



OPEN ACCESS

EDITED BY
Predrag S. Stanimirovic,
University of Niš, Serbia

REVIEWED BY
Liangming Chen,
Chinese Academy of Sciences (CAS), China
Yang Shi,
Yangzhou University, China

*CORRESPONDENCE
Zhengtai Xie
✉ xzt1xzt@vip.qq.com

RECEIVED 01 February 2024

ACCEPTED 22 February 2024

PUBLISHED 20 March 2024

CITATION

Wen L and Xie Z (2024) A data-driven acceleration-level scheme for image-based visual servoing of manipulators with unknown structure. *Front. Neurobot.* 18:1380430. doi: 10.3389/fnbot.2024.1380430

COPYRIGHT

© 2024 Wen and Xie. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A data-driven acceleration-level scheme for image-based visual servoing of manipulators with unknown structure

Liuyi Wen^{1,2} and Zhengtai Xie^{1,3*}

¹The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Qinghai Normal University, Xining, China, ²School of Arts, Lanzhou University, Lanzhou, China, ³School of Information Science and Engineering, Lanzhou University, Lanzhou, China

The research on acceleration-level visual servoing of manipulators is crucial yet insufficient, which restricts the potential application range of visual servoing. To address this issue, this paper proposes a quadratic programming-based acceleration-level image-based visual servoing (AIVS) scheme, which considers joint constraints. Besides, aiming to address the unknown problems in visual servoing systems, a data-driven learning algorithm is proposed to facilitate estimating structural information. Building upon this foundation, a data-driven acceleration-level image-based visual servoing (DAIVS) scheme is proposed, integrating learning and control capabilities. Subsequently, a recurrent neural network (RNN) is developed to tackle the DAIVS scheme, followed by theoretical analyses substantiating its stability. Afterwards, simulations and experiments on a Franka Emika Panda manipulator with eye-in-hand structure and comparisons among the existing methods are provided. The obtained results demonstrate the feasibility and practicality of the proposed schemes and highlight the superior learning and control ability of the proposed RNN. This method is particularly well-suited for visual servoing applications of manipulators with unknown structure.

KEYWORDS

recurrent neural network (RNN), image-based visual servoing (IBVS), data-driven technology, acceleration level, learning and control

1 Introduction

Robots can accurately perform complex tasks and have become a vital driving force in industrial production (Agarwal and Akella, 2024). Among industrial robots, redundant robots, equipped with multiple degrees of freedom (DOFs), have gained significant recognition and favor due to their exceptional flexibility and automation capabilities (Tang and Zhang, 2022; Zheng et al., 2024). Therefore, numerous control schemes are designed to extend the application range of redundant robots, such as medical services (Zeng et al., 2024) and visual navigation (Wang et al., 2023). Furthermore, in these application scenarios, information on the external environment and the robot's status is acquired from various sensors, especially for the image capture of visual information (Jin et al., 2023). Therefore, unknown situations inevitably exist caused by sensor limitations, environmental variability, and robot modification, which hinder the evolution of robot applications. To address this issue, intelligent algorithms based on data-driven technology are exploited to process the acquired information and convert it into knowledge to drive the regular operation of the robot system (Na et al., 2021; Xie et al., 2022). Yang et al. (2019) construct a robot learning system by improving the adaptive ability of a robot with the information interaction between the robot and environment, which enhances the safety and reliability of robot applications in reality Peng et al. (2023). Li et al. (2019) investigate a

model-free control method to cope with the unknown Jacobian problems inside the robot system. On this basis, the dynamic estimation method of robot parameters is researched in the study by Xie and Jin (2023). However, the aforementioned methods primarily operate at the joint velocity level and cannot directly be applicable to robots driven by joint acceleration.

As a crucial robot application, visual servoing simulates the bionic system of human eyes, which can obtain information about real objects through optical devices, thus dynamically responding to a visible object. The fundamental task of visual servoing is to impose the error between the corresponding image feature and the desired static reference to approach zero (Zhu et al., 2022). According to the spatial position or image characteristics of the robot, the visual servoing system can be categorized into two types: position-based visual servoing (PBVS) system (Park et al., 2012), which utilizes 3-D position and orientation information to adjust the robot's state, and image-based visual servoing (IBVS) system, which utilizes 2-D image information for guidance (Van et al., 2018). Recently, the research on visual servoing has achieved many unexpected results (Hashimoto et al., 1991; Malis et al., 2010; Zhang et al., 2017; Liang et al., 2018). For instance, visual servoing is applied to bioinspired soft robots in the underwater environment with an adaptive control method, which extends the scope of visual servoing (Xu et al., 2019). Based on the neural network method, a resolution scheme for IBVS is developed at the velocity level. This enables the manipulator to accurately track fixed desired pixels, resulting in fast convergence (Zhang and Li, 2018). However, the aforementioned methods are difficult to deal with the emergence of unknown conditions, such as focal length change, robot abrasion, or parameter variation. This is because these methods rely on accurate structural information of the robot vision system. To tackle this challenge, this study focuses on data-driven control of visual servoing for robots with an unknown Jacobian matrix.

Neural networks have gained significant recognition as powerful tools for solving challenging problems, such as automatic drive (Jin et al., 2024), mechanism control (Xu et al., 2023), and mathematical calculation (Zeng et al., 2003; Stanimirovic et al., 2015). In robot redundancy analysis, neural networks have shown superior performance. In recent decades, numerous control laws based on neural networks have been developed to harness the potential of redundant manipulators (Zhang and Li, 2023). One specific application of the neural network approach addresses the IBVS problem. In this context, the IBVS problem is formulated as a quadratic programming scheme and tackled using a recurrent neural network (RNN). The RNN drives the robot vision system's feature to rapidly converge toward the desired point (Zhang et al., 2017). Additionally, Li et al. (2020) investigate an inverse-free neural network technique to deal with the IBVS task, ensuring that the error approaches zero within a finite time while considering the manipulator's physical constraints.

Most control schemes accomplish the given task at the velocity level, especially for visual servoing applications (Hashimoto et al., 1991; Malis et al., 2010; Zhang et al., 2017; Liang et al., 2018; Van et al., 2018; Zhang and Li, 2018; Xu et al., 2019; Li et al., 2020). These velocity-level schemes control redundant robots via joint velocities. However, when confronted with acceleration or torque-driven robots, the velocity-level schemes exhibit limitations

and cannot provide precise control. Furthermore, the velocity-level scheme may yield abrupt joint velocities that are impractical in real-world applications. Consequently, research on acceleration-level visual servoing for robot manipulators has become crucial (Keshmiri et al., 2014; Anwar et al., 2019). Motivated by the issues above, this study investigates the application of visual servoing in robots at the acceleration level. The technical route of this study is shown in Figure 1. As illustrated, the contributions of this study are shown as follows:

- An acceleration-level image-based visual servoing (AIVS) scheme is designed, taking into account multiple joint constraints.
- Considering potential unknown factors in the visual servoing system, a data-driven acceleration-level image-based visual servoing (DAIVS) scheme is developed, enabling simultaneous learning and control.
- RNNs are proposed to solve the AIVS scheme and DAIVS scheme, enabling visual servoing control of the manipulator. Theoretical analyses guarantee the stability of the RNNs.

In addition, the feasibility of the proposed schemes is demonstrated through simulative and experimental results conducted on a Franka Emika Panda manipulator with an eye-in-hand structure.

Before concluding this section, the remaining sections of the study are shown as follows. Section 2 presents the robot kinematics of visual servoing and introduces the data-driven learning algorithm, formulating the problem at the acceleration level. Section 3 constructs an AIVS scheme with the relevant RNN. Subsequently, considering the unknown factors, a DAIVS scheme and corresponding RNN are proposed, and theoretical analyses proved the learning and control ability of the RNN, as shown in Section 4. Section 5 provides abundant simulations and performance comparisons, embodying the proposed method's validity and superiority. Section 6 displays physical experiments on a real manipulator. Finally, Section 7 briefly concludes this study.

2 Preliminaries

In this section, the robot visual servoing kinematics and data-driven learning algorithm are introduced as the preliminaries. Note that this study specifically tackles the problem at the acceleration level.

2.1 Robot visual servoing kinematics

The forward kinematics, which contains the transformation between the joint angle $\phi(t) \in \mathbb{R}^m$ of a robot and the end-effector position and posture $\mathbf{s}(t) \in \mathbb{R}^6$, can be expressed as follows:

$$f(\phi(t)) = \mathbf{s}(t), \quad (1)$$

where $f(\cdot)$ is the non-linear mapping related to the structure of the robot. In view of strongly non-linear and redundant characteristics

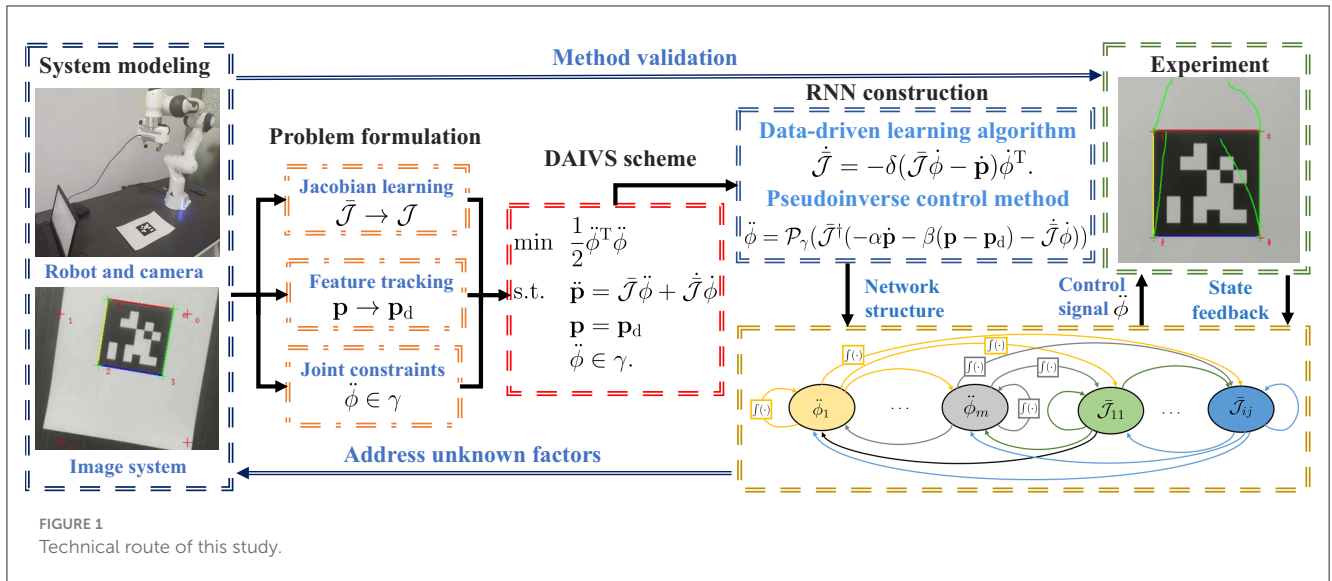


FIGURE 1
Technical route of this study.

