Adapted Mapping Estimator in Visual Servoing Control for Model-Free Robotics Manipulator

Jun Tian[1(✉)], Xungao Zhong[2,3(✉)], Jiaguo Luo[2,3], and Xiafu Peng[1]

[1] School of Aerospace Engineering, Xiamen University, Xiamen 361005, China
`tianjun256@163.com`
[2] School of Electrical Engineering and Automation, Xiamen University of Technology, Xiamen 361024, China
`zhongxungao@163.com`
[3] Xiamen Key Laboratory of Frontier Electric Power Equipment and Intelligent Control, Xiamen 361024, China

Abstract. This paper presents a model-free approach to visual servoing control of a robotic manipulator operated in unknown environments. A mapping estimator with the learning network is applied to visual servoing control of model-free robotic manipulator, which can online estimate the vision-motor relationship in a stochastic environment without knowledge of noise statistics. The dynamic mapping identification problems are solved by incorporating the improved Kalman filtering (KF) and network learning techniques, moreover, an observation correlation updating method is used to estimate the variance of the noises via online learning. Various grasping positioning experiments are conducted to verify the proposed approach by using an eye-in-hand robotic manipulator without calibration.

Keywords: Robot manipulation · Visual servoing · Mapping estimator · Model-free robotics system

1 Introduction

Visual-based feedback control is one promising solution for robots to operate in unstructured environments through image features [1–3]. There are many successful implementation of visual feedback control in robots, such as a position-based visual servoing (PBVS) [4] and image-based visual servoing (IBVS) methods are popular tracked [5]. In PBVS, the task is defined in the Cartesian space and the systems retrieve the 3D information about the scene by using the geometric model. Then the pose of the target was estimated with respect to the robotics coordinate system [6, 7]. Thus, this servoing system is inevitably associated with the robotic hand-eye calibration. In consequence, the PBVS is more sensitive to the calibration errors and the image features may disappear from the field of view (FOV) of camera [8, 9]. In contrast, the IBVS method regulates the robot dynamic behavior by using image features from its visual sensor [10–13]. This method does not require the 3D reconstruction of target, but requires the camera calibration and the image depth information. IBVS more suitable for preventing the image feature from leaving the FOV. However, it cannot keep the robot movement insider its workspace, particularly when a large displacement of positioning is required.

It is clear that the robotics calibration model and the image depth information should be provided for mentioned visual servoing (VS) methods for computing the dynamic mapping between the vision space and the robot workspace. Thus, to overcome the difficulties regarding in mapping calculation, we propose a new VS method with mapping estimator, which is treated as a stochastic state estimation problem.

The Kalman filter (KF) is one of the best linear state estimators, and its filtering gain is evaluated from the Gaussian white noise statistics of plant. As the noises are unknown in most real-world applications, the optimality of Kalman filter is never guaranteed. Therefore, some colored noise handling solutions have been presented, such as the dimension extension of Kalman filter [14, 15], LMS-based adaptive filtering [16], adaptive Wiener filtering [17], and others such neural network techniques [18, 19]. However, most of these filtering approaches require the noise variance parameters that are difficult to be derived in most situations.

This paper proposes a model-free approach to visual servoing control of a robotic manipulator operated in unknown noise variance environments. The visual-motor mapping and online estimation are conducted using adaptive Kalman filter with network learning techniques, in which the Kalman filtering model is built by adopting an equivalent observation equation for universal non-Gaussian noise. An observation correlation updating method is used to estimate the variance of the measurement noise via online learning. The network learning adjusts the network weights and enables the noise variances to be dynamically estimated.

The proposed mapping estimator does not require systems calibration and the depth information, and the 2D image measurements are directly used to estimate the desired movement of the robotic manipulator. The grasping positioning tasks are performed by reducing the image error between a set of current and desired image features in the image plane, providing highly flexibility for robot to operate in unknown environments. Extensive experiments are conducted on challenging tasks to verify the significant performance of the proposed approach.

2   The Problem Descriptions

The visual servoing control should firstly estimate a mapping matrix $J(k)$ to describes the dynamic differential relationship between the visual space S and the robotic workspace P, and then construct a controller to derive the end-effector moving $U(k)$ needed to minimize the errors of the image features $F(k)$.

We consider a model-free system without the hand-eye calibration, let $J(k) = \frac{\partial F(k)}{\partial U(k)} \in R^{r \times l}$ be the robot visual-motor mapping matrix. It then be formulated as the state estimation problem with KF techniques, in which the system state vector formed by concatenations of the row and the column elements of $J(t)$, $i.e.$ $J(k) \subset R^{r \times l} \rightarrow X(k) \subset R^n$, $n = r \cdot l$. Assume that the state and observation equations of a robotic system are as follows:

$$X(k) = \varphi X(k-1) + \Gamma \xi(k) \tag{1}$$

$$Z(k) = hX(k) + \upsilon(k) \tag{2}$$

where $Z(k) \in R^m$ is the observation vector, $\varphi \in R^{n \times n}$ and $h \in R^{m \times n}$ are the state transformation and state observation matrix, respectively.

Assume that the system processing noise $\xi(k) \in R^n$ and observation noise $\upsilon(k) \in R^m$ are the random Gaussian white noise sequences with zero mean, and the covariance matrices that are designated by Q_ξ and R_υ respectively. According to the KF equations [20], the system state estimation is based on the following recurrence equations:

$$\hat{X}(k/k-1) = \varphi\hat{X}(k-1/k-1) \tag{3}$$

$$\hat{X}(k/k) = \hat{X}(k/k-1) + K(k/k)(Z(k/k) - C\hat{X}(k/k-1)) \tag{4}$$

$$K(k/k) = P(k/k-1)h^T(hP(k/k-1)h^T + R_\upsilon)^{-1} \tag{5}$$

$$P(k/k-1) = \varphi P(k/k)\varphi^T + Q_\xi \tag{6}$$

$$P(k/k) = (E - K(k)h)P(k/k-1) \tag{7}$$

The mapping online estimation value can be recovered from the system state, *i.e.* $\hat{X}(k/k) \subset R^n \to \hat{J}(k) \subset R^{r \times l}$. Since the observation noise of sensors is not the standard Gaussian white noise sequences, the noise variances are very difficult to be determined. Therefore, the observation Eq. (2) needs to be adjusted, and the noises variances Q_ξ and R_υ should be online estimated before using KF in our visual servoing control system.

3   The Mapping Estimator with Online Learning

Considering the universal non-Gaussian noise model is stationary that can be generated by passing a white noise through a filter:

$$\upsilon(k) = \lambda\upsilon(k-1) + \eta(k-1) \tag{8}$$

where λ is the transition coefficient and $\eta(k)$ is the random white noise sequences with zero mean and the covariances R_η.

Based on Eqs. (1), (2) and (8), we can derive the observation vector:

$$\begin{aligned}Z(k+1) &= hX(k+1) + \upsilon(k+1) \\ &= (h\varphi - \lambda h)X(k) + \lambda Z(k) + h\xi(k) + \eta(k)\end{aligned} \tag{9}$$

Equation (9) is considered as equivalent observation equation, and can be deformed to:

$$\underbrace{(Z(k+1) - \lambda Z(k))}_{Z^*(k)} = \underbrace{(h\varphi - \lambda h)}_{h^*} X(k) + \underbrace{h\Gamma\xi(k) + \eta(k)}_{\upsilon^*(k)} \tag{10}$$

The variance of the observation noise $\upsilon^*(k)$ is computed as:

$$\begin{aligned}R_{\upsilon^*}(k) &= E\left\{(h\Gamma\xi(k) + \eta(k))(h\Gamma\xi(k) + \eta(k))^T\right\} \\ &= h\Gamma Q_\xi(k)(h\Gamma)^T + R_\eta(k)\end{aligned} \tag{11}$$

Then, an observation correlations approach is conducted for online estimation of Q_ξ and R_η in Eq. (11).

According to Eq. (2) and Eq. (10), we have:

$$Z^*(k) = h^*\varphi^i X(k-i) + h^*\Gamma\xi(k) + \upsilon^*(k) \tag{12}$$

Assume that X(k) and $\upsilon^*(k)$ are not correlated, the random series $\{Z^*(k)\}$ is stationary and ergodic. Then the auto-correlation function $C_{z^*}(i)$ of new observation series $\{Z^*(k)\}$ can be derived as follows:

$$\begin{aligned}
C_{Z^*}(i) &= E\left[Z^*(k)Z^{*T}(k-i)\right] \\
&= E\left\{\left[h^*\varphi^i X(k-i) + h^*\Gamma\xi(k) + \upsilon^*(k)\right] \times \left[h^*X(k-i) + \upsilon^*(k-i)\right]\right\} \\
&= h^*\varphi^i \varepsilon_X h^{*T} + h\Gamma Q_\xi (h\Gamma)^T + R_\eta, i \geq 1
\end{aligned} \tag{13}$$

where $\varepsilon_X = E\left[X(k)X(k)^T\right]$. However, $C_{Z^*}(i)$ cannot be computed as Q_ξ and R_η are unknown.

Figure 1 shows an online single layer network learning used to estimate the $C_{Z^*}(i)$ simultaneously. Let the input vector of the network be:

$$X(i) = \left(h^*\varphi^i \varepsilon_X h^{*T}, (h\Gamma)^T, I\right)^T \tag{14}$$

According to Eq. (13), the network output for the i-th training sample is given by:

$$C_{Z^*}(i) = W^T X(i) \tag{15}$$

where W is the network's weight vector shown below:

$$W = \left(I, h\Gamma Q_\xi, R_\eta\right)^T \tag{16}$$

After $Z^*(0),\dots,Z^*(k)$ have been obtained, we can get the estimate of the auto-correlation function $C_{Z^*}(i)$ below:

$$\begin{aligned}
\hat{C}_{Z^*}(i) &= \frac{1}{k+1} \sum_{m=i}^{k} Z(m)Z^T(m-i) \\
&= \frac{1}{k+1}\left(k\hat{C}_{k-1}(i) + Z(k)Z^T(k-i)\right), 1 \leq i \leq k, k \geq 1
\end{aligned} \tag{17}$$

It is obvious that (17) is updated sequentially and $\hat{C}_{Z^*}(i)$ can be estimated online from the previous and new measurements. The network training updates the weights W to minimum the error between $C_{Z^*}(i)$ and $\hat{C}_{Z^*}(i)$, with the training cost function below:

$$E = \frac{1}{k} \sum_{i=1}^{k} \left|\hat{C}_{Z^*}(i) - C_{Z^*}(i)\right|^2 \tag{18}$$

The weight vector is updated according to:

$$W(j+1) = W(j) + \frac{\gamma}{k} \sum_{i=1}^{k} \left(\hat{C}_{z^*}(i) - W^T X(i)\right) X(i) \tag{19}$$

where j and γ are the time instant and the learning rate, respectively.

Thus, we can obtain the noises variances Q_ξ and R_η from weights W. The non-Gaussian noise Kalman filtering structure with the online learning algorithm which called mapping estimator is shown in the Fig. 1.
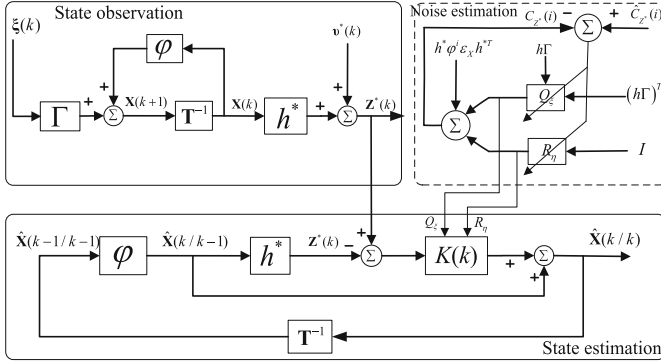

Fig. 1. The structure of mapping estimator based on KF and the learning network, in which the network is used to estimate the variance of the noises.

4 Model-free Robotics Visual Servoing Base on Mapping Estimator

This section gives a model-free visual servoing scheme base on the mapping estimator. The image error function in image plane is defined as follows:

$$e_F(k) = F(k) - F^d \tag{20}$$

where $F(k)=(f_1(k), ..., f_n(k)) \in R^n$ and $F^d=\left(f_1^d, ..., f_n^d\right) \in R^n$ are the n-D current image features and desired, respectively. As the desired features F^d are constant, the derivation of error function (20) is:

$$\dot{e}_F(k) = \frac{d}{dk}\left(F(k) - F^d\right) = \dot{F}(k) \tag{21}$$

For a discretionary manipulation task, the association of the time change of the image feature $F(k)$ with the robot's motion $U(k)$ is done by [8]:

$$\dot{F}(k)=J(k)U(k) \tag{22}$$

Substituting (22) into (21), we have:

$$\dot{e}_F(k) = J(k)U(k) \tag{23}$$

where $U(k)=\left(V(k)\ W(k)\right)^{\mathbf{T}}$ is the robotic control variable, in which $V(k)$ and $W(k)$ are linear and angular velocity of the end-effector, respectively.

There is nonzero constant ρ to make the following equation:

$$\dot{e}_F(k) = -\rho e_F(k) \tag{24}$$

Then substituting (24) into (23), we have the control law:

$$U(k) = -\rho J^+(k)e_F(k) \tag{25}$$

where ρ is the control rate, and $J^+(k)$ is the inverse mapping matrix as follows:

$$J^+(k) = J(k)^T \left(J(k)J(k)^T \right)^{-1} \tag{26}$$

where the mapping matrix $J(k)$ is estimated by mapping estimator shown in Fig. 2.

The steps of the model-free visual servoing are detailed below:

1) Given the desired feature F^d, control rate ρ, mapping matrix $J(0)$, and initial state vector $J(0) \rightarrow X(0/0)$.
2) At the k time, updating the system state $X(k)$ for time k by Eq. (1), and then calculate the observation vector $Z^*(k)$ by Eq. (10).
3) The state estimate $\hat{X}(k/k)$ at k time can be obtained by using the mapping estimator shown in Fig. 1.
4) To conduct mapping estimation $\hat{X}(k/k) \rightarrow \hat{J}(k)$.
5) To control the robot motion by Eq. (25).
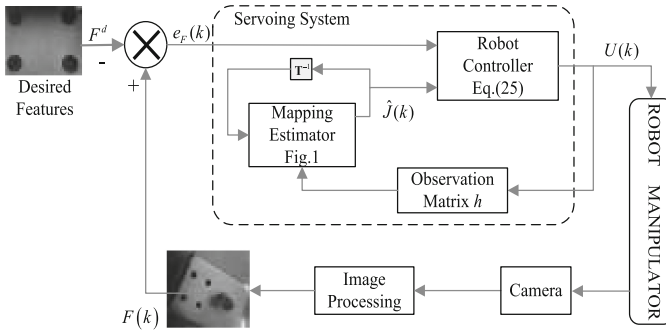6) $k \leftarrow k + 1$, go to Step 2).

Fig. 2. The schema of model-free robotics visual servoing

5   Results and Discussions

For simplicity but without loss of generality, the system state transformation matrix and the noises drive matrix in state Eq. (1) are given by $\Gamma = \varphi = I$. Let the observation vector $Z(k) = \Delta F(k) = J(k)U(k)$, and the observation matrix in Eq. (2) is

$$h = \begin{pmatrix} U(k) & & 0 \\ & \ddots & \\ 0 & & U(k) \end{pmatrix}.$$ The observation noise model is described by the filter (8) λ=0.5.

Thus, the new equivalent observation equation can be obtained by using (10). The noise variances are estimated simultaneously by the observation correlation-based algorithm with a learning network mentioned in Fig. 1.

The real experiment has been carried out by an eye-in-hand robot in our lab, which is shown in Fig. 3. The task is to control the manipulator from an arbitrary initial pose to the desired grasping pose by using the proposed model-free visual servoing with close feedback of image features. The center points of small circular disks on the object board are used as image feature. The control rate ρ in Eq. (25) is selected as 0.25.

In test 1, the initial pose of the robot is far from the desired pose, and the test results are shown in Fig. 4. In test 2, the initial features and desired features are located for the initial pose of the robot relative to the desired pose is rotation combined with translation positioning, the test results are shown in Fig. 5.
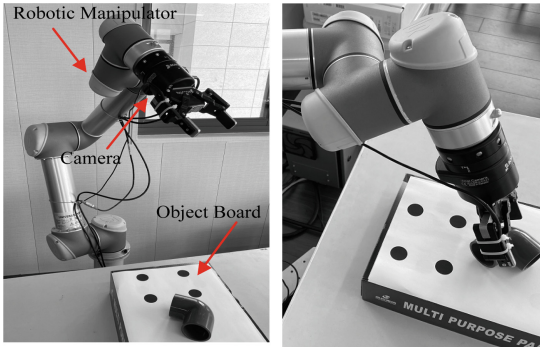

Fig. 3. The eye-in-hand robot system without any calibration parameters, the robot current pose and desired pose.

The FOV of the camera is 640×480 pixels in above two groups of positioning tests. The initial and the desired feature points are set as close as possible to the edge of the FOV. It can be seen from Fig. 4 (a) and Fig. 5 (a) that the motion trajectories of image features are smooth and stable within the FOV, and there are no feature points deviate from the image plane. On the other hand, it can be seen from Fig. 4 (b) and Fig. 5 (b) that the end-effector had the stabile motion without retreat and vibration in the process of the robot positioning. The robot trajectories in Cartesian space are almost the straight line from initial pose to the desired pose, with no conflict among the robot joints.

Figure 6 shows image errors for two tests, in which the error of the feature points uniformly converges, and the positioning steady-state error is within 10 pixels. It is clear that the model-free visual serving controller can provide high positioning precision.

In the following tests, we verify the performance of the mapping estimator and the traditional KF method in Eq. (3) to Eq. (7). Two kinds of estimation approaches are applied to the model-free visual servoing controller. We chose the Gaussian distribution
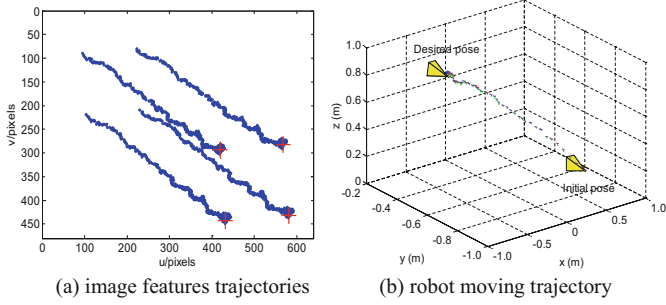
(a) image features trajectories   (b) robot moving trajectory

Fig. 4. Experimental results for Test 1 by using model-free visual servoing.

(a) image features trajectories   (b) robot moving trajectory
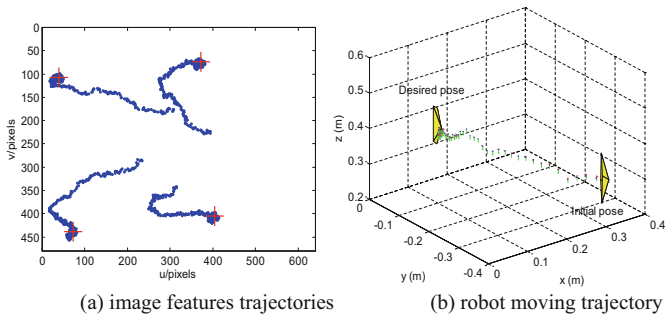
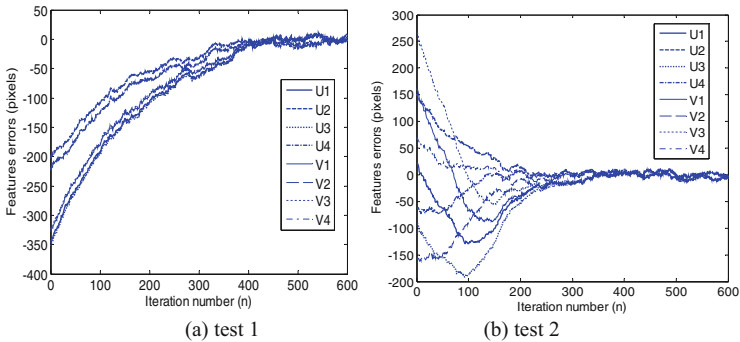**Fig. 5.** Experimental results for Test 2 by using model-free visual servoing.

(a) test 1   (b) test 2

Fig. 6. The image errors during testing

white noise with zero mean. The system noise variance is $Q_\xi=0.02$ and the observation noise variance is $R_\upsilon=0.15$ for the KF method. The mapping estimator uses the proposed approach and the learning network to estimate the variance of the noise online. The robot motion from initial pose to the desired pose has large range moving and a rotation of X and Z axis.

Figure 7 shows the results of mapping estimator using in model-free visual servoing and Fig. 8 shows the results of the traditional KF method. Comparing Fig. 7 (a) with Fig. 8 (a), it can be seen that the image feature trajectories by using our estimator are smoother, short and stable than the feature trajectories by the KF method. From Fig. 7 (b) and Fig. 8 (b), we can see that the robot trajectories in Cartesian space by using proposed estimator is stable without oscillation, while the KF estimation method has large motion oscillation, retreat and serious detour in the same tasks. As can be seen from Fig. 9 the steady-state positioning error for our estimator is smaller than the KF method.

To sum up, the model-free visual servoing with the mapping estimator has eliminated the requirement for the system calibration and target modeling. Also, it has the capability of online estimating the visual-moto mapping in a stochastic environment without the knowledge of noise statistics, and the performances of the robotics is improved greatly.
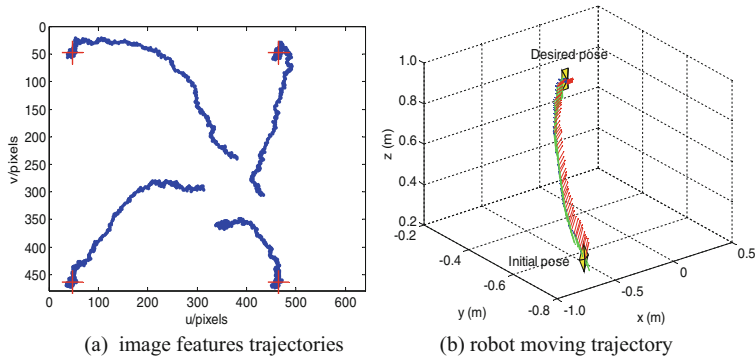

(a) image features trajectories   (b) robot moving trajectory

Fig. 7. Experimental results by using the mapping estimator.

(a) image features trajectories   (b) robot moving trajectory
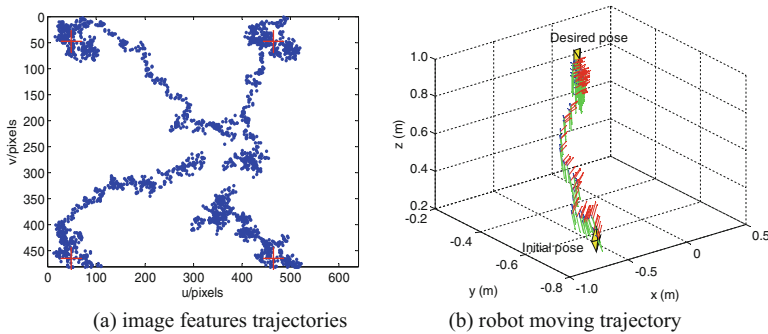
Fig. 8. Experimental results by using the KF method.

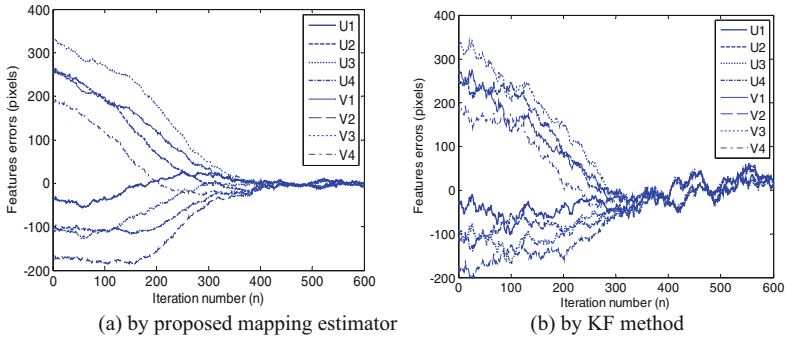(a) by proposed mapping estimator   (b) by KF method

Fig. 9. The image errors

6 Conclusion

In this work, a mapping estimator and a model-free robotic visual servoing scheme have been investigated for robotic grasping manipulation. The proposed mapping estimator can be used in the visual servo system without the need of calibration and the image depth information. Moreover, the mapping identification problems were solved by incorporating KF and network learning techniques. The proposed approach is able to online estimate the vision-motor differential relationship in unknown environments without noise statistical information. Various experiments were conducted by using both the KF and our methods. The results clearly show that the proposed visual servoing with mapping estimator approach outperform the traditional approach in terms of the trajectories of the image features and robot movement in the Cartesian space.

References

1. Said, H.B., Marie, R., Stéphant, J., Labbani-Igbida, O.: Skeleton-based visual servoing in unknown environments. IEEE/ASME Trans. Mechatron. **23**(6), 2750–2761 (2018)
2. Mao, H., Xiao, J.: Real-time conflict resolution of task-constrained manipulator motion in unforeseen dynamic environments. IEEE Trans. Rob. **35**(5), 1276–1283 (2019)
3. Yang, C., Peng, G., Li, Y., Cui, R., Cheng, L., Li, Z.: Neural networks enhanced adaptive admittance control of optimized robot–environment interaction. IEEE Trans. Cybern. **49**(7), 2568–2579 (2019)
4. Josip, M., Mirjana, B., Mojmil, C.: Comparison of uncalibrated model-free visual servoing methods for small-amplitude movements: a simulation study. Int. Jour. Adv. Robot. Syst. **108**(11), 1–16 (2014)
5. Farrokh, J.S., Deng, L., Wilson, W.J.: Comparison of basic visual servoing methods. IEEE/ASME Trans. Mechatron. **16**(5), 967–983 (2011)

6. Chaumette, F., Hutchinson, S.: Visual servo control. part I: basic approaches. IEEE Robot. Autom. Mag. **13**(4), 82–90 (2006)
7. Jiang, P., Cheng, Y., Wang, X., Feng, Z.: Unfalsified visual servoing for simultaneous object recognition and pose tracking. IEEE Trans. Cybern. **46**(12), 3032–3046 (2016)
8. Chaumette, F., Hutchinson, S.: Visual servo control. part II: Advanced approaches. IEEE Robo. Auto. Maga. **14**(1), 109–118 (2007)
9. Do-Hwan, P., Jeong-Hoon, K., In-Joong, H.: Novel position-based visual servoing approach to robust global stability under field-of-view constraint. IEEE Trans. Ind. Electron. **59**(12), 4735–4752 (2012)
10. Keshmiri, M., Xie, W.: Image-based visual servoing using an optimized trajectory planning technique. IEEE/ASME Trans. Mech. **22**(1), 359–370 (2017)
11. Shi, H., Sun, G., Wang, Y., Hwang, K.S.: Adaptive image-based visual servoing with temporary loss of the visual signal. IEEE Trans. Ind. Infor. **15**(4), 1956–1965 (2019)
12. Liang, X., Wang, H., Liu, Y., You, B., Liu, Z., Chen, W.: Calibration-free image-based trajectory tracking control of mobile robots with an overhead camera. IEEE Trans. Autom. Sci. Eng. **17**(2), 933–946 (2020)
13. Xie, H., Lynch, A.F., Low, K.H., Mao, S.: Adaptive output-feedback image-based visual servoing for quadrotor unmanned aerial vehicles. IEEE Trans. Control Syst. Technol. **28**(3), 1034–1041 (2020)
14. Gibson, J.D., Koo, B., Gray, S.D.: Filtering of colored noise for speech enhancement and coding. IEEE Trans. Signal Process. **39**(8), 1732–1742 (1991)
15. Alanis, A.Y., Sanchez, E.N., Loukianov, A.G., Perez-Cisneros, M.A.: Real-time discrete neural block control using sliding modes for electric induction motors. IEEE Trans. Control Syst. Technol. **18**(1), 11–21 (2010)
16. Soni, T., Rao, B.D., Zeidler, J.R., Ku, W.H.: Enhancement of images using the 2-D LMS adaptive algorithm. In: Proceedings of IEEE International Conference on Acoustic Speech and Signal Processing, pp. 3029–3032. IEEE, Toronto, ON, Canada (1991)
17. Stephanakis, I.M., Kollias, S.D.: Wavelet-based approach to adaptive wiener filtering of images in the presence of colored noise. Opt. Eng. **39**(5), 1353–1363 (2000)
18. Khanesar, M.A., Kayacan, E., Teshnehlab, M., Kaynak, O.: Extended Kalman filter based learning algorithm for type-2 fuzzy logic systems and its experimental evaluation. IEEE Trans. Ind. Electron. **59**(11), 4443–4455 (2012)
19. Zhan, R., Wan, J.: Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation. IEEE Signal Process. Lett. **13**(7), 445–448 (2006)
20. Kalman, R.E.: A new approach to linear filtering and prediction problems. Trans. ASME-J. Basic Eng. **82**(1), 34–45 (1960)