Neurocomputing 151 (2015) 268–277

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Robots visual servo control with features constraint employing Kalman-neural-network filtering scheme

Xungao Zhong, Xunyu Zhong*, Xiafu Peng

Department of Automation, Xiamen University, Siming South Road, 361005 Xiamen, China

ARTICLE INFO

Article history:
Received 13 January 2014
Received in revised form
19 July 2014
Accepted 22 September 2014
Communicated by Cheng-Wu Chen
Available online 7 October 2014

Keywords:
Robots manipulation
Features constraint
Image Jcobian estimation
Kalman filter
Neural network

ABSTRACT

This paper presents an image-based servo control approach with a Kalman-neural-network filtering scheme for robots manipulation in uncalibrated environment. The image Jacobian on-line identification problems are firstly addressed by introducing the state estimation techniques, which have been incorporated neural network assists Kalman filtering (NNAKF). In fact, this is, the neural network (NN) can serve to play exactly the role of the error estimator, has the task of compensate the errors of Kalman filtering (KF). Then, by employing the NNAKF scheme, the proposed image-based servo control approach has guaranteed the robustness with respect to destabilized system attached dynamic noises, as well as the image features are constrained in field-of-view (FOV) of the camera. Furthermore, it is without requiring the intrinsic and extrinsic parameters of the camera during visual servoing tasks. To demonstrate further the validity and practicality of proposed approach, various simulation and experimental results have been presented using a six-degree-of-freedom robotic manipulator with eye-in-hand configurations.

1. Introduction

It is necessary that guaranteed the robust stability along with feature points kept within the FOV of the camera to successful robot manipulation by visual servo control. The visual-sensor-based robot manipulation is depends mainly on visual feedback to control the positioning or motion of a manipulator [1], [2]. There are two general categories, may be due to the difference definition of the feedback information [3], are position-based visual servoing (PBVS), image-based visual servoing (IBVS), and their hybrid visual servoing [4–10].

The visual sensors used in PBVS are to provide 3D position to regulate the pose of robot's end-effector relative to the object in the Cartesian space. This method is suitable for most industrial robotic manipulators owe to its characteristic of the global asymptotic stability [11]. However, this servo system is inevitably associated with the "hand-eye" calibration model. In consequence, the PBVS be more sensitive, in practical, with respect to the calibration errors and the depth information of the object [12], further with the possibility that the image features disappear from FOV [13–15].

In IBVS, there is direct control of the feature points on the image plane for robot manipulation, and the image Jacobian matrix is used for the description of the differential relationship between image features and end-effector moving [2]. This characteristic has simplified the computation of PBVS, therefore, the IBVS has been catch more attention in recently [6–8]. While, there is, the most pressing of issues presented in uncalibrated IBVS is the calculation of image Jacobian matrix, which as a local and linear approximation to this nonlinear and highly coupled mapping between visual-motor spaces.

As far as the image Jacobian matrix calculation is concerned, some existing works actually considered it as a dynamic parameters identification problem, the solution for this issue are broadly with on-line techniques, such as Broyden-based method and the family of Broyden updating formulas can be defined to estimates the Jacobian matrix [16], [17], and in [18] the exponentially weighted recursive least square update method is used for Jacobian matrix estimation. However, these techniques may actually depend on the system configurations or the tasks to be accomplished, and ill-suited for dealing with the dynamic noises of servo system. While a statistically robust M-estimator is proposed in [19], this method uses visual-motor memory to gradually increase the quality of the Jacobian estimate, which does not require the system model parameters. In [20], it was presented a method for local calculation of the image Jacobian through training without the need for depth estimation. And in recent years, the literature with Kalman–Bucy filter (KBF) for Jacobian matrix estimation is proposed in [21]. This KF-based method assumes that the filtering parameters are known and the

* Corresponding author. Tel.: +86 13720885429.
E-mail addresses: zhongxungao@163.com (X. Zhong), zhongxunyu@xmu.edu.cn (X. Zhong), pengxiafu@126.com (X. Peng).

http://dx.doi.org/10.1016/j.neucom.2014.09.043
0925-2312/

constructed system with the observed states of Jacobian elements is unvaried. However, this assumption unsuited to some unknown dynamic noises environments, especially the situation with serious changing of observation and state models. As an improvement, the [22] investigated the application of KF in the state space model with variable noise parameters. In [23], an iterative adaptive extended Kalman filter (EKF) is proposed by integrating mechanisms for noise adaptation and iterative-measurement linearization in visual servoing tasks. In [24], the transition matrix of the KF is adjusted to address the problem of where the system model is unavailable, and then the performance of this adaptive KF predictor has been evaluated on a visual servoing application.

The famous Kalman filtering structure is a minimum-variance state estimator for ideal linear dynamic systems with strongly Gaussian white noises [25], which is widely used for real-time state estimation, and parameter identification under the least squares condition [26], [27]. In most practices, however, the processing noises and the observation noises of visual sensors rather than the simple Gaussian white noises. In this case, the limitation to the use of KF for state estimation is the suboptimum, meanwhile the filter will be easy divergent. Some solutions to the nonwhite noise handling have been presented, among which include the dimension extension of KF [28], LMS-based adaptive filtering [29], wavelet-based adaptive Wiener filtering [30]. In recently, many works shed new light on adaptive KF approaches by introduction the concept of neural network [31–35]. In [31], The KF gain was replaced by a feed-forward neural network, the filter with robust capability for estimating the states of the plant in a stochastic environment without knowledge of noise statistics. In [32], it was thinking that the estimation accuracy of the tracking Kalman filter (TKF) is degraded due to the uncertainties which can't be expressed by the linear state-space model, and then presents a method for improving the TKF's estimation accuracy by using a multilayered neural network. In [33], the KF with backpropagation network was proposed to track the maneuvering targets, which the inaccuracies in model were corrected by the neural network and the tracking accuracy can be improved.

As mentioned above, those KF-based techniques have the same properties as recursive-least-squares-based estimation processing, which rely on "known" noise statistics, and good approximation of linear state-space model. Deviations from such assumptions usually lead to degraded Jacobian matrix estimation during visual servoing, that failure to have end-effector positioning in cases of large displacement between the initial and desired poses, and the image features out of FOV is also a risk, *i.e.* the results are only accurate in a small subspace of robot workspace, and non-robustness with respect to destabilized system model attached dynamic noise.

In this paper, we present a robust NNAKF state estimator for image Jacobian on-line identification, without requiring the intrinsic and extrinsic parameters of the camera and the depth information of the target, and then a new image-based servo control approach based on NNAKF is proposed for model-free robot manipulation. In practices, the KF for Jacobian matrix estimation was often suboptimum, in that the traditional KF lacks the adaptive ability to the maneuvering model and the dynamic feature of noises. Thus there are at least two error-elements should be taken into account to improve the KF for best state-estimation: (1) modeling error, due to the perturbation of robot visual servoing system, it is equivalent that the linear time invariant observation model and state model intrinsic contains the nonlinear approximation errors. (2) Statistics error, the precise statistic knowledge of processing noise and observation noise is difficult to be definitely determined in actual environment, since the covariance of those noises may be dynamic changing. So in our considering, a NN was adopted to play exactly the role of the error estimator for assisting the KF algorithm, experiments showed that the proposed NN could improve the robustness of KF for dynamic noises and uncertainty of the system model, and then the NNAKF could construct a robust state estimator for image Jacobian on-line identification with high precise. Finally, we have design a new IBVS framework by employing NNAKF. In our finding, the image Jcobian matrix is dynamic estimated with nothing to do with the camera calibration error and target's 3D-modeling errors. In additional, the proposed IBVS is different from the traditional PBVS methods, with the merits of robust stability no matter the system destabilized with odious dynamic noises, also the servo controller could guaranteed the image features are constrained on the FOV of the camera in global workspace of the robotics.

The rest of the paper is organized as follows: the background of the visual servoing for model-free robot manipulation is represented in the next section. The image Jacobian identification problem with state estimation techniques is presented in Section 3. The NNAKF algorithm and a new IBVS scheme based on NNAKF are proposed in Sections 4 and 5, respectively. The simulation and experimental results are discussed in Section 6, followed by conclusions in Section 7.

2. Visual servo control for robot manipulation

First, we given an eye-in-hand robotics configuration system, the goal of image-based visual servoing is to drive the end-effector from the current pose, to the desired pose. Herein, let us define a image error $\mathbf{e}_s(t)(\mathbf{e}_s(t) \subset \Re^n)$, according to

$$\mathbf{e}_s(t) = \mathbf{S}\left(p_i(t), C\right) - \mathbf{S}'\left(p_i'(t), C\right) \tag{1}$$

where $\mathbf{S}(t)$, $\mathbf{S}'(t)$ are belong to n-dimensional current image features and desired image features, respectively. $p_i(t)$, for $i = 1,2,3,\ldots,\ n$ represents a set of 3D position of n feature points of the object in the camera frame, and $p'_i(t)$ denote the 3D coordinates of desired feature points. The coefficient $\boldsymbol{C}$ has includes the calibration parameters of the camera and the depth information of the feature points.