of $f(\cdot)$, it is difficult to obtain the desired angle information directly from the desired end-effector information $\mathbf{s}_d(t)$, i.e., $\mathbf{s}(t) = \mathbf{s}_d(t)$. By taking the time derivative of both sides of Equation (1), one can deduce

$$J_{ro}\dot{\phi}(t) = \dot{\mathbf{s}}(t), \quad (2)$$

where $\dot{\phi}(t)$ denotes the joint velocity; $\dot{\mathbf{s}}(t)$ covers the joint velocity and translational velocity of the end-effector; $J_{ro} = \partial f(\phi(t))/\partial \phi(t) \in \mathbb{R}^{6 \times m}$ stands for the robot Jacobian matrix. Owing to the physical properties of manipulators, output control signals based on design formulas and intelligent calculations may not be suitable for the normal operation of real robots. Therefore, to ensure the protection of the robot, it is crucial to take into account the following joint restrictions:

$$\begin{aligned} \phi^- &\leq \phi \leq \phi^+ \\ \dot{\phi}^- &\leq \dot{\phi} \leq \dot{\phi}^+ \\ \ddot{\phi}^- &\leq \ddot{\phi} \leq \ddot{\phi}^+, \end{aligned}$$

where ϕ^- , $\dot{\phi}^-$, and $\ddot{\phi}^-$ signify the lower bounds of joint angle, joint velocity, and joint acceleration, respectively; ϕ^+ , $\dot{\phi}^+$ and $\ddot{\phi}^+$ denote the upper bounds of joint angle, joint velocity, and joint acceleration, respectively. Utilizing the special conversion techniques (Zhang and Zhang, 2012; Xie et al., 2022), the joint restrictions would be integrated into the acceleration level as $\ddot{\phi} \in \gamma$, where $\gamma = \{g \in \mathbb{R}^m, \gamma^- \leq g \leq \gamma^+\}$ is the safe range of joints with γ^- and γ^+ denoting the lower bound and upper bound of γ , respectively. In detail, the i -th elements of γ^- and γ^+ are designed as

$$\begin{aligned} \gamma_i^- &= \max\{\mu(\phi_i^- + \theta_i - \phi_i), \nu(\dot{\phi}_i^- - \dot{\phi}_i), \ddot{\phi}_i^-\} \\ \gamma_i^+ &= \min\{\mu(\phi_i^+ - \theta_i - \phi_i), \nu(\dot{\phi}_i^+ - \dot{\phi}_i), \ddot{\phi}_i^+\}, \end{aligned}$$

where $i = 1, 2, 3, \dots, m$; $\mu > 0$ and $\nu > 0$ are designed to select the feasible region for different levels; θ_i is the margin to ensure that the acceleration has a sufficiently large feasible region (Xie et al., 2022). Then, a brief introduction to the visual servoing system is presented

as follows. Regarding visual servoing tasks, the number of features determines the complexity of a visual servoing system. Simply considering a visual servoing system with one feature, a miniature camera is mounted on the end-effector of the manipulator and moves with the end-effector. Figure 2 illustrates the geometric transformation in different coordinate systems. Three-dimensional space with O_{ca} as the original point and $[X, Y, Z]$ as the coordinate axis is called the camera system with the internal coordinate point $\mathbf{q} = [x, y, z]^T$. Relatively, with O_{im} as the center point, the image system is the two-dimensional space with the projection pixel point of \mathbf{q} being $[p_x, p_y]^T$ and the pixel coordinate being $\mathbf{p} = [p_u, p_v]^T \in \mathbb{R}^2$. According to the similar triangle, it can be readily obtained in the study by Zhang et al. (2017) and Zhang and Li (2018):

$$\begin{bmatrix} p_x \\ p_y \end{bmatrix} = \frac{l}{z} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

and

$$\begin{aligned} p_u &= u_0 + a_x p_x \\ p_v &= v_0 + a_y p_y, \end{aligned} \quad (4)$$

with l standing for the focal length of the camera; u_0 and v_0 denoting the pixel coordinate of principle point; and $[a_x, a_y]^T$ standing for the conversion scale. Based on Equations (3, 4), the image Jacobian matrix $J_{im}(\mathbf{p}, z) \in \mathbb{R}^{2 \times 6}$ is defined using the following relationship (Liang et al., 2018):

$$J_{im}(\mathbf{p}, z)\dot{\mathbf{s}} = \dot{\mathbf{p}}, \quad (5)$$

where $\dot{\mathbf{p}}$ stands for the movement velocity of the pixel coordinate and

$$J_{im}(\mathbf{p}, z) = H \begin{bmatrix} -\frac{l}{z} & 0 & \frac{lp_x}{z} & \frac{p_x p_y}{l} & -\frac{p_x^2 + l^2}{l} & p_y \\ 0 & -\frac{l}{z} & \frac{lp_y}{z} & -\frac{p_y^2 + l^2}{l} & -\frac{p_x p_y}{l} & -p_x \end{bmatrix},$$

with

$$p_x = \frac{p_u - u_0}{a_x}, \quad p_y = \frac{p_v - v_0}{a_y}, \quad H = \begin{bmatrix} a_x & 0 \\ 0 & a_y \end{bmatrix}.$$

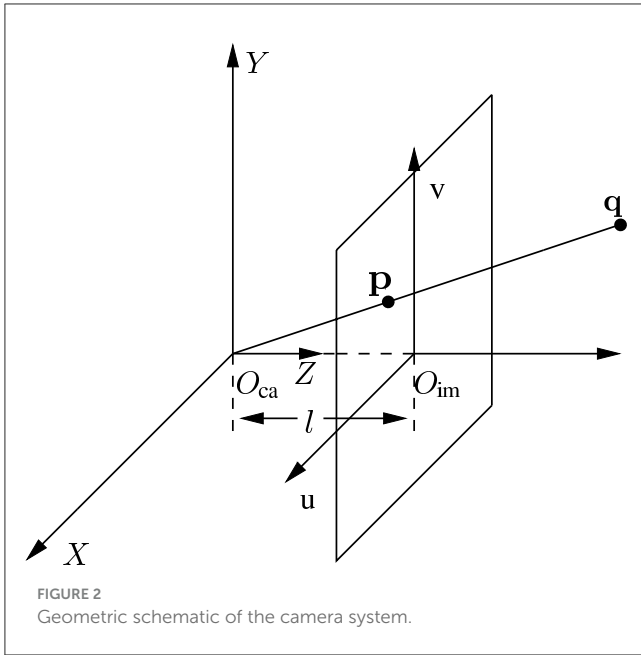


FIGURE 2 Geometric schematic of the camera system.

For the sake of convenience, Equations (2, 5) can be combined as follows:

$$\mathcal{J}\dot{\phi} = \dot{\mathbf{p}}, \tag{6}$$

with $\mathcal{J} = J_{im}(\mathbf{p}, z)J_{ro} \in \mathbb{R}^{2 \times m}$ defined as the visual Jacobian matrix. The relationship between joint space and image space is established directly by Equation (6) at the velocity level. Taking the time derivatives of both sides of Equation (6) generates

$$\dot{\mathcal{J}}\dot{\phi} + \mathcal{J}\ddot{\phi} = \ddot{\mathbf{p}}, \tag{7}$$

where $\dot{\mathcal{J}}$ is the time derivative of \mathcal{J} ; $\ddot{\phi}$ denotes the joint acceleration; and $\ddot{\mathbf{p}}$ stands for the movement acceleration of the pixel coordinate. When it comes to a complicated situation with more features, the above analyses still hold under the requirements of appropriate dimensions. It is worth noting that a single feature is analyzed as an example for simple illustration. When the number of features increases, the principle of coordinate transformation remains unchanged along with the increase in dimension.

2.2 Data-driven learning algorithm

However, unknown conditions may exist in the robot visual servoing system, such as focal length changes or robot modifications. In this regard, it could not control the robot accurately to execute the IBVS task based on $\dot{\mathcal{J}}$. Hence, motivated by this issue, a data-driven learning algorithm is designed as follows. To begin with, a virtual IBVS system is established, incorporating the virtual visual Jacobian matrix $\bar{\mathcal{J}} \in \mathbb{R}^{2 \times m}$ and the following relationship:

$$\bar{\mathcal{J}}\dot{\phi} = \bar{\dot{\mathbf{p}}},$$

where $\bar{\dot{\mathbf{p}}} \in \mathbb{R}^2$ is the virtual pixel velocity determined by the virtual robot and $\dot{\phi}$ is the joint velocity measured in real time from

the robot. Beyond dispute, the goal of the data-driven learning algorithm is to guarantee that $\bar{\dot{\mathbf{p}}}$ can rapidly converge to the real pixel velocity $\dot{\mathbf{p}}$. Thereout, an error function is devised as $\ell = \|\bar{\dot{\mathbf{p}}} - \dot{\mathbf{p}}\|_2^2/2$, where $\|\cdot\|_2$ is the Euclidean norm of a vector. On the basis of the gradient descent method (Stanimirovic et al., 2015) to minimize the error function along the negative gradient direction, one can get

$$\dot{\bar{\mathcal{J}}} = -\delta \frac{\partial \ell}{\partial \bar{\mathcal{J}}} = -\delta(\bar{\mathcal{J}}\dot{\phi} - \dot{\mathbf{p}})\dot{\phi}^T, \tag{8}$$

where $\dot{\bar{\mathcal{J}}}$ is the time derivation of $\bar{\mathcal{J}}$; $\delta > 0$ denotes the coefficient that controls the convergence rate. Hereinafter, $\dot{\bar{\mathcal{J}}}$ and $\bar{\mathcal{J}}$ are used to replace the calibrated parameter $\dot{\mathcal{J}}$ and \mathcal{J} to deal with the unknown situations. This method directly explores the relationship between joint space and image space without the utilization of $\dot{\mathcal{J}}$ and \mathcal{J} . It is worth highlighting that Equation (8) does not involve real structural information and estimates structural information from the joint velocity $\dot{\phi}$ and velocity of the pixel coordinate $\dot{\mathbf{p}}$ measured by sensors, which belongs to the core idea of the data-driven learning algorithm.

