Adaptive Neuro-Filtering Based Visual Servo Control of a Robotic Manipulator

XUNGAO ZHONG[1], (Member, IEEE), **XUNYU ZHONG**[2],
HUOSHENG HU[3], (Senior Member, IEEE), AND **XIAFU PENG**[2]

[1]School of Electrical Engineering and Automation, Xiamen University of Technology, Xiamen 361024, China
[2]School of Aerospace Engineering, Xiamen University, Xiamen 361005, China
[3]School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K.

Corresponding author: Xunyu Zhong (zhongxunyu@xmu.edu.cn)

ABSTRACT This paper focuses on the solutions to flexibly regulate robotic by vision. A new visual servoing technique based on the Kalman filtering (KF) combined neural network (NN) is developed, which need not have any calibration parameters of robotic system. The statistic knowledge of the system noise and observation noise are first given by Gaussian white noise sequences, the nonlinear mapping between robotic vision and motor spaces are then on-line identified using standard Kalman recursive equations. In real robotic workshops, the perfect statistic knowledge of the noise is not easy to be derived, thus an adaptive neuro-filtering approach based on KF is also studied for mapping on-line estimation in this paper. The Kalman recursive equations are improved by a feedforward NN, in which the neural estimator dynamic adjusts its weights to minimize estimation error of robotic vision-motor mapping, without the knowledge of noise variances. Finally, the proposed visual servoing based on adaptive neuro-filtering has been successfully implemented in robotic pose regulation, and the experimental results demonstrate its validity and practicality for a six-degree-of-freedom (DOF) robotic system which the hand-eye without calibrated.

INDEX TERMS Robotics regulation, visual servo control, mapping estimation, adaptive filtering, neural network.

I. INTRODUCTION

Robotic manipulators are increasingly deployed in unstructured environments to provide services to people in needs. This new generation of robotic manipulators should reliably interact with people and explore their environments instead of traditionally following a predefined path in industry, which is a very challenging task. Also they should adapt their motion in real time to parameter changes and external disturbance. Therefore, traditional robotic controllers that are based on kinematics or dynamics of the robotic system cannot satisfy the requirement of new challenges [1], [2].

Visual servoing is one promising solution to regulate robotic manipulators to physically interact with their environments through vision features. The successful

The associate editor coordinating the review of this manuscript and approving it for publication was Long Wang.

implementation of visual servoing has been shown in several works [3]–[6]. Both position-based visual servoing (PBVS) and image-based visual servoing (IBVS) methods are popular, with one or two eye-in-hand or eye-to-hand configurations [7], [8]. The PBVS is based on the computation of 3D Cartesian errors from an end-effector to the object. This solution requires a perfect modeling of the object and calibration of eye-in-hand robotic systems. Therefore, the computation is conducted at each control cycle, and the 3D Cartesian errors may lead to a strong sensitivity with respect to noise perturbations [9].

In IBVS, the feature points on the image plane are used to regulate the robotic pose, and a mapping is conducted to describe the differential relationship between visual features and end-effector velocities in the Cartesian space [10]. IBVS has no need for 3D target restructuring and its computation is simpler than PBVS. Therefore, it has attracted much attention

recently [11]–[14]. But the most pressing issue of IBVS is how to accurately calculate a mapping matrix, *i.e.* a local and linear approximation to the nonlinear and highly coupled interaction between vision and motor spaces.

In order to provide real time adaptation to the parameters change of robotic manipulator, the visual serving controller should be endowed with a variable mapping estimator for motion generation. Some existing works actually consider the mapping estimation issue as a dynamic identification problem, and the solution includes on-line techniques, such as Broyden-based method [15], decoupling static Broyden method [16], dynamic Broyden and exponential weighted recursive least squares (RLS) [17], Gauss-Newton and recursive RLS Jacobian estimation methods [18].

Hao and Sun proposed universal state-space approach to uncalibrated visual serving [19], in which Broyden-Gauss-Newton and Broyden recursive RLS are unified into the state space. Qian and Su conducted online estimation of image Jacobian matrix by using KF for uncalibrated stereo vision feedback [20]. KF has reliability and high estimation accuracy due to the optimality of its gain that depends on the knowledge of noise statistics [21]. However, their method is unsuitable to some stochastic environments where the knowledge of noise statistics are unknown, especially for the situation that has serious change of observation and state models [22], [23]. In view of those problems, some works proposed the adaptive KF approaches using learning technology of neural network [24]–[27].

Some solutions to deal with the dynamic noises have been proposed, including the dimension extension of KF [28], [29], least mean square (LMS) based adaptive filtering [30], wavelet-based adaptive Wiener filtering [31]. Lv and Huang investigated the application of KF in the state space model with variable noise parameters [32]. An iterative adaptive extended Kalman filtering (EKF) was proposed by integrating mechanisms for noise adaptation and iterative-measurement linearization in robotic pose estimation tasks [33]. Wira and Urban adjusted the transition matrix of the KF to address the problem of the unavailable system model, and then its performance was evaluated in visual serving [34].

It is clear that the robot manipulators that work in a stochastic noisy environment without calibration should be endowed with the ability to adapt their motion in real time. They also should have self-adaption capabilities to operate robustly beyond a set of settled calibrated or modeling parameters. Therefore, this paper presents a new visual serving framework which is based on KF. It is a robust and adaptive neuro-filtering approach that allows a robotic manipulator to safely operate in uncalibrated environments. Our research conducted in this paper made the following three contributions.

- The NN have been deployed for on-line minimization of the error-estimation variance in order to obtain a robust state estimation. As a result, the neural estimator has the capability of improve the robustness of KF.

- An adaptive neuro-filtering approach is proposed, and has been successfully applied in on-line identification and mapping between robotic vision and motor spaces, without the need for the knowledge of noise statistics.

- We also have design a un-calibration visual servoing framework by employing neuro-filtering. It does not require hand-eye calibration and camera's parameters. In our finding, the mapping matrix is dynamically estimated, which is not affected by the errors of hand-eye calibration. In additional, it differs from the traditional IBVS and KF methods, with the merits of robust stability under the dynamic noises.

The rest of the paper is organized as follows. Section II outlines the preliminaries of the robotic visual servoing without parameters. Section III presents a neuro-filtering schema based on KF technique for robotic mapping estimation. Then a new visual servoing framework with neuro-filtering is proposed in Section IV. The simulation and experimental results are presented in Section V to show the feasibility and performance of the proposed approach. Finally, a brief conclusion is given in Section VI.

II. PRELIMINARIES

In this study, we design a robotic visual serving framework without parameters, which includes an adaptive estimator for the real-time mapping, and a control law for robotic pose regulation. The problem statement is presented as follows and the main notations used in this paper are summarized in Table 1.

A. ROBOTICS VISUAL SERVING WITHOUT PARAMETERS

Visual serving techniques consist of using the image data provided by camera in order to regulate the pose of a robotic system [7]. The camera can be mounted on the robotic end-defector (eye-in-hand), or other alternatively configuration, the goal of visual serving is to drive the robotic end-effector from the current pose to the desired pose through the minimization of an error in the vision space.

In this paper, it is assumed that an eye-in-hand robotic system without any calibrated parameters, and an image error function is defined by:

$$\mathbf{e}_s(k) = \mathbf{S}(k) - \mathbf{S}^d \qquad (1)$$

where $\mathbf{S}(k) \in \Re^n$ is current image feature, and $\mathbf{S}^d \in \Re^n$ is desired image feature, $\mathbf{S}(k)$ and $\mathbf{S}^d$ are obtained by camera which fixed on the end-effector of robotic.

The camera projection model is illustrated in Fig. 1, where $C\{O^c\text{-}X^cY^cZ^c\}$ is camera coordinate system, and $I\{O^I\text{-}UV\}$ is imaging plane. The target's position in C is $P = (x^c, y^c, z^c)$, and the camera coordinate origin is O^c. The image plane I is intersected at point S.

