AIMS Mathematics, 10(11): 27364–27380.
DOI: 10.3934/math.20251203
Received: 13 August 2025
Revised: 15 October 2025
Accepted: 21 October 2025
Published: 25 November 2025

https://www.aimspress.com/journal/Math

Research article

A calibration-free visual control approach for table tennis training robots using a positional monocular camera

Quanyu Song[1], Rong Lu[2], Yaobo Long[1] and Xiaobing Zheng[3,*]

[1] Guangdong University of Science and Technology, Dongguan 523083, China
[2] College of Artificial Intelligence, Dongguan Polytechnic, Dongguan 523808, China
[3] Department of Physical Education, Guangzhou Xinhua University, Dongguan 523133, China

* **Correspondence:** Email: zhengxiaobing0073@xhsysu.edu.cn.

Abstract: Developing low-cost and easy-to-deploy table tennis training robots is a significant challenge, largely due to the stringent, time-consuming, and error-prone camera calibration required by traditional visual servoing systems. This paper directly addressed this problem by proposing a robust, calibration-free visual control framework that enables a robotic manipulator to perform training tasks using a single monocular camera with completely unknown intrinsic and extrinsic parameters. This is a depth-independent visual feedback controller that directly translates the pixel error between the racket and the ball into control actions. To handle the unknown camera projection model, a computationally efficient adaptive law was designed. This law utilizes a regression matrix to isolate the unknown parameters and updates them in real time, completely eliminating the need for any pre-calibration process. The stability of the entire closed-loop system was rigorously proven via the Lyapunov method, guaranteeing the convergence of the tracking error. Extensive simulations on a manipulator validated the method's effectiveness and practicality. The results demonstrated that the system achieves rapid and precise tracking, with the image error converging completely to zero in under 0.8 second. The controller's robustness was further confirmed in scenarios with varying target positions and continuous multi-stroke sequences, demonstrating its suitability for dynamic and realistic training environments.

Keywords: table tennis training robot; calibration-free; visual control technology; adaptive technology
Mathematics Subject Classification: 93C85, 93C40

1. Introduction

Table tennis is a sport where rapid reactions and precise techniques are critical [1–3]. Ball speeds often exceed 100 kilometers per hour, with complex spins and trajectories that require players to make

split-second adjustments. Athletes dedicate extensive time to training [4], to improve shot accuracy, anticipate opponent actions, and maintain performance during competitive matches. Traditional table tennis training has notable limitations. Human partners experience fatigue—reducing practice quality over time—and often have limited availability, which disrupts training routines [5]. Moreover, the cost of hiring a dedicated coach is undoubtedly substantial, which conflicts with the goal of widely accessible training [6]. Therefore, the development of a simple and low-cost table tennis practice robot technology is a pressing need [7].

For a table tennis training robot to effectively simulate real match scenarios, it must first accurately perceive the movement of the ball. The position, speed, and spin of the ball during play are dynamic variables that directly determine how the robot should respond with a return shot. To capture these variables in real time, a reliable visual system is essential [8–10]. Visual sensors, as the simplest and most direct means of feedback, are the most commonly adopted solution in existing research [11,12]. A dual-ANN (Artificial Neural Network) method was proposed to predict table tennis trajectories as two parabolas, achieving better accuracy (mean error: 39.6 mm) than single-network models (42.9 mm) and physical models (57.9 mm) with fewer parameters and shorter training time, though it still has prediction errors in physical experiments [13]. A trajectory prediction model combining simple physical motion constraints and dual LSTM (Long Short Term Memory) neural network correction based on binocular vision was developed to improve trajectory extraction and prediction accuracy with a certain hitting success rate, but exhibits low trajectory recognition accuracy when the target is far from the camera or when the image resolution is excessively high [14]. While the aforementioned articles proposed highly innovative methods for the visual recognition of table tennis motion states, integrating these perception methods with robotic control remains a key challenge.

To address the integration challenge mentioned above, visual feedback-based robotic systems have emerged as a critical solution in bridging visual recognition and robotic control [15, 16]. In response to this challenge, a deep reinforcement learning approach with spin velocity estimation capability was developed to learn ball stroke strategies, ensