

Visual Servo System Based on Cubature Kalman Filter and Backpropagation Neural Network

Wenjuan Huang¹, Wei Zhao^{2*}, Jin Zhang²

¹Tangshan Vocational and Technical College, Tangshan 063004, China

²Intelligence and information Engineering college, Tangshan University, Tangshan 063000, China

Corresponding Author Email: zhwei19800@163.com

<https://doi.org/10.18280/i2m.180208>

ABSTRACT

Received: 13 January 2019

Accepted: 20 March 2019

Keywords:

visual servo system, cubature Kalman filter, neural network, image Jacobian matrix

This paper attempts to enhance the estimation accuracy of image Jacobian matrix, despite the impacts from unknown statistical features of process noise and measurement noise. For this purpose, the singular value decomposition aided Cubature Kalman filter (SVDCKF) was coupled with a noise compensator for visual servo systems based on the backpropagation neural network (BPNN). The proposed algorithm improves the accuracy of the Kalman filter and compensates for the two noise covariances through the nonlinear approximation function of the BPNN. The proposed Jacobian matrix estimation algorithm was subjected to a binocular stereo visual servoing experiment on the Motoman UP6 robot experimental platform. The experimental results show that the proposed algorithm is capable of completing servo positioning.

1. INTRODUCTION

Thanks to the development of computer vision, the camera (visual sensor) has been widely adopted to serve as the “eyes” of robots [1]. The camera mainly captures images on the external environment. The information from the images was taken as the feedback signal to realize the closed-loop control of robot [2]. This process, known as visual servoing, has been applied in medical rescue and many other fields [3].

There are many ways to classify the visual servo systems in actual application [4, 5]. By system structure, the visual servo systems can be divided into eye-in-hand system and eye-to-hand system. By the number of cameras, the visual servo systems can be categorized as monocular, binocular and multi-eye systems. By feedback signal, the visual servo systems can be classified into position-based visual servoing (PBVS), image-based visual servoing (IBVS) and hybrid visual servoing.

Among them, the binocular stereo visual servo system with parallel optical axis stands out for its easy access to stereo information and excellent system performance. Meanwhile, the IBVS system is robust to camera calibration for the direct control of 2D images, eliminating the need to consider 3D reconstruction. As a result, the integration between the two types of systems has become increasingly popular in recent years [6].

The online estimation of image Jacobian matrix is fundamental to visual servo systems. Many scholars have applied Kalman filtering algorithm to estimate the matrix online using robot kinematics. For example, Reference [7] designs a Kalman filtering algorithm for online estimation of Jacobian matrix or homogenous transformation matrix of composite images, which can reduce the impact of noise. Reference [8] develops an adaptive Kalman filter to estimate the image Jacobian matrix of eye-in-hand servo system. Reference [9] constructs an improved Kalman filter for online estimation of Jacobian matrix. The Kalman filters in these

studies boast strong applicability and recursiveness.

The Kalman filter can work perfectly in stochastic linear systems, provided that the model is accurate and that the system statistics and noises belong to Gaussian white noise sequences with known variance. However, visual servo systems will inevitably suffer from random noises, especially when the noise statistics are unknown. In this case, the traditional Kalman filter will perform poorly and even cause divergence [10].

To solve the problem, the Kalman filter can be fused with the neural network in two ways. One is to use the Kalman filter to improve the classification performance of the neural network. For instance, Reference [11] constructs a linear model based on the Kalman filter, and relies on it to improve the classification effect of the neural network through post-processing; the linear model ensures that the neural network predicts near-perfect output through linear combination of object features and the predicted output. This approach reduces the error and enhances the classification performance of the neural network.

The other is to better the filtering effect using the nonlinear approximation of the neural network. For example, Reference [12] adopts extended Kalman filter assisted artificial neural network to enhance the self-sensing capability of shape-memory alloy actuator. Specifically, the artificial neural network was employed to bridge the gap between the estimation by the extended Kalman filter and the actual stretch of the spring. Reference [13] investigates the data fusion of asynchronous, multi-rate and multi-sensor linear systems, and estimates the state vector using the neural network integrating multiple Kalman filters; The research results show that the integrated method outperformed the other ways of data fusion.

In our previous research [14], the singular value decomposition-aided cubature Kalman filter with neural network (NNSVDCKF) is compared with the singular value decomposition-aided cubature Kalman filter (SVDCKF) and standard Kalman filter (KF) in terms of the estimation of the

image Jacobian matrix in an uncalibrated binocular stereo visual servo system for a 6DOF PUMA-560 robot. In this paper, the binocular stereo visual model and the NNSVDCKF are introduced in details, the image Jacobian matrix is estimated by the SVDCKF coupled with a backpropagation neural network (BPNN) noise compensator, and the proposed Jacobian matrix estimation algorithm is verified on the experimental platform for a Motoman UP6 robot.

2. CONFIGURATION OF BINOCULAR STEREO VISUAL MODEL

2.1 Jacobian matrix

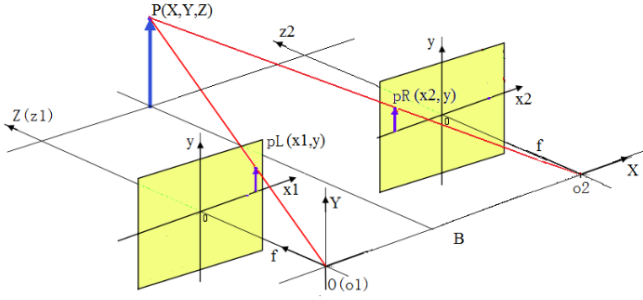


Figure 1. The binocular stereo visual model

As shown in Figure 1, the binocular stereo visual model is an eye-in-hand servo system. It is assumed that the two cameras have the same parameters, and their focal length f equals object distance; the origin O of the global coordinate system coincides with the optical center o_1 of the left camera, which is parallel to the right one o_2 ; the length between the two lenses is denoted as B . In keeping with human habits, the imaging point is set at the front of the camera center. In addition, the left and right image planes are defined as x_1 - y_1 and x_2 - y_2 , respectively. Since the two lenses are parallel to the installation, neither have a significant y -axis displacement. Hence, y_1 can be considered as equal to y_2 , and both are denoted as y in this paper. According to the geometric relations in Figure 1, the change relationship between the position of feature points in the left and right image planes, and the position and pose of the end-effector [14]:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{y} \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \frac{x_2 - x_1}{B} & 0 & \frac{x_1(x_1 - x_2)}{fB} & + \frac{x_1 y}{f} & - \frac{f^2 + x_1^2}{f} & + y \\ 0 & \frac{x_2 - x_1}{B} & \frac{y(x_1 - x_2)}{fB} & + \frac{f^2 + y^2}{f} & - \frac{x_1 y}{f} & - x_1 \\ \frac{x_2 - x_1}{B} & 0 & \frac{x_2(x_1 - x_2)}{fB} & + \frac{y x_2}{f} & - \frac{f^2 + x_1 x_2}{f} & + y \end{bmatrix} \begin{bmatrix} \dot{S}_x \\ \dot{S}_y \\ \dot{S}_z \\ \dot{\theta}_x \\ \dot{\theta}_y \\ \dot{\theta}_z \end{bmatrix} \quad (1)$$

where S_x , S_y and S_z are the position coordinates of the origin of the end-effector coordinate system; θ_x , θ_y and θ_z are the space angles of the origin of the end-effector coordinate system.

let $p_{single} = [x_1, y, x_2]^T$ be the position matrix of the image feature points for a single object in the left and right image planes, and $U = [S_x, S_y, S_z, \theta_x, \theta_y, \theta_z]^T$ be the position and pose matrix of the end-effector in the global coordinate system. Then, Equation (1) can be extended to accommodate for multiple feature points:

$$\dot{p} = \begin{bmatrix} \dot{p}_{single,1} \\ \dot{p}_{single,2} \\ \dots \\ \dot{p}_{single,m} \end{bmatrix} = \begin{bmatrix} J_{single,1}(p_{single,1}) \\ J_{single,2}(p_{single,2}) \\ \dots \\ J_{single,m}(p_{single,m}) \end{bmatrix} \dot{U} = J \dot{U} \quad (2)$$

