated by solving the open-loop finite-horizon optimal control problem. The first element of the sequence is treated as the input $w(k)$. This process is repeated at the next sampling instant. The sequential quadratic programing (SQP) algorithm is used to solve the above optimal control problem.

B. UNCERTAINTY AND LEARNING FOR IBVS

In most of MPC-based IBVS schemes, the prediction models are on the basis of the traditional image Jacobian matrix [28]–[31]. The traditional Jacobian matrix maps the visual signals onto the camera Cartesian space and the depth parameter appears nonlinearly in the traditional image Jacobian matrix. Generally, the camera values and the depth parameter which cannot be measured from the image should be known. Here, the prediction model of the proposed MPC method for IBVS is formulated on the basis of the depth-independent interaction matrix which maps the image errors onto the joint space of the manipulator directly, and the depth can be linearly parameterized by the constant camera parameters and 3-D position values of the feature, then, the unknown parameters can be represented as a linear form by using the perspective projection model, so that the unknown parameters can be learned based on the measurements of image features and robot joints by using the system identification algorithm. Moreover, under the depth-independent interaction matrix framework, the unknown parameters in the prediction model and the perspective projection model are the same, then the parameter estimates $\hat{\theta}$ updated from the identification algorithm form the basis for the prediction model of the feature at each time instant.

The perspective projection model (1) is rewritten as

$$Z^c s_m = \begin{bmatrix} p_1^T \\ p_2^T \end{bmatrix} T \begin{bmatrix} y \\ 1 \end{bmatrix}, \tag{18}$$

and the depth Z^c of the feature point (4) can be represented as a linear form of products of unknown camera parameters and feature position parameters

$$Z^c = \gamma^T \theta_3, \tag{19}$$

where $\bar{\gamma} = (r_{11}, r_{21}, r_{31}, r_{12}, r_{22}, r_{32}, r_{13}, r_{23}, r_{33}, t_1, t_2, t_3)^T$, $\gamma = (\bar{\gamma}^T, 1)^T$. R and t are the rotation matrix and the translational vector of the forward kinematics matrix T, and r_{ij} is the (i, j) component of R and t_i is the ith component of t. p_{ij} is the (i, j) component of the camera intrinsic and extrinsic matrix P. Denote $\bar{\theta}_i = (p_{i1}y_1, p_{i2}y_1, p_{i3}y_1, p_{i1}y_2, p_{i2}y_2, p_{i3}y_2, p_{i1}y_3, p_{i2}y_3, p_{i3}y_3, p_{i1}, p_{i2}, p_{i3})^T$ and $\theta_i = (\bar{\theta}_i^T, p_{i4})^T$. Substituting (19)

into (18), we can obtain the parameter estimation model

$$s_m(k) = \psi^T(k)\theta, \tag{20}$$

where $\psi(k)$ denotes the measurement matrix on the basis of the measurements of image coordinates and joint angles, which is expressed as

$$\psi(k) = \begin{bmatrix} \gamma^T & 0_{1\times 13} & -u_m(k)\bar{\gamma}^T \\ 0_{1\times 13} & \gamma^T & -v_m(k)\bar{\gamma}^T \end{bmatrix}^T, \tag{21}$$

and $\theta = \bar{\theta}/p_{34}, \bar{\theta} = (\theta_1^T, \theta_2^T, \bar{\theta}_3^T)^T$. The row vectors in the depth-independent image Jacobian matrix (3) can be expressed as

$$p_i^T \frac{\partial}{\partial q} \left(T \begin{pmatrix} y \\ 1 \end{pmatrix} \right)$$

$$= \frac{\partial}{\partial q} (r_{11}p_{i1}y_1 + r_{21}p_{i2}y_1 + r_{31}p_{i3}y_1 + r_{12}p_{i1}y_2$$

$$+ r_{22}p_{i2}y_2 + r_{32}p_{i3}y_2 + r_{13}p_{i1}y_3 + r_{23}p_{i2}y_3$$

$$+ r_{33}p_{i3}y_3 + t_1p_{i1} + t_2p_{i2} + t_3p_{i3})$$

$$= \frac{\partial}{\partial q} (\bar{\gamma}\bar{\theta}_i), \tag{22}$$

then the updated parameter estimates $\hat{\theta}$ can be used to update the estimated image Jacobian matrix $\frac{1}{\hat{Z}^c}\hat{D}$ together with the prediction model. From (20), the recursive least square method is employed to estimate the unknown parameters θ, the criterion function to be minimized at time N is defined as

$$\hat{\theta} = \arg\min_{\theta} \sum_{k=1}^{N} \varepsilon^{N-k}(s_m(k) - \psi^T(k)\theta)^2, \tag{23}$$

where $0 < \varepsilon \leq 1$ represents the exponential forgetting factor. The parameter estimates are updated by

$$\hat{\theta}(k) = \hat{\theta}(k-1) + Y(k)(s_m(k) - \psi^T(k)\hat{\theta}(k-1)) \tag{24}$$

$$Y(k) = H(k-1)\psi(k)(\varepsilon + \psi^T(k)H(k-1)\psi(k))^{-1} \tag{25}$$

$$H(k) = (I - Y(k)\psi^T(k))H(k-1). \tag{26}$$

The image errors are defined as

$$e(k) = s_m(k) - s_d = \psi^T(k)\theta - s_d, \tag{27}$$

since the forward kinematics components $r_{ij}(i = 1, 2, 3, j = 1, 2, 3)$ and $t_i(i = 1, 2, 3)$ are bounded, we have

$$\|\psi_i(k)\| \leq m_1 + m_2 s_{mi}(k), \tag{28}$$

where $0 < m_1 < \infty$, $0 < m_2 < \infty$, and $\psi_i(k)$ and $s_{mi}(k)$ are the ith row of $\psi(k)$ and $s_m(k)$, respectively.

We assume that the desired trajectory s_d satisfies $|s_d| \leq m_3 < \infty$, then we can obtain the linear boundedness condition

$$\|\psi_i(k)\| \leq m_1 + m_2(m_3 + |e_i(k)|) = C_1 + C_2 |e_i(k)|, \tag{29}$$
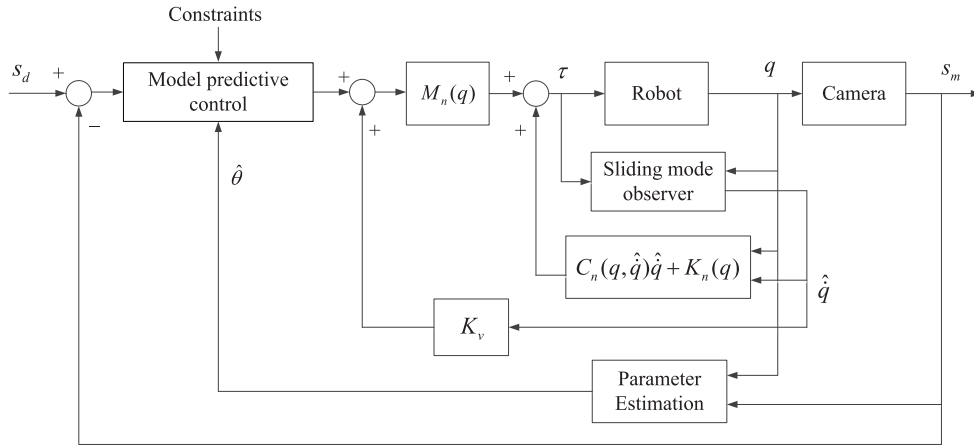
where $0 < C_1 < \infty, 0 < C_2 < \infty$.

FIGURE 1. The control system diagram.

When θ is unknown, we have $\tilde{\theta}(k) = \hat{\theta}(k) - \theta$ and the function $V(k) = \tilde{\theta}^T(k)H^{-1}(k)\tilde{\theta}(k)$, from the matrix inversion lemma, we have

$$H^{-1}(k) = H^{-1}(k-1) + \frac{1}{\varepsilon}\psi(k)\psi^T(k), \qquad (30)$$

using (26) and the matrix inversion lemma, we obtain

$$H^{-1}(k)\tilde{\theta}(k) = H^{-1}(k-1)\tilde{\theta}(k-1), \qquad (31)$$

using (24) and (31), we have

$$V(k) - V(k-1) = \left[\tilde{\theta}(k)-\tilde{\theta}(k-1)\right]^T H^{-1}(k-1)\tilde{\theta}(k-1)$$

$$= -\frac{\left[\tilde{\theta}^T(k-1)\psi(k)\right]^2}{\varepsilon + \psi^T(k)H(k-1)\psi(k)} \leq 0, \quad (32)$$

thus $V(k)$ is a bounded, nonnegative and nonincreasing function and hence it can converge, if the condition $\psi^T(k)\hat{\theta} = s_d$ is satisfied, then it can be obtained from [33] that

$$\lim_{k\to\infty} \frac{|e_i(k)|^2}{1 + 2(\alpha_{max}[H(k-1)])\|\psi_i(k)\|^2} = 0. \qquad (33)$$

From the matrix inversion lemma, we can obtain

$$\alpha_{min}[H^{-1}(k)] \geq \alpha_{min}[H^{-1}(k-1)] \geq \alpha_{min}[H^{-1}(0)], \quad (34)$$

where $\alpha_{min}[H^{-1}(k)]$ denotes the minimum eigenvalue of the matrix $H^{-1}(k)$.