3 Acceleration-level IBVS solution

In this section, an AIVS scheme is proposed with joint constraints considered. Subsequently, we propose a corresponding RNN and provide theoretical analyses. Note that the presented method requires an accurate visual Jacobian matrix.

3.1 AIVS scheme

It is worth pointing out that there are few acceleration-level robot control schemes for dealing with IBVS problems. None of the existing acceleration-level solutions take joint constraints into account (Keshmiri et al., 2014; Anwar et al., 2019). In this regard, considering joint constraints, acceleration control, and visual servoing kinematics, the AIVS scheme is constructed as a quadratic programming problem, taking the following form:

$$\text{minimize } \frac{1}{2}\ddot{\phi}^T \ddot{\phi} \tag{9a}$$

$$\text{subject to } \ddot{\mathbf{p}} = \mathcal{J}\ddot{\phi} + \dot{\mathcal{J}}\dot{\phi} \tag{9b}$$

$$\mathbf{p} = \mathbf{p}_d \tag{9c}$$

$$\ddot{\phi} \in \gamma, \tag{9d}$$

where \mathbf{p}_d denotes the desired pixel coordinate. As a result, the goal of AIVS scheme (9) is to make the end-effector track the desired pixel point. In addition, according to robot Jacobian matrix J_{ro} and the image Jacobian matrix J_{im} , the visual Jacobian matrix \mathcal{J} and its time derivative $\dot{\mathcal{J}}$ are determined by the structure and parameters of the robot and the parameter settings inside the camera. Hence, if there are any changes in the internal parameters or structures, leading to an unknown state, the accuracy of \mathcal{J} and $\dot{\mathcal{J}}$ may be compromised, potentially leading to a decline in performance. In contrast to velocity-level visual servoing schemes (Hashimoto et al., 1991; Malis et al., 2010; Zhang et al., 2017; Liang et al., 2018; Van et al., 2018; Zhang and Li, 2018; Xu et al.,

2019; Li et al., 2020), the proposed AIVS scheme (9) offers two advantages. First, it utilizes joint acceleration as the control signal, resulting in continuous joint velocities. This helps mitigate the issues associated with excessive and discontinuous joint velocities. Second, AIVS scheme (9) takes into account the equality and inequality constraints at the acceleration level. This allows for a more comprehensive consideration of constraints, expanding the range of applications.

3.2 RNN solution and theoretical analysis

For the AIVS scheme (9), the pseudoinverse method is applied to generate the relevant RNN solution (Cigliano et al., 2015; Li et al., 2020). Primarily, as reported in the study by Zhang and Zhang (2012) and Xie et al. (2022), one can readily extend pixel coordinate error $\mathbf{p} - \mathbf{p}_d$ into the acceleration level by neural dynamics method (Liufu et al., 2024) as

$$\ddot{\mathbf{p}} - \ddot{\mathbf{p}}_d = -\alpha(\dot{\mathbf{p}} - \dot{\mathbf{p}}_d) - \beta(\mathbf{p} - \mathbf{p}_d), \quad (10)$$

where the design parameter $\alpha > 0$ and $\beta > 0$; $\dot{\mathbf{p}}_d$ and $\ddot{\mathbf{p}}_d$ are the desired velocity and the desired acceleration of the pixel coordinates, respectively. It is worth pointing out that the desired pixel coordinates \mathbf{p}_d is a constant, thus $\dot{\mathbf{p}}_d = \ddot{\mathbf{p}}_d = 0$. As a result, Equation (10) can be rearranged as

$$\ddot{\mathbf{p}} = -\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d). \quad (11)$$

Substituting Equation (11) into Equation (9b), it could be obtained:

$$\dot{\mathcal{J}}\dot{\phi} + \mathcal{J}\ddot{\phi} = -\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d).$$

In light of the pseudoinverse method, the joint acceleration can be minimized with the following formula:

$$\ddot{\phi} = \mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi}), \quad (12)$$

where superscript \dagger denotes the pseudoinverse operation of a matrix with $\mathcal{J}^\dagger = \mathcal{J}^T(\mathcal{J}\mathcal{J}^T)^{-1}$. It is deserved to note that Equation (12) is employed in the study by Keshmiri et al. (2014) and Anwar et al. (2019) to generate the acceleration command for a manipulator. However, the research in the study by Keshmiri et al. (2014) and Anwar et al. (2019) does not consider joint constraints of the manipulator. To address this problem, the RNN corresponding to the AIVS scheme (9) is derived as

$$\dot{\phi} = \mathcal{P}_\gamma(\mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi})), \quad (13)$$

where projection function $\mathcal{P}_\gamma(c) = \operatorname{argmin}_{b \in \gamma} \|b - c\|_2$. Furthermore, theoretical analyses regarding the convergence of RNN (13) are presented as follows.

Theorem 1: The pixel error $\xi = \mathbf{p} - \mathbf{p}_d$ driven by AIVS scheme (9) assisted with RNN (13) globally converges to a zero vector.

Proof: According to Equations (7, 13), one has

$$\dot{\mathbf{p}} = \dot{\mathcal{J}}\dot{\phi} + \mathcal{J}\ddot{\phi} = \mathcal{P}_\gamma(\mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi})) + \dot{\mathcal{J}}\dot{\phi}.$$

Due to the fact that \mathbf{p}_d is a fixed feature, error function ξ can be readily derived as

$$\dot{\xi} = \mathcal{P}_\gamma(\mathcal{J}^\dagger(-\alpha\dot{\xi} - \beta\xi - \dot{\mathcal{J}}\dot{\phi})) + \dot{\mathcal{J}}\dot{\phi}.$$

By considering the projection function, a substitution matrix \mathcal{H} is designed to replace $\mathcal{P}_\gamma(\cdot)$, leading to

$$\dot{\xi} = \mathcal{H}\mathcal{J}^\dagger(-\alpha\dot{\xi} - \beta\xi - \dot{\mathcal{J}}\dot{\phi}) + \dot{\mathcal{J}}\dot{\phi}, \quad (14)$$

of which

$$\mathcal{H} = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_m \end{bmatrix} \in \mathbb{R}^{m \times m}$$

and

$$h_i = \frac{(\mathcal{P}_\gamma(\mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi})))_i}{(\mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi}))_i} \in (0, 1].$$

By matrix decomposition, structural analyses of matrix $\mathcal{H}\mathcal{J}^\dagger = [a_{11}, a_{12}; a_{21}, a_{22}] \in \mathbb{R}^{2 \times 2}$ are given as follows:

$$\mathcal{H}\mathcal{J}^\dagger = \mathcal{L}L^T\mathcal{J}^T(\mathcal{J}\mathcal{J}^T)^{-1},$$

where $L = \sqrt{\mathcal{H}}$. In this regard, matrix $\mathcal{H}\mathcal{J}^\dagger$ can be viewed as the product of two positive definite matrices. It is evident that the eigenvalues of $\mathcal{H}\mathcal{J}^\dagger$ are greater than zero and $\det(\mathcal{H}\mathcal{J}^\dagger) = \det(\mathcal{L}L^T\mathcal{J}^T)\det((\mathcal{J}\mathcal{J}^T)^{-1}) > 0$ with $\det(\cdot)$, denoting the determinant of a matrix. According to the properties of the diagonal elements of the matrix, it can be concluded that the diagonal elements of $\mathcal{H}\mathcal{J}^\dagger$ are greater than zero ($a_{11} > 0, a_{22} > 0$). Furthermore, Equation (12) can be rewritten as

$$\begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} -\alpha\dot{\xi}_1 - \beta\xi_1 - (\dot{\mathcal{J}}\dot{\phi})_1 \\ -\alpha\dot{\xi}_2 - \beta\xi_2 - (\dot{\mathcal{J}}\dot{\phi})_2 \end{bmatrix} + \begin{bmatrix} (\dot{\mathcal{J}}\dot{\phi})_1 \\ (\dot{\mathcal{J}}\dot{\phi})_2 \end{bmatrix},$$

and further we get

$$\dot{\xi}_1 + a_{11}\alpha\dot{\xi}_1 + a_{11}\beta\xi_1 = -a_{12}(\alpha\dot{\xi}_2 + \beta\xi_2 + (\dot{\mathcal{J}}\dot{\phi})_2) + (1 - a_{11})(\dot{\mathcal{J}}\dot{\phi})_1$$

and

$$\dot{\xi}_2 + a_{22}\alpha\dot{\xi}_2 + a_{22}\beta\xi_2 = -a_{21}(\alpha\dot{\xi}_1 + \beta\xi_1 + (\dot{\mathcal{J}}\dot{\phi})_1) + (1 - a_{22})(\dot{\mathcal{J}}\dot{\phi})_2,$$

which can be regarded as a perturbed second-order constant coefficient differential equation with respect to ξ . In conclusion, pixel error ξ is able to converge exponentially. To illustrate the steady state of the system (Equation 14), further derivations continue to be given. As the pixel error decreases, all joint properties return to the interior of joint constraints. In this sense, joint properties, i.e., $\dot{\phi}$, ϕ and ϕ , are all inside the joint limits with $h_i = 1$. Therefore, Equation (14) can be reorganized as

$$\ddot{\xi}(t) + \alpha\dot{\xi}(t) + \beta\xi(t) = 0. \quad (15)$$

It is worth mentioning that Equation (15) can be regarded as a second-order constant coefficient differential equation with regard to ξ . Moreover, the solutions of Equation (15) can be segmented into three subcases on account of different settings of α and β , given the original state $\xi(0) = \mathbf{p} - \mathbf{p}_d$.

Subcase I: As for $\alpha^2 - 4\beta > 0$, the characteristic roots could be obtained simply as $\mathcal{R}_1 = (-\alpha + \sqrt{\alpha^2 - 4\beta})/2$ and $\mathcal{R}_2 = (-\alpha - \sqrt{\alpha^2 - 4\beta})/2$ with real number $\mathcal{R}_1 \neq \mathcal{R}_2$. Therefore, one can readily deduce

$$\xi(t) = \xi(0)(D_1 \exp(\mathcal{R}_1 t) + D_2 \exp(\mathcal{R}_2 t)),$$

with $D_1 = \alpha/(2\sqrt{\alpha^2 - 4\beta}) + 1/2$ and $D_2 = 1/2 - \alpha/(2\sqrt{\alpha^2 - 4\beta})$

Subcase II: As to $\alpha^2 - 4\beta = 0$, calculating characteristic roots generates $\mathcal{R}_1 = \mathcal{R}_2 = -\alpha/2$. Hence, it can be readily obtained:

$$\xi(t) = \xi(0)\exp(-\alpha/2t)(1 + \alpha/2t).$$

Subcase III: As to $\alpha^2 - 4\beta < 0$, we get two complex number roots as $\mathcal{R}_1 = \zeta + i\eta$ and $\mathcal{R}_2 = \zeta - i\eta$. Accordingly, it is evident that

$$\xi(t) = \xi(0)\exp(-\zeta t)(\cos(\eta t) - \zeta \sin(\eta t)/\eta).$$

The above three subcases indicate that the pixel error $\xi = \mathbf{p} - \mathbf{p}_d$ converges to zero over time globally. The proof is complete.