In order to conduct the robot finishes one manipulation task by visual feedback, the visual servo system needs employing a control law $\mathbf{U}(t)$ $(\mathbf{U}(t) \subset \Re^m)$, to minimize the cost function as follows:

$$F(t) = \frac{1}{2}\mathbf{e}_s(t)^{\mathbf{T}}\mathbf{e}_s(t) \tag{2}$$

Generally, a slightly location movement of end-effector will lead to the nonlinear complex change of many features on image plane. There is a practical solution for describing the change-relationship between the image features and the pose of end-effector, by adopting an image Jacobian matrix which was original proposed in [2]. The association of the time change of the image error $\mathbf{e}_s(t)$ with the time derivative of end-effector's spatial velocity $\mathbf{V}_e(t)$ $(\mathbf{V}_e(t) \subset \Re^m)$ is done assuming linearity through the Jacobian matrix $\mathbf{J}_t$, as follows:

$$\dot{\mathbf{e}}_s(t) = \mathbf{J}_t(\mathbf{S}(t), \mathbf{V}_e(t))\dot{\mathbf{V}}_e(t) \tag{3}$$

where $\dot{\mathbf{V}}_e(t) = [\dot{v}(t), \dot{w}(t)]^{\mathbf{T}}$, $v(t)$ and $w(t)$ are the linear velocity and angular velocity of the end-effector, respectively. Considering the desired image feature $\mathbf{S}'(t)$ is constant parameter due the fixed goal pose, then substituting Eq. (1) into Eq. (3), we have

$$\dot{\mathbf{S}}(t) = \mathbf{J}_t\mathbf{S}(\mathbf{S}(t), \mathbf{V}_e(t))\dot{\mathbf{V}}_e(t) \tag{4}$$

where the $\mathbf{J}_t$ is defined by

$$\mathbf{J}_t(\mathbf{S}(t), \mathbf{V}_e(t)) = \left[\frac{\partial \mathbf{S}(t)}{\partial \mathbf{V}_e(t)}\right]_{m \times n} \tag{5}$$

Above Eq. (5) denotes that the camera's parameters and the depth information of target have been implied in $\mathbf{J}_t$, due to S(t) is the function of $p_i(t)$ and $\mathbf{C}$, *i.e.* it is means that the camera-robot model must be firstly calibrated for the robot manipulation.

In this paper, we suggest a dynamic formation of the image Jacobian matrix $\mathbf{J}_t$, without the parameters of camera-robot model and the depth information of feature points, as follows:

$$\mathbf{J}_t(\mathbf{S}(t),\mathbf{V}_e(t)) = \begin{bmatrix} \frac{\partial s^1(t)}{\partial v_e^1(t)} & \cdots & \frac{\partial s^n(t)}{\partial v_e^1(t)} \\ \vdots & \ddots & \vdots \\ \frac{\partial s^1(t)}{\partial v_e^m(t)} & \cdots & \frac{\partial s^n(t)}{\partial v_e^m(t)} \end{bmatrix} = \begin{bmatrix} j_{11} & \cdots & j_{1n} \\ \vdots & \ddots & \vdots \\ j_{m1} & \cdots & j_{mn} \end{bmatrix}_{m\times n} \tag{6}$$

The element j_{mn} of $\mathbf{J}_t$ indicates the derivative-relationship between the image features $\mathbf{S}(t)$ and the end-effector spatial velocity $\mathbf{V}_e(t)$. The Jacobian on-line identification problems then were solved by introduces in state-space infrastructure, which has been incorporated the state stimation techniques.

3. Dynamic Jacobian identification problems

One of the essential problems of image-based visual servo control is precise calculating the image Jacobian matrix, for robotics visual-motor spaces mapping. The image Jacobian identification can be formulated as state estimation problem with Kalman filtering techniques, considering a linear discrete-time dynamical system to represent the state model of robotic system, as follows:

$$\mathbf{X}_{t+1} = \mathbf{I}\mathbf{X}_t + \mathbf{W}_t \tag{7}$$

$$\mathbf{Z}_t = \mathbf{H}_t\mathbf{X}_t + \mathbf{V}_t \tag{8}$$

where $\mathbf{W}_t$ ($\mathbf{W}_t \subset \Re^{mn}$) is the processing noise with zero mean, and the variances is $\mathbf{Q}(t)$. $\mathbf{V}_t$ ($\mathbf{V}_t \subset \Re^m$) is the observation noise with zero mean, and the variances is $\mathbf{R}(t)$. $\mathbf{X}_t$ is the state vector of robotic system, which formed by concatenations of the row and the column elements of Jacobian matrix $\mathbf{J}_t$, *i.e.*

$$\mathbf{X}_t = [j_{11} \quad j_{12} \quad \cdots \quad j_{mn}]^{\mathbf{T}}_{mn\times 1} \tag{9}$$

In Eq. (8), $\mathbf{Z}_t$ ($\mathbf{Z}_t \subset \Re^m$) is the observation vector which provided by plant outputting, gives

$$\mathbf{Z}_t = \mathbf{S}_{t+1} - \mathbf{S}_t = \mathbf{J}_t\dot{\mathbf{V}}_e(t) \tag{10}$$

Thus, the observation matrix $\mathbf{H}_t$ can be formed as

$$\mathbf{H}_t = \begin{bmatrix} \dot{\mathbf{V}}_e(t)^{\mathbf{T}} & & 0 \\ & \ddots & \\ 0 & & \dot{\mathbf{V}}_e(t)^{\mathbf{T}} \end{bmatrix}_{m\times mn} \tag{11}$$

The state model (7) and the observation model (8) with the white noise $\mathbf{W}_t$ and $\mathbf{V}_t$ are satisfied with the conditions of traditional KF's [25], thus the KF algorithm can be derived for dynamic Jcobian estimation in state-space. The KF includes the estimation and observation update two basic processes, the filter bring the comparing between the values of the actual and predicted state of the robotics, the state errors of prediction are fed back in an optimal way to correct the state estimates by multiplying the error by the filtering gains. Immediately before the next observation, the robot states and covariance are predicted to the next sample instant using the systems model. Note that, the KF structure is the optimal state estimator under the least squares condition, only for ideal linear dynamic systems with "known" Gaussian white noise. To our best knowledge, however, the KF applied in actual robot manipulation is very sensitive to the statistic characteristics of noise and the errors of system modeling, that due to in actual workshop environment the processing noises and the observation noises of visual sensors are not the simple Gaussian white ones. In next section, we propose a practical filtering schema by employing a neural compensator for assisting the KF to improve the robust performance in dynamic environment.

4. The KF with neural network for Jacobian identification

The KF was originally designed to provide the best state-estimation of a linear time invariant dynamic system with "known" Gaussian white noises, if the observation vector is strictly generated by the linear observation model (8) and state model (7), the best estimation of robot's state vector $\mathbf{X}_t$ will be optimal, in the sense that $E\{\mathbf{X}_t\} = \hat{\mathbf{X}}_t$ is minimized. However, due to the perturbation of observation model and state model, the linear time invariant system intrinsic contains the nonlinear approximation errors. On the other hand, the KF's equation gives the filtering gain as the function of noises statistics, it is contradictory that the precise statistic knowledge of noises is difficult to be definitely determined in the actual environment, due to in the covariance $\mathbf{Q}(t)$ and $\mathbf{R}(t)$ of noises $\mathbf{W}_t$ and $\mathbf{V}_t$ may be dynamic changing, *i.e.* there are errors of statistics uncertainty implied in KF processing. Therefore the best state-estimation $\hat{\mathbf{X}}'_{t/t}$ one should designed according to

$$\hat{\mathbf{X}}'_{t/t} = \hat{\mathbf{X}}_{t/t} + \mathbf{e}_{\hat{X}_{t/t}} \tag{12}$$

where $\hat{\mathbf{X}}_{t/t}$ is the suboptimal state, which estimated by KF.$\mathbf{e}_{\hat{X}_{t/t}}$ is the estimation error of KF, which caused by system's modeling error and statistic errors of the processing noise $\mathbf{W}_t$ and observation noise $\mathbf{V}_t$. In this paper, we propose a method to eliminate $\mathbf{e}_{\hat{X}_{t/t}}$ by employing a neural network. In fact, there is, the NN can serve to play exactly the role of the error estimator, with the task of minimizing the error between the desired state $\hat{\mathbf{X}}'_{t/t}$ and the estimated state $\hat{\mathbf{X}}_{t/t}$. The NN with two-layer structure, which the output of each neuron of the hidden layer is, defined as

$$n_i^1 = f\left(\|\mathbf{W}^1 - \mathbf{G}\| \bullet \theta_i^1\right) \tag{13}$$

$$f(x) = \exp(-x^2) \tag{14}$$

$$\|\mathbf{W}^1 - \mathbf{G}\| = \sqrt{(\mathbf{W}^1 - \mathbf{G}^{\mathbf{T}})(\mathbf{W}^1 - \mathbf{G}^{\mathbf{T}})^{\mathbf{T}}} = \sqrt{\sum_{i=1}^{k} (w_i^1 - g_i)^2} \tag{15}$$

where $\mathbf{G} \in R^{m\times n}$ is the input vector with the ith element denoted as g_i, $\mathbf{W}^1 \in R^{n\times m}$ is the neuron weight of input vector to hidden layer, θ_i^1 is the threshold of ith neuron, $f(\cdot)$is a Gaussian basis function.