2.2. The estimation model for image Jacobian matrix

Based on the binocular stereo visual model in Section 2.1, the image Jacobian matrix should be solved before realizing the servoing of the mechanical arm. Here, the elements of the image Jacobian matrix are estimated by the positions of the image feature points. This calls for a process equation for the position of the state vector of the image Jacobian matrix and the measurement equation for the position of the image feature points.

Equation (2) can be discretized into:

$$(p_{k+1} - p_k) / \Delta t = J \cdot \Delta U_k / \Delta t \quad (3)$$

where ΔU_k is the changes of the position and pose matrix of the end-effector between moments $k-1$ and k . Then, the state vector χ_k and measurement vector Z_k in each element of the image Jacobian matrix can be expressed as:

$$\chi_k = [J_{1,1}, J_{1,2}, \dots, J_{1,6}, J_{2,1}, \dots, J_{m,6}]^T_{18m \times 1} \quad (4)$$

$$Z_k = [p_{k+1} - p_k]_{3m \times 1} \quad (5)$$

Then, the process equation and the measurement equation of the linear discrete dynamic system can be established as:

$$\chi_{k+1} = \chi_k + \omega \quad (6)$$

$$Z_k = H_k \cdot \chi_k + v \quad (7)$$

where $\omega_{18m \times 1}$ is the process noise vector, an uncorrelated zero-mean Gaussian white noise sequence with covariance $Q_{18m \times 18m}$; $V_{3m \times 1}$ is the measurement vector, whose covariance $R_{3m \times 3m}$ is defined the same as that of ω ; H_k is the measurement matrix:

$$H_k = \begin{bmatrix} \Delta U_k' & \dots & 0 \\ \dots & \Delta U_k' & \dots \\ 0 & \dots & \Delta U_k' \end{bmatrix} \quad (8)$$

3. JACOBIAN MATRIX ESTIMATION ALGORITHM

3.1 SVDCKF

Consider a nonlinear discrete random state-space model in the following form:

$$x_{k+1} = f(x_k) + w_k \quad (9)$$

$$Z_{k+1} = h(x_{k+1}) + v_{k+1} \quad (10)$$

where x_k is the n -dimensional system state vector; Z_{k+1} is the m -dimensional measurement vector; $f(\cdot)$ is the state equation; $h(\cdot)$ is the measurement equation; w_k is the zero-mean white noise with covariance matrices Q ; v_{k+1} is the zero-mean white noise with covariance matrices R . w_k and v_{k+1} are assumed to be uncorrelated with each other, and are not related to the initial state.

Then, the complete process of the SVDCKF algorithm can be explained as:

Step 1. Initialization

$$\hat{x}_0 = E(x_0) \text{ and } \hat{P}_{x,0} = E((x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T) \quad (11)$$

Step 2. Time update ($k=0, 1, 2, \dots$)

$$\hat{P}_{x,k} = U_k \begin{bmatrix} S_k^* & 0 \\ 0 & 0 \end{bmatrix} V_k^T \quad (12)$$

$$\hat{X}_{i,k} = U_k \sqrt{S_k^*} \xi_i + \hat{x}_k \quad i = 1, 2, \dots, 2n \quad (13)$$

$$\xi_i = \sqrt{n}[1]_i \quad (14)$$

where $[1]_i$ is the i -th column of the point set [1].

$$\hat{X}_{i,k+1/k} = f(X_{i,k}) \quad (15)$$

$$\hat{x}_{k+1/k} = \frac{1}{2n} \sum_{i=1}^{2n} \hat{X}_{i,k+1/k} \quad (16)$$

$$\hat{P}_{x,k+1/k} = \left(\frac{1}{2n} \sum_{i=1}^{2n} \hat{X}_{i,k+1/k} \hat{X}_{i,k+1/k}^T - \hat{x}_{k+1/k} \hat{x}_{k+1/k}^T \right) + Q \quad (17)$$