4 DAIVS solution

The existing IBVS schemes, including the AIVS scheme (9), often require a detailed knowledge of the robot visual servoing system. However, in a non-ideal state, many unknown cases often exist, which can disturb the precise control of the robot, thus resulting in large errors. Recalling the data-driven learning algorithm (Equation 8), virtual visual Jacobian matrix $\bar{\mathcal{J}}$ is exploited to solve this issue.

4.1 DAIVS scheme and RNN solution

Based on the virtual visual Jacobian matrix, a DAIVS scheme (8) would be designed as

$$\begin{aligned} & \text{minimize} \quad \frac{1}{2} \ddot{\phi}^T \ddot{\phi} \\ & \text{subject to} \quad \dot{\mathbf{p}} = \bar{\mathcal{J}}\dot{\phi} + \dot{\bar{\mathcal{J}}}\phi \\ & \quad \mathbf{p} = \mathbf{p}_d \\ & \quad \ddot{\phi} \in \gamma. \end{aligned}$$

It is a remarkable fact that the DAIVS scheme does not involve the visual structure of the real robot. Instead, the virtual visual Jacobian matrix $\bar{\mathcal{J}}$ conveys the transformation relationship between the joint space and image space to deal with possible unknowns in the structure of the robot system. Compared with acceleration-level visual servoing schemes (Keshmiri et al., 2014; Anwar et al., 2019), the proposed DAIVS scheme offers two distinct advantages. First, it prioritizes the safety aspect by considering

joint limits. Second, the DAIVS scheme takes into account the uncertainty of the robot vision system and employs the virtual visual Jacobian matrix for robot control, enhancing the fault tolerance ability. The existing acceleration-level visual servoing schemes (Keshmiri et al., 2014; Anwar et al., 2019) cannot accurately implement visual servoing tasks when the Jacobian matrix lacks precision. Furthermore, combining Equations (8, 13) generates

$$\ddot{\phi} = \mathcal{P}_\gamma(\bar{\mathcal{J}}^T(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\bar{\mathcal{J}}}\phi)) \quad (16a)$$

$$\dot{\bar{\mathcal{J}}} = -\delta(\bar{\mathcal{J}}\dot{\phi} - \dot{\mathbf{p}})\dot{\phi}^T. \quad (16b)$$

It is worth pointing out that the RNN (16) is divided into the inner cycle and outer cycle, i.e., the learning cycle and control cycle. Subsystem (Equation 16a), which can be viewed as the outer cycle, mainly generates the control signal to adjust the joint properties via virtual visual Jacobian matrix $\bar{\mathcal{J}}$. In return, inner cycle (Equation 16b) with learning ability can explore the relationship between end-effector motion and joint motion, thus producing virtual visual Jacobian matrix $\bar{\mathcal{J}}$ to simulate the movement process of real robots. From a control point of view, the inner cycle (Equation 16b) must converge faster than the outer cycle (Equation 16a). In this sense, $\delta \gg \alpha$ is a necessary condition for the normal operation of the system.

Note that both RNN (13) and RNN (16) involve the use of pseudo-inverse operations. As a result, various existing methods can be employed to mitigate singularity issues, such as the damped least squares method. Specifically, $\bar{\mathcal{J}}^\dagger$ can be calculated via $\bar{\mathcal{J}}^\dagger = \bar{\mathcal{J}}^T(\bar{\mathcal{J}}\bar{\mathcal{J}}^T + h\mathcal{I})^{-1}$ with h being a tiny constant and \mathcal{I} being an identity matrix. The additional item $h\mathcal{I}$ ensures that all eigenvalues of $\bar{\mathcal{J}}\bar{\mathcal{J}}^T + h\mathcal{I}$ are never zero during the inversion process, thereby preventing singular issues. In addition, RNN (16) relies on the virtual visual Jacobian matrix and estimates the real Jacobian matrix using Equation (16b). This enables a robust handling of the visual system's uncertainty. However, RNN (13) relies on the real visual Jacobian matrix, leading to potential inaccuracies in the robot control process.

4.2 Stability analyses of RNN

The learning and control performance of the proposed DAIVS scheme aided with RNN (16) are proved by the following theorem.

Theorem 2: The Jacobian matrix error $E = \bar{\mathcal{J}} - \mathcal{J}$ and pixel error $\xi = \mathbf{p} - \mathbf{p}_d$ produced by RNN (16) converges to zero, given a large enough δ .

Proof: The proof is segmented into two parts: (1) proving learning convergence; (2) proving control convergence.

Part 1: Proving learning convergence. Design the i -th system of Jacobian matrix error as $E_i = \bar{\mathcal{J}}_i - \mathcal{J}_i$ ($i = 1, 2$) where $\bar{\mathcal{J}}_i$ and \mathcal{J}_i denote the i -th row of $\bar{\mathcal{J}}$ and \mathcal{J} and set the Lyapunov candidate $\mathcal{V}_i = (\bar{\mathcal{J}}_i - \mathcal{J}_i)(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T$. Calculating the time derivative of \mathcal{V}_i leads

to

$$\begin{aligned} \dot{\mathcal{V}}_i &= (\dot{\mathcal{J}}_i - \dot{\mathcal{J}}_i)(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T \\ &= -\delta(\dot{\mathcal{J}}_i\dot{\phi} - \dot{\mathbf{p}}_i)\dot{\phi}^T(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T - \dot{\mathcal{J}}_i(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T \\ &= -\delta(\dot{\mathcal{J}}_i\dot{\phi} - \dot{\mathcal{J}}_i\dot{\phi})\dot{\phi}^T(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T - \dot{\mathcal{J}}_i(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T \\ &\leq -\delta\Pi(\dot{\phi}\dot{\phi}^T)(\bar{\mathcal{J}}_i - \mathcal{J}_i)(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T - \dot{\mathcal{J}}_i(\bar{\mathcal{J}}_i - \mathcal{J}_i)^T, \end{aligned}$$

where $\dot{\mathbf{p}}_i$ represents the i -th element of $\dot{\mathbf{p}}$, and $\Pi(\dot{\phi}\dot{\phi}^T)$ denotes the least eigenvalue of matrix $\dot{\phi}\dot{\phi}^T$. When the manipulator is tracking the feature, the value of $\Pi(\dot{\phi}\dot{\phi}^T)$ is always greater than zero. In this case, we substitute $E_i = \bar{\mathcal{J}}_i - \mathcal{J}_i$ into the above equation, resulting in the following expression:

$$\begin{aligned} \dot{\mathcal{V}}_i &\leq -\delta\Pi(\dot{\phi}\dot{\phi}^T)E_iE_i^T - \dot{\mathcal{J}}_iE_i^T \\ &\leq -\delta\Pi(\dot{\phi}\dot{\phi}^T)\|E_i\|_2^2 + \|\dot{\mathcal{J}}_i\|_2\|E_i\|_2 \\ &= \|E_i\|_2(\|\dot{\mathcal{J}}_i\|_2 - \delta\Pi(\dot{\phi}\dot{\phi}^T))\|E_i\|_2. \end{aligned}$$

For further analysis, we consider three cases based on the above equation:

- If $\|E_i\|_2 > \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$, we observe $\dot{\mathcal{V}}_i < 0$ and $\mathcal{V}_i > 0$. This indicates that in this case, E_i converges until $\|E_i\|_2 = \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$.
- If $\|E_i\|_2 = \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$, we find $\dot{\mathcal{V}}_i \leq 0$ and $\mathcal{V}_i > 0$. This implies that E_i will continue to converge or remain at the state with $\|E_i\|_2 = \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$.
- If $\|E_i\|_2 < \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$, we have two possibilities: either $\dot{\mathcal{V}}_i > 0$ and $\mathcal{V}_i > 0$, or $\dot{\mathcal{V}}_i \leq 0$ and $\mathcal{V}_i > 0$. In the former possibility, the error will increase until $\|E_i\|_2 = \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$. In the latter possibility, the error will continue to converge or remain constant.

Combining the above three cases, it can be summarized that $\lim_{t \rightarrow +\infty} \|E_i\|_2 \leq \|\dot{\mathcal{J}}_i\|_2/\delta\Pi(\dot{\phi}\dot{\phi}^T)$. Furthermore, it can be deduced that the Jacobian matrix error $E = \bar{\mathcal{J}} - \mathcal{J}$ produced by RNN (16a) globally approach zero, given a sufficiently large value of δ .

Part 2: Proving control convergence.

According to the proof in *Part 1*, we take advantage of the LaSalle's invariant principle (Khalil, 2001) again to conduct the convergence proof on Equation (16b). In other words, the following formula is provided by replacing $\bar{\mathcal{J}}$ and $\dot{\mathcal{J}}$ with \mathcal{J} and $\dot{\mathcal{J}}$:

$$\ddot{\phi} = \mathcal{P}_\gamma(\mathcal{J}^\dagger(-\alpha\dot{\mathbf{p}} - \beta(\mathbf{p} - \mathbf{p}_d) - \dot{\mathcal{J}}\dot{\phi})), \quad (17)$$

which is equivalent to Equation (13). In consequence, the proof on the convergence of the pixel error $\mathbf{p} - \mathbf{p}_d$ in Equation (17) has been discussed in Theorem 1 and thus omitted here. The proof is complete.

5 Simulation verifications

In this section, simulations are conducted on a Franka Emika Panda manipulator with 7 DOFs for completing a visual servoing task, which are synthesized by the proposed AIVS scheme (9) and the proposed DAIVS scheme. Note that the AIVS scheme (9) is able

to drive the redundant manipulator to perform the visual servoing task with a given visual Jacobian matrix, and that, the DAIVS scheme can deal with the unknown situation in the robot system dynamically in the absence of the visual Jacobian matrix. For the simulations, this study utilizes a computer with an Intel Core i7-12700 processor and 32 GB RAM. The simulations are performed using MATLAB/Simulink software version R2022a.