The neuron of output layer with linear sum of the outputs of hidden neurons, the linear function $g(x)$ is approximated as

$$n_i^2 = g(n^1) = \sum_{i=1}^{n} w_i^2 n_i^1 + \theta_i^2 \tag{16}$$

where w_i^2 and θ_i^2 are the neuron weight and threshold of output layer, respectively.

The learning algorithm have been applied to determine the structure of neural network, and the learning goal is obtain the convergence of the connection weights $\mathbf{W}^1$ and $\mathbf{W}^2$ of the input vector to hidden layer and the hidden layer to output layer, respectively. The law of weight updating is given by

$$w_{ij}^l(k) = w_{ij}^l(k-1) - \alpha\frac{\partial E}{\partial w_{ij}^l(k-1)} \tag{17}$$

$$E = \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{M}(e^{nm}_{\hat{X}'_{t/t}} - e_o^{nm})^2 \tag{18}$$

where $e^{nm}_{\hat{X}'_{t/t}}$ is the output of the mth neurons of output layer with nth input sample, e_o^{nm} is the desired output of output layer, α is the learning rate, l=1, 2 refers to the hidden layer and output layer, respectively.

As illustrated in Fig. 1, the NN was incorporated into the Kalman filter, where $\mathbf{Z}^{-1}$ represents a unit of time delay, the inputs of the NN including three elements, as follows:

The Kalman filtering gain error

$$\mathbf{e}_{\tilde{k}_{t/t-1}} = \mathbf{K}_t - \mathbf{Z}^{-1}\mathbf{K}_t \tag{19}$$

The state estimation error

$$\mathbf{e}_{\tilde{X}_{t/t-1}} = \hat{\mathbf{X}}_{t/t} - \mathbf{Z}^{-1}\hat{\mathbf{X}}_{t/t} \tag{20}$$

The observation error:

$$\mathbf{e}_{\tilde{Z}_{t/t-1}} = \mathbf{Z}_t - \mathbf{H}_t\hat{\mathbf{X}}_{t-1/t-1} \tag{21}$$

where $\mathbf{K}_t$ is the Kalman filtering gain, $\hat{\mathbf{X}}_{t/t}$ is the value of state-estimation by KF.

The goal of employing NN is compensate the filtering error of KF in each iteration time, so the outputs of network is the estimation error $\mathbf{e}_{\hat{X}_{t/t}}$ for KF, gives

$$\mathbf{e}_{\hat{X}_{t/t}} = \sum_{i=1}^{n} w_i^2 n_i^1 + \theta_i^2 \tag{22}$$

Finally, the best state-estimation $\hat{\mathbf{X}}'_{t/t}$ was obtained by NNAKF, as shown in Eq. (12).

5. Development of visual servo control schema based on NNAKF

In this section, we present an image-based visual servoing (IBVS) approach by employing the NNAKF schema for robot manipulation in uncalibrated environment. In our finding, there is unnecessary considering the kinematics model of the robotic manipulator, and the intrinsic and extrinsic parameters of the camera.

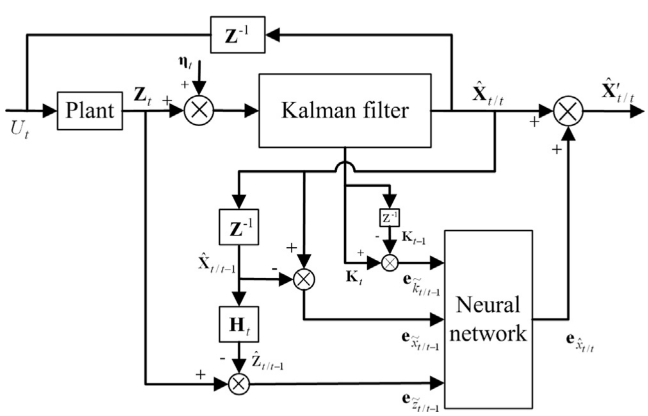


Fig. 1. The structure of neural network assists Kalman filtering (NNAKF).

The schematic of the proposed IBVS is shown in Fig. 2. The NNAKF state estimator is provided with currently image features which captured by CCD camera. In this paper, the desired feature vector $\mathbf{S}'_T$ does not change over time, and therefore can be calculated before the main control loop of the experiment. On the other hand, the currently feature vector $\mathbf{S}_t$ is not constant due to the camera motion, and is obtained in each time instant by camera, gives

$$\mathbf{S}_t = [s_1 \quad s_2 \quad \cdots \quad s_n]^{\mathbf{T}}$$
$$= [u_1 \quad v_1 \quad u_2 \quad v_2 \quad \cdots \quad u_n \quad v_n]^{\mathbf{T}}_{2n\times 1} \tag{23}$$

where $s_i = [u_i \ v_i]$ is the feature point. Supposing $\mathbf{V}_t = [v_x, v_y, v_z]^{\mathbf{T}}$ is the linear velocity vector of the end-effector, $\mathbf{W}_t = [w_x, w_y, w_z]^{\mathbf{T}}$ is the angular velocity vector, and $\mathbf{U}_t = [\mathbf{V}_t, \mathbf{W}_t]^{\mathbf{T}}_{6\times 1}$ is the control variable, according to Eq. (5), the image Jacobian matrix is given by

$$\mathbf{J}_t = \left[\frac{\partial \mathbf{S}_t}{\partial \mathbf{U}_t}\right]_{2n\times 6} \tag{24}$$

According to Eqs. (9) and (11), the state-vector of robotic instrumental system and the observation matrix are given, respectively, by

$$\mathbf{X}_t = \left[j_{11} \quad j_{12} \quad j_{13} \quad \cdots \quad j_{ik}\right]^{\mathbf{T}}_{(2n*6)\times 1} \tag{25}$$

$$\mathbf{H}_t = \left[\begin{pmatrix} \mathbf{V}_t\Delta t & & 0 \\ & \ddots & \\ 0 & & \mathbf{V}_t\Delta t \end{pmatrix}_{2n\times(3*2n)} \quad \begin{pmatrix} \mathbf{W}_t\Delta t & & 0 \\ & \ddots & \\ 0 & & \mathbf{W}_t\Delta t \end{pmatrix}_{2n\times(3*2n)} \right] \tag{26}$$

where j_{ik} refers to ith row and kth column of the $\mathbf{J}_t$.

Note that, the state initialization is very important for robot with robust stability manipulation, the popular method is use a neural network to provide the initial join guess. While for the initial state guess, we chosen a common method obtained by introduce the robot probe moving at the neighborhood of its initial position n times $\partial p^{j-n+1} \ldots \partial p^j$, and observe the corresponding features displacements $\partial f^{j-n-1} \ldots \partial f^j$ in image frame. The initial state vector $\mathbf{X}(0)$ then could be obtained reasonably by

$$\mathbf{X}(0) = [\partial f^{j-n-1} \ldots \partial f^j][\partial p^{j-n+1} \ldots \partial p^j]^{\mathbf{T}} \tag{27}$$

The control law should be employed to drive the end-effector to reach its desired pose. The image error as shown in Eq. (1) at the time instant t, rewritten as

$$\mathbf{e}_s(t) = \mathbf{S}(t) - \mathbf{S}'(t) \tag{28}$$

Then according to Eq. (3) the control law can be formed by

$$\mathbf{U}(t) = -\lambda \hat{\mathbf{J}}_t^{+} \mathbf{e}_s(t) \tag{29}$$

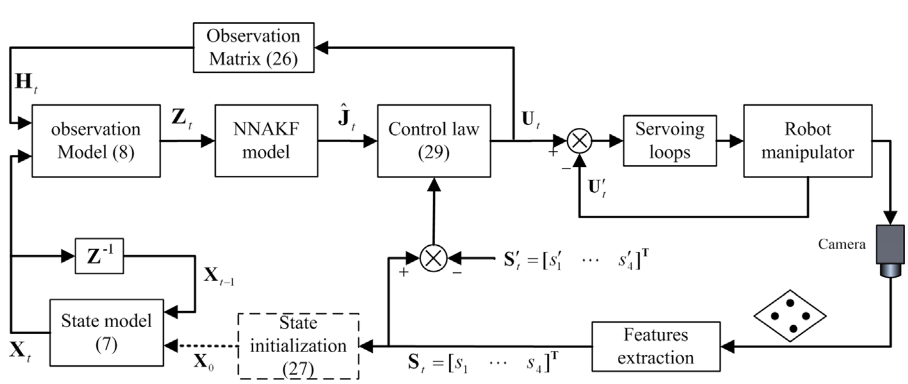


Fig. 2. The schema of proposed IBVS based on NNAKF.