Since $\alpha_{min}[H^{-1}(k)]$ is a non-decreasing function and it has a lower bound $\alpha_{min}[H^{-1}(0)] > 0$. Then, from (33), the uniform boundedness condition is satisfied

$$0 < a_1 = 1 < \infty, \quad 0 < a_2 = 2\alpha_{max}[H(k-1)] < \infty. \quad (35)$$

Based on the linear boundedness condition (29) and the uniform boundedness condition (35), from (33) we can obtain

$$\lim_{k\to\infty} e_i(k) = 0. \qquad (36)$$

C. THE SLIDING MODE OBSERVER DESIGN

The velocity measurements are usually contaminated by noises, which will result in the degradation performances of the control system, or even lead to servoing failure. Under the condition that the robot and camera parameters are unknown, in order to overcome the inabilities without the velocity measurements, the sliding mode observer is designed to estimate the joint velocities for the controller design. The overall control system diagram is illustrated in Figure 1.

The functions M, C and K are assumed uncertain, while the corresponding nominal functions M_n, C_n and K_n are supposed to be known. The control input τ is supposed to be given by the known feedback functions. Introducing new variables $x_1 = q$, $x_2 = \dot{q}$, $u = \tau$, the model (5) can be represented by a second order model

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = g(t, x_1, x_2, u) + \eta(t, x_1, x_2, u)$$
$$y = x_1, \qquad (37)$$

where the nominal part of the system dynamics is expressed as $g(t, x_1, x_2, u) = M_n^{-1}(x_1)(u - C_n(x_1, x_2)x_2 - K_n(x_1))$, M_n, C_n and K_n are the known nominal functions and the uncertainties are lumped in the term $\eta(t, x_1, x_2, u)$. To apply the internal state feedback and to perform system prediction (13)−(14), the knowledge of joint velocities is required, while the joint velocity measurements $\dot{q}$ are unavailable, hence, the following sliding mode observer is proposed to estimate the joint velocities

$$\dot{\hat{x}}_1 = \hat{x}_2 + \kappa|x_1 - \hat{x}_1|^{1/2}\frac{x_1 - \hat{x}_1}{|x_1 - \hat{x}_1| + \xi}$$

$$\dot{\hat{x}}_2 = g(t, x_1, \hat{x}_2, u) + \sigma\frac{x_1 - \hat{x}_1}{|x_1 - \hat{x}_1| + \xi}, \qquad (38)$$

where $\hat{x}_1$ and $\hat{x}_2$ are the estimation values, and κ, σ and ξ are the observer parameters. Defining the estimation errors
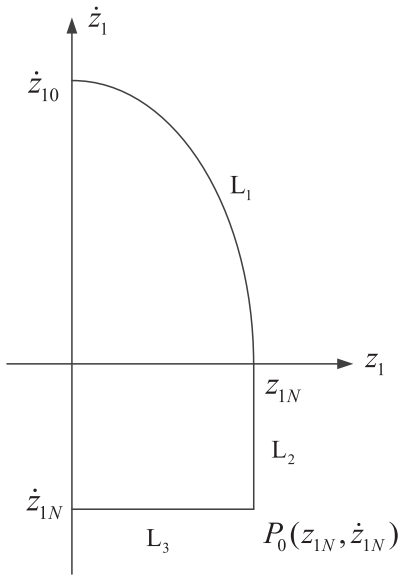
FIGURE 2. Majorant curve for the sliding mode observer.
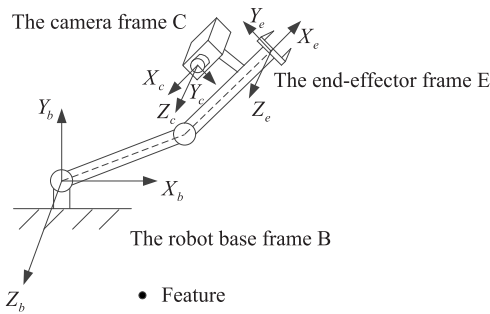

FIGURE 3. The structure diagram of the 2-DOF robot manipulator with eye-in-hand camera configuration.
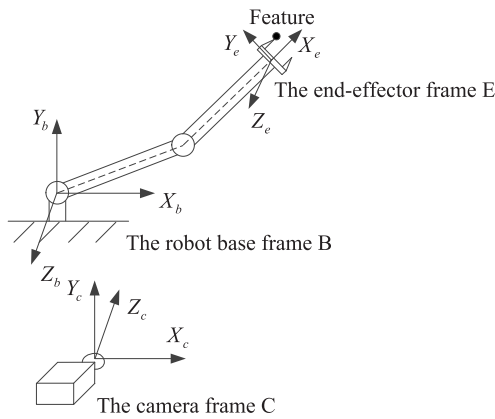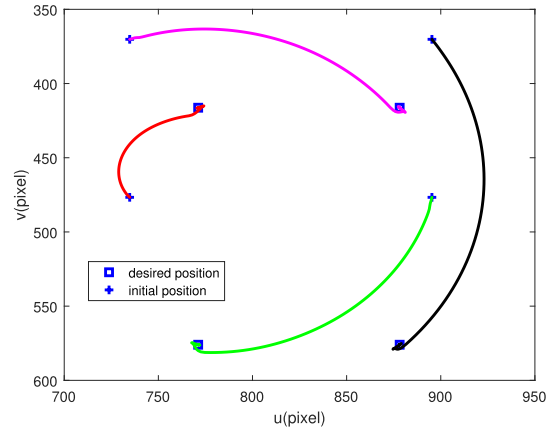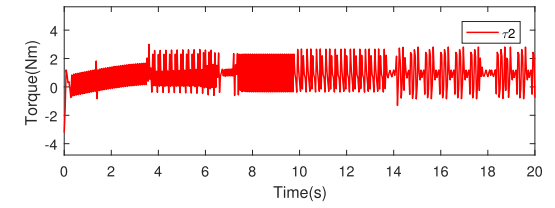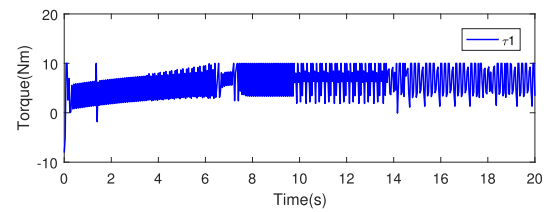

FIGURE 4. The structure diagram of the 2-DOF robot manipulator with eye-to-hand camera configuration.

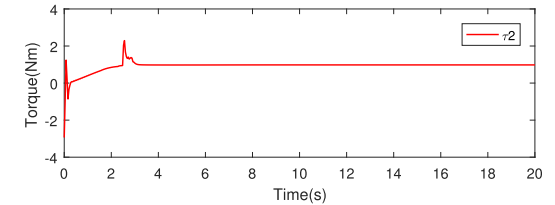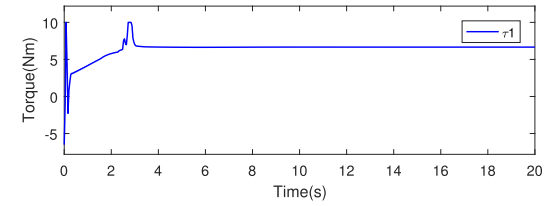$z_1 = x_1 - \hat{x}_1$, $z_2 = x_2 - \hat{x}_2$, and the error functions can be expressed as

$$\dot{z}_1 = z_2 - \kappa|z_1|^{1/2}\frac{z_1}{|z_1| + \xi}$$

$$\dot{z}_2 = G(t, x_1, x_2, \hat{x}_2) - \sigma\frac{z_1}{|z_1| + \xi}, \qquad (39)$$

FIGURE 5. Simulation results for 2-DOF robot manipulator with eye-in-hand camera configuration. (a) Image trajectory by the proposed method. (b) The chattering phenomenon caused by the traditional sliding mode observer. (c) Torque by the proposed method.

where $G(t, x_1, x_2, \hat{x}_2) = g(t, x_1, x_2, u(t, x_1, x_2)) - g(t, x_1, \hat{x}_2, u(t, x_1, x_2)) + \eta(t, x_1, x_2, u(t, x_1, x_2))$. Suppose that there exists a positive constant, such that

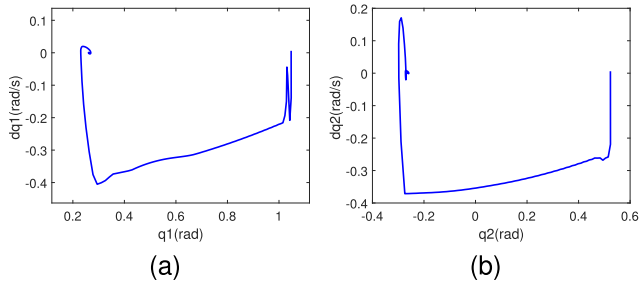$$|G(t, x_1, x_2, \hat{x}_2)| < g^+, \qquad (40)$$