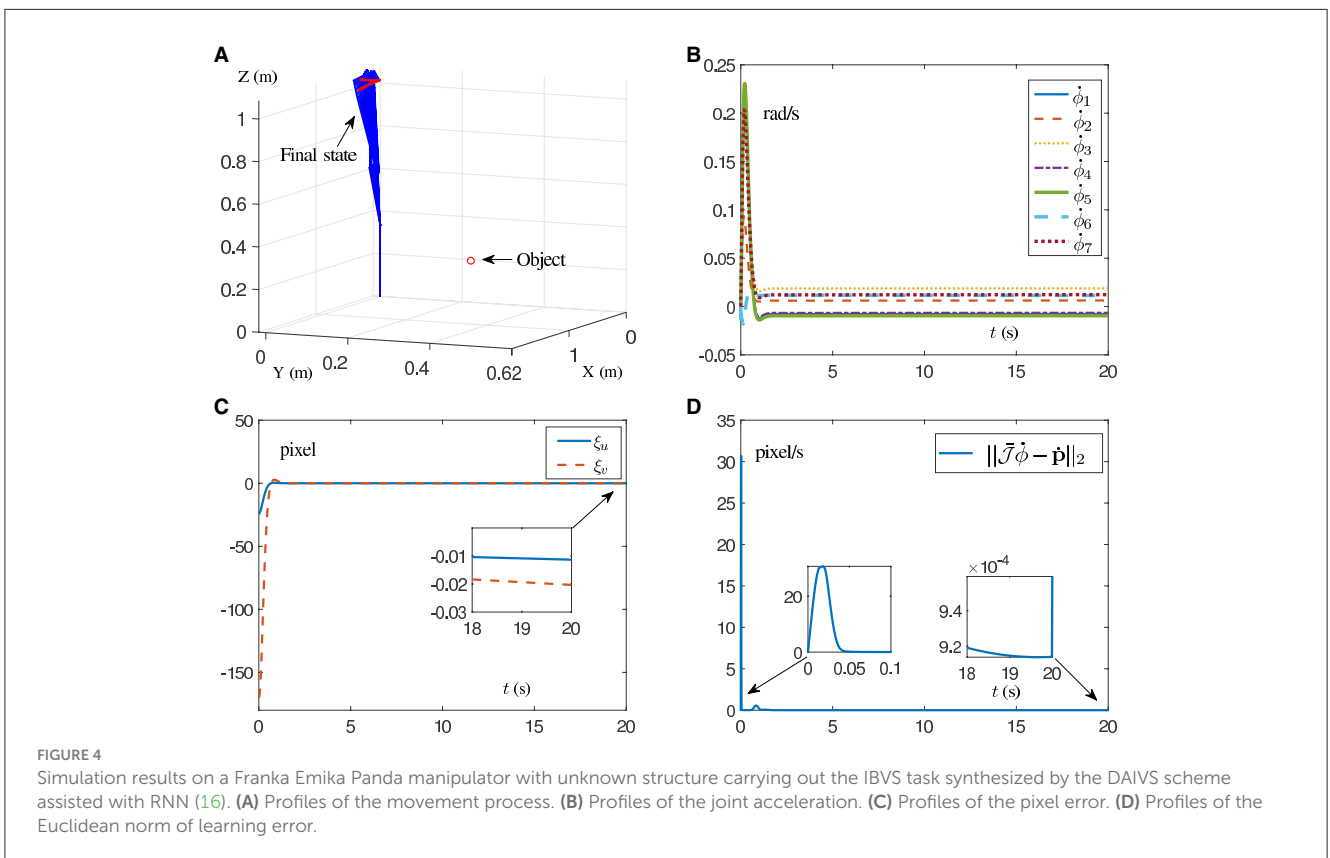
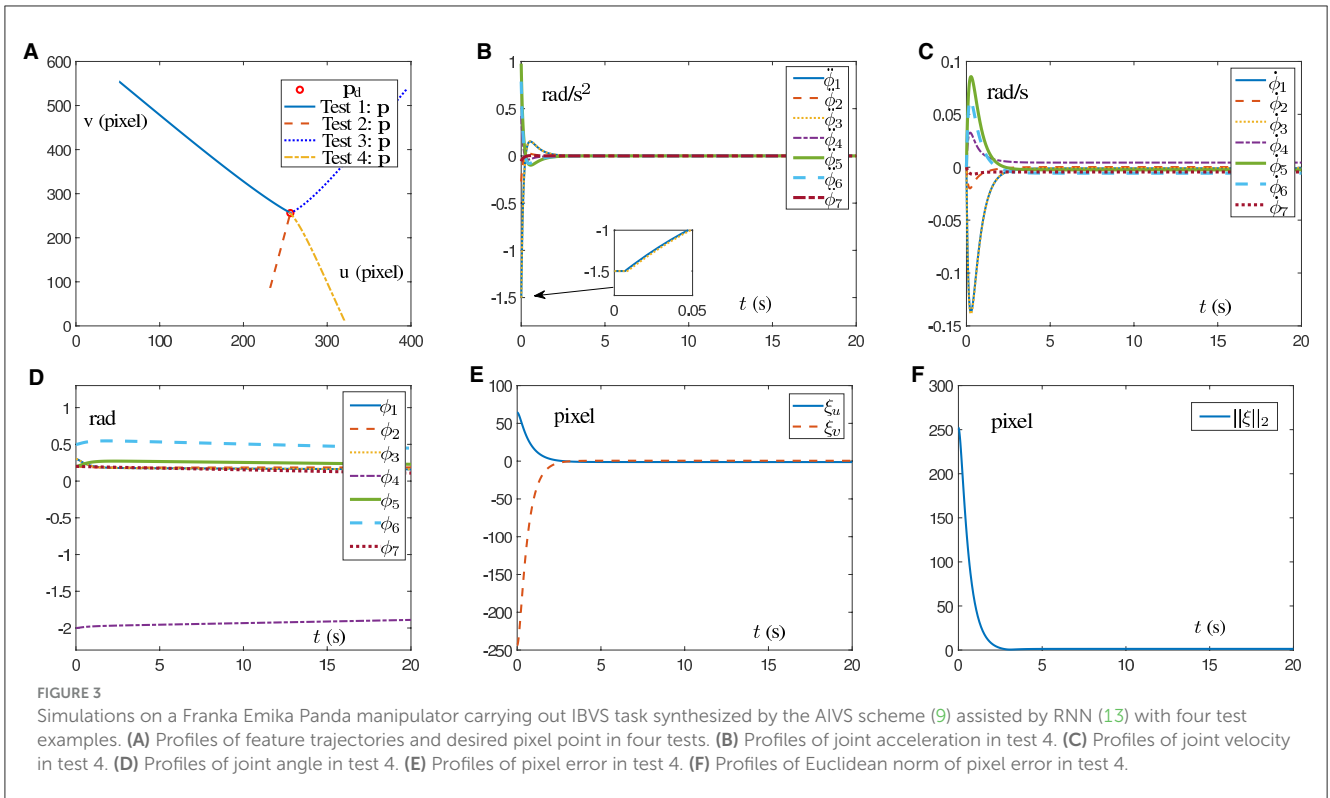
First, some necessary information and parameter settings about the manipulator and camera structure are given below. The Franka Emika Panda manipulator is a 7-DOF redundant manipulator (Gaz et al., 2019), with a camera mounted on its end-effector. In addition, we set $l = 8 \times 10^{-3}$ m, $u_0 = v_0 = 256$ pixel, $a_x = a_y = 8 \times 10^4$ pixel/m, and design $\mu = \nu = 20$ with $z = 2$, task execution time $T = 20$ s and $\mathbf{p}_d = [256, 256]^T$ pixel. In addition, the joint limits are set as $\ddot{\phi}^+ = -\ddot{\phi}^- = [2]_{7 \times 1}$ rad/s², $\dot{\phi}^+ = -\dot{\phi}^- = [0.6]_{7 \times 1}$ rad/s, $\phi^+ = -\phi^- = [2.5]_{7 \times 1}$ rad and $\theta = [0.076]_{7 \times 1}$ rad. It is noteworthy that the parameters can be divided into two categories: structural parameters and convergence parameters. Structural parameters, such as l , u_0 , v_0 , a_x , and a_y , are dependent on the configuration of the visual servo system. On the other hand, the convergence parameters, namely, μ , ν , α , β , and δ , play a vital role in adjusting the convergence behavior of RNN (16). These convergence parameters are set to values greater than zero, and their specific values can be determined through the trial and error method.