TABLE 1. Main notation definitions.

| Symbol	Definition
$\mathbf{S}(k)$	Image feature at k
$\mathbf{S}^d$	Desired image feature
$\mathbf{e}_s(k)$	Image error at k
P	Target position
S	Target imaging point
η	Projection matrix
Π	Camera intrinsic parameters
$\mathbf{U}(k)$	Robot control variable at k
$\Delta\mathbf{S}(k)$	Variation of image feature at k
$\mathbf{J}_s(k)$	Mapping matrix at k
$\mathbf{X}(k)$	State vector at k
$\xi(k)$	System noise at k
φ	State transformation matrix
Γ	Noise drive matrix
Q	Variance of system noise
$\mathbf{Z}(k)$	Observation vector at k
$\upsilon(k)$	Observation noise at k
R	Variance of observation noise
$\mathbf{H}(k)$	Observation matrix at k
$\mu(k)$	Cost function at k
$e(k)$	Estimation error at k
$\mathbf{W}_{ij}^{(l)}(k)$	Weight between the i-th node in the l-th layer and the j-th node in the $(l-1)$-th layer at k
$\mathbf{O}(k)$	Output of the NN at k
$\mathbf{O}'(k)$	Output of the hidden neuron at k
λ	Control rate
$\mathbf{J}_s^+(k)$	Inverse mapping matrix at k
$L(k)$	Lyapunov function at k
$V(k)$	Linear velocity at k
$W(k)$	Angular velocity at k
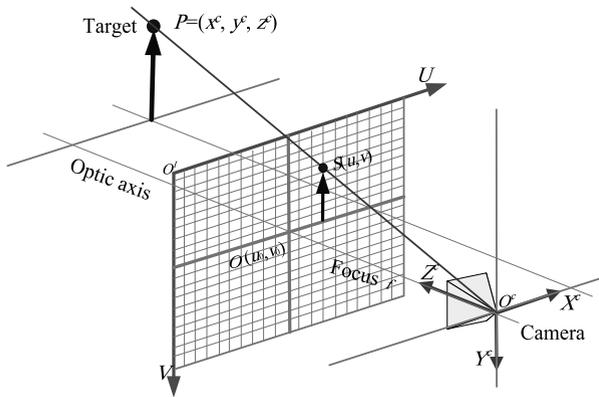
FIGURE 1. The perspective projection of camera.

Considering the perspective projection and epipolar geometry, the target imaging point $S = (uv)$, given by:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z^c} \eta \begin{pmatrix} x^c \\ y^c \\ z^c \\ 1 \end{pmatrix} \qquad (2)$$

where $\eta = \left(\Pi \mid_{3\times3} \ 0 \mid_{3\times1} \right) \in \Re^{3\times4}$ is the projection matrix, $\Pi \in \Re^{3\times3}$ is the camera's intrinsic parameters, and z^c is the depth information of target.

Assume $\mathbf{U}(k) \in \Re^m$ is the control variable of the robotic, and $\Delta\mathbf{S}(k) = \mathbf{S}(k) - \mathbf{S}(k-1)$ is the time variation of image feature, the relationship between robotic control variable $\mathbf{U}(k)$ and the time variation of image feature $\Delta\mathbf{S}(k)$ can be described, as follows [35]:

$$\Delta\mathbf{S}(k) = \mathbf{J}_s(k)\mathbf{U}(k) \qquad (3)$$

where $\mathbf{J}_s(k) \in \Re^{n\times m}$ is a nonlinear mapping matrix, which is depended on the robotic system calibration parameters and the depth information of target [7]. Hence, the mapping matrix is unknown as the robotic work in an un-calibrated environment.

B. SYSTEM STATE AND MEASUREMENT EQUATIONS

One of essential issue of un-calibrated visual servoing is precisely calculating the mapping matrix $\mathbf{J}_s(k)$, in this paper, we formulate it as a state estimation problem, and consider the non-explicit mapping below:

$$\begin{aligned} \mathbf{J}_s(k) &= \frac{\partial \mathbf{S}(k)}{\partial \mathbf{U}(k)} \\ &= \begin{pmatrix} j_{11}(k) & j_{12}(k) & \cdots & j_{1m}(k) \\ \vdots & \vdots & \cdots & \vdots \\ j_{n1}(k) & j_{n2}(k) & \cdots & j_{nm}(k) \end{pmatrix} \in \Re^{n\times m} \quad (4) \end{aligned}$$

Let variable $\mathbf{c}(k)$ be formed by concatenations of the row and the column elements of $\mathbf{J}_s(k)$, *i.e.*

$$\mathbf{c}(k) = \begin{pmatrix} j_{11}(k) & j_{12}(k) & \cdots & j_{nm}(k) \end{pmatrix} \in \Re^{nm\times1} \quad (5)$$

Considering the following system dynamic equation:

$$\mathbf{c}(k+1) = \mathbf{c}(k) + \Delta t \dot{\mathbf{c}}(k) + \frac{1}{2}(\Delta t)^2 \xi(k) \qquad (6)$$

where $\dot{\mathbf{c}}(k) = \mathbf{c}(k+1) - \mathbf{c}(k)$, $\xi(k)$ is system noise and Δt is sampling interval.

The system state equation can be established by choosing the state variables $\mathbf{x}_1(k) \in \Re^{nm\times1}$, $\mathbf{x}_2(k) \in \Re^{nm\times1}$ as follows:

$$\begin{cases} \mathbf{x}_1(k) = \mathbf{c}(k) \\ \mathbf{x}_2(k) = \dot{\mathbf{c}}(k) \end{cases} \qquad (7)$$

Let the state vector be:

$$\mathbf{X}(k) = \begin{pmatrix} \mathbf{x}_1(k) \\ \mathbf{x}_2(k) \end{pmatrix} \in \Re^{2nm\times1} \qquad (8)$$

According to Eq. (6), the system state equation can be described as:

$$\begin{aligned} \mathbf{X}(k+1) &= \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1(k) \\ \mathbf{x}_2(k) \end{pmatrix} + \begin{pmatrix} \frac{1}{2}(\Delta t)^2 \\ \Delta t \end{pmatrix} \xi(k) \\ &= \varphi \mathbf{X}(k) + \Gamma \xi(k) \end{aligned} \qquad (9)$$

where φ and Γ are state transformation matrix and noise drive matrix, respectively. Let $\xi(k)$ be Gaussian white noise sequences with zero mean, and the variance is Q.

In this paper, the observation vector $\mathbf{Z}(k) \in \Re^{n\times1}$ is given by:

$$\mathbf{Z}(k) = \Delta\mathbf{S}(k)$$

$$= \mathbf{S}(k) - \mathbf{S}(k-1)$$
$$= \mathbf{J}_s(k)\mathbf{U}(k) \tag{10}$$

And let the linear system observation equation is:

$$\mathbf{Z}(k) = \mathbf{H}(k)\mathbf{X}(k) + \boldsymbol{v}(k) \tag{11}$$

where $\boldsymbol{v}(k)$ is observation white noise sequences with zero mean, and the variance is $\boldsymbol{R}$, $\mathbf{H}(k)$ is dynamic observation matrix, as follows:

$$\mathbf{H}(k) = \begin{pmatrix} \mathbf{U}(k) & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \mathbf{U}(k) \end{pmatrix} \in \Re^{n \times 2nm} \tag{12}$$

The system state Eq. (9) and the observation Eq. (11) are suited to the standard KF. So the Kalman recursive equations can be derived for mapping on-line estimation, the algorithm flow is as follows [21]:

Algorithm 1 KF for Mapping Estimation

Initial: $Q \in \Re^{2nm \times 2nm}$, $R \in \Re^{n \times n}$, $P(0) \in \Re^{2nm \times 2nm}$
$\mathbf{X}(0) \in \Re^{2nm \times 1}$
for $k = 1: q$
$\mathbf{X}(k-1) \leftarrow \mathbf{J}_s(k-1)$
$\hat{\mathbf{X}}(k/k-1) = \boldsymbol{\varphi}\hat{X}(k-1/k-1)$
$\mathbf{P}(k/k-1) = \boldsymbol{\varphi}\boldsymbol{P}(k-1)\boldsymbol{\varphi}(k/k-1)^{\mathbf{T}} + Q$
$\mathbf{K}(k) = \mathbf{P}(k/k-1)\mathbf{H}(k)^{\mathbf{T}} \left(\mathbf{H}(k)\mathbf{P}(k/k-1)\mathbf{H}(k)^{\mathbf{T}} + R\right)^{-1}$
$\hat{\mathbf{X}}(k/k) = \hat{\mathbf{X}}(k/k-1) + \mathbf{K}(k)\left(\mathbf{Z}(k) - \mathbf{H}(k)\hat{\mathbf{X}}(k/k-1)\right)$
$\mathbf{P}(k) = (\mathbf{E} - \mathbf{K}(k)\mathbf{H}(k))\,\mathbf{P}(k/k-1)\,(\mathbf{E} - \mathbf{K}(k)\mathbf{H}(k))^{\mathbf{T}}$
$\qquad + \mathbf{K}(k)R\mathbf{K}(k)^{\mathbf{T}}$
$\mathbf{J}_s(k) \leftarrow \hat{\mathbf{X}}(k/k)$
End for
End

In a real environment, the system dynamic noise $\boldsymbol{\xi}(k)$ and the sensor observation noise $\boldsymbol{v}(k)$ are generally difficult to be derived, *i.e.* their statistic parameters Q and R are unknown. Thus in the next section, a neural estimator based on the KF is chosen as an alternative, where we give an adaptive neuro-filtering schema in the light of unknown statistic knowledge of the noises.

III. ADAPTIVE NEURO-FILTERING FOR MAPPING ON-LINE ESTIMATION

Considering the perturbation on both observation and system state models, the linear time invariant system intrinsic contains the nonlinear approximation errors. On the other hand, the filtering gain of the KF equations is the function of noises statistics. However, it is difficult to obtain the statistic knowledge of noises in an actual environment. Hence, on-line minimization of the error-estimation variance should be conducted to obtain a robust Kalman state estimation.

Fig. 2 shows the schema of adaptive neuro-filtering, where the Kalman gain is improved by a feedforward network. The NN is used to minimize the estimation error $\mathbf{e}(k) \in \Re^{n \times 1}$

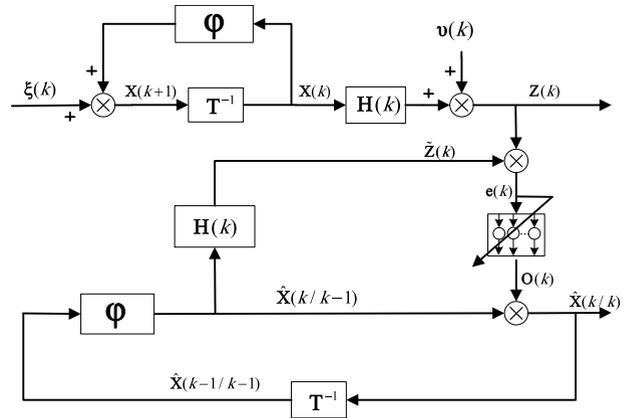**FIGURE 2.** The schema of adaptive neuro-filtering.

between observation $\mathbf{Z}(t) \in \Re^{n \times 1}$ and estimation $\hat{\mathbf{Z}}(t) \in \Re^{n \times 1}$. In real environment, the NN's output vector of desired values is not known. Thus, the on-line learning will be applied in this paper, the goal of learning algorithm is to minimize the cost function in each iteration step, and no need obtaining the convergence of the NN weights. The cost function becomes:

$$\boldsymbol{\mu}(k) = \frac{1}{2}\mathbf{e}^T(k)\mathbf{e}(k) \tag{13}$$

where the estimation error is:

$$\mathbf{e}(k) = \mathbf{Z}(k) - \hat{\mathbf{Z}}(k) \tag{14}$$

The learning law of the network is given by:

$$\mathbf{W}_{ij}^{(l)}(k+1) = \mathbf{W}_{ij}^{(l)}(k) + \gamma\mathbf{e}^T(k)\frac{\partial\hat{\mathbf{Z}}(k)}{\partial\mathbf{W}_{ij}^{(l)}} \tag{15}$$

where k is the time instant, l represents the layer number of the NN, *i.e.* $l = 1$ is the hidden layer and $l = 2$ is the output layer. $i = 1,\ldots,n_l$, $j = 1,\ldots,n_l$ are the number of the nodes in the l-th layer. w_{ij}^l is the weight connection between the i-th node in the l-th layer and the j-th node in the $(l-1)$-th layer. γ is the learning rate.

Consider the variation of the system output with respect to $\frac{\partial\hat{\mathbf{Z}}(k)}{\partial\mathbf{W}_{ij}^{(l)}}$, we have:

$$\hat{\mathbf{Z}}(k) = \mathbf{H}(k)\hat{\mathbf{X}}(k/k-1) \tag{16}$$

and

$$\hat{\mathbf{X}}(k/k-1) = \boldsymbol{\varphi}^{k-1}\hat{\mathbf{X}}(0) + \sum_{m=1}^{k-1}\boldsymbol{\varphi}^{m-1}\mathbf{O}(k-m) \tag{17}$$

where $\mathbf{O}(k-m)$ is the output of the NN in the $(k-m)$-th iteration.
Hence, we have:

$$\frac{\partial\hat{\mathbf{Z}}(k)}{\partial\mathbf{W}_{ij}^{(l)}} = \mathbf{H}(k)\sum_{m=1}^{k-1}\boldsymbol{\varphi}^{m-1}\frac{\partial\mathbf{O}(k-m)}{\partial\mathbf{W}_{ij}^{(l)}} \tag{18}$$

Then, Eq. (15) becomes:

$$\mathbf{W}_{ij}^{(l)}(k+1) = \mathbf{W}_{ij}^{(k)} + \gamma\mathbf{e}^T(k)\mathbf{H}(k)\sum_{m=1}^{k-1}\boldsymbol{\varphi}^{m-1}\frac{\partial\mathbf{O}(k-m)}{\partial\mathbf{W}_{ij}^{(l)}} \tag{19}$$

The NN has 150 neurons in the hidden layer and 96 neurons in the output layer. The output of each hidden neuron with S-function, as follows:

$$\mathbf{O}'(k) = \frac{1}{1 + e^{-\mathbf{h}(k)}} \tag{20}$$

where

$$\mathbf{h}(k) = \mathbf{W}_{ij}^1 \mathbf{e}(k) \tag{21}$$

and the output of the NN is:

$$\mathbf{O}(k) = \mathbf{W}_{ij}^2 \mathbf{O}'(k) \tag{22}$$

The next section presents a visual servoing framework based on this adaptive neuro-filtering for an eye-in-hand robotics system.

IV. DEVELOP A VISUAL SERVOING CONTROL FRAMEWORK WITHOUT PARAMETERS

As the robotic desired pose is appointed, and the desired image features are constant, the derivation of error function Eq. (1) becomes:

$$\dot{\mathbf{e}}(k) = \frac{d}{dk} \left(\mathbf{S}(k) - \mathbf{S}^d \right)$$
$$= \dot{\mathbf{S}}(k)$$
$$= \mathbf{J}_s(k)\mathbf{U}(k) \tag{23}$$

There is a nonzero constant to establish:

$$\dot{\mathbf{e}}_s(k) = -\lambda \mathbf{e}_s(k) \tag{24}$$

Then substituting Eq. (24) into Eq. (23), we have:

$$-\lambda \mathbf{e}_s(k) = \mathbf{J}_s(k)\mathbf{U}(k) \tag{25}$$

Generally, the dimension of image features n should opted greater than the controlling variable m. Thus the Eq. (25) can be transformed to:

$$\mathbf{U}(k) = -\lambda \mathbf{J}_s^+(k)\mathbf{e}_s(k) \tag{26}$$

Eq. (26) is the robotic servo control law, λ is control rate, and the inverse mapping matrix $\mathbf{J}_s^+(k)$ is:

$$\mathbf{J}_s^+(k) = \mathbf{J}_s(k)^T \left(\mathbf{J}_s(k)\mathbf{J}_s(k)^\mathbf{T} \right)^{-1} \tag{27}$$