where λ is the control rate, and $\hat{\mathbf{J}}_t^+$ is the inverse Jacobian matrix of $\hat{\mathbf{J}}_t$, given by

$$\hat{\mathbf{J}}_t^+ = \hat{\mathbf{J}}_t^{\mathbf{T}} \left(\hat{\mathbf{J}}_t \hat{\mathbf{J}}_t^{\mathbf{T}} \right)^{-1} \tag{30}$$

As illustrated in Fig. 2, the main task of NNAKF is play a state estimator for dynamic Jacobian identification, and then the servo loops drive the end-effector to reach the desired pose by employing a control law, the algorithm of IBVS based on NNAKF can be simply depicted as follows. First, the servo system should been initialize as Eq. (27), *i.e.* we can get the robot's initialization state vector $\mathbf{X}(0)$, then according to currently control variable $\mathbf{U}_{t-1}$, compute the observation matrix $\mathbf{H}_{t-1}$ as Eq. (26). The updating of the state vector from time $t-1$ to time t can be realized by using the state equation (7), and the output vector $\mathbf{Z}_t$ can be calculated by using observation equation (8). The best state-estimation $\hat{\mathbf{X}}'_{t/t-1}$ of the state vector $\mathbf{X}_{t-1/t-1}$ at $t-1$ time can be obtained by NNAKF, in this step we will get the best identification of the image Jacobian matrix ($\hat{\mathbf{J}}_t \leftarrow \hat{\mathbf{X}}'_{t/t-1}$). Finally, calculate the robot's control variable $\mathbf{U}_t$ by employing control law equation (29), then the servoing loops drives the robot from currently pose to the next pose, if the image error $e_S(t)=0$, the iteration loop will be end, otherwise go to the next iteration time ($t \leftarrow t+1$).

6. Results and discussions

To evaluate the performance of the proposed IBVS method which based on NNAKF, the simulation and experimental results have been carried out using eye-in-hand configurations. In simulation, the activities of the camera cover the linear and rotational movement in robot's workspace. Herein, the most difficult servo tasks which consist of two cases have been performed. In case 1, we demonstrate simulation that our IBVS base don NNAKF scheme ensures the image features were constrained in FOV, by considering the case where the feature points of the object at both the initial and the desired poses are very close to the FOV boundary. In case 2, we demonstrate simulation the robust stability property of proposed IBVS with respect to destabilized system and dynamic noises, by considering the case where translational and rotational displacements from the initial pose to the desired pose are large.

In experiment, we have set up an eye-in-hand robotic system (see Fig. 3(a)), the experimental task is drive the robot's end-effector from the initial pose (example see Fig. 3(b)) to the desired pose (example see Fig. 3(c)) by applying proposed IBVS. Our visual servoing system consists of a *DENSO* RC7M-VSG6BA robotic controller, a computer with Intel Corei5 2.67 GHz CPU, 4 GBs RAM for image processing, the computer communicate with controller by RS232C serial interface, and a *DENSO VS*-6556GM six-DoF robotic manipulator with an Basler scA1300-32fm/fc camera mounted at its end-effector. The object is an A4 paper with four black-colored small circular disks on it, the object images are captured by the camera at the rate of 30 Hz, resolution is 640×480, and the center points of the small circular disks are used as feature point.

In simulation and experiment, the input samples set for training the NN are from KF cycle, thus, before the training phase, the KF without adding NN for robot manipulation are necessary, then the knowledge of the filtering gain error $\mathbf{e}_{\tilde{k}_{t/t-1}}$, the state estimation error $\mathbf{e}_{\tilde{x}_{t/t-1}}$, and the observation error $\mathbf{e}_{\tilde{z}_{t/t-1}}$ were obtained at each KF iteration time, on the other hand, the output samples are from the errors between desired state and estimation state of KF. In this paper, total 601 input–output samples pair are used for training of the NN, The learning laws with the gradient descent method as shown in Eq. (17), the finally training results show that the hidden layer with 295 neurons, and output layer with 48 neurons, the training reaches the best validation performances with the minimum sum squared error (MSE) is 1.3×10^{-6}, it is seen that the test output of network is very close to the training output.

Simulation case 1, in this case, the goal is to test the performances of image features constraint between traditional PBVS, KF and our proposed IBVS methods. Setting the initial features $\mathbf{S}(k)=$ (378.6, 235.7, 311.9, 289.4, 363.6, 354.5, 429.4, 301.4), desired features $\mathbf{S}^*(k)=$(47.7, 47.7, 47.7, 464.3, 464.3, 464.3, 464.3, 47.7), which both are very close to the FOV boundary. The results are illustrated in Fig. 4, among the first row of Fig. 4. Fig. 4(a-1) shown that the feature trajectories on image plane which obtained by our IBVS method are constrained on the FOV, but the PBVS and KF methods for this task the feature trajectories easy leaves the FOV (see Fig. 4(b-1) for PBVS, and Fig. 4(c-1) for KF). In addition, the comparison between Figs. 4(a-2), (b-2) and (c-2) in second row of Fig. 4, shown that the camera trajectory in the Cartesian space, which obtained by our IBVS moves with smooth from the initial pose toward the desired pose. For the same task, however, the PBVS method converges to the desired pose as the camera motion becomes slight odd, in other words, the robot's end-effector with

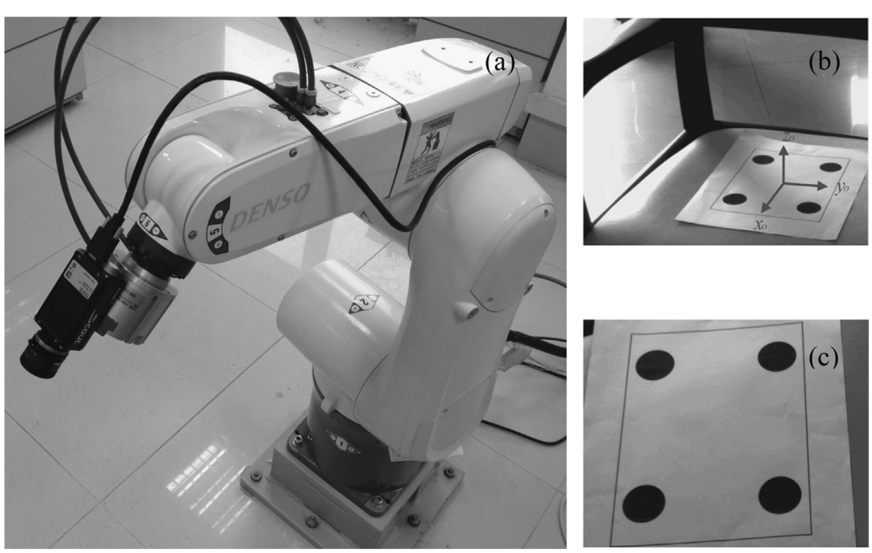


Fig. 3. (a) Experimental environment with eye-in-hand configurations, (b) initial features, and (c) desired features.