FIGURE 6. The phase portrait by the proposed method for 2-DOF robot manipulator with eye-in-hand camera configuration. (a) q_1, $\dot{q}_1$. (b) q_2, $\dot{q}_2$.

for any t, x_1, x_2. Calculating the derivative of $\dot{z}_1$, we can obtain

$$\ddot{z}_1 = \dot{z}_2 - \kappa |z_1|^{1/2} \dot{z}_1 \left(\frac{1}{2} \frac{1}{|z_1|} \frac{z_1}{|z_1| + \xi} \, signz_1 + \frac{1}{z_1} \left(\frac{z_1}{|z_1| + \xi} - \left(\frac{z_1}{|z_1| + \xi} \right)^2 signz_1 \right) \right), \quad (41)$$

where $d |z_1| / dt = \dot{z}_1 signz_1$. We have $\dot{z}_2 \in [-g^+, +g^+] - \sigma \frac{z_1}{|z_1| + \xi}$, and $\frac{z_1}{|z_1| + \xi} \in [-1, 1]$, denote $\rho = \frac{z_1}{|z_1| + \xi}$, then we can obtain

$$\ddot{z}_1 \in [-g^+, +g^+] - \left(\sigma \rho + \kappa |z_1|^{1/2} \dot{z}_1 \left(\frac{1}{2} \frac{1}{|z_1|} \rho signz_1 + \frac{1}{z_1} \left(\rho - \rho^2 signz_1 \right) \right) \right). \quad (42)$$

At the initial instant, $z_1 = 0$ and $z_2 = x_2 - 0 = x_2$, let $z_1 > 0$, $\dot{z}_1 > 0$, then the trajectory is confined between the axis $z_1 = 0$, $\dot{z}_1 = 0$ and $\ddot{z}_1 = -(\sigma \rho - g^+)$. The majorant curve for the sliding mode observer is shown in Figure 2. As shown in Figure 2 line $(L1)$, let z_{1N} be the intersection of the majorant curve and the axis $\dot{z}_1 = 0$, we have $\dot{z}_{10}^2 = 2(\sigma \rho - g^+) z_{1N}$, where $\dot{z}_{10} > 0$ is the intersection of the curve and the axis $z_1 = 0$, let σ satisfies the inequality $\sigma \rho > g^+$, for $z_1 > 0$, $\dot{z}_1 > 0$, we have

$$\ddot{z}_1 \leq g^+ - \sigma \rho - \kappa |z_1|^{1/2} \dot{z}_1 \left(\frac{1}{2} \frac{1}{|z_1|} \rho + \frac{1}{z_1} \left(\rho - \rho^2 \right) \right) < 0. \quad (43)$$

Hence, the trajectory approaches the axis $\dot{z}_1 = 0$. For $z_1 > 0$, $\dot{z}_1 \leq 0$, the majorant curve is composed of two parts, for the first part, when the trajectory enters the half plane $\dot{z}_1 \leq 0$, $\dot{z}_1$ keeps decreasing to point P_0 until $\ddot{z}_1 = 0$, the majorant curve is a vertical line (see Line $(L2)$ in Figure 1), when the right hand side of (42) is equal to zero in the worst case, we have

$$-g^+ - \left(\sigma \rho + \kappa |z_1|^{1/2} \dot{z}_1 \left(\frac{1}{2} \frac{1}{|z_1|} \rho + \frac{1}{z_1} \left(\rho - \rho^2 \right) \right) \right) = 0, \quad (44)$$

the coordinate of the point P_0 is $(z_{1N}, \dot{z}_{1N})$, then from (44), we can obtain

$$\dot{z}_{1N} = \left(-2 \big/ (\kappa (3\rho - 2\rho^2))) (\sigma \rho + g^+) z_{1N}^{1/2}. \quad (45)$$

For the second part, the majorant curve is the horizontal line between the points $(z_{1N}, (-2 / (\kappa (3\rho - 2\rho^2))) (\sigma \rho + g^+) z_{1N}^{1/2})$ and $(0, \dot{z}_{1N})$, as line $L3$ depicted in Figure 2. Then, we can obtain

$$\frac{|\dot{z}_{10}|}{|\dot{z}_{1N}|} = \frac{\kappa (3\rho - 2\rho^2) \sqrt{2(\sigma \rho - g^+)}}{2(\sigma \rho + g^+)}, \quad (46)$$

we have $3\rho - 2\rho^2 < 2$, then if the following inequality is satisfied

$$\kappa > \frac{(\sigma \rho + g^+)}{\sqrt{2(\sigma \rho - g^+)}}, \quad (47)$$

we can obtain

$$\frac{|\dot{z}_{10}|}{|\dot{z}_{1N}|} > 1, \quad (48)$$

the series $\{\dot{z}_{1i}\} = \dot{z}_{10}, \dot{z}_{11}, \cdots, \dot{z}_{1i}, \cdots$ are the consequent crossing points of the system (37) starting at $(0, \dot{z}_{10})$ with axis $z_1 = 0$, (48) means that $|\dot{z}_{1i}|$ becomes smaller and we can obtain the convergence of the variable $(0, \dot{z}_{1i})$ to $z_1 = \dot{z}_1 = 0$.

IV. SIMULATION RESULTS
A. 2-DOF ROBOT MANIPULATOR
In order to illustrate the effectiveness of the proposed method via the joint velocity estimation in the presence of constraints and unknown parameters, the simulations on a 2-DOF robot manipulator with both eye-in-hand and eye-to-hand