Step 3. Measurement update

$$\hat{P}_{x,k+1/k} = U_k \begin{bmatrix} S_k^* & 0 \\ 0 & 0 \end{bmatrix} V_k^T \quad (18)$$

$$X_{i,k+1/k}^* = U_k \sqrt{S_k^*} \xi_i + \hat{x}_{k+1/k} \quad i = 1, 2, \dots, 2n \quad (19)$$

$$\hat{Y}_{i,k+1/k} = h(X_{i,k+1/k}^*) \quad (20)$$

$$\hat{Z}_{k+1/k} = \frac{1}{2n} \sum_{i=1}^{2n} \hat{Y}_{i,k+1/k} \quad (21)$$

$$\hat{P}_{zZ,k+1/k} = \frac{1}{2n} \sum_{i=1}^{2n} X_{i,k+1/k}^* (\hat{Y}_{i,k+1/k})^T - \hat{x}_{k+1/k} (\hat{Z}_{k+1/k})^T \quad (22)$$

$$\hat{P}_{Z,k+1/k} = \frac{1}{2n} \sum_{i=1}^{2n} \hat{Y}_{i,k+1/k} (\hat{Y}_{i,k+1/k})^T - \hat{Z}_{k+1/k} (\hat{Z}_{k+1/k})^T + R \quad (23)$$

$$K_{k+1} = \hat{P}_{zZ,k+1/k} \hat{P}_{Z,k+1/k}^{-1} \quad (24)$$

$$\hat{x}_{k+1} = \hat{x}_{k+1/k} + K_{k+1} (Z_{k+1} - \hat{Z}_{k+1/k}) \quad (25)$$

$$\hat{P}_{x,k+1} = \hat{P}_{x,k+1/k} - K_{k+1} \hat{P}_{Z,k+1/k} K_{k+1}^T \quad (26)$$

3.2 The BPNN noise compensator

A two-layer BPNN was selected as the noise compensator, and used to compensate for the covariance of system and measurement noises before each filtering process.

The inputs of the BPNN are the elements of the state estimation error Δe_χ and the innovation V_k between the current moment k and moment $(k-1)$:

$$\Delta e_\chi = \hat{\chi}_k - \hat{\chi}_{k-1} \quad (27)$$

$$V_k = Z_k - H_k \hat{\chi}_{k-1} \quad (28)$$

The outputs of the BPNN are the compensated values p of Q and R at moment k , which are denoted as ΔQ_k and ΔR_k , respectively. The two values can be obtained by training the BPNN with Sage-Husa method [15]. The values of f Q and R applied in Kalman filter at each moment can be calculated as:

$$Q_k(i, i) = Q_{k-1}(i, i) + \Delta Q_k, \quad i = 1, 2, \dots, mn \quad (29)$$

$$R_k(i, i) = R_{k-1}(i, i) + \Delta R_k, \quad i = 1, 2, \dots, n \quad (30)$$

where

$$\Delta Q_k = \text{mean} \left(\sum_{i=1}^{mn} (\hat{Q}_{k-1}(i, i) - Q_{k-1}(i, i)) \right) \quad (31)$$

$$\Delta R_k = \text{mean} \left(\sum_{i=1}^n (\hat{R}_{k-1}(i, i) - R_{k-1}(i, i)) \right) \quad (32)$$

3.3 The binocular stereo visual servo system with SVDCKF and BPNN noise compensator

The binocular stereo visual servo system was modeled as shown in Figure 2. For simplicity, the following assumptions were put forward: (1) The internal and external parameters of the cameras are known. (2) The feature points of the object are always within the camera horizon. (3) The terminals of the robot arm coincide with the origin of the plane coordinate system of the left camera and the origin of the object coordinate system.