To insure the control system stability, we give the Lyapunov-candidate-function based on the image error Eq. (1):

$$L(k) = \frac{1}{2} \|\mathbf{e}_s(k)\|^2 = \frac{1}{2} \left(\mathbf{e}_s(k) \right)^\mathbf{T} \mathbf{e}_s(k) \tag{28}$$

whose derivation is:

$$\dot{L}(k) = \frac{d}{dk} \left(\frac{1}{2} \left(\mathbf{e}_s(k) \right)^T \mathbf{e}_s(k) \right) = \left(\mathbf{e}_s(k) \right)^T \dot{\mathbf{e}}_s(k) \tag{29}$$

Substituting Eq. (23) and Eq. (26) into Eq. (29), we have:

$$\dot{L}(k) = \left(\mathbf{e}_s(k) \right)^\mathbf{T} \mathbf{J}_s(k)\mathbf{U}(k)$$
$$= -\lambda \left(\mathbf{e}_s(k) \right)^T \mathbf{J}_s(k)\mathbf{J}_s^+(k)\mathbf{e}_s(k) \tag{30}$$
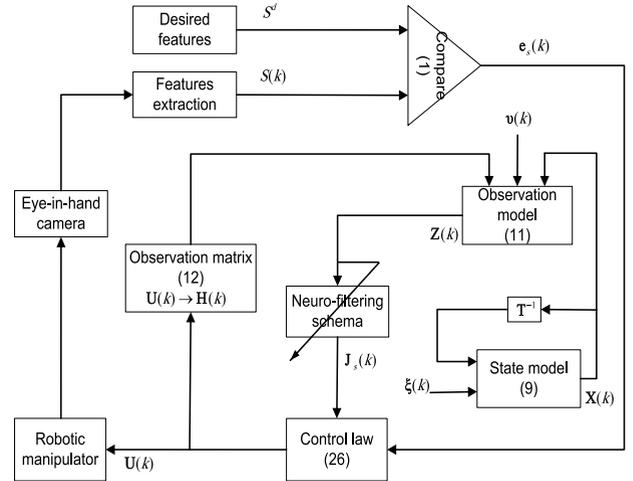

FIGURE 3. The block diagram for the overall visual servoing system.

Algorithm 2 Visual Servoing Control Algorithm

Step 1: system initialization, given the desired image feature $\mathbf{S}^d$, control rate λ, and $\mathbf{J}_s(0)$.

Step 2: at k time, the current image features $\mathbf{S}(k)$ are captured by camera, as shown in Fig. 3, then compare $\mathbf{S}(k)$ and $\mathbf{S}^d$ by Eq. (1). If the image error $\mathbf{e}_s(k)$ converges to zero, the iteration loop ends, otherwise go to the next step.

Step 3: as shown in Fig. 3, calculate the robot control variable $\mathbf{U}(k)$ using control law Eq. (26), and drive the end-effector from the current pose to the next pose, the observation matrix $\mathbf{H}(k)$ is obtained from $\mathbf{U}(k)$ by Eq. (12).

Step 4: the state vector $\mathbf{X}(k)$ in Eq. (8) is constructed, then the state transition from time k to time $k+1$ can be realized using system state Eq. (9).

Step 5: as shown in Fig. 3, the state vector $\mathbf{X}(k)$ and observation matrix $\mathbf{H}(k)$ are used as inputs of the observation model, then the state observation value $\mathbf{Z}(k)$ can be calculated using observation Eq. (11).

Step 6: the state estimation $\hat{\mathbf{X}}(k+1)$ for $k + 1$ time can be obtained using the neuro-filtering schema which shown in Fig. 2, and the best identification of the robotic mapping matrix can be obtained by $\mathbf{J}_S(k + 1) \leftarrow \hat{\mathbf{X}}(k + 1)$

Step 7: time update $k \leftarrow k + 1$, go to step 2).

As $n > m$, $\mathbf{J}_s(k)\mathbf{J}_s^+(k)$ is positive semidefinite, and $(\mathbf{e}_s(k))^\mathbf{T} \mathbf{J}_s(k)\mathbf{J}_s^+(k)\mathbf{e}_s(k) \geq 0$. When the control rate $\lambda > 0$, we have $\dot{L}(k) \leq 0$. According to Lyapunov-candidate-function, the control system is stable.

Fig. 3 shows our overall visual servo system, the algorithm flow is described as follows:

V. RESULTS AND DISCUSSIONS

In this section, the proposed neuro-filtering based visual servoing framework is tested in simulated and real eye-in-hand six-DOF robotic manipulator. The task is to control the manipulator from arbitrary initial pose to the desired pose
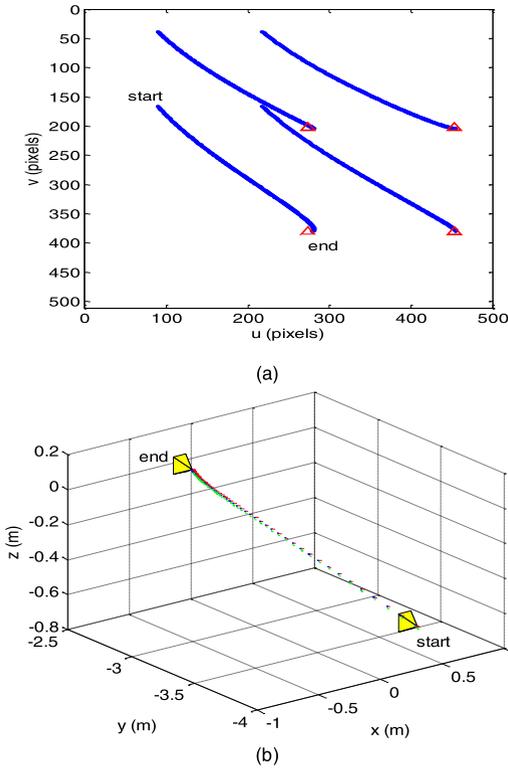
FIGURE 4. Simulation results obtained by the proposed visual servoing without calibrated parameters for task 1. (a) image feature trajectories. (b) robot moving trajectory.

by using the close feedback of image features. The image features are consist of four-image-points for robotic manipulation tests. The desired features $\mathbf{S}^d$ does not change over time, and the current image feature $\mathbf{S}(k)$ is not constant due to the camera with the end-defector movement. The current image features is obtained through:

$$\mathbf{S}(k) = \big(s_1(k) \quad s_2(k) \quad s_3(k) \quad s_4(k) \big) \in \Re^{8\times 1} \quad (31)$$

where

$$s_i = \big(u_i \quad v_i \big), \quad i = 1, 2, 3, 4 \quad (32)$$

Let the robotic control variable:

$$\mathbf{U}(k) = \big(V(k) \quad W(K) \big)^{\mathbf{T}} \in \Re^{6\times 1} \quad (33)$$

where

$$V(k) = \big(v_x(k) \quad v_y(k) \quad v_z(k) \big) \quad (34)$$
$$W(k) = \big(w_x(k) \quad w_y(k) \quad w_z(k) \big) \quad (35)$$