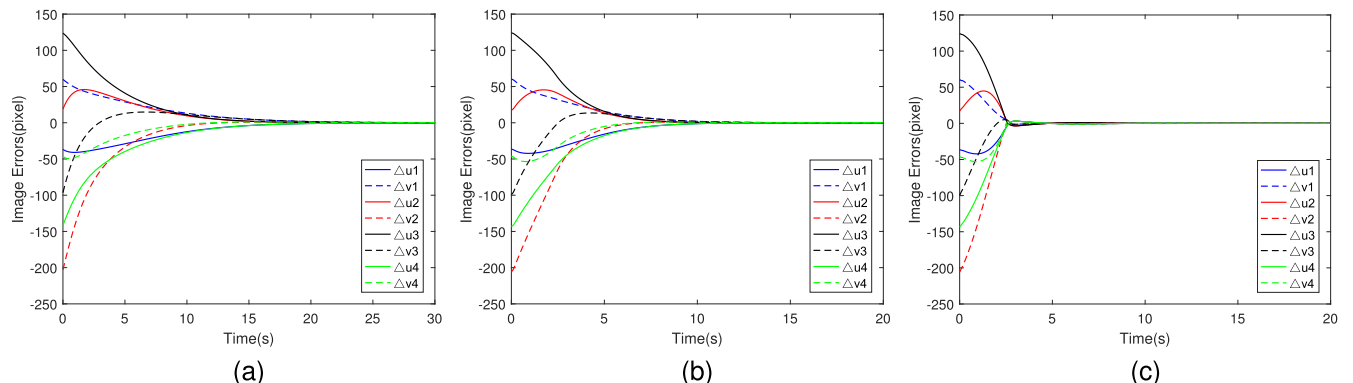

FIGURE 7. Simulation results for 2-DOF robot manipulator with eye-in-hand camera configuration. (a) Image errors of the classical IBVS scheme. (b) Image errors of the visual predictive control scheme. (c) Image errors of the proposed method.
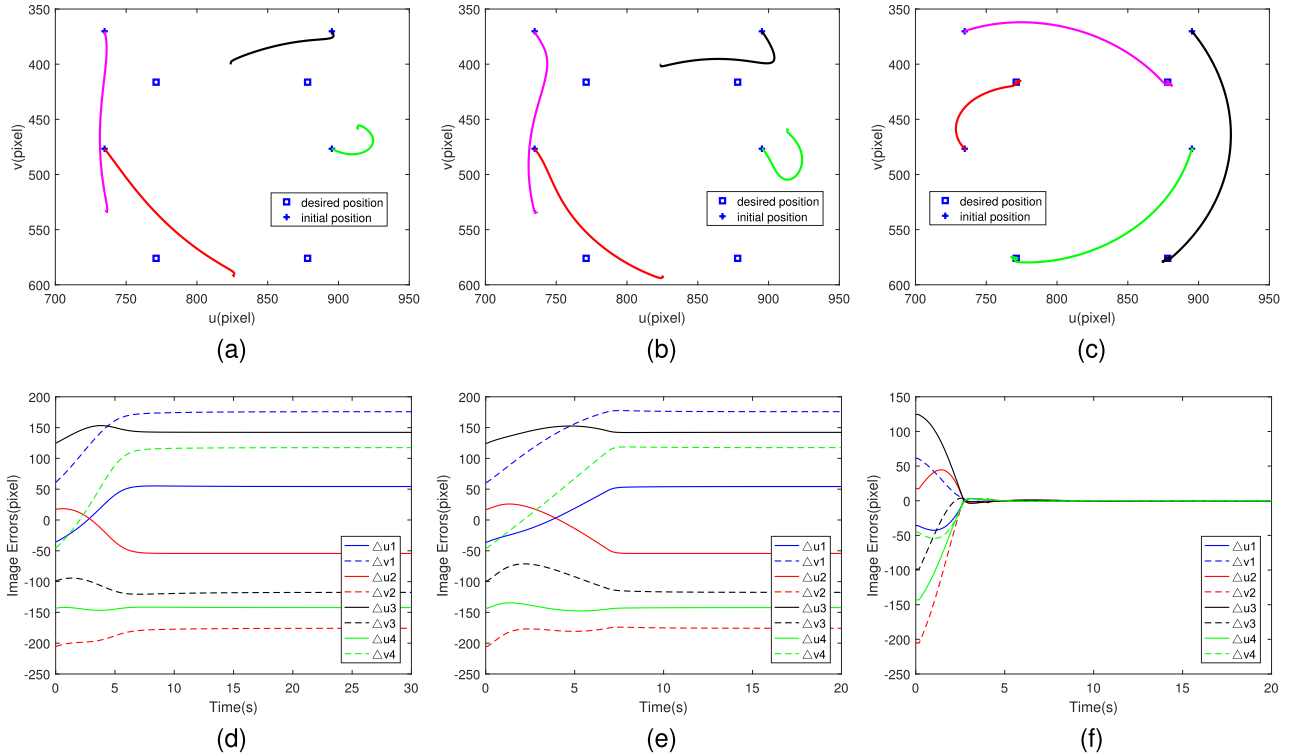
FIGURE 8. Simulation results for 2-DOF robot manipulator with eye-in-hand camera configuration. (a) Image trajectory of the classical IBVS scheme. (b) Image trajectory of the visual predictive control scheme. (c) Image trajectory of the proposed method. (d) Image errors of the classical IBVS scheme. (e) Image errors of the visual predictive control scheme. (f) Image errors of the proposed method.

camera configurations are presented. The structure diagrams of the 2-DOF robot manipulator with eye-in-hand camera configuration and eye-to-hand camera configuration are shown in Figure 3 and Figure 4, respectively. The real camera intrinsic parameters are as follows: the focal length $f = 0.005$ m, the coordinates of the principal point $(u_0, v_0) = (646, 482)$ pixels, the scale factors along the u axis $\alpha_u = 269167$ pixels/m and the scale factors along the v axis $\alpha_v = 267778$ pixels/m. The image has a 1292×964 pixels resolution. The lengths of the first and second links are $l_1 = 0.18$ m and $l_2 = 0.15$ m, the mass centers of first and second links are $l_{c1} = 0.091$ m, $l_{c2} = 0.105$ m, the masses of the first and second links are $m_1 = 23.9$ kg and $m_2 = 4.44$ kg, and the inertia of the first and second links are $I_1 = 1.27$ kgm^2, $I_2 = 0.24$ kgm^2. The sampling period of the controller is 40 ms, which corresponds to an usual camera capturing rate with 25 frames/second. The visibility constraint is performed as

$$\begin{bmatrix} u_{min} = 0 \\ v_{min} = 0 \end{bmatrix} \le s_m \le \begin{bmatrix} u_{max} = 1292 \\ v_{max} = 964 \end{bmatrix}. \quad (49)$$

Moreover, actuator saturation limits of -10 Nm $\le \tau \le 10$ Nm are considered. The simulations are conducted under MATLAB on a desktop with a 3.2 GHz Intel Core i7. The SQP algorithm (fmincon function from optimization Matlab toolbox) is used for solving the optimization problems. The rough camera intrinsic parameters are $\hat{u}_0 = 500$ pixels, $\hat{v}_0 = 500$ pixels, $\hat{\alpha}_u = 250000$ pixels/m, $\hat{\alpha}_v = 250000$ pixels/m.

The rough robot parameters are $\hat{l}_1 = 0.05$ m, $\hat{l}_2 = 0.05$ m, $\hat{l}_{c1} = 0.05$ m, $\hat{l}_{c2} = 0.05$ m, $\hat{m}_1 = 10$ kg, $\hat{m}_2 = 2$ kg, $\hat{I}_1 = 1$ kgm^2, and $\hat{I}_2 = 0.2$ kgm^2.

The simulation results of the IBVS system under the proposed control method is compared with those under the classical IBVS scheme in [4] and the visual predictive control scheme in [29].

In the eye-in-hand camera configuration, the simulations are carried out to take four feature points from initial positions to their desired positions on the image plane. The 3-D position coordinates of the feature points with respect to the robot base frame are $(0.1, 0.15, 2.5)^T$ m, $(-0.1, -0.15, 2.5)^T$ m, $(0.1, -0.15, 2.5)^T$ m and $(-0.1, 0.15, 2.5)^T$ m, respectively. The transformation matrix from the robot end-effector frame to the camera frame is

$$T_e^c = \begin{bmatrix} -1 & 0 & 0 & 0.01 \\ 0 & -1 & 0 & -0.02 \\ 0 & 0 & 1 & 0.015 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (50)$$

The initial roughly estimated camera extrinsic parameter matrix $\hat{T}_e^c(0)$ is

$$\hat{T}_e^c(0) = \begin{bmatrix} -1 & 0 & -0.3 & 0.001 \\ 0 & -1 & 0 & 0.03 \\ 0 & 0 & 0.95 & 0.001 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (51)$$

The initial roughly estimated 3-D position parameters of the feature points in the robot base frame are $(0.09,$

$0.01, 0.05)^T$ m, $(0.09, 0.01, 0.05)^T$ m, $(0.09, 0.01, 0.05)^T$ m and $(0, 0, 0.05)^T$ m, respectively. The initial locations of the features points on the image plane are $(734.8, 476.7)^T$ pixels, $(895.3, 370.2)^T$ pixels, $(895.3, 476.7)^T$ pixels and $(734.8, 370.2)^T$ pixels, respectively. The desired locations of the features points on the image plane are $(771.1, 416.3)^T$ pixels, $(878.2, 576)^T$ pixels, $(771.1, 576)^T$ pixels and $(878.2, 416.3)^T$ pixels, respectively. The weighting matrices are $F = 10I_{8 \times 8}$, $G = I_{2 \times 2}$. To balance the computational time and the control efficiency, the prediction horizon and the control horizon are chosen as $N_p = 3$ and $N_c = 2$. The initial covariance matrix is chosen as $H(0) = 10^{10}I$, the forgetting factor is set as $\varepsilon = 0.995$, $K_v = -5I_{2 \times 2}$, and the sliding mode observer parameters are $\kappa = 100$, $\sigma = 80$, $\xi = 2$. The true values of the camera intrinsic and extrinsic parameters and the depth information are not necessary. Only the initial rough camera intrinsic and extrinsic parameters and feature position parameters are used for the controller design. Figure 5(a) depicts the 2-D trajectory of the feature points on the image plane under the control of the proposed method, which illustrates that the proposed controller takes the feature points to their desired values successfully. Figure 5(b) shows the chattering phenomenon caused by the traditional sliding mode observer, the severe chattering in the robot system is very harmful. Then, the proposed sliding mode observer is used here to eliminate the chattering effects, and the corresponding result is given in Figure 5(c), from which we can observe that the chattering effect is completely eliminated, and the torque constraint of joint 1 is quickly reached to guarantee a fast motion. With the initial conditions $(q_1(0), \dot{q}_1(0))^T = (\pi/3, 0.006)^T$, $(q_2(0), \dot{q}_2(0))^T = (\pi/6, 0.006)^T$, the phase-portrait of the system is depicted in Figure 6 to indicate system dynamic behavior. The computational time in solving the optimization problem under the proposed approach is about 18 ms per sampling period. We can observe that the proposed control approach can reject the model uncertainty well in the case when the joint velocity measurements are unavailable, while the visibility constraint and the torque constraint can be considered.