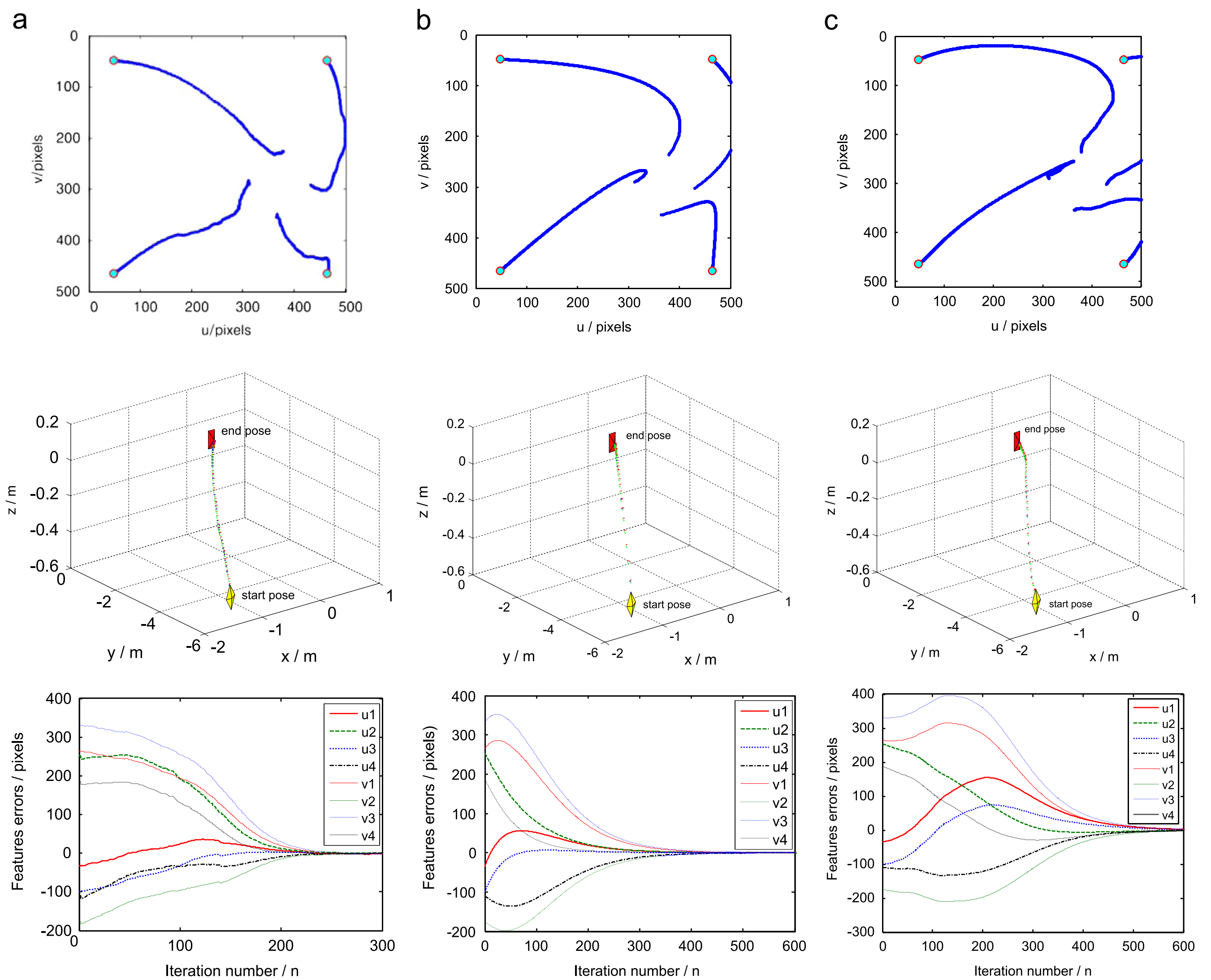


Fig. 4. Simulation results of case 1. 1st, 2nd and 3rd columns are the results obtained by (a) our proposed IBVS, (b) PBVS, and (c) KF methods, respectively. 1st, 2nd and 3rd rows corresponding to (1) image feature trajectories, (2) camera moving trajectories, and (3) image feature errors, respectively.

abrupt movement in initial phase. And the KF method converges to the desired pose, the camera with slight retreat moving, that means both PBVS and KF methods' result are not perfect compare with our approach's. Among the third row of Fig. 4, Figs. 4(a-3), (b-3) and (c-3), shown that the image error are converge to zero at finally, *i.e.* the robot has finishes the manipulation task. Note that, in actual experiments, the PBVS and the KF methods will lead to the robot lost the abilities of positioning, for the cause that the image features lost in FOV. Therefore, compared with IBVS and KF methods, our approach takes the advantages in both to constrain the feature trajectories on image plane, and to optimize the camera motion trajectory in the Cartesian space, that all own to the neural network plays a crucial role of error compensation to improve the performances of robot manipulation for keeping the feature points within FOV of the camera.

Simulation case 2, this case is to examine the property of robust stability in odious dynamic noisy environment between our IBVS based on NNAKF and KF methods. We setting the initial features $\mathbf{S}(k)=(272.2, 336.3, 186.9, 380.7, 230.5, 465.7, 315, 420.7)$, desired features $\mathbf{S}^*(k)=(89.3, 89.3, 89.3, 422.7, 422.7, 422.7, 422.7, 89.3)$. The uniformly distributed random noises were added to the NNAKF model, at each time instant. In order to illustrate the robust property of our IBVS scheme, we further have divided this case into three sub-cases, by considering different statistic characteristics of the noise. In sub-case 1, the noises with zero mean, and the variance is 3×10^{-8} were added to the NNAKF model, the results are presented in Fig. 5, compare with the KF, the camera moving in the Cartesian space and the features trajectory on the image plane are more smoother and preferable produced by our approach.

When the noises with zero mean, and the variance is 9×10^{-8}, the performances of camera moving and the features trajectory for this sub-case are shown in Fig. 6. It is obvious that the results of KF method are altered by the influence of noises, the end-effector could reaches the desired position and zeroing the image errors are accomplished, but the camera with unnecessary retreat moving. While our IBVS method be used this task not only the camera can converges on the desired pose but also with the property of robust stability.

On the other hand, when the noises with zero mean, and the variance increases to 2×10^{-7}, Fig. 7 shown that the results of the KF method will serious altered by the influence of noises, almost lost the position performance, since the camera with large retreat moving in Cartesian space, which it is easy reaching the limitation of robot join. In the same sub-case, however, by our proposed IBVS the camera always can converge to the desired pose with smooth

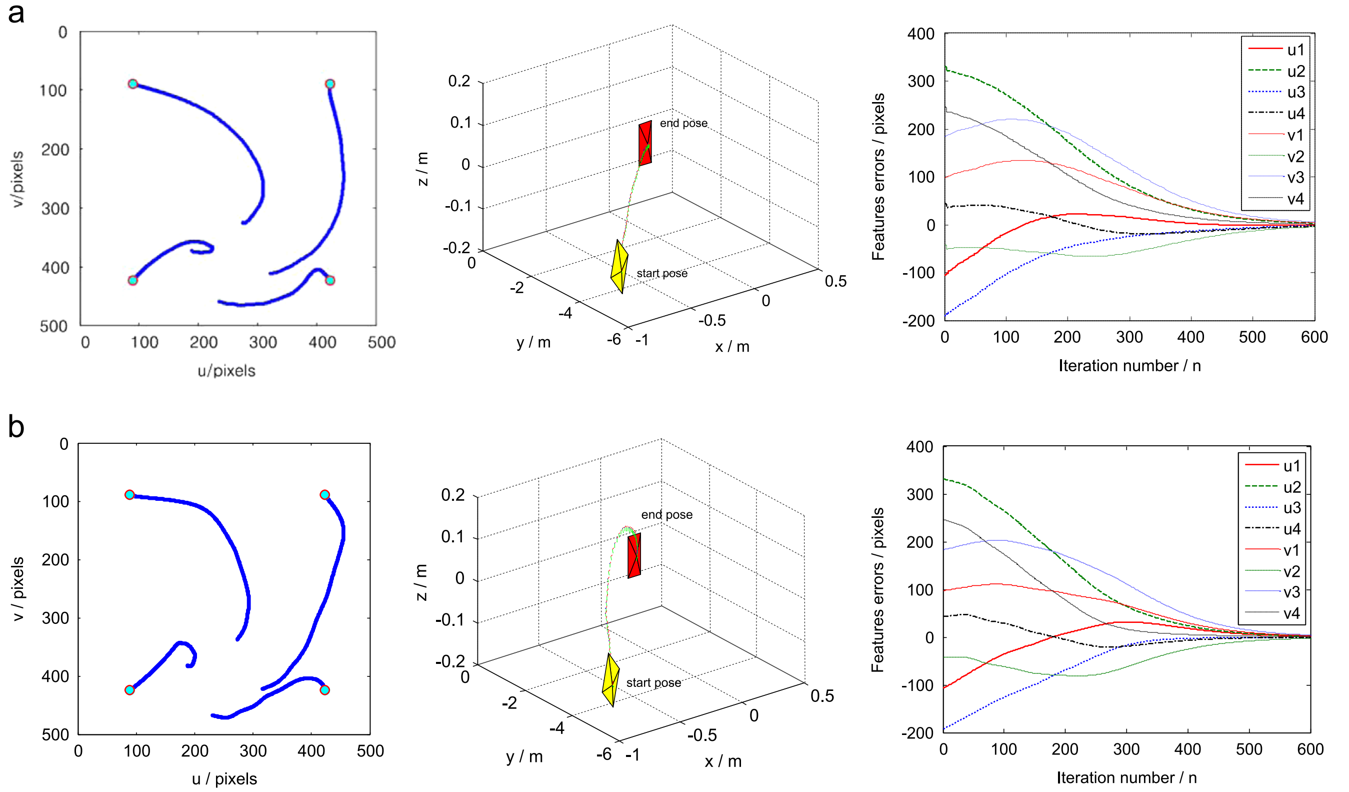


Fig. 5. Simulation results of sub-case 1 in case 2. 1st and 2nd, rows are the results obtained by (a) our IBVS based on NNAKF and (b) KF methods, respectively. 1st, 2nd and 3rd columns corresponding to (1) image feature trajectories, (2) camera moving trajectories, and (3) image feature errors, respectively.