Eq. (34) and (35) are the end-effector's linear and angular velocity in workspace respectively. Thus, the size of mapping matrix in Eq. (4) is 8×6, and the size of the system state in Eq. (8) is 96×1. The control rate λ is selected 0.25. The intrinsic camera parameters for traditional IBVS are chosen as, image center $u_0 = v_0 = 256$, and focal length $f = 1000$.
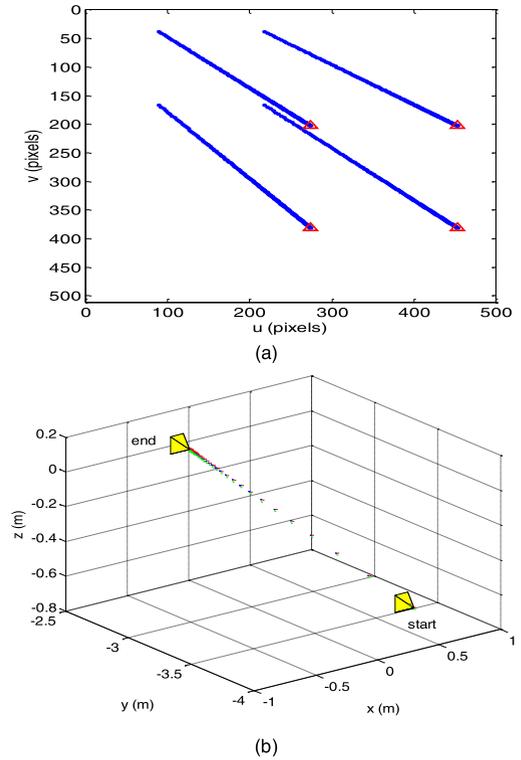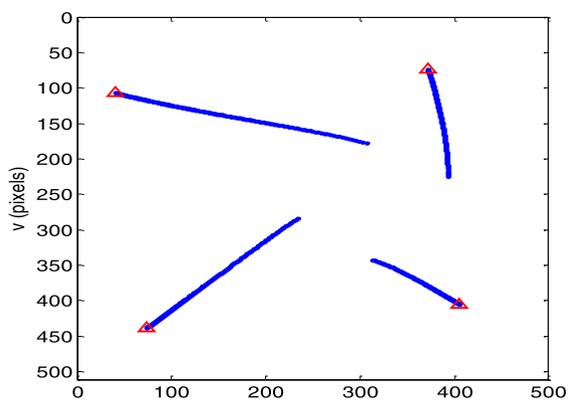

FIGURE 5. Simulation results obtained by IBVS for task 1. (a) image feature trajectories. (b) robot moving trajectory.
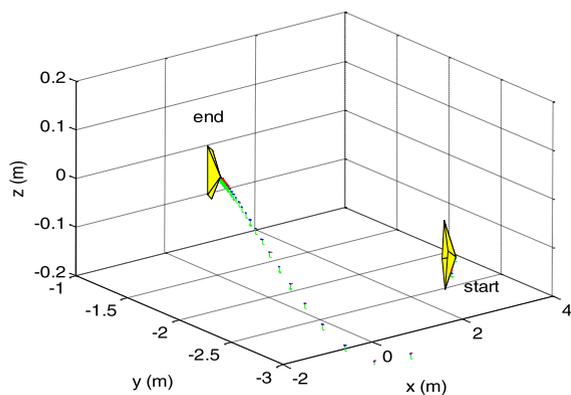
A. SIMULATION
The simulation is conducted using robotic simulation system in MATLAB, in which two different tests are conducted. The robotic manipulator conducts linear and combination rotational movements to evaluate the performances of our visual serving method.

In the first test, the camera simply conducts translation movements. We set $\mathbf{S}(0) = (89.33, 38.05, 89, 166.3, 217.5, 166.3, 217.5, 38.05)^{\mathbf{T}}$ and $\mathbf{S}^d = (273.9, 202.4, 273.9, 381.7, 452.4, 381, 452.4, 202.4)^{\mathbf{T}}$, respectively. Fig. 4 shows the results of the proposed method. More specifically, the feature trajectories in the image plane are shown in Fig. 4(a), in which the trajectory ends at the position as the $\triangle$ sign. The image features moving in almost straight lines from initial to desired position with our method. Fig. 4(b) shows the 3-D trajectory of the robotic end-effector in the Cartesian space, which is a straight line from the initial pose to the desired pose.

For the same test, the results performed by IBVS are shown in Fig. 5, the feature trajectories in Fig. 5(a) and the robotic end-effector trajectory in Fig. 5(b) are similar to results obtained by our proposed method. Note that the IBVS method needs the camera calibration parameters to calculate the mapping matrix, while our method is on-line identified techniques without any calibrated parameters, on the other hand, for the same translation movements visual serving task the performance of the two methods are very similar, thus the robotic mapping matrix on-line estimated by the proposed neuro-filtering approach is accurate.
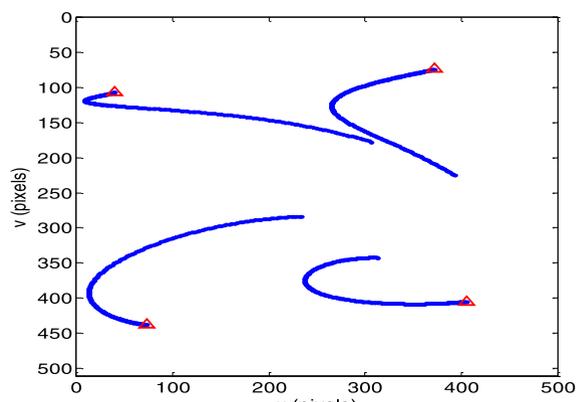
FIGURE 6. Simulation results obtained by IBVS for task 2. (a) image feature trajectories. (b) robot moving trajectory.

FIGURE 7. Simulation results obtained by the proposed visual servoing without calibrated parameters for task 2. (a) image feature trajectories. (b) robot moving trajectory.
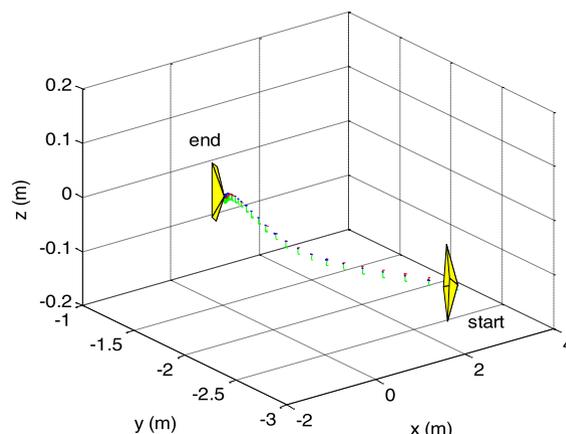
In the second test, the camera implements a combination of translational with rotational movements so that a long distance visual servoing task is performed. The initial and the desired features are located in far away each other, *i.e.* $\mathbf{S}(0) =$ $(360.9, 177.5, 234.1, 283.8, 313.5, 342.5, 393.7, 225.3)^{\mathbf{T}}$ and $\mathbf{S}^d = (40.2, 106.8, 73.5, 438.5, 405.1, 405.2, 371.9, 73.5)^{\mathbf{T}}$. For the same visual servoing task, Fig. 6 shows the results obtained by IBVS and Fig. 7 shows the results obtained by our visual servoing method, respectively.

By comparing Fig. 6(a) with Fig. 7(a), we can see that the performances of image features are same constrained on the camera field-of-view (FOV). Fig. 6(b) shows that the trajectory of the robotic end-effector has large retreat movements that are beyond the 3-D workspace. However, the trajectory obtained by our method has no retreat, *i.e.* almost straight line from the initial pose to the desired pose, as shown in Fig. 7(b).

As the same long distance visual servoing task, the robotic performance of the two methods are different, it is clear that the drawbacks of the IBVS is hard to keep the end-effector inside the workspace due to the robotic retreat. While our visual servoing method could provide better results for the same task because of the estimator with adaptive ability for accurate predict the mapping between robotic vision and motor spaces.

B. REAL EXPERIMENT

The real experiments have been carried out using a robotic manipulator with an eye-in-hand configuration. As shown in Fig. 8, our visual servoing system consists of a DENSOR C7M-VSG6BA controller, a computer with an Intel Core i5 2.67-GHz CPU, 4GBs of RAM for image processing. The robotic controller and image processing computer can communicate through RS-232 serial interface, and a DENSO six-DOF robotic manipulator with a Microsoft Kinect camera mounted at its end-effector. The object is an A4 paper with four black-colored small circular disks on it. The resolution of object images is 640×480, and the center points of the small circular disks are used as image features.