The image errors of the classical IBVS scheme, the visual predictive control scheme and the proposed control approach are illustrated in Figure 7. Compared with the classical IBVS scheme and the visual predictive control scheme, the settling time for response of the IBVS system under the proposed control method (about 4 s) is shorter than that under the visual predictive control scheme (about 11 s) and the classical IBVS scheme (about 20 s). Furthermore, if the initial estimates of the camera values and position coordinates of features are too rough, the classical IBVS scheme and the visual predictive control scheme will fail in the control tasks. The initial estimated camera extrinsic values are set as

$$\hat{T}_e^c(0) = \begin{bmatrix} 1 & 0 & -0.3 & 0.001 \\ 0 & -1 & 0 & 0.03 \\ 0 & 0 & 0.95 & 0.001 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (52)$$
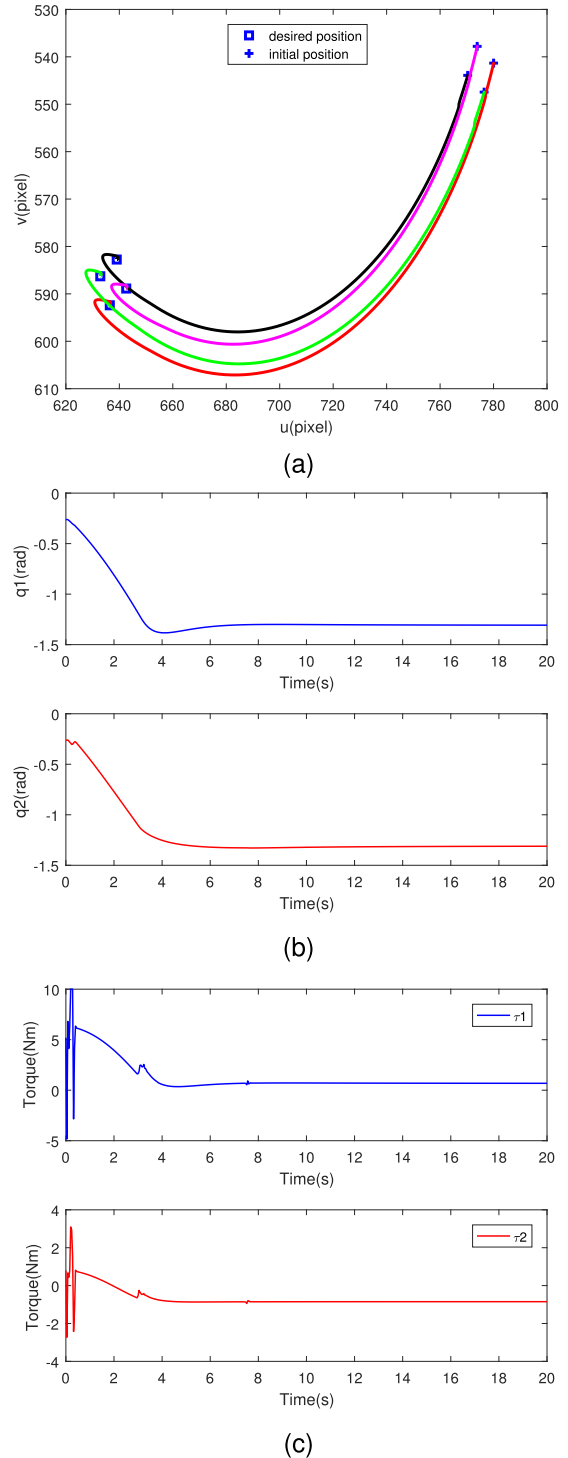

FIGURE 9. Simulation results by the proposed method for 2-DOF robot manipulator with eye-to-hand camera configuration. (a) Image trajectory. (b) Robot joint angles. (c) Torque.

then the comparison results of the three methods are illustrated in Figure 8, it can be seen from the Figure 8 that only the proposed method successfully fulfills the visual servoing tasks, the classical IBVS scheme and the visual predictive control scheme cannot fulfill
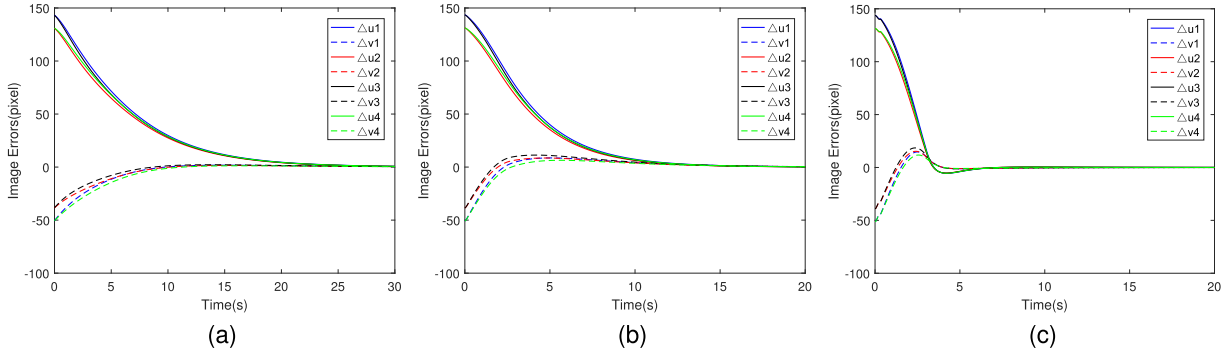
FIGURE 10. Simulation results for 2-DOF robot manipulator with eye-to-hand camera configuration. (a) Image errors of the classical IBVS scheme. (b) Image errors of the visual predictive control scheme. (c) Image errors of the proposed method.
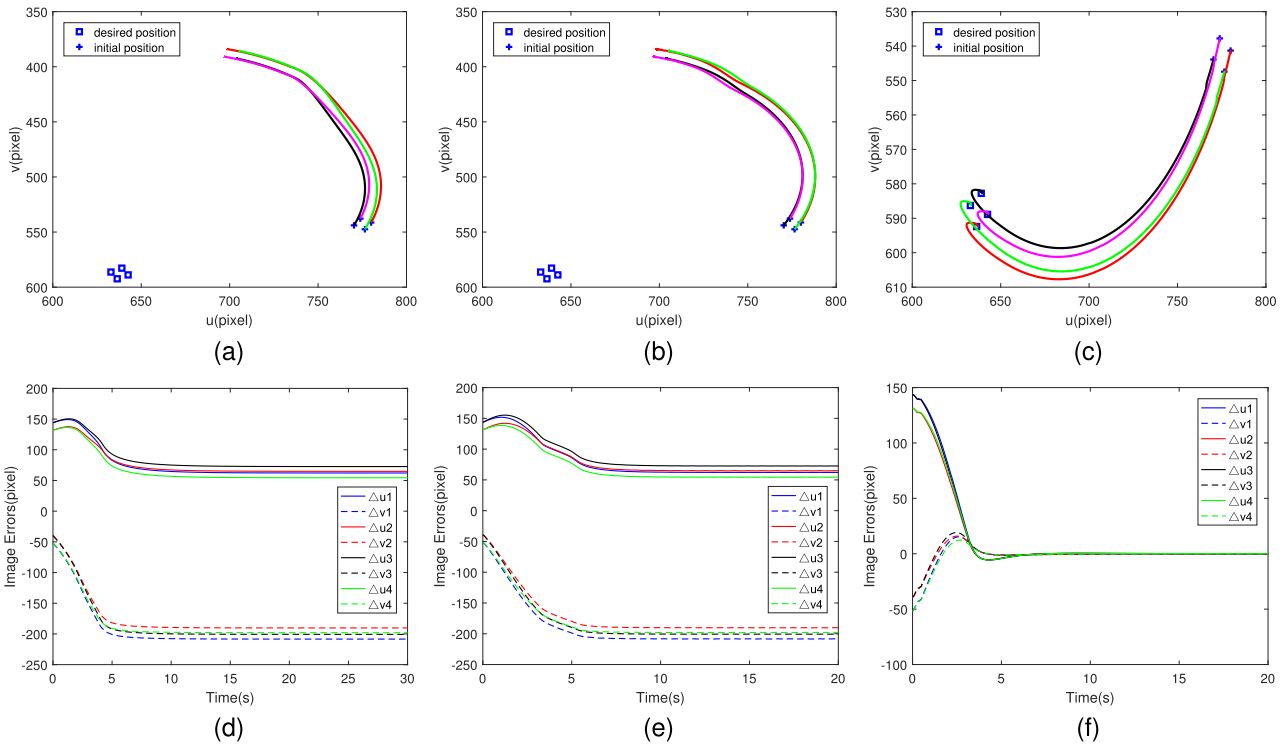

FIGURE 11. Simulation results for 2-DOF robot manipulator with eye-to-hand camera configuration. (a) Image trajectory of the classical IBVS scheme. (b) Image trajectory of the visual predictive control scheme. (c) Image trajectory of the proposed method. (d) Image errors of the classical IBVS scheme. (e) Image errors of the visual predictive control scheme. (f) Image errors of the proposed method.

the visual servoing tasks because of the large model uncertainty.