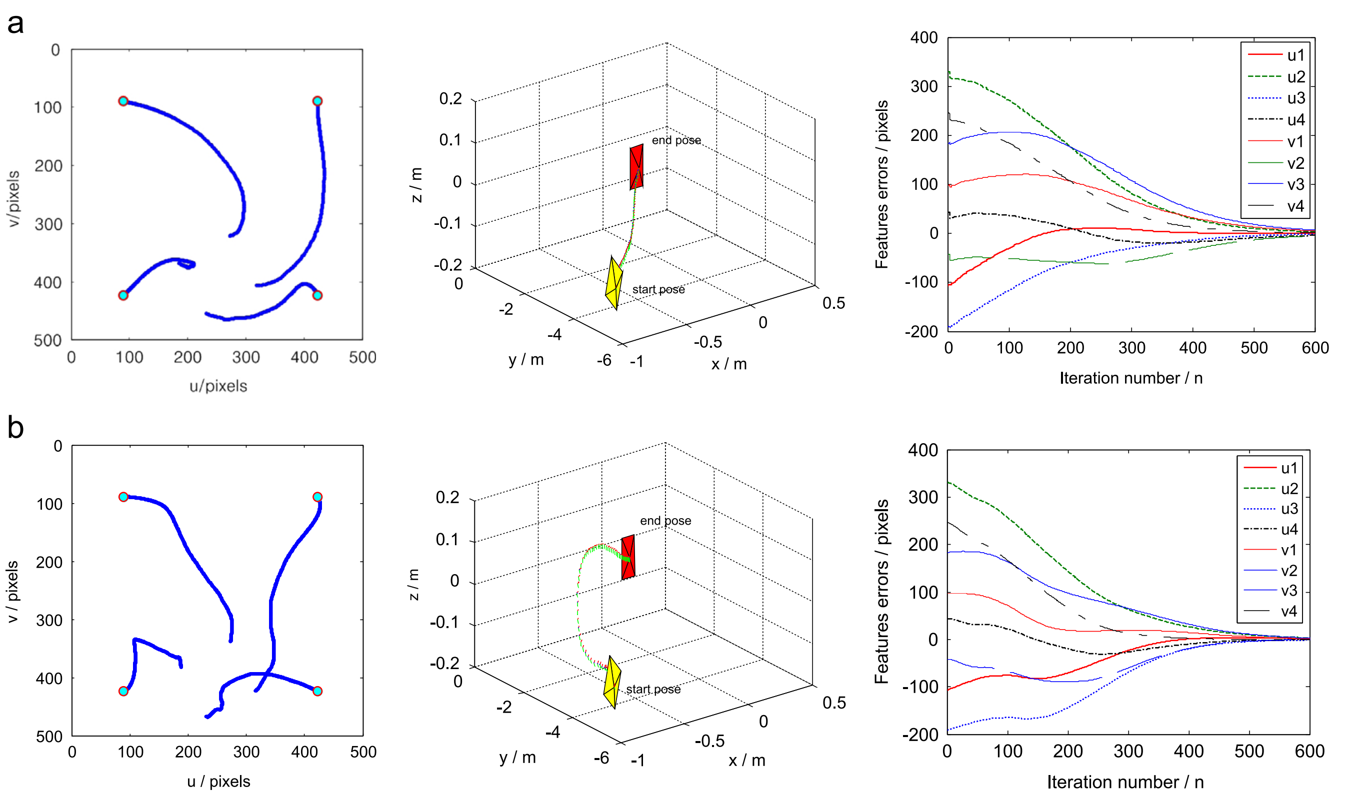


Fig. 6. Simulation results of sub-case 2 in case 2. 1st and 2nd, rows are the results obtained by (a) our IBVS based on NNAKF and (b) KF methods, respectively. 1st, 2nd and 3rd columns corresponding to (1) image feature trajectories, (2) camera moving trajectories, and (3) image feature errors, respectively.

moving. It is clear that our method with robust stability performances no matter the destabilized system attached with odious dynamic noises.

Furthermore, many more serious values of noise statistic characteristics have been considered for different simulations, the results tell us that when the covariance of noises is changing

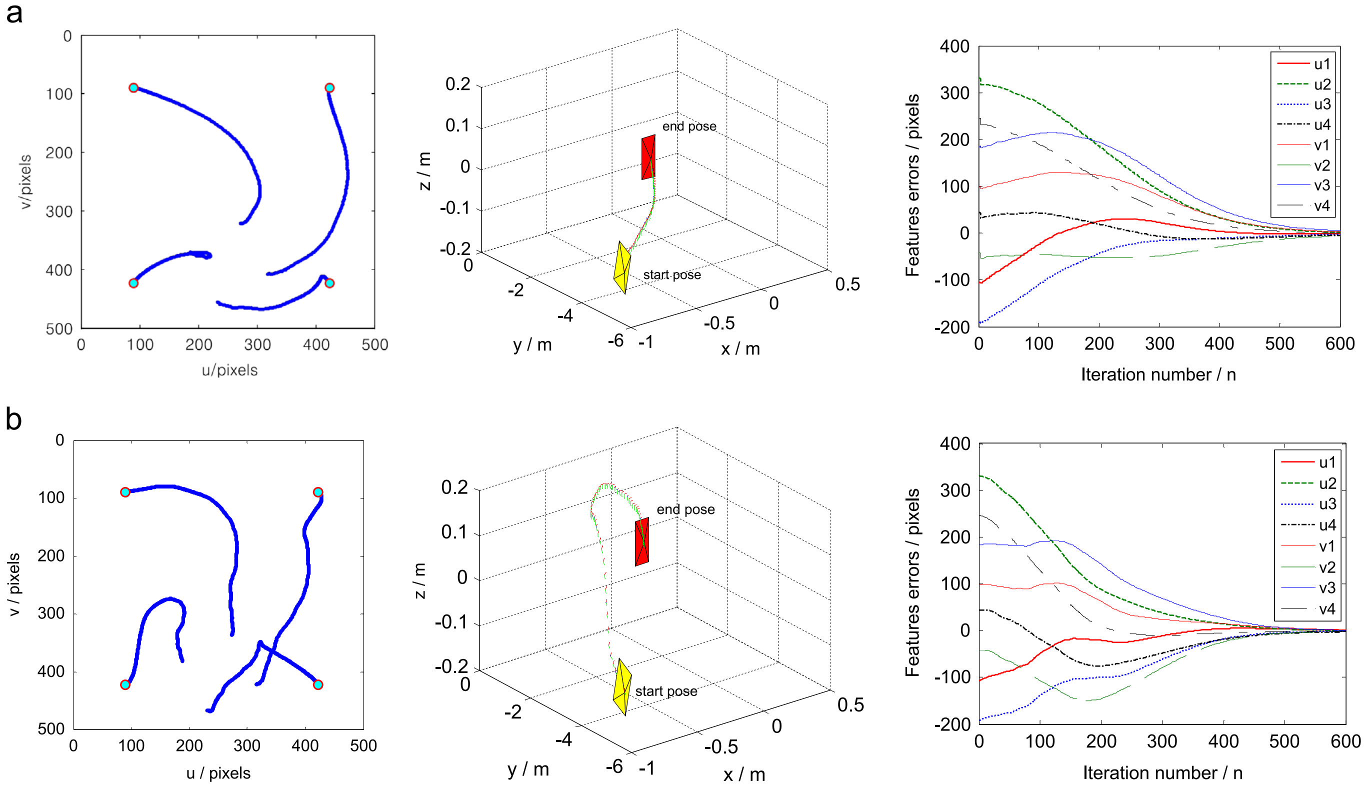


Fig. 7. Simulation results of sub-case 3 in case 2. 1st and 2nd, rows are the results obtained by (a) our IBVS based on NNAKF and (b) KF methods, respectively. 1st, 2nd and 3rd columns corresponding to (1) image feature trajectories, (2) camera moving trajectories, and (3) image feature errors, respectively.

slightly, the results of the KF method will alter largely even lose its positioning ability, since the camera with random moving in Cartesian space, and the image features in image plan are unknown. When the NN was selected as an assistant for KF, our proposed IBVS method always works well, although the system instability caused by dynamic noise with changing in larger region. This means that the proposed NNAKF can guarantee the robust stability for our IBVS approach.

Finally, we give some experimental results with 3 cases to illustrate the performances of our IBVS based on NNAKF. In the experimental case 1, the camera with rotational movement around *Z*-axis, and in the experimental case 2, the camera with pure linear movement from initial pose toward desired pose, while in the experimental case 3, the camera with a combination of the translational and rotational movement, note that, comparing the simulation, the camera with lower values of the rotation angle in this two experimental cases, due to the limitation of the joint angle and the robot's workspace. The experimental results are shown in Fig. 8, among of them, the performances of the feature trajectories are constrained on image plane, and the image error are minimized with the initial features converging toward the desired features, the steady-state errors in the image space are about 5 pixels for all three cases. In addition, the experimental results of the camera trajectory in the Cartesian space shown that all the servo tasks for 3 cases are completed with robust stability without camera retreat, along with feature points kept within the FOV. In order further to test the robust performances of the proposed IBVS, the indoor light condition have been changed to create a artificial environment to simulate different noise models, in this situation we have only consider a difficult servo task same as experimental case 3 the camera converging toward desired pose with a combination of translational and rotational movement, the experimental results are shown in Fig. 9, which very similar to those results in Figs. 8(c-1) and (c-2). It is means that our IBVS method indeed with robust global stability for image processing noises.

Note that, the simulation results shown that the KF and PBVS methods easy leading to the image features leave the FOV, so in actual experimental, those two approaches will fail to control the robot's end-effector reaching the desired pose, due to the controller lacks the visual feedback information. On the other hand, the PBVS method needs the camera-robot calibration parameters, so this approach unfit our model-free robotics configuration system, since our approach is direct control of the image features for robot converging toward the desired pose, the feature errors converging toward zero that meaning is actually equivalent to robot achieving the task manipulation. Furthermore, for the traditional servo control method, there is sensitive to the camera calibration error. In some works, the influences of the calibration parameters on the system is investigated [36], and as a solution to this problem using intelligent hybrid control laws, by introducing neural network reinforcement learning [37]. In our method, the image-Jcobian is online estimation without requiring any calibration parameters of camera-robot and the depth information of the target, therefore our IBVS method is independent from the robot's kinematic model and camera calibration, which avoids the corrupted performances caused by calibration and modeling errors.

7. Conclusions

In this work, a new IBVS based on NNAKF scheme have been discussed for model-free robots manipulation. The proposed NNAKF state estimator could guarantee the visual servo system with robust stability, and constrain the image features within FOV of the camera. In addition, the image Jacobian on-line identification problems were solved by incorporating state estimation techniques, which without requiring the camera-robot calibration parameters. Therefore, the

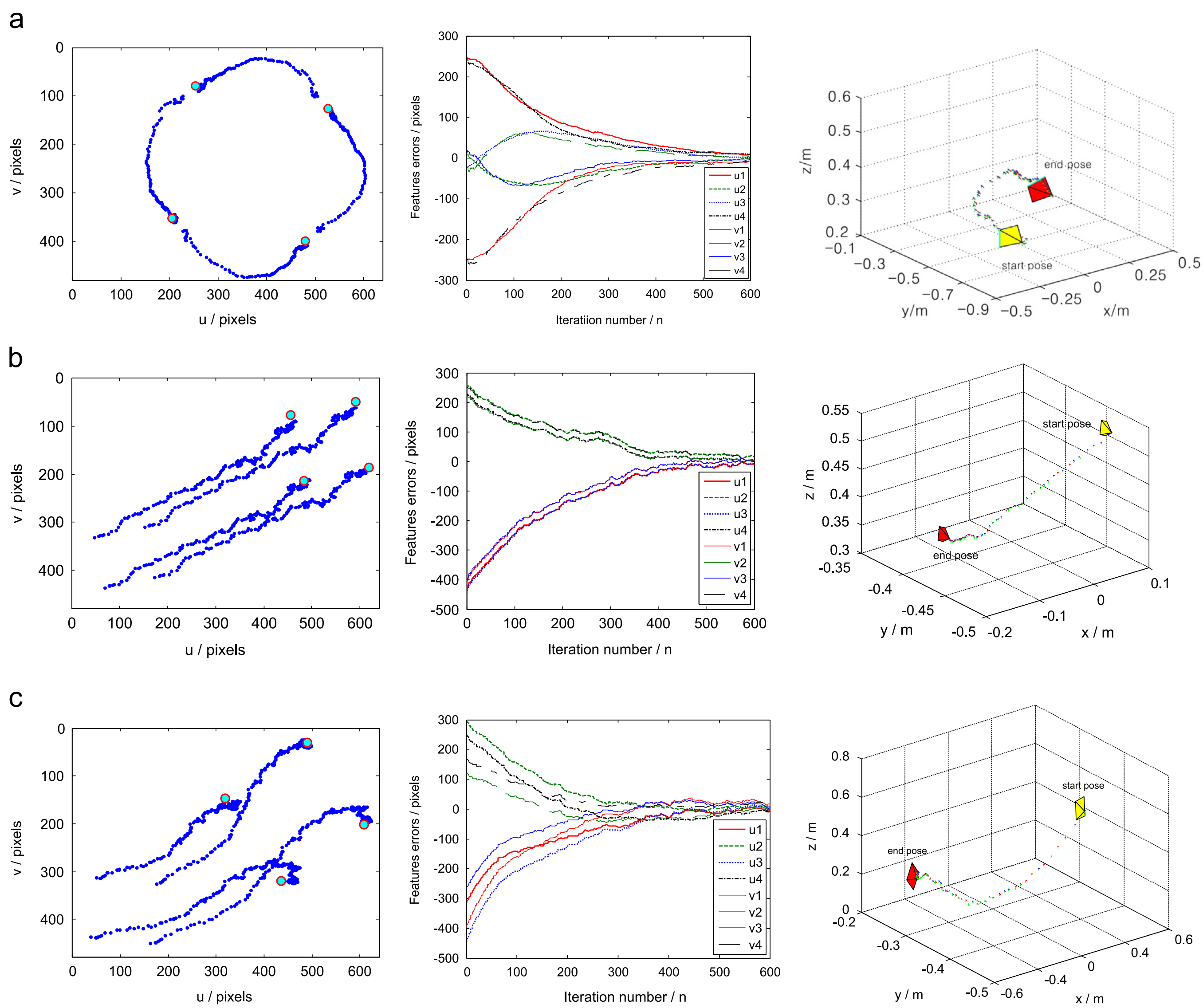


Fig. 8. Experimental results obtained by our IBVS method. The 1st, 2nd and 3rd rows corresponding to (a) case 1, the camera with rotational movement, (b) case 2, the camera with linear movement, and (c) case 3, the camera with a combination of translational and rotational movement. The 1st, 2nd and 3rd columns corresponding to (1) image feature trajectories, (2) image feature errors, and (3) camera moving trajectories, respectively.

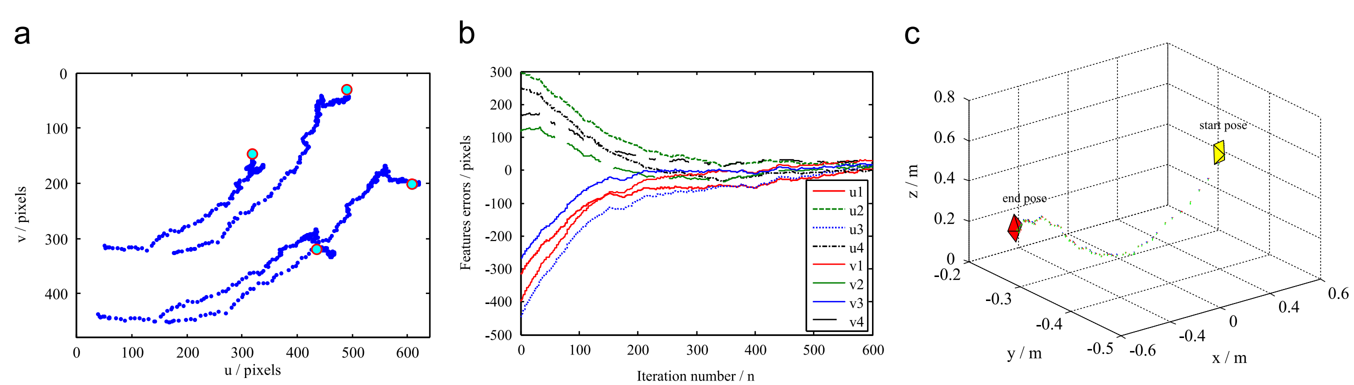


Fig. 9. Experimental results obtained by our IBVS method of case 3, the camera with a combination of translational and rotational movement with different noises by change indoor light condition. (a) The image feature trajectories, (b) the image feature errors, and (c) the camera moving trajectories.

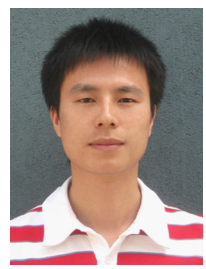
proposed IBVS method has avoids further the corrupted performances caused by calibration error. Through various simulation and experimental results by PBVS, KF and our IBVS methods, shown that, the proposed visual servo control approach takes the perfect advantages in both trajectories of the image features on image plane and camera movement in the Cartesian space.