Experimemt 1, the experiment aims to examine the performances of visual servoing framework based on neuro-filtering and traditional KF methods in a real un-calibrated environment. As the traditional KF method needs the statistic knowledge of environment noise in advance, we chose the uniformly distributed random noise with zero mean and the system noise variance is $Q = 0.001 \times$ E', in which $E' \in \Re^{96 \times 96}$ is unit matrix of size 96×96, and the observation noise variance is $R = 0.001 \times E''$,
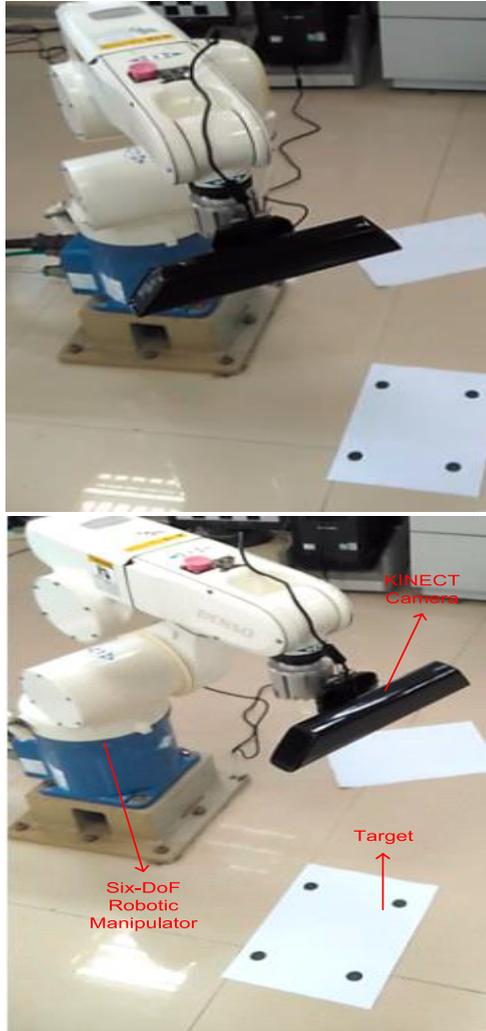
$E'' \in \Re^{8 \times 8}$ is unit matrix of size 8×8. In contrast, our method is self-adaptive to the dynamic noise and does not require the statistic parameters of environment noise.

The task is for the end-effector of the robot conducts translational and rotational movements from the initial pose to the desired pose. We set the initial features $\mathbf{S}(0) = (303, 179, 234, 284, 311, 344, 396, 226)^{\mathbf{T}}$, and desired features given $\mathbf{S}^d = (255, 123, 255, 417, 549, 417, 549, 123)^{\mathbf{T}}$, this task requires the robot to conduct a complicated motion in 3-D Cartesian space. Fig. 9 shows experimental results obtained by the proposed visual servoing with KF. More specifically, Fig. 9(a) shows the image feature trajectories that were obtained by the traditional KF method and influenced by noises. Fig. 9(b) shows that the end-effector moved towards the desired pose and reduced the image errors. Fig. 9(c) shows that the end-effector has unnecessary retreat and serious vibration.

Fig. 10 shows experimental results obtained by the proposed visual servoing with neuro-filtering. More specifically, Fig. 10(a) shows that the image features smoothly converged to the desired position, and Fig. 10(b) shows the image errors
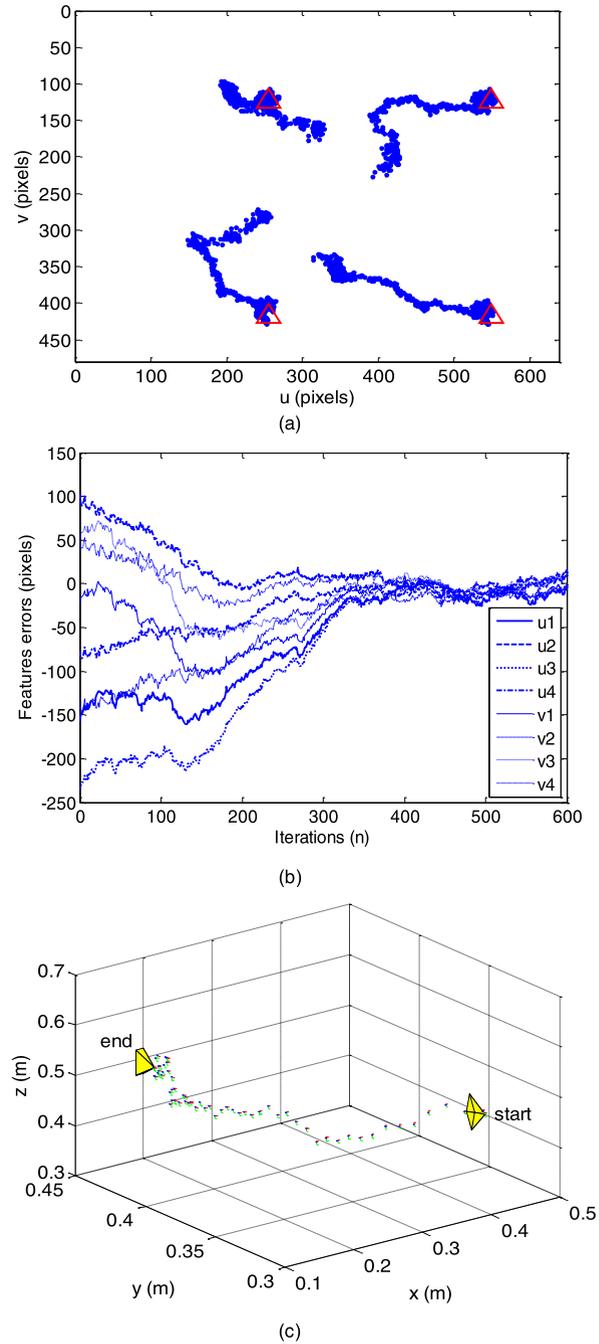

FIGURE 9. Experimental results obtained by the proposed visual servoing with KF. (a) image feature trajectories. (b) image errors between current features and desired. (c) robot moving trajectory.

converging in image plant. Fig. 10(c) shows that the robot end-effector had the stabile motion without retreat and vibration. It is clear that the proposed neuro-filtering method outperformed the traditional KF method, *i.e.* the smooth motion in the Cartesian space and the stable feature trajectories on the image plane.

Experiment 2, the experiment aims to test the robustness of visual servoing framework based on traditional KF and neuro-filtering. The KF optimal filtering gain depends on the
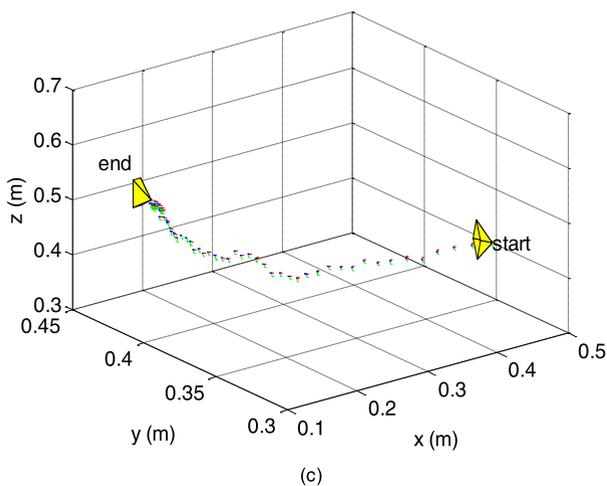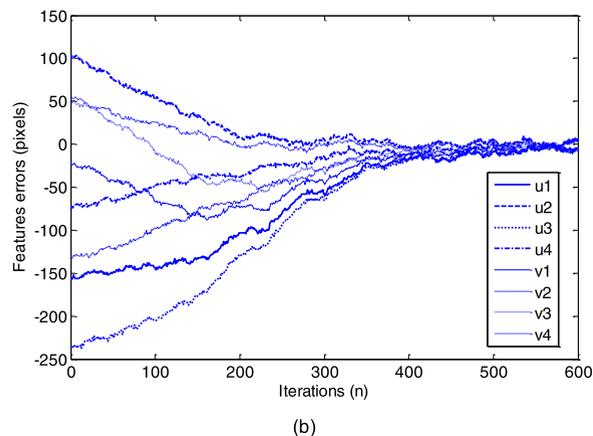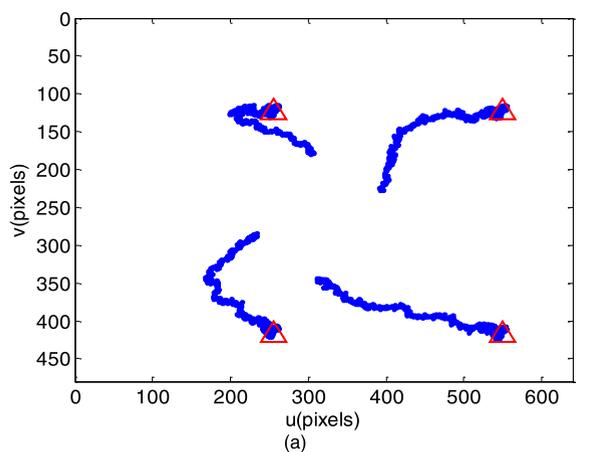
FIGURE 10. Experimental results obtained by the proposed visual servoing with neuro-filtering. (a) image feature trajectories. (b) image errors between current features and desired. (c) robot moving trajectory.