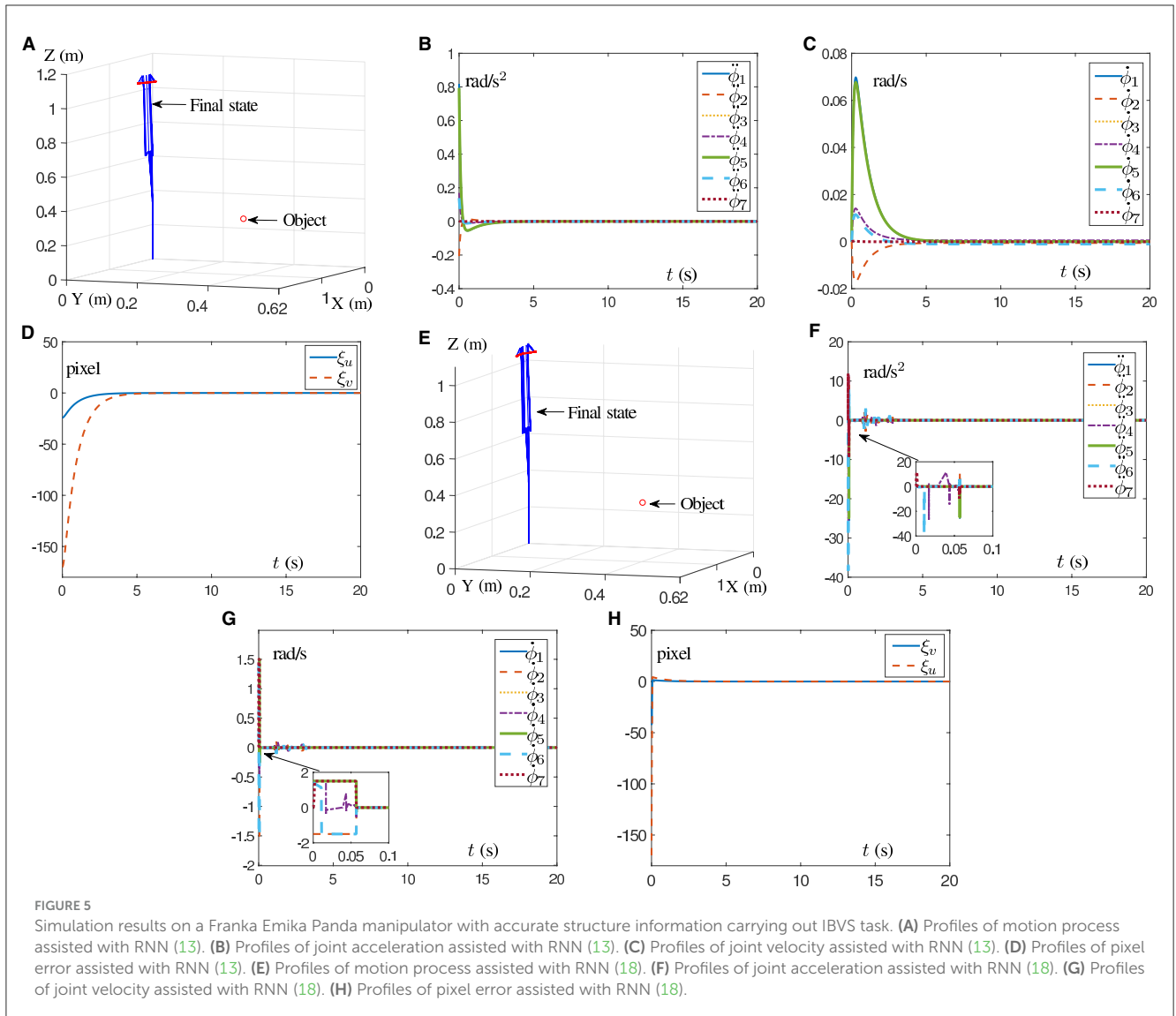
5.1 Simulation of AIVS scheme

In this subsection, in order to prove the feasibility of the AIVS scheme (9), four simulations with different initial position states of the Franka Emika Panda manipulator are conducted to trace one desired feature with results shown in Figure 3. Simply design $\alpha = 10$ and $\beta = 10$. It would be readily discovered from Figure 3A that four test examples from four different directions are straightforward to successfully pursue the desired pixel. With test 4 as an example, detailed joint data and pixel errors are shown in Figure 3B through Figure 3F, which illustrate that the joint angle, joint velocity, and joint acceleration are all kept inside the joint limit and that the pixel error can converge to zero within 5 s. The above descriptions well verify the validity of the proposed AIVS scheme (9) in the case of the known visual servoing Jacobian matrix to solve the visual servoing problem at the acceleration level.

5.2 Simulation of DAIVS scheme

This subsection indicates the feasibility and capability of the pixel error convergence of the DAIVS scheme aided with the RNN (16) by providing simulation results, as shown in Figure 4. Furthermore, we choose $\delta = 2 \times 10^4$, $\alpha = 10$ and $\beta = 40$. Notably, the virtual visual Jacobian matrix is exploited with random initial values, instead of the real visual Jacobian matrix to facilitate system operation. The end-effector of the robotic arm is oriented toward the object, as shown in Figure 4A. In addition, the joint acceleration is shown in Figure 4B, which is confined to the joint limit and maintain the normal operation. As shown in Figure 4C, the Franka





Emika Panda manipulator successfully traces the desired feature with pixel error converging to zero and maintaining the order of 10^{-2} pixel. As for the learning ability, Figure 4D illustrates that the virtual robot manipulator can learn the movement of the real robot manipulator with the learning error approaching to zero in 0.05 s and maintaining the order of 10^{-4} pixel/s. In short, the simulation results in Figure 4 highlight the simultaneous learning and control ability of RNN (16).

5.3 Comparisons of proposed schemes

This subsection offers simulation comparison results between the proposed schemes aided with the corresponding RNNs and the IBVS method presented in the study by Zhang and Li (2018). In this regard, the RNN provided in the study by Zhang and Li (2018) is shown as

$$\dot{\phi} = \mathcal{P}_\gamma(-\kappa_1 \mathcal{J}^T(\mathbf{p} - \mathbf{p}_d) - \kappa_2 \mathcal{J}^T \int_0^t (\mathbf{p} - \mathbf{p}_d) dt), \quad (18)$$

where parameters $\kappa_1 > 0$ and $\kappa_2 > 0$ determine the rate of error convergence. It is worth pointing out that the IBVS method in the study by Zhang and Li (2018) assisted with RNN (18) is constructed from the viewpoint of the velocity level, and that, RNN (18) requires exact structural information \mathcal{J} to maintain the normal operation.

In the first place, simulations are conducted on the Franka Emika Panda manipulator for IBVS task with Figures 5A–D synthesized by RNN (13) and Figures 5E–H synthesized by RNN (18). Notably, the results in Figure 5 are carried out on the premise of known structural information \mathcal{J} with parameters $\kappa_1 = \kappa_2 = 2$, $\alpha = 10$, and $\beta = 10$. As shown in Figures 5A, E, the manipulator’s end-effector is controlled to point toward the object. In Figure 5B, the joint acceleration generated by RNN (13) is safely confined within the joint limits, while the joint acceleration generated by RNN (18) exists a sudden change of ~ 38 rad/s² in Figure 5F, which may cause damage to the robot. Furthermore, in contrast to Figure 5G, the joint velocity shown in Figure 5C is smaller and exhibits smoother changes, making it more suitable for real-world scenarios. Figures 5D, H demonstrate that both RNN (13) and RNN

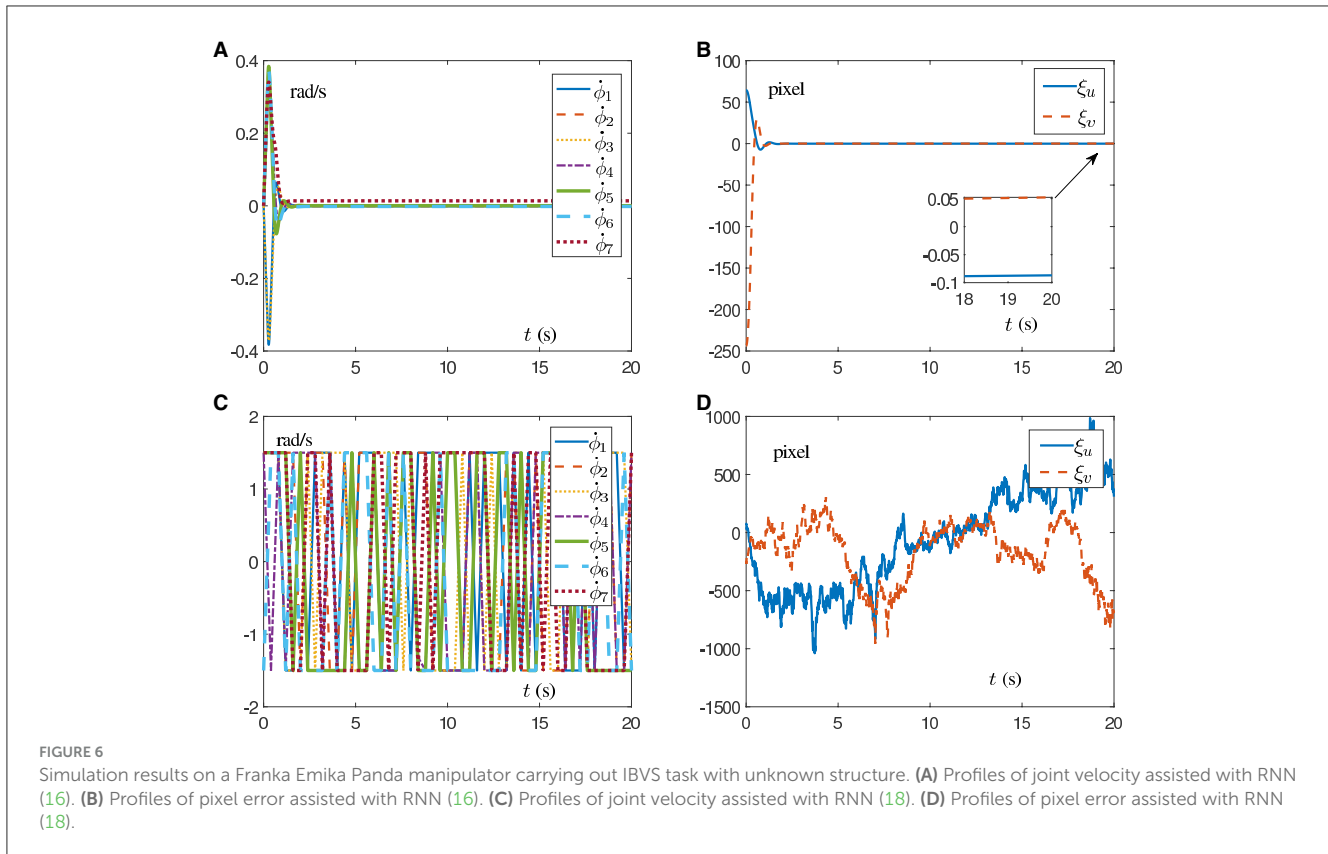


TABLE 1 Comparisons among different approaches for visual servoing of robot manipulators.

	Visual servoing	Scheme level	Velocity constraints	Acceleration constraints	Structure information	Jacobian matrix learning
RNN (13)	Yes	Acceleration	Yes	Yes	Unnecessary	Yes
RNN (16)	Yes	Acceleration	Yes	Yes	Necessary	No
Van et al. (2018)	Yes	Velocity	No	No	Necessary	No
Hashimoto et al. (1991)	Yes	Velocity	No	No	Necessary	No
Zhang et al. (2017)	Yes	Velocity	Yes	No	Necessary	No
Zhang and Li (2018)	Yes	Velocity	Yes	No	Necessary	No
Li et al. (2020)	Yes	Velocity	Yes	No	Necessary	No
Keshmiri et al. (2014)	Yes	Acceleration	No	No	Necessary	No
Anwar et al. (2019)	Yes	Acceleration	No	No	Necessary	No
Zhu et al. (2022)	Yes	Torque	No	No	Necessary	No

(18) are able to quickly propel pixel errors to zero. Therefore, it is concluded from the above results that AIVS scheme (9) aided by RNN (13) is able to guarantee a better safety performance when controlling the manipulator.