In the eye-to-hand camera configuration, four feature points are marked on the robot end-effector. The control objective is to regulate the feature points from initial locations to desired locations on the image plane. The transformation matrix from the robot base frame to the camera frame is

$$T_b^c = \begin{bmatrix} 1 & 0 & 0 & 0.06 \\ 0 & -1 & 0 & 0.05 \\ 0 & 0 & -1 & 3.8 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (53)$$

The initial roughly estimated camera extrinsic matrix $\hat{T}_b^c(0)$ is

$$\hat{T}_b^c(0) = \begin{bmatrix} 1 & -0.2 & -0.3 & 0.01 \\ 0 & -0.98 & 0 & 0.025 \\ 0 & 0 & -0.95 & 0.02 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (54)$$

The initial roughly estimated position values of the feature points in the end-effector frame are $(0.01, -0.01, 0.01)^T$ m, $(0.3, -0.1, 0.042)^T$ m, $(-0.1, 0.003, 0)^T$ m, $(-0.049, 0, 0.1)^T$ m, respectively. The initial positions of the features points on the image plane are $(780, 541.3)^T$ pixels, $(770.3, 544)^T$ pixels, $(776.5, 547.4)^T$ pixels and $(773.9, 537.8)^T$ pixels,
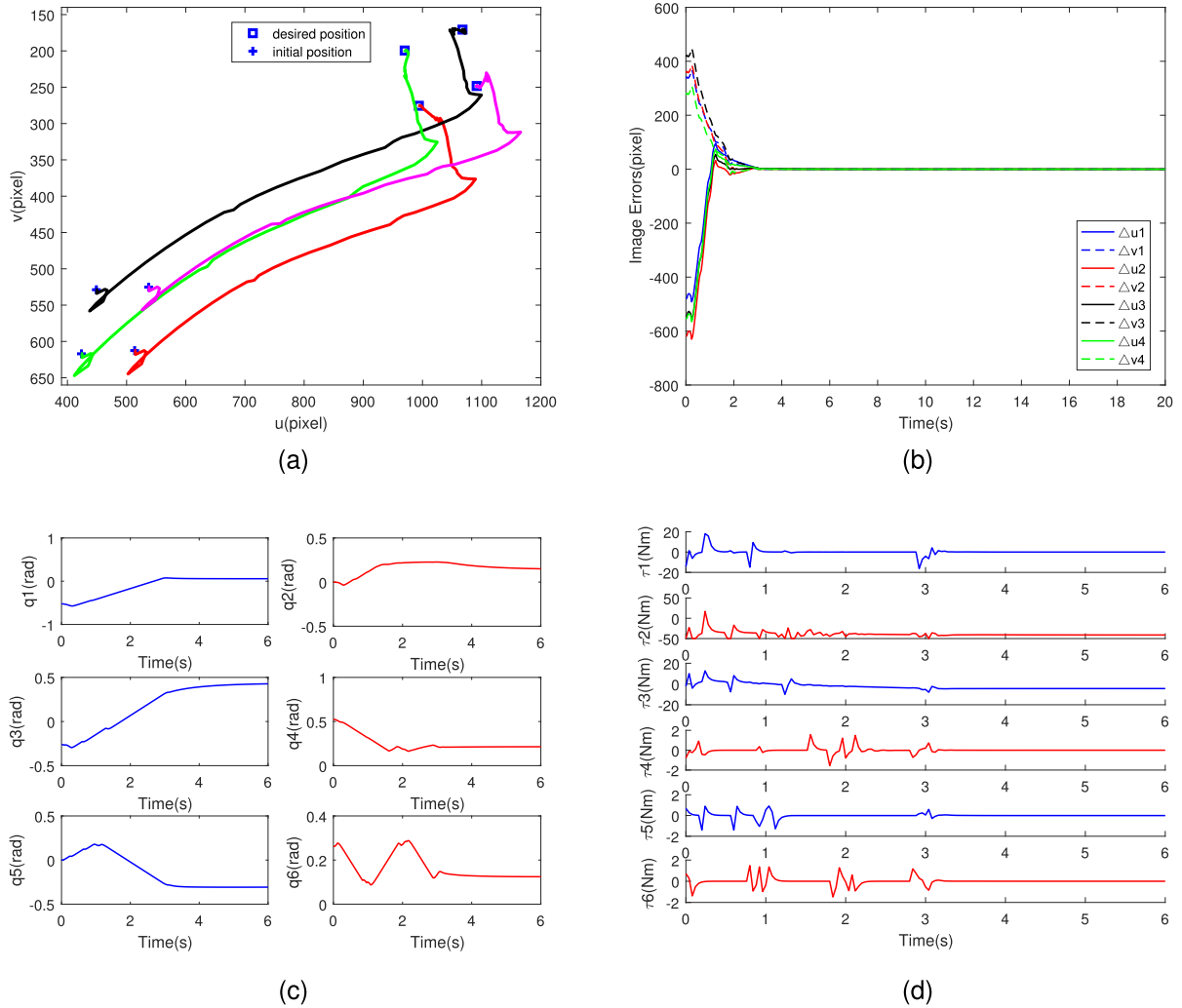
FIGURE 12. Simulation results by the proposed method for 6-DOF robot manipulator with eye-in-hand camera configuration. (a) Image trajectory. (b) Image errors. (c) Robot joint angles. (d) Torque.

respectively. The desired positions of the features points on the image plane are $(636.4, 592.4)^T$ pixels, $(639, 582.8)^T$ pixels, $(632.9, 586.3)^T$ pixels and $(642.6, 588.9)^T$ pixels, respectively. The prediction horizon, the control horizon, the initial covariance matrix, and the forgetting factor are set the same values to the previous simulation, the weighting matrices are $F = 10I_{8\times8}$, $G = 10I_{2\times2}$, $K_v = -20I_{2\times2}$, and the sliding mode observer parameters are $\kappa = 60$, $\sigma = 50, \xi = 2$. The simulation results of the proposed control approach are presented in Figure 9. The image trajectory of the feature points are illustrated in Figure 9(a), the result shows that the proposed controller can make the feature points reach to the desired positions while keeping the feature points within the camera's field of view. The changes of the joint angles in the visual servoing task are illustrated in Figure 9(b). Figure 9(c) shows that the torques of joint 1 rapidly reach to the saturation constraint, ensuring a fast response. The computational time in solving the optimization problem under the proposed approach is about 22 ms per

sampling period. We can observe that the model uncertainty can be rejected under the situation that the joint velocity measurements are unavailable in the proposed control method, while the visibility constraint and the torque constraint can be taken into account during the visual servoing task.

The image errors of the three approaches are demonstrated in Figure 10. It can be seen from Figure 10 that the settling time for response of the IBVS system under the proposed control method is about 6 s, which is much shorter than that under the visual predictive control scheme (about 12 s) and the classical IBVS scheme (about 22 s). Moreover, if the model uncertainty is too large, the classical IBVS scheme and the visual predictive control scheme will fail in the visual servoing tasks. The initial estimated transformation matrix from the robot base frame to the camera frame is set as

$$\hat{T}_b^c(0) = \begin{bmatrix} -1 & -0.8 & -0.3 & 0.01 \\ 0 & 0.6 & 0 & 0.025 \\ 0 & 0 & -0.95 & 0.02 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (55)$$
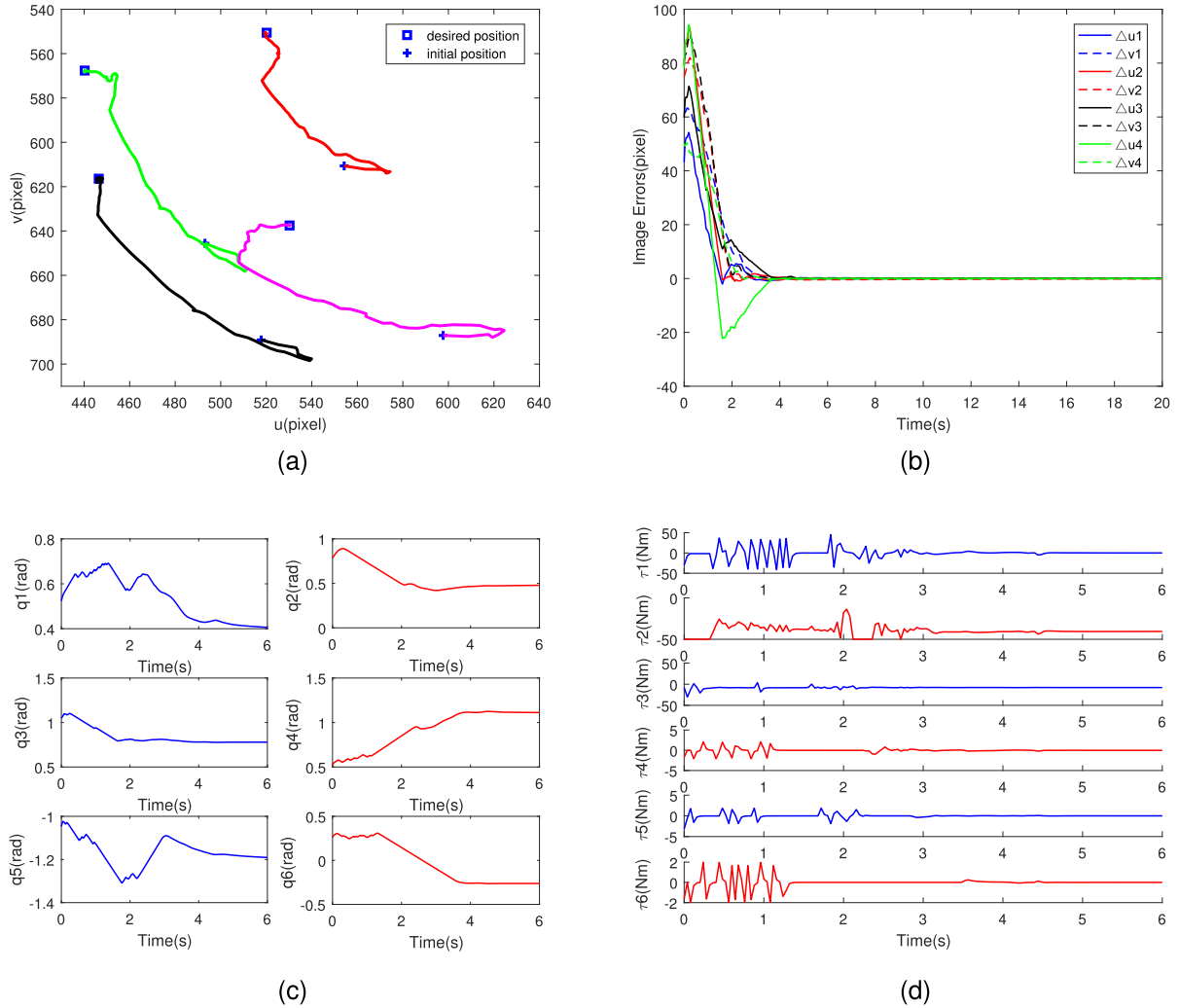
FIGURE 13. Simulation results by the proposed method for 6-DOF robot manipulator with eye-to-hand camera configuration. (a) Image trajectory. (b) Image errors. (c) Robot joint angles. (d) Torque.

then the comparison results of the three methods are shown in Figure 11, from Figure 11, we can see that only the proposed method succeeds in completing the visual servoing tasks, which validates the effectiveness of the proposed method.