FIGURE 11. Result obtained by the proposed visual servoing with KF. (a) robot moving trajectory. (b) image feature trajectories

noise statistics, thus, the change of the variances of system noise or observation noise would made the KF method easy out of convergence. In this test we chose the system noise variance is $Q = 0.02 \times E'$, and the observation noise variance is $R = 0.15 \times E''$. Fig. 11 and Fig. 12 shows a comparison of the robot moving trajectory and image feature trajectories in a real environment.
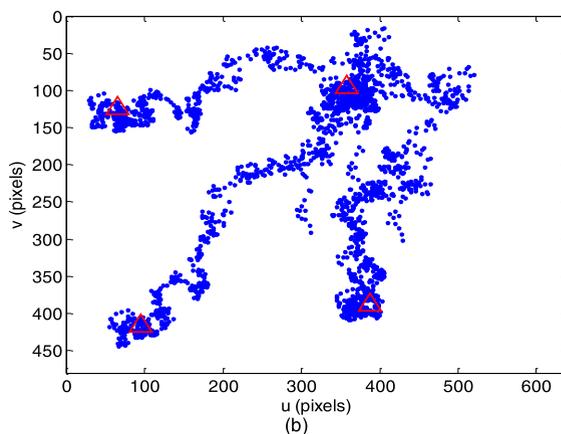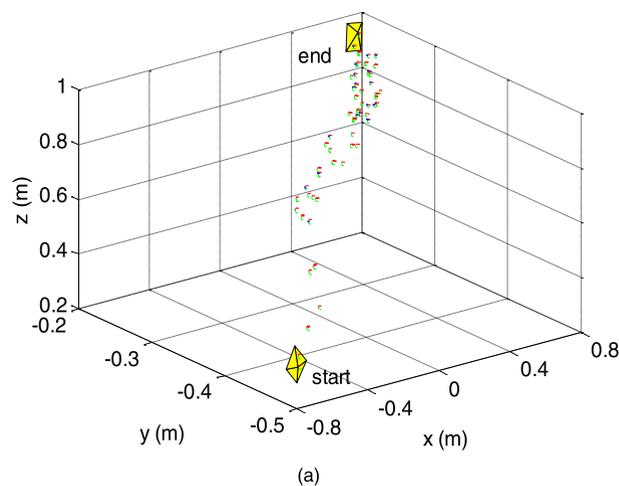
More specifically, Fig. 11 shows the result obtained by the visual servoing with KF. The camera has almost random motion in the Cartesian space, and the image features in image plan are scattered near the camera FOV limit, those results are illustrated in Fig. 11(a) and Fig. 11(b), respectively. In contrast, Fig. 12 shows the result obtained by the visual servoing with proposed neuro-filtering. The robot end-effector has a smooth moving trajectory and a good stability, and the image trajectories could also converged to the desired position, one can find in Fig. 12(a) and Fig. 12(b).

This means that the KF method has risk leads the robot to large retreat motion in the Cartesian space, and is easily reaching the limitation of workspace, also it made the image features move near to the camera FOV limit. In the contrast, our proposed neuro-filtering approach is adaptive to dynamic noises made the robot always converge to the desired pose with smooth trajectories and robust stability performances in servoing tasks.

The simulation and experimental results show that the robot moving in the 3D Cartesian space and the feature trajectories on the 2D image plane both are very smooth and stable. The image errors are reduced by the servoing controller. The proposed visual servoing based neuro-filtering method is
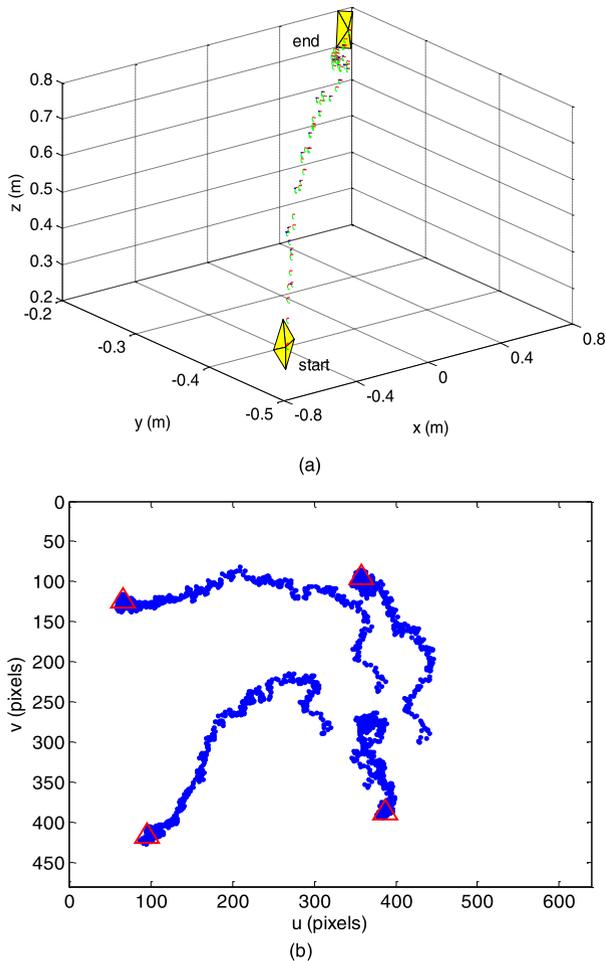
FIGURE 12. Result obtained by the proposed visual servoing with neuro-filtering. (a) robot moving trajectory. (b) image feature trajectories.

efficient for robot manipulation in un-calibrated environment with unknown noise.

VI. CONCLUSIONS

In this work, we have proposed a novel visual servoing framework for regulation of a robotic manipulator operated in an un-calibrated environment, which is based on neuro-filtering. The NN based KF approach has been proposed to identify the mapping of robotic vision space and motor space. One of advantages of neuro-filtering is its state estimating ability without the statistics knowledge of environment noise. Apart from simulation, we have applied this neuro-filtering schema in a real robotic manipulator to avoid the complex calibration process of its eye-in-hand system. Several experiments have been conducted to compare the performance of our proposed system with IBVS and KF methods. Results show that the proposed visual servoing approach has better performance than these existing methods.

REFERENCES

[1] M. A. V. J. Muthugala and A. G. B. P. Jayasekara, "A review of service robots coping with uncertain information in natural language instructions," *IEEE Access*, vol. 6, pp. 12913–12928, Mar. 2018.

[2] S.-H. Tseng, F.-C. Liu, and L.-C. Fu, "Active learning on service providing model: Adjustment of robot behaviors through human feedback," *IEEE Trans. Cognitive Develop. Syst.*, vol. 10, no. 3, pp. 701–711, Sep. 2018.

[3] C. Teulière and E. Marchand, "A dense and direct approach to visual servoing using depth maps," *IEEE Trans. Robot.*, vol. 30, no. 5, pp. 1242–1249, Oct. 2014.

[4] K. Wang, Y. Liu, and L. Li, "Vision-based tracking control of underactuated water surface robots without direct position measurement," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 6, pp. 2391–2399, Nov. 2015.

[5] D. Zheng, H. Wang, J. Wang, S. Chen, W. Chen, and X. Liang, "Image-based visual servoing of a quadrotor using virtual camera approach," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 972–982, Apr. 2017.

[6] F. Ke, Z. Li, H. Xiao, and X. Zhang, "Visual servoing of constrained mobile robots based on model predictive control," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 47, no. 7, pp. 1428–1438, Jul. 2017.

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

[8] M. Josip, B. Mirjana, and C. Mojmil, "Comparison of uncalibrated model-free visual servoing methods for small-amplitude movements: A simulation study," *Int. J. Adv. Robotic Syst.*, vol. 11, no. 7, pp. 1–16, Jul. 2014.

[9] 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.

[10] H. Wang, "Adaptive visual tracking for robotic systems without image-space velocity measurement," *Automatica*, vol. 55, pp. 294–301, May 2015.

[11] F. Yan, B. Li, W. Shi, and D. Wang, "Hybrid visual servo trajectory tracking of wheeled mobile robots," *IEEE Access*, vol. 6, pp. 24291–24298, May 2018.

[12] W. Pan, M. Lyu, K.-S. Hwang, M.-Y. Ju, and H. Shi, "A neuro-fuzzy visual servoing controller for an articulated manipulator," *IEEE Access*, vol. 6, pp. 3346–3357, 2018.

[13] 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.

[14] Q. Bateux and E. Marchand, "Histograms-based visual servoing," *IEEE Robot. Autom. Lett.*, vol. 2, no. 1, pp. 80–87, Jan. 2017.

[15] J. A. Piepmeier, G. V. Mcmurray, and H. Lipkin, "A dynamic quasi-Newton method for uncalibrated visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, Piscataway, NJ, USA, May 1999, pp. 1595–1600.

[16] M. Asada, T. Tanaka, and K. Hosoda, "Visual tracking of unknown moving object by adaptive binocular visual servoing," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, Taipei, Taiwan, Aug. 1999, pp. 249–254.

[17] J. A. Piepmeier, G. V. McMurray, and H. Lipkin, "Uncalibrated dynamic visual servoing," *IEEE Trans. Robot. Autom.*, vol. 20, no. 1, pp. 143–147, Feb. 2004.

[18] J. A. Piepmeier and H. Lipkin, "Uncalibrated eye-in-hand visual servoing," *Int. J. Robot. Res.*, vol. 1, nos. 10–11, pp. 805–819, Oct. 2002.

[19] M. Z. Hao Sun, "A universal state-space approach to uncalibrated model-free visual servoing," *IEEE/ASME Trans. Mech.*, vol. 17, no. 5, pp. 833–846, Oct. 2012.

[20] J. Qian and J. Su, "Online estimation of image Jacobian matrix by Kalman–Bucy filter for uncalibrated stereo vision feedback," in *Proc. IEEE Int. Conf. Robot. Automat.*, Washington, DC, USA, May 2002, pp. 562–567.

[21] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 34–45, 1960.

[22] H. Mosbah and M. E. El-Hawary, "Optimized neural network parameters using stochastic fractal technique to compensate Kalman filter for power system-tracking-state estimation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 379–388, Feb. 2019.

[23] R. Zhan and J. Wan, "Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation," *IEEE Signal Process. Lett.*, vol. 13, no. 7, pp. 445–448, Jul. 2006.
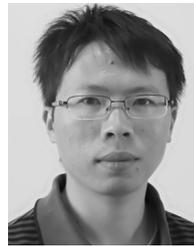
[24] L. Zhang and P. B. Luh, "Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 59–66, Feb. 2005.

[25] M. Choi, R. Sakthivel, and W. K. Chung, "Neural network-aided extended Kalman filter for SLAM problem," in *Proc. IEEE Conf. Robot. Automat.*, Roma, Italy, Apr. 2007, pp. 1686–1690.

[26] S. Li, W. Ma, J. Liu, and H. Chen, "A Kalman gain modify algorithm based on BP neural network," in *Proc. 16th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Qingdao, China, Sep. 2016, pp. 1–4.

[27] S.-H. Hur, "Estimation of useful variables in wind turbines and farms using neural networks and extended Kalman filter," *IEEE Access*, vol. 7, pp. 24017–24028, Feb. 2019.

[28] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1732–1742, Aug. 1991.

[29] A. Y. Alanis, E. N. Sanchez, A. G. Loukianov, and M. A. Perez-Cisneros, "Real-time discrete neural block control using sliding modes for electric induction motors," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 1, pp. 11–21, Jan. 2010.

[30] T. Soni, B. D. Rao, J. R. Zeidler, and W. H. Ku, "Enhancement of images using the 2-D LMS adaptive algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech. Signal Process.*, Toronto, ON, Canada, Apr. 1991, pp. 3029–3032.

[31] I. M. Stephanakis and S. D. Kollias, "Wavelet-based approach to adaptive Wiener filtering of images in the presence of colored noise," *Opt. Eng.*, vol. 39, no. 5, pp. 1353–1363, 2000.

[32] X. Lv and X. Huang, "Fuzzy adaptive Kalman filtering based estimation of image Jacobian for uncalibrated visual servoing," in *Proc. IEEE/RSJ/GI Conf. Intell. Robot. Syst.*, Beijing, China, Oct. 2006, pp. 2167–2172.

[33] 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.

[34] P. Wira and J. P. Urban, "A new adaptive Kalman filter applied to visual servoing tasks," in *Proc. IEEE Conf. Knowl.-Based Intell. Eng. Syst. Allied Technol.*, Brighton, U.K., Aug./Sep. 2000, pp. 1–4.

[35] H. Shi, X. Li, K.-S. Hwang, W. Pan, and G. Xu, "Decoupled visual servoing with fuzzy Q-learning," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 241–252, Jan. 2018.

XUNYU ZHONG received the M.E. degree in mechatronics engineering from Harbin Engineering University, Harbin, China, in 2007, and the Ph.D. degree in control theory and control engineering from Harbin Engineering University, in 2009. He is currently an Associate Professor with the School of Aerospace Engineering, Xiamen University, Xiamen, China. He is an Academic Visitor of the School of Computer Science and Electronic Engineering, University of Essex, U.K., for 1 year, from September 2017. His current research interests include robot motion planning, visual servo, and autonomous robots.

HUOSHENG HU (SM'01) received the M.Sc. degree in industrial automation from Central South University, China in 1982, and the Ph.D. degree in robotics from the University of Oxford, U.K., in 1993. He is currently a Professor with the School of Computer Science and Electronic Engineering, University of Essex, U.K., leading the Robotics Group. He has authored over 500 research articles published in journals, books, and conference proceedings. His research interests include autonomous robots, human–robot interaction, multi-robot collaboration, embedded systems, pervasive computing, sensor integration, intelligent control, and networked robots. He is a Fellow of the Institute of Engineering and Technology, Fellow of the Institution of Measurement and Control, and a Chartered Engineer in the U.K. He currently serves as the Editor-in-Chief for the *International Journal of Automation and Computing*, the Editor-in-Chief of the *MDPI Robotics Journal*, and an Executive Editor for the *International Journal of Mechatronics and Automation*.

XUNGAO ZHONG (M'19) received the B.Eng. degree in electronic information engineering from Nanchang University, in 2007, the M.Sc. degree in electro-mechanical engineering from the Guangdong University of Technology in 2011, and the Ph.D. degree in control theory and engineering from Xiamen University, in 2014. He is currently an Associate Professor with the School of Electrical Engineering and Automation, Xiamen University of Technology at Xiamen, China. His current research interests include machine learning and robotic visual servoing. He is selected as distinguished young scientific research talent of Fujian province, China, in 2018.

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

• • •