Beyond that, in the case of the unknown visual system, corresponding comparison simulations are driven by the DAIVS scheme aided with the RNN (16) and the IBVS method in the study by Zhang and Li (2018) assisted with RNN (18). The results are shown in Figure 6 with parameters $\kappa_1 = \kappa_2 = 2, \alpha = 10,$

$\beta = 40,$ and $\delta = 2 \times 10^4$. To simulate the unknown visual system, $\bar{\mathcal{J}}$ in Equation (16) and \mathcal{J} in Equation (18) are random matrices of constants with the absolute value of each element < 100 . Figures 6A, B well embody that, when encountering unknown structural information, the DAIVS scheme assisted with RNN (16) controls the Franka Emika Panda manipulator to preferably complete IBVS task with the pixel error converging to zero. Nevertheless, the generated joint velocity in Figure 6C changes dramatically within the joint limit in a mess. Even worse, the

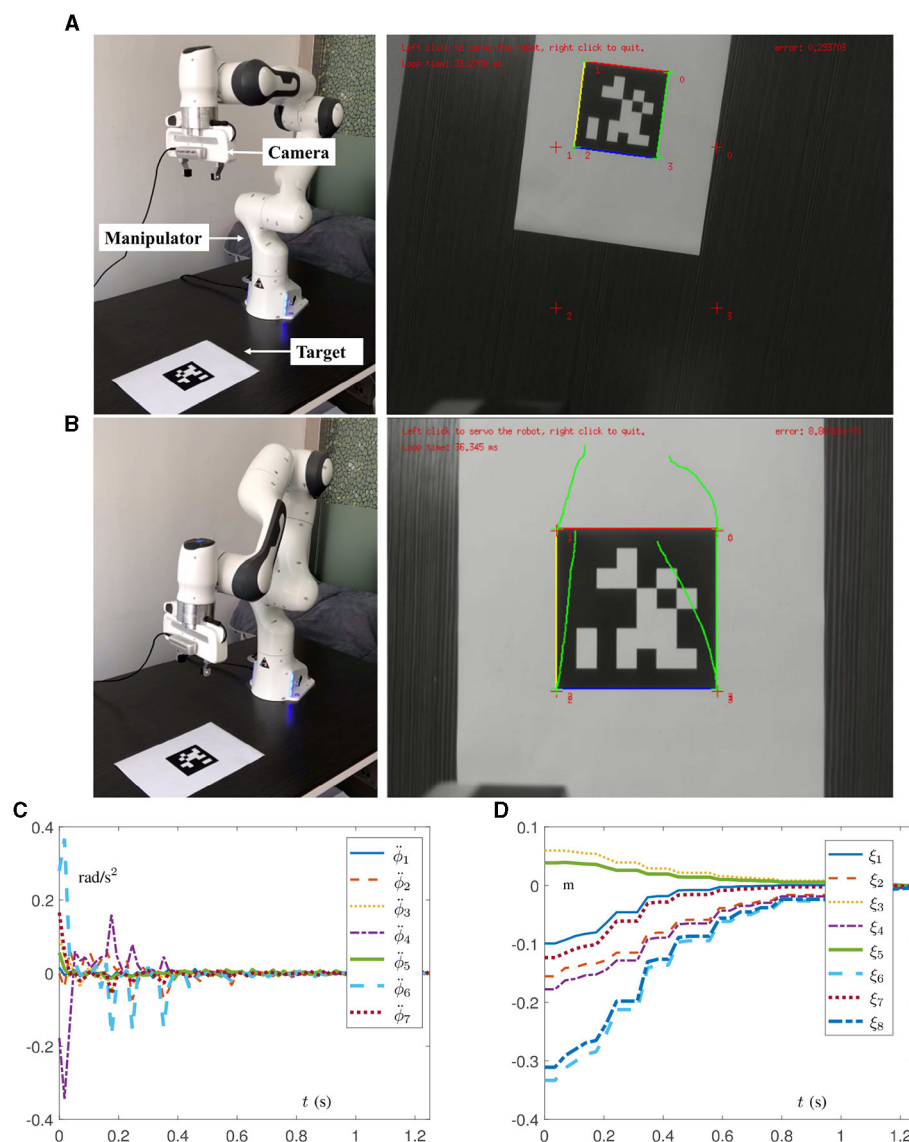


FIGURE 7 Physical experiments on a Franka Emika Panda manipulator assisted with RNN (16) for carrying out IBVS task with a fixed target. (A) Initial states of the manipulator and camera. (B) Final states of the manipulator and camera. (C) Profiles of joint acceleration. (D) Profiles of tracking error.

pixel error driven by RNN (18) does not converge and maintain a diffused state, which indicates the failure of the IBVS task. In conclusion, the proposed DAIVS scheme is able to deal with the unknown structural information in the robot system and fulfill the visual servo control with simultaneous learning and control performance.

Furthermore, comparison results among different existing approaches (Hashimoto et al., 1991; Keshmiri et al., 2014; Zhang et al., 2017; Van et al., 2018; Zhang and Li, 2018; Anwar et al., 2019; Li et al., 2020; Zhu et al., 2022) for visual servoing of robot manipulators are presented in Table 1. It is worth emphasizing that, compared with the prior art, the proposed RNN (13) and RNN (16) are the first acceleration-level work, considering the multiple levels of joint constraints, and RNN (16) is the first study to dispose the unknown situations in the robot visual system with simultaneous learning and control ability. As a

result, the above two points are the innovative contributions of this study.

6 Experiments on real manipulators

To verify the effectiveness and practicability of the proposed DAIVS scheme, physical experiments on a real manipulator are conducted in this section, which are driven by the DAIVS scheme aided with RNN (16). Specifically, the experiments essentially rely on C++ and the visual servoing platform (ViSP) for embedding algorithms and control (Marchand et al., 2005), which are built on ubuntu 16.04 LTS operating system. In addition, the experiment platform consists of a Franka Emika Panda manipulator, an Intel RealSense Camera D435i, a personal computer, and an AprilTag (target). It is worth mentioning that the acceleration control

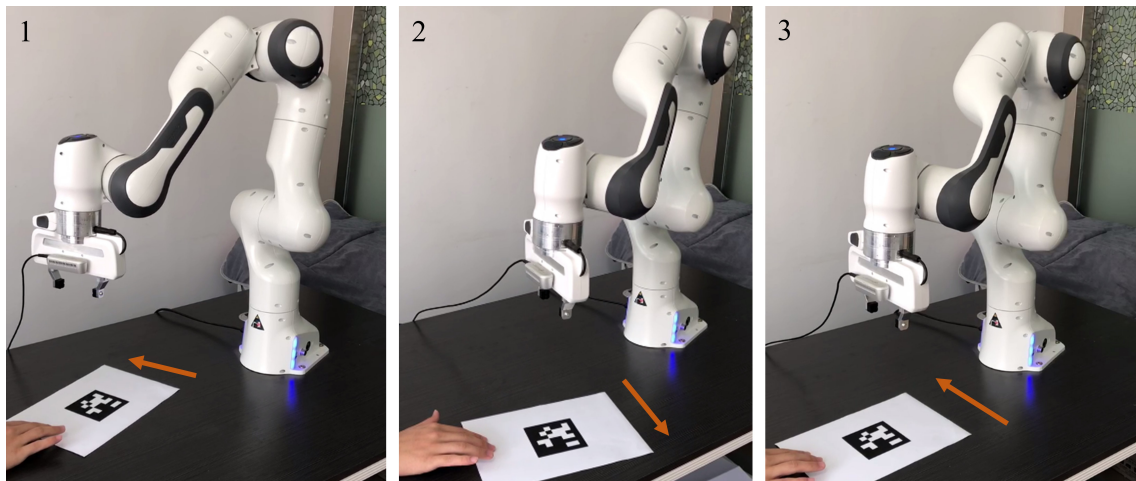


FIGURE 8
Physical experiments on a Franka Emika Panda manipulator assisted with RNN (16) for carrying out IBVS task with a moving target.

commands generated by the proposed RNN (16) are transmitted in a discrete form with a frequency of 1,000 Hz, and parameter settings of RNN (16) are designed as follows. We choose $\alpha = 10$, $\beta = 10$, $\delta = 10^6$, $\mu = \nu = 20$, $\ddot{\phi}^+ = -\ddot{\phi}^- = [15, 7.5, 10, 12.5, 15, 20, 20]^T$ rad/s², $\dot{\phi}^+ = -\dot{\phi}^- = [2.1, 2.1, 2.1, 2.1, 2.6, 2.6, 2.6]^T$ rad/s, $\phi^+ = [2.8, 1.7, 2.8, -0.1, 2.9, 3.7, 2.8]^T$ rad, $\phi^- = [-2.8, -1.7, -2.8, -3.0, -2.8, -0, -2.8]^T$ rad, and $\mathcal{J}(0) = \mathcal{J}(0)$. As for the parameter settings of the camera and pixel coordinates, they can be directly referenced to ViSP (Marchand et al., 2005). Different from the previous simulations, the physics experiments set the target as an AprilTag containing four features. As a result, the physical parameters associated with the features are expanded to 8 instead of 2.

Experiment results on the Franka Emika Panda manipulator tracking the fixed target are shown in Figures 7, 8 with $\mathbf{p}_d = [-0.06, -0.06, 0.06, -0.06, 0.06, 0.06, -0.06, 0.06]^T$ m for the given task in the camera system. It is worth mentioning that the robot manipulator adjusts the joint state to recognize and approach the target, and when the pixel error reaches the order of 10^{-5} pixel, the task automatically completes. It is important that the whole process of learning and control does not involve the real Jacobian matrix to simulate the situation of the unknown structure. In Figures 7A, B, the initial and final states of the manipulator and camera indicate that the visual servoing task is successfully realized by the DAIVS scheme with execution time of 1.25 s. Specifically, the joint acceleration in Figure 7C varies normally within the joint constraints. In the meantime, the tracking errors ξ of four features are presented in Figure 7D, which illustrate the precise control ability of the DAIVS scheme with global convergence to zero.

Beyond that, experiments on the Franka Emika Panda manipulator tracking the moving target are conducted to demonstrate the feasibility of the DAIVS scheme. In Figure 8, the AprilTag is moved artificially by the hand toward the left and right and simultaneously the manipulator constantly adjusts joint states

to achieve the characteristics of real-time visual tracking. More vividly, the experiment videos corresponding to Figures 7, 8 are available at <https://youtu.be/6uw35bidVcw>.