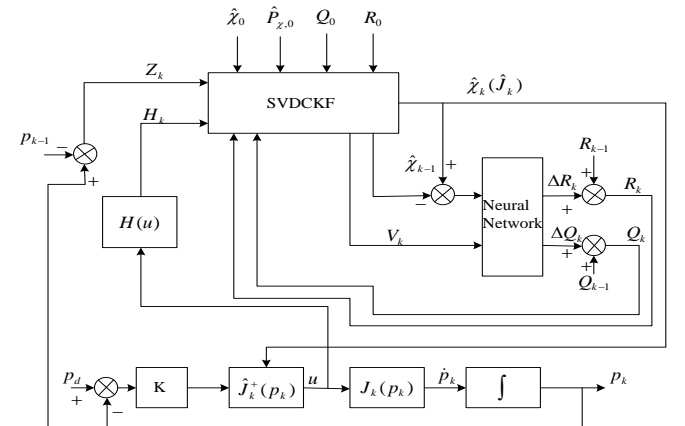


Figure 2. The structure of the proposed Jacobian matrix estimation algorithm

Theorem 1. Considering control plan in Figure 2 and kinematic servoing, the stable system control law under positioning control can be defined as:

$$u = J(p)^+ Ke = J(p)^+ K(p_d - p) \quad (33)$$

where K is the positive definite matrix; p_d is the current position; p is the desired position; e is the image error; $J(p)^+$ is the pseudo inverse of image Jacobian matrix.

Proof: The Lyapunov function can be defined as:

$$V(t) = \frac{1}{2}(e^T e) \quad (34)$$

Taking the derivative of Equation (34), we have:

$$\dot{V}(t) = e^T \dot{e} \quad (35)$$

The following can be derived from Equation (33):

$$\dot{e} = \dot{p}_d - \dot{p} = -\dot{p} = -J(p)u \quad (36)$$

$$\dot{V}(t) = -e^T J(p)u = -e^T J(p)J(p)^+ Ke = -e^T Ke \quad (37)$$

This means the proposed visual servo system is asymptotically stable in its workspace under the control law. In this case, the target image in the camera image plane coordinate system moves towards the desired position, and eventually stabilizes in the desired position. Thus, Theorem 1 is proved.

4. EXPERIMENTAL VERIFICATION

The proposed Jacobian matrix estimation algorithm was verified through a binocular stereo visual servoing experiment on the Motoman UP6 robot experimental platform.

The main hardware of the platform includes a Motoman UP6 arm, an XRC robot control cabinet, a control computer, two charge-coupled device (CCD) cameras and a DH-CG300 image acquisition card. The two cameras were mounted parallel to the end effector of the arm, forming the binocular vision model in Section 2.1, and moved along with the actuator.

The software of the platform was programmed under the VC++6.0 environment. The image acquisition module was implemented using the SDK in the DH-CG300 acquisition card. The feature matching module calls the function `cvGoodFeaturesToTrack()` provided by OpenCV. The image Jacobian matrix was estimated by the proposed SVDCKF with BPNN noise compensator. The generalized inverse matrix was calculated by the SVD function `InversionSingularValue()` in the class of matrix. The robot remote control module was accomplished by MotoCom32 library.

As shown in Figure 3, the experimental platform has a cube (side length: 5.4 cm) as the tracking object. The top surface of the cube was black and the other surfaces were in white, making it easy to match the feature points. The top four vertices of the cube were taken as feature points.



Figure 3. Experimental platform of binocular stereo visual servo system

The parameters of the two CCD cameras are as follows: lens size of 2.8 mm 1/3, general focal length of 303 pixels, the length between the two lens $B=110$ mm, the proportional gain $K=1.25$ and the image resolution of 400×300 . The left camera was set as the origin of the arm coordinate system. In the robot coordinate system, the desired position at the end of the arm can be expressed as (1,012, 222, 609) (mm), and the initial angle as (158, -3, 28) ($^\circ$). Meanwhile, the pixel coordinates of the feature points of the desired object are set as:

$$\begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -16 & 34 & -16 & 35 \\ 116 & 112 & 75 & 71 \\ -89 & -41 & -99 & -48 \\ 121 & 116 & 61 & 77 \end{bmatrix}$$

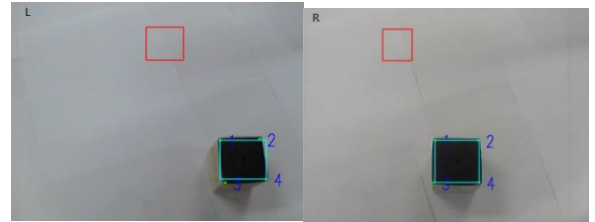


Figure 4. The desired and initial image positions of the object in binocular stereo visual servo system

Thus, we have:

$$p_d = \begin{bmatrix} x_1 \\ y \\ x_2 \end{bmatrix} = \begin{bmatrix} -16 & 34 & -16 & 35 \\ 116 & 112 & 75 & 71 \\ -89 & 41 & -99 & -48 \end{bmatrix}$$

At the initial time, the position of the end of the arm in the robot coordinate system was (1,057, 76, 611) (mm); the initial angle was (158, 2, 30) ($^\circ$).

Then, the pixel coordinates of the binocular image feature points are as follows:

$$\begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 85 & 139 & 92 & 114 \\ 39 & -37 & -97 & -92 \\ -16 & 45 & -17 & 46 \\ -34 & -34 & -93 & -93 \end{bmatrix}$$

Thus, we have:

$$p = \begin{bmatrix} x_1 \\ y \\ x_2 \end{bmatrix} = \begin{bmatrix} 85 & 139 & 92 & 114 \\ 39 & -37 & -97 & -92 \\ -16 & 45 & -17 & 46 \end{bmatrix}$$

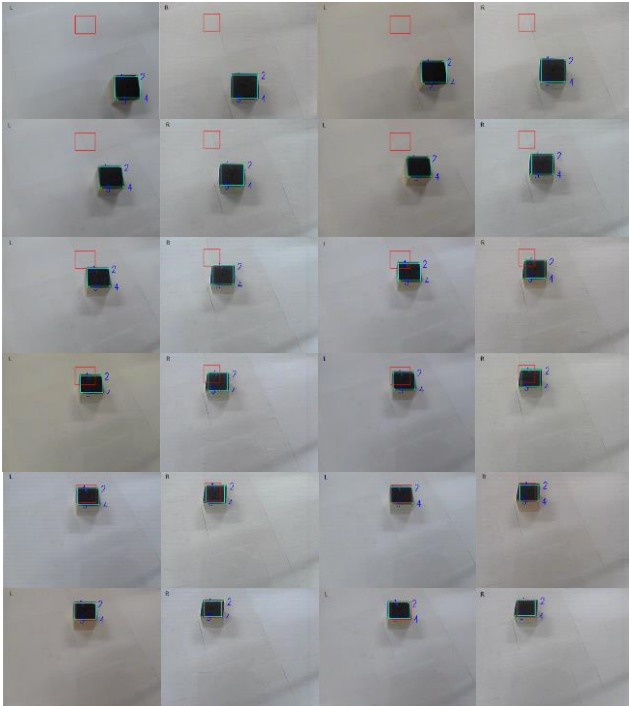


Figure 5. The image sequence captured by the binocular CCD cameras demonstrating the whole process of the visual servoing

The desired target image position (red square frame) and the initial image (green square frame) are shown in Figure 4. In the process of the visual servoing, the successive approximation images for the gradually approaching object are recorded in Figure 5. Obviously, the proposed algorithm can position the servo system.

5. CONCLUSIONS

Unlike monocular vision system, the binocular vision servo system eliminates the need for depth information. In practice, however, the estimation accuracy of image Jacobian matrix is affected by the unknown statistical features of process noise and measurement noise. In this paper, the SVDCKF is adopted to achieve a high filtering accuracy, and the BPNN is taken as a noise compensator to reduce the influence of the statistical features of uncertain noise to the filter. The proposed algorithm was proved as capable of completing servo positioning through an experiment on the Motoman UP6 robot experimental platform.