It can be concluded that the proposed approach can achieve satisfactory control performance by considering robot dynamics for both the eye-in-hand and eye-to-hand camera configurations without the use of joint velocity measurements in the presence of the visibility constraint, the torque constraint and the model uncertainty.

B. 6-DOF ROBOT MANIPULATOR

To further demonstrate the performance of the proposed algorithm and verify that the method can cope with the 3D motion problem of robot manipulators, simulation results on a 6-DOF PUMA560 robot manipulator are presented in this subsection. The dynamic model and the dynamic parameters are given in [34].

In the eye-in-hand-camera configuration, the feature position coordinates with respect to the robot base frame are $(0.15, 0.15, 4)^T$ m, $(-0.15, -0.15, 4)^T$ m, $(0.15, -0.15, 4)^T$ m and $(-0.15, 0.15, 4)^T$ m, respectively. The transformation matrix from the robot end-effector frame to the camera frame is

$$T_e^c = \begin{bmatrix} 1 & 0 & 0 & 0.01 \\ 0 & 1 & 0 & -0.02 \\ 0 & 0 & 1 & 0.015 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (56)$$

the initial roughly estimated camera extrinsic matrix $\hat{T}_e^c(0)$ is

$$\hat{T}_e^c(0) = \begin{bmatrix} 1 & 0 & -0.3 & 0.001 \\ 0 & 1 & 0 & 0.03 \\ 0 & 0 & 0.95 & 0.001 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (57)$$

the initial rough estimations of camera intrinsic values and the rough estimations of feature position coordinates are
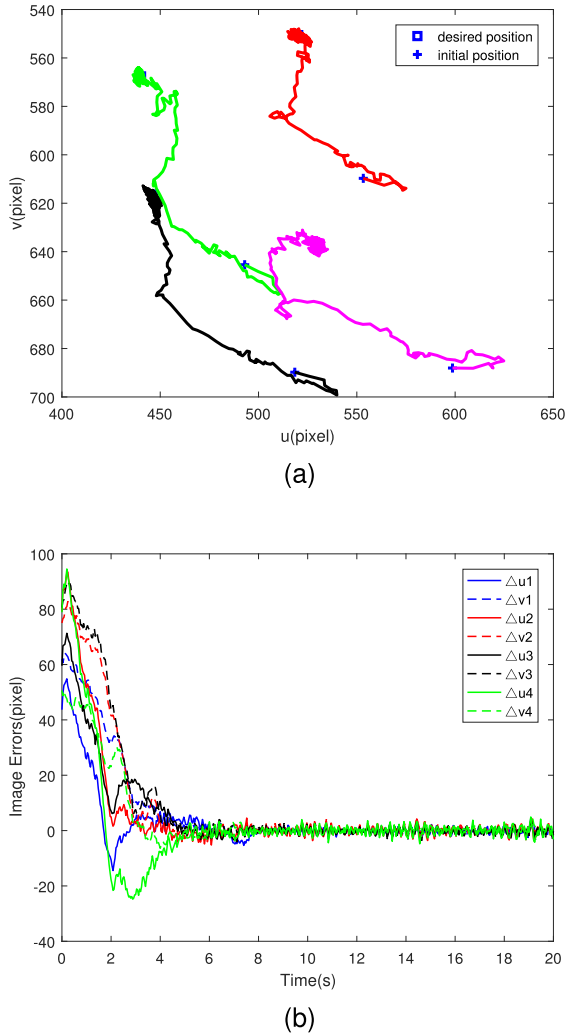
FIGURE 14. Simulation results by the proposed method for 6-DOF robot manipulator with eye-to-hand camera configuration in the case with noises. (a) Image trajectory. (b) Image errors.

set the same initial values to the simulation of the 2-DOF robot manipulator with eye-in-hand camera configuration. The initial positions of the features points on the image plane are $(513.6, 612.6)^T$ pixels, $(448.9, 528.8)^T$ pixels, $(423.7, 617)^T$ pixels and $(537.2, 525.2)^T$ pixels, respectively. The desired locations of the features points on the image plane are $(993.3, 275.5)^T$ pixels, $(1067.6, 170.7)^T$ pixels, $(969.9, 199.5)^T$ pixels and $(1091.5, 248.3)^T$ pixels, respectively. The objective is to control the motion of the robot manipulator by employing image information from the camera to take the feature points to their desired locations on the image plane. The actuator saturation limits of $-50\,\text{Nm} \leq \tau \leq 50\,\text{Nm}$ are considered. The weighting matrices are $F = 10I_{8\times8}$, $G = 0.1I_{6\times6}$, the prediction horizon and the control horizon are set as $N_p = 2$ and $N_c = 1$, the initial covariance matrix is $H(0) = 10^{10}I$, the forgetting factor is $\varepsilon = 0.995$, $K_v = -10I_{6\times6}$, and the sliding mode observer parameters are $\kappa = 6, \sigma = 1, \xi = 2$. The simulation results of the proposed method are shown in Figure 12, from the figure, we can see

that the proposed method can take the features to their desired locations for the control of the 6-DOF robot manipulator. The computational time in solving the optimization problem under the proposed approach is about 32 ms per sampling period.

In the eye-to-hand-camera configuration, the feature points are fixed on the robot end-effector. The transformation matrix from the robot base frame to the camera frame is

$$T_b^c = \begin{bmatrix} 0 & 0 & 1 & 0.05 \\ -1 & 0 & 0 & 0.2 \\ 0 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{58}$$

and the initial roughly estimated transformation matrix $\hat{T}_b^c(0)$ is

$$\hat{T}_b^c = \begin{bmatrix} 0 & -0.2 & 0.95 & 0.01 \\ -1 & 0 & 0 & 0.25 \\ 0 & -0.98 & -0.3 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{59}$$

the initial roughly estimated camera intrinsic parameters and the roughly estimated feature position parameters are set the same initial values to the simulation of the 2-DOF robot manipulator with eye-to-hand camera configuration. The initial locations of the features points on the image plane are $(554.2, 610.6)^T$ pixels, $(517.8, 689.1)^T$ pixels, $(493.1, 645.5)^T$ pixels and $(597.7, 687.1)^T$ pixels, respectively. The desired positions of the features points on the image plane are $(520.2, 550.6)^T$ pixels, $(446.4, 616.4)^T$ pixels, $(440.2, 567.6)^T$ pixels and $(530.2, 637.5)^T$ pixels, respectively. The actuator saturation limits, the weighting matrices, the prediction horizon, the control horizon, the initial covariance matrix are set the same values to the previous simulation, the forgetting factor is $\varepsilon = 0.994$, $K_v = -20I_{6\times6}$, and the sliding mode observer parameters are $\kappa = 20, \sigma = 10, \xi = 2$. The simulation results of the proposed algorithm are given in Figure 13, from which we can observe that the satisfactory control performance can be achieved for the control of the 6-DOF robot manipulator. The computational time in solving the optimization problem under the proposed approach is about 33 ms per sampling period. Moreover, to test the robustness of the proposed control approach, random noises with amplitude of ± 3 pixels are then added to the system output to simulate image processing errors, the visual servoing results are illustrated in Figure 14. It can be seen from Figure 14 that the control tasks are still satisfied. Furthermore, from Figures $12 - 14$, we can also see that the proposed algorithm can cope with the 3D motion problem of robot manipulators.

V. CONCLUSION
In visual servoing, the constraints are usually either ignored or treated at the kinematic level, and many existing schemes including MPC-based IBVS methods have limitations in the presence of model uncertainty. In this paper,

a new composite MPC-based IBVS design method by considering the robot dynamics is proposed to simultaneously deal with the constraints, the unknown camera intrinsic and extrinsic parameters and unknown depth information under the situation that the joint velocity measurements are not available. The design of the MPC scheme based on identification algorithm and sliding mode observer for the IBVS system has been investigated. The proposed approach can be applied to both the eye-in-hand and eye-to-hand camera configurations. By using MPC strategy, the control signal is computed through minimizing the cost function based on the image errors, and the constraints due to the visibility constraint and the torque constraint can be explicitly taken into account. Under the depth-independent image Jacobian matrix framework, the iterative identification algorithm is incorporated in the MPC design method for obtaining the model parameters at each time step. Moreover, to avoid performance decaying caused by measurement errors of the joint velocity, the sliding mode observer is used to provide the estimated values of the joint velocities. The simulation results show that the IBVS system under the composite MPC method can achieve satisfactory control performance.