7 Conclusion

This study has proposed an AIVS scheme for robot manipulators, taking into account joint limits at multiple levels. On this basis, incorporating data-driven techniques, a DAIVS scheme has been proposed to handle potential unknown situations in the robot visual system. Furthermore, RNNs have been exploited to generate the online solution corresponding to the proposed schemes with theoretical analyses, demonstrating the simultaneous learning and control ability of the proposed DAIVS scheme. Then, numerous simulations and experiments have been carried out on a Franka Emika Panda manipulator to track the desired feature. The results validate the theoretical analyses, demonstrate the feasibility of the AIVS scheme, and showcase the fast convergence and robustness of the DAIVS scheme. Compared with the method in the study by Zhang and Li (2018), the DAIVS scheme exhibits superior learning capability and achieves visual servoing control with the unknown Jacobian matrix.

In summary, this study provides a data-driven approach for the precise manipulation of robots in IBVS tasks, addressing unknown situations that could affect the robot's Jacobian matrix. In the future, we aim to expand our research to incorporate dynamic factors, utilizing joint torque as control signals and considering dynamic uncertainties.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LW: Investigation, Methodology, Resources, Software, Writing – original draft. ZX: Data curation, Investigation, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Key Laboratory of Tibetan Information Processing, Ministry of Education, under Grant 2024-Z-002, and supported by the Supercomputing Center of Lanzhou University.

References

- Agarwal, S., and Akella, S. (2024). Line coverage with multiple robots: algorithms and experiments. *IEEE Transact. Robot.* 40, 1664–1683. doi: 10.1109/TRO.2024.3355802
- Anwar, A., Lin, W., Deng, X., Qiu, J., and Gao, H. (2019). Quality inspection of remote radio units using depth-free image-based visual servo with acceleration command. *IEEE Transact. Ind. Electron.* 66, 8214–8223. doi: 10.1109/TIE.2018.2881948
- Cigliano, P., Lippiello, V., Ruggiero, F., and Siciliano, B. (2015). Robotic ball catching with an eye-in-hand single-camera system. *IEEE Transact. Cont. Syst. Technol.* 23, 1657–1671. doi: 10.1109/TCST.2014.2380175
- Gaz, C., Cognetti, M., Oliva, A., Robuffo Giordano, P., and De Luca, A. (2019). Dynamic identification of the Franka Emika Panda robot with retrieval of feasible parameters using penalty-based optimization. *IEEE Robot. Automat. Lett.* 4, 4147–4154. doi: 10.1109/LRA.2019.2931248
- Hashimoto, K., Kimoto, T., Ebine, T., and Kimura, H. (1991). “Manipulator control with image-based visual servo,” in *Proceedings IEEE International Conference on Robotics and Automation* (New York, NY: Sacramento, CA: IEEE), 2267–2271.
- Jin, P., Lin, Y., Song, Y., Li, T., and Yang, W. (2023). Vision-force-fused curriculum learning for robotic contact-rich assembly tasks. *Front. Neurobot.* 17:1280773. doi: 10.3389/fnbot.2023.1280773
- Jin, L., Liu, L., Wang, X., Shang, M., and Wang, F.-Y. (2024). Physical-informed neural network for MPC-based trajectory tracking of vehicles with noise considered. *IEEE Transact. Intell. Vehicl.* doi: 10.1109/TIV.2024.3358229
- Keshmiri, M., Xie, W.-F., and Mohebbi, A. (2014). Augmented image-based visual servoing of a manipulator using acceleration command. *IEEE Transact. Ind. Electron.* 61, 5444–5452. doi: 10.1109/TIE.2014.2300048
- Khalil, H. K. (2001). *Nonlinear Systems*. Englewood Cliffs, NJ: Prentice Hall.
- Li, S., Shao, Z., and Guan, Y. (2019). A dynamic neural network approach for efficient control of manipulators. *IEEE Transact. Syst. Man Cybernet. Syst.* 49, 932–941. doi: 10.1109/TSMC.2017.2690460
- Li, W., Chiu, P. W. Y., and Li, Z. (2020). An accelerated finite-time convergent neural network for visual servoing of a flexible surgical endoscope with physical and RCM constraints. *IEEE Transact. Neural Netw. Learn. Syst.* 31, 5272–5284. doi: 10.1109/TNNLS.2020.2965553
- Liang, X., Wang, H., Liu, Y.-H., Chen, W., and Jing, Z. (2018). Image-based position control of mobile robots with a completely unknown fixed camera. *IEEE Trans. Automat. Contr.* 63, 3016–3023. doi: 10.1109/TAC.2018.2793458
- Liufu, Y., Jin, L., Shang, M., Wang, X., and Wang, F.-Y. (2024). ACP-incorporated perturbation-resistant neural dynamics controller for autonomous vehicles. *IEEE Transact. Intell. Vehicl.* doi: 10.1109/TIV.2023.3348632
- Malis, E., Mezouar, Y., and Rives, P. (2010). Robustness of image-based visual servoing with a calibrated camera in the presence of uncertainties in the three-dimensional structure. *IEEE Transact. Robot.* 26, 112–120. doi: 10.1109/TRO.2009.2033332
- Marchand, E., Spindler, F., and Chaumette, F. (2005). ViSP for visual servoing: a generic software platform with a wide class of robot control skills. *IEEE Robot. Automat. Mag.* 12, 40–52. doi: 10.1109/MRA.2005.1577023
- Na, J., Yang, J., Wang, S., Gao, G., and Yang, C. (2021). Unknown dynamics estimator-based output-feedback control for nonlinear pure-feedback systems. *IEEE Transact. Syst. Man Cybernet. Syst.* 51, 3832–3843. doi: 10.1109/TSMC.2019.2931627
- Park, D.-H., Kwon, J.-H., and Ha, I.-J. (2012). Novel position-based visual servoing approach to robust global stability under field-of-view constraint. *IEEE Transact. Ind. Electron.* 59, 4735–4752. doi: 10.1109/TIE.2011.2179270
- Peng, G., Chen, C. L. P., and Yang, C. (2023). Robust admittance control of optimized robot-environment interaction using reference adaptation. *IEEE Transact. Neural Netw. Learn. Syst.* 34, 5804–5815. doi: 10.1109/TNNLS.2021.3131261
- Stanimirovic, P. S., Zivkovic, I. S., and Wei, Y. (2015). Recurrent neural network for computing the Drazin inverse. *IEEE Transact. Neural Netw. Learn. Syst.* 26, 2830–2843. doi: 10.1109/TNNLS.2015.2397551
- Tang, Z., and Zhang, Y. (2022). Refined self-motion scheme with zero initial velocities and time-varying physical limits via Zhang neurodynamics equivalency. *Front. Neurobot.* 16:945346. doi: 10.3389/fnbot.2022.945346
- Van, M., Ge, S. S., and Ceglarek, D. (2018). Fault estimation and accommodation for virtual sensor bias fault in image-based visual servoing using particle filter. *IEEE Transact. Ind. Inf.* 14, 1312–1322. doi: 10.1109/TII.2017.2723930
- Wang, X., Sun, Y., Xie, Y., Bin, J., and Xiao, J. (2023). Deep reinforcement learning-aided autonomous navigation with landmark generators. *Front. Neurobot.* 17:1200214. doi: 10.3389/fnbot.2023.1200214
- Xie, Z., Jin, L., Luo, X., Hu, B., and Li, S. (2022). An acceleration-level data-driven repetitive motion planning scheme for kinematic control of robots with unknown structure. *IEEE Transact. Syst. Man Cybernet. Syst.* 5152, 5679–5691. doi: 10.1109/TSMC.2021.3129794
- Xie, Z., and Jin, L. (2023). A fuzzy neural controller for model-free control of redundant manipulators with unknown kinematic parameters. *IEEE Transact. Fuzzy Syst.* 32, 1589–1601. doi: 10.1109/TFUZZ.2023.3328545
- Xu, C., Sun, Z., Wang, C., Wu, X., Li, B., and Zhao, L. (2023). An advanced bionic knee joint mechanism with neural network controller. *Front. Neurobot.* 17:1178006. doi: 10.3389/fnbot.2023.1178006
- Xu, F., Wang, H., Wang, J., Au, K. W. S., and Chen, W. (2019). Underwater dynamic visual servoing for a soft robot arm with online distortion correction. *IEEE ASME Transact. Mechatron.* 24, 979–989. doi: 10.1109/TMECH.2019.2908242
- Yang, C., Chen, C., He, W., Cui, R., and Li, Z. (2019). Robot learning system based on adaptive neural control and dynamic movement primitives. *IEEE Transact. Neural Netw. Learn. Syst.* 30, 777–787. doi: 10.1109/TNNLS.2018.2852711
- Zeng, D., Liu, Y., Qu, C., Cong, J., Hou, Y., and Lu, W. (2024). Design and human-robot coupling performance analysis of flexible ankle rehabilitation robot. *IEEE Robot. Automat. Lett.* 9, 579–586. doi: 10.1109/LRA.2023.3330052
- Zeng, Z., Wang, J., and Liao, X. (2003). Global exponential stability of a general class of recurrent neural networks with time-varying delays. *IEEE Transact. Circ. Syst. I Fund. Theory Appl.* 50, 1353–1358. doi: 10.1109/TCSI.2003.817760
- Zhang, Y., and Li, S. (2018). A neural controller for image-based visual servoing of manipulators with physical constraints. *IEEE Transact. Neural Netw. Learn. Syst.* 29, 5419–5429. doi: 10.1109/TNNLS.2018.2802650

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Zhang, Y., and Li, S. (2023). Kinematic control of serial manipulators under false data injection attack. *IEEE CAA J. Automat. Sin.* 10, 1009–1019. doi: 10.1109/JAS.2023.123132

Zhang, Y., Li, S., Liao, B., Jin, L., and Zheng, L. (2017). “A recurrent neural network approach for visual servoing of manipulators,” in *2017 IEEE International Conference on Information and Automation (ICIA)* (New York, NY: Macao: IEEE), 614–619.

Zhang, Z., and Zhang, Y. (2012). Acceleration-level cyclic-motion generation of constrained redundant robots tracking different paths. *IEEE Transact.*

Syst. Man Cybernet. Part B 42, 1257–1269. doi: 10.1109/TSMCB.2012.2189003

Zheng, X., Liu, M., Jin, L., and Yang, C. (2024). Distributed collaborative control of redundant robots under weight-unbalanced directed graphs. *IEEE Transact. Ind. Inf.* 20, 681–690. doi: 10.1109/TII.2023.3268778

Zhu, M., Huang, C., Qiu, Z., Zheng, W., and Gong, D. (2022). Parallel image-based visual servoing/force control of a collaborative delta robot. *Front. Neurobot.* 16:922704. doi: 10.3389/fnbot.2022.922704