Acknowledgments

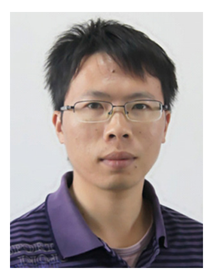
This work is supported by National Natural Science Foundation of China under Grant 61305117.

References

[1] L.E. Weis, A.C. Sanderson, C.P. Neuman, Dynamic sensor-based control of robots with visual feedback control, IEEE J. Robot. Autom. 5 (3) (1987) 404–417.
[2] Y. Shirai, H. Inoue, Guiding a robot by visual feedback in assembling tasks, Pattern Recognit. 5 (2) (1973) 99–108.
[3] F. Chaumette, S. Hutchinson, Visual servo control. part I: basic approaches, IEEE Robot. Autom. Mag. 4 (13) (2006) 82–90.
[4] B. Thuilot, P. Martinet, L. Cordesses, J. Gallice, Position based visual servoing: keeping the object in the field of vision, in: Proceedings of the 2002 IEEE International Conference on Robotics and Automation, 2002, pp. 1624–1629.
[5] W. Wilson, C. Hulls, G. Bell, Relative end-effector control using Cartesian position-based visual servoing, IEEE J. Robot. Autom. 5 (12) (1996) 684–696.
[6] P. Jorge, G. Pablo, T. Fernando, Visual control of robots using range images, Sensors 10 (2010) 7303–7322.
[7] X.G. Zhong, X.Y. Zhong, X.F. Peng, Robust Kalman filtering cooperated Elman neural network learning for vision-sensing-based robotic manipulation with global stability, Sensors 13 (2013) 13464–13486.
[8] N. Garcia-Aracil, C. Perez-Vidal, et al., Robust and cooperative image-based visual servoing system using a redundant architecture, Sensors 11 (2011) 11885–11900.
[9] E. Malis, F. Chaumette, S. Boudet, 2 1/2-D visual servoing, IEEE J. Robot. Autom. 2 (15) (1999) 238–250.
[10] L. Deng, F. Janabi-Sharifi, W.J. Wilson, Hybrid motion control and planning strategies for visual servoing, IEEE Trans. Ind. Electron. 4 (52) (2005) 1024–1040.
[11] M. Baumann, S. Leonard, et al., Path planning for improved visibility using a probabilistic road map, IEEE Trans. Robot. 1 (26) (2010) 195–200.
[12] F. Chaumette, Potential problems of stability and convergence in image-based and position based visual servoing, Conflu. Vis. Control Lect. Notes Control Inf. Sci. 237 (1998) 66–78.
[13] N. Garcia-Aracil, E. Malis, R. Aracil-Santonja, C. Perez-Vidal, Continuous visual servoing despite the changes of visibility in image features, IEEE Trans. Robot. 6 (21) (2005) 1214–1220.
[14] G. Chesi, K. Hashimoto, D. Prattichizzo, A. Vicino, Keeping features in the field of view in eye-in-hand visual servoing: a switching approach, IEEE Trans. Robot. 5 (20) (2004) 908–913.
[15] A. Comport, M. Pressigou, E. Marchand, F. Chaumette, A visual servoing control law that is robust to image outliers, IEEE Int. Conf. Intell. Robots Syst. (2003) 492–497.
[16] M. Jagersand, Visual servoing using trust region methods and estimation of the full coupled visual-motor Jacobian, in: Proceedings of IASTED Applications of Robotics and Control, 1996, pp. 105–108.
[17] J.A. Piepmeier, H. Lipkin, Uncalibrated eye-in-hand visual servoing, Int. J. Robot. Res. 22 (10–11) (2003) 805–819.
[18] K. Hosoda, M. Asada, Versatile visual servoing without knowledge of true Jacobian, in: Proceedings of the IEEE/RSJ/GI Conference Intelligent Robots and Systems, 1994, pp. 186–193.
[19] S. Azad, F. Amir-massoud, J. Martin, Robust Jacobian Estimation for Uncalibrated Visual Servoing, IEEE Int. Conf. Robot. Autom. (2010) 5564–5569.
[20] D.I. Kosmopoulos, Robust Jacobian matrix estimation for image-based visual servoing, Robot Comput.-Integr. Manuf. 27 (2011) 82–87.
[21] J. Qian, J. Su, Online estimation of image Jacobian matrix by Kalman-Bucy filter for uncalibrated stereo vision feedback, in: Proceedings of the 2002 IEEE International Conference on Robotics and Automation, 2002, pp. 562–567.
[22] X. Lv, X. Huang, Fuzzy adaptive Kalman filtering based estimation of image Jacobian for uncalibrated visual servoing, in: Proceedings of the IEEE/RSJ/GI Conference Intelligent Robots and Systems, 2006, pp. 2167–2172.
[23] F. Janabi-Sharifi, M. Marey, A. Kalman-Filter-Based, Method for pose estimation in visual servoing, IEEE Trans. Robot. 5 (26) (2010) 939–947.
[24] P. Wira, J.P. Urban, A new adaptive Kalman filter applied to visual servoing tasks, Proc. IEEE Conf. Knowl.-Based Intell. Eng. Syst. Allied Technol. 1 (2000) 267–270.
[25] R.E. Kalman, A new approach to linear filtering and prediction problems, Trans. ASME – J. Basic Eng. 82 (1960) 34–45.
[26] K.L. Lo, Y. Rathamarit, State estimation of a boiler model using the unscented Kalman filter, IET Generation, Transm. Distrib. 2 (6) (2008) 917–931.
[27] S. Thomas, H. Wilfried, W. Ralf, Improving operational performance of active magnetic bearings using Kalman filter and state feedback control, IEEE Trans. Ind. Electron. 59 (2) (2012) 821–829.
[28] J.D. Gibson, B. Koo, S.D. Gray, Filtering of colored noise for speech enhancement and coding, IEEE Trans. Signal Process. 8 (39) (1991) 1732–1742.
[29] T. Soni, B.D. Rao, J.R. Zeidler, W.H. Ku, Enhancement of images using the 2-D LMS adaptive algorithm, in: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing vol. 4, 1991, pp. 3029–3032.
[30] I.M. Stephanakis, S.D. Kollias, Wavelet-based approach to adaptive Wiener filtering of images in the presence of colored noise, Opt. Eng. 5 (39) (2000) 1353–1363.
[31] S.S. Xiong, Z.Y. Zhou, Neural filtering of colored noise based on Kalman filter structure, IEEE Trans. Instrum. Meas. 3 (52) (2003) 742–747.
[32] K. Takaba, Y. Iiguni, H. Tokumaru, An improved tracking Kalman filter using a multilayered neural network, Math. Comput. Model. 23 (1–2) (1996) 119–128.
[33] V. Vaidehi, N. Chitra, M. Chokkalingam, C.N. Krishnan, Neural network aided Kalman filtering for multitarget tracking applications, Comput. Electr. Eng. 27 (2001) 217–228.
[34] R.H. Zhan, J.W. Wan, Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation, IEEE Signal Process. Lett. 7 (13) (2006) 445–448.
[35] A.Y. Alanis, E.N. Sanchez, A.G. Loukianov, Real-time discrete back stepping neural control for induction motors, IEEE Trans. Control Syst. Technol. 2 (19) (2011) 359–366.
[36] Z. Ceren, E. Altug, Image based and hybrid visual servo control of an unmanned aerial vehicle, J. Intell. Robot. Syst. 65 (1–4) (2012) 325–344.
[37] M. Zoran, M. Marko, L. Mihailo, B. Bojan, Neural network reinforcement learning for visual control of robot manipulators, Expert Syst. Appl. 40 (2013) 1721–1736.



Xungao Zhong received the B.E. degree in electronic information engineering from Nanchang University, Nanchang, China, in 2007, the M.S. degree in electromechanical engineering from Guangdong University of Technology, Guangzhou, China, in 2011. He is currently working toward the Ph.D. degree with the Department of Automation, Xiamen University. His current research interests include machine vision and robotics visual servoing control.



Xunyu Zhong received the M.E. degree in mechatronics engineering from Harbin Engineering University, Harbin, Heilongjiang, China, in 2007 and the Ph.D. degree in control theory and control engineering from Harbin Engineering University, in 2009. He is currently an Assistant Professor with the Department of Automation, Xiamen University at Xiamen, China. His current research interests include robot motion planning, mobile and autonomous robotics.

Xiafu Peng received the M.S. and Ph.D. degrees in control science from the Harbin Engineering University, in 1994 and 2001, respectively. He is currently a Professor with the Department of Automation, Xiamen University at Xiamen. His current research interests include the navigation and motion control of robots.

Prof. Peng is a Fellow of the Fujian Association for the advancement of Automation and Power, and a Senior Member of the Chinese Institute of Electronics. He is the recipient of the provincial/ministerial Scientific and Technological Progress Award.