REFERENCES

[1] Nelson, B.J., Papanikolopoulos, N.P., Khosla, P.K. (2015). Robotic visual servoing and robotic assembly tasks. *IEEE Robotics & Automation Magazine*, 3(2): 23-31. <https://doi.org/10.1109/100.511777>

[2] Mohebbi, A., Keshmiri, M., Xie, W. (2016). A Comparative study of eye-in-hand image-based visual servoing: Stereo vs. Mono. *Journal of Integrated Design & Process Science*, 19(3): 25-54. <http://dx.doi.org/10.3233/jid-2015-0006>

[3] Hom, B.K.P. (1996). *Robot vision*. Cambridge, Massachusetts: MIT Press.

[4] Pomares, J., Gil, P., Torres, F. (2010). Visual control of robots using range images. *Sensors*, 10(8): 7303-7322. <https://doi.org/10.3390/s100807303>

[5] Diosi, A., Segvic, S., Remazeilles, A., Chaumette, F. (2011). Experimental evaluation of autonomous driving based on visual memory and image-based visual servoing. *IEEE Transactions on Intelligent Transportation Systems*, 12(3): 870-883. <https://doi.org/10.1109/TITS.2011.2122334>

[6] Zhong, X.G., Zhong, X.Y., Peng, X.P. (2013). Robust Kalman filtering cooperated Elman neural network learning for vision-sensing-based robotic manipulation with global stability. *Sensors*, 13(10): 13464-13486. <https://doi.org/10.3390/s131013464>

[7] Wang, J.P., Liu, A., Tao, X.D., Chao, H. (2008). Microassembly of micropeg and -hole using uncalibrated visual servoing method. *Precision Engineering*, 32(3): 173-181.

[8] Xin, J., Bai, L., Liu, D. (2014). Adaptive Kalman filter-based robot 6DOF uncalibrated vision positioning. *Journal of System Simulation*, 26(3): 586-591.

[9] Wang, H.B., Li, P. (2009). On-line estimation for image Jacobin matrix based on improved Kalman filter. *Journal of Wuhan University of Technology*, 31(14): 145-148.

[10] Lv, X.D., Huang, X.H. (2006). Fuzzy adaptive Kalman filtering based estimation of image Jacobian for uncalibrated visual servoing. *International Conference on Intelligent Robots and Systems, Beijing, China*, pp. S2167-S2172. <https://doi.org/10.1109/IROS.2006.282555>

[11] Siswanto, J., Prabuwo, A.S., Abdullah, A., Idrus, B. (2016). A linear model based on Kalman filter for improving neural network classification performance. *Expert Systems with Applications an International Journal*, 49(C): 112-122. <https://doi.org/10.1016/j.eswa.2015.12.012>

[12] Gurung, H., Banerjee, A. (2016). Self-sensing SMA actuator using extended Kalman filter and artificial neural network. *Procedia Engineering*, 144: 629-634. <https://doi.org/10.1016/j.proeng.2016.05.054>

[13] Safari, S., Shabani, F., Simon, D. (2014). Multirate multisensor data fusion for linear systems using Kalman filters and a neural network. *Aerospace Science & Technology*, 39: 465-471. <https://doi.org/10.1016/j.ast.2014.06.005>

[14] Zhao, W., Li, H.G., Zou, L.Y. (2016). Singular value decomposition aided cubature Kalman filter with neural network in uncalibrated binocular stereo visual servoing system. *International Journal of Control and Automation*, 9(12): 357-372.

[15] Shang, L., Liu, G.H., Zhang, R., Li, G.T. (2013). An adaptive Kalman filtering algorithm for autonomous orbit determination based-on BP neural network. *Journal of Astronautics*, 34(7): 926-931. <http://dx.chinadoc.cn/10.3873/j.issn.1000-1328.2013.07.006>