REFERENCES

[1] D.-H. Park, J.-H. Kwon, and I.-J. Ha, "Novel position-based visual servoing approach to robust global stability under field-of-view constraint," *IEEE Trans. Ind. Electron.*, vol. 59, no. 12, pp. 4735–4752, Dec. 2012.

[2] B. Thuilot, P. Martinet, L. Cordesses, and J. Gallice, "Position based visual servoing: Keeping the object in the field of vision," in *Proc. IEEE Int. Conf. Robot. Automat.*, Washington, DC, USA, May 2002, pp. 1624–1629.

[3] V. Lippiello, B. Siciliano, and L. Villani, "Position-based visual servoing in industrial multirobot cells using a hybrid camera configuration," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 73–86, Feb. 2007.

[4] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[5] F. Chaumette and S. Hutchinson, "Visual servo control. II. Advanced approaches [Tutorial]," *IEEE Robot. Autom. Mag.*, vol. 14, no. 1, pp. 109–118, Mar. 2007.

[6] A. Chan, S. Leonard, E. A. Croft, and J. J. Little, "Collision-free visual servoing of an eye-in-hand manipulator via constraint-aware planning and control," in *Proc. Amer. Control Conf.*, San Francisco, CA, USA, Jun./Jul. 2011, pp. 4642–4648.

[7] X. Liang, H. Wang, W. Chen, and Y.-H. Liu, "Uncalibrated image-based visual servoing of rigid-link electrically driven robotic manipulators," *Asian J. Control*, vol. 16, no. 3, pp. 714–728, May 2014.

[8] X. Liang, H. Wang, W. Chen, D. Guo, and T. Liu, "Adaptive image-based trajectory tracking control of wheeled mobile robots with an uncalibrated fixed camera," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 6, pp. 2266–2282, Nov. 2015.

[9] N. R. Gans and S. A. Hutchinson, "Stable visual servoing through hybrid switched-system control," *IEEE Trans. Robot.*, vol. 23, no. 3, pp. 530–540, Jun. 2007.

[10] V. Lippiello, J. Cacace, A. Santamaria-Navarro, J. Andrade-Cetto, M. Á. Trujillo, Y. R. Esteves, and A. Viguria, "Hybrid visual servoing with hierarchical task composition for aerial manipulation," *IEEE Robot. Autom. Lett.*, vol. 1, no. 1, pp. 259–266, Jan. 2016.

[11] Y. Wang, H. Lang, and C. W. D. Silva, "A hybrid visual servo controller for robust grasping by wheeled mobile robots," *IEEE/ASME Trans. Mechatronics*, vol. 15, no. 5, pp. 757–769, Oct. 2010.

[12] F. Janabi-Sharifi and M. Marey, "A Kalman-filter-based method for pose estimation in visual servoing," *IEEE Trans. Robot.*, vol. 26, no. 5, pp. 939–947, Oct. 2010.

[13] A. Anwar, W. Lin, X. Deng, J. Qiu, and H. Gao, "Quality inspection of remote radio units using depth-free image based visual servo with acceleration command," *IEEE Trans. Ind. Electron.*, to be published.

[14] S. Haobin, H. Kao-Shing, L. Xuesi, and C. Jialin, "A learning approach to image-based visual servoing with a bagging method of velocity calculations," *Inf. Sci.*, vol. 481, pp. 244–257, May 2019.

[15] Y. Mezouar and F. Chaumette, "Path planning for robust image-based control," *IEEE Trans. Robot. Autom.*, vol. 18, no. 4, pp. 534–549, Aug. 2002.

[16] G. Chesi, "Visual servoing path planning via homogeneous forms and LMI optimizations," *IEEE Trans. Robot.*, vol. 25, no. 2, pp. 281–291, Apr. 2009.

[17] Y. Mezouar and F. Chaumette, "Optimal camera trajectory with image-based control," *Int. J. Robot. Res.*, vol. 22, nos. 10–11, pp. 781–803, Oct. 2003.

[18] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision," *IEEE Trans. Robot. Autom.*, vol. 9, no. 1, pp. 14–35, Feb. 1993.

[19] K. Hashimoto and H. Kimura, "LQ optimal and nonlinear approaches to visual servoing," in *Visual Servoing* (World Scientific Series in Robotics and Intelligent Systems), vol. 7, K. Hashimoto, Ed. Singapore: World Scientific, 1993, pp. 165–198.

[20] K. Hashimoto, T. Ebine, and H. Kimura, "Visual servoing with hand-eye manipulator-optimal control approach," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 766–774, Oct. 1996.

[21] P. Danès and D. Bellot, "Towards an LMI approach to multicriteria visual servoing in robotics," *Eur. J. Control*, vol. 12, no. 1, pp. 86–110, 2006.

[22] M. U. Khan, I. Jan, A. Ahmed, M. A. Azad, and N. Iqbal, "Uncalibrated visual servo control with multi-constraint satisfaction," in *Proc. IEEE Int. Conf. Robot. Automat.*, Shanghai, China, May 2011, pp. 6318–6323.

[23] O. Kermorgant and F. Chaumette, "Avoiding joint limits with a low-level fusion scheme," in *Proc. EEE/RSJ Int. Conf. Intell. Robots Syst.*, San Francisco, CA, USA, Sep. 2011, pp. 768–773.

[24] O. Kermorgant and F. Chaumette, "Dealing with constraints in sensor-based robot control," *IEEE Trans. Robot*, vol. 30, no. 1, pp. 244–257, Feb. 2014.

[25] G. Chesi, K. Hashimoto, D. Prattichizzo, and A. Vicino, "Keeping features in the field of view in eye-in-hand visual servoing: A switching approach," *IEEE Trans. Robot.*, vol. 20, no. 5, pp. 908–914, Oct. 2004.

[26] G. Lopez-Nicolas, S. Bhattacharya, J. J. Guerrero, C. Sagues, and S. Hutchinson, "Switched homography-based visual control of differential drive vehicles with field-of-view constraints," in *Proc. IEEE Int. Conf. Robot. Automat.*, Roma, Italy, Apr. 2007, pp. 4238–4244.

[27] Y. Zhang and S. Li, "A neural controller for image-based visual servoing of manipulators with physical constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5419–5429, Nov. 2018.

[28] M. Sauvée, P. Poignet, and E. Dombre, "Ultrasound image-based visual servoing of a surgical instrument through nonlinear model predictive control," *Int. J. Robot. Res.*, vol. 27, no. 1, pp. 25–40, Jan. 2008.

[29] G. Allibert, E. Courtial, and F. Chaumette, "Predictive control for constrained image-based visual servoing," *IEEE Trans. Robot.*, vol. 26, no. 5, pp. 933–939, Oct. 2010.

[30] T. T. Wang, W. F. Xie, G. D. Liu, and Y. M. Zhao, "Quasi-min-max model predictive control for image-based visual servoing with tensor product model transformation," *Asian J. Control*, vol. 17, no. 2, pp. 402–416, Mar. 2015.

[31] A. Hajiloo, M. Keshmiri, W.-F. Xie, and T.-T. Wang, "Robust online model predictive control for a constrained image-based visual servoing," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2242–2250, Apr. 2016.

[32] J. Davila, L. M. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1785–1789, Nov. 2005.

[33] G. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete-time multivariable adaptive control," *IEEE Trans. Autom. Control*, vol. 25, no. 3, pp. 449–456, Jun. 1980.

[34] B. Armstrong, O. Khatib, and J. Burdick, "The explicit dynamic model and inertial parameters of the PUMA 560 arm," in *Proc. IEEE Int. Conf. Robot. Automat.*, San Francisco, CA, USA, Apr. 1986, pp. 510–518.

ZHOUJINGZI QIU received the M.S. degree in control engineering from the Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China, in 2014. She is currently pursuing the Ph.D. degree with the School of Aeronautics and Astronautics, Shanghai Jiao Tong University.

Her scientific interests include visual servoing, predictive control, and computer vision.

SHIQIANG HU received the Ph.D. degree in electronics and information technology from the Beijing Institute of Technology, in 2002.

He is currently a Professor and the Associate Dean of the School of Aeronautics and Astronautics, Shanghai Jiao Tong University. His research areas include intelligent information processing, image understanding, and nonlinear filtering.

XINWU LIANG received the B.S. and Ph.D. degrees in control engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2011, respectively.

He was a Postdoctoral Fellow of the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, from 2011 to 2014, where he is currently an Associate Professor with the School of Aeronautics and Astronautics. He was a Postdoctoral Fellow of the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong, from 2014 to 2015. His current research interests include robot control, visual servoing, adaptive control, and computer vision.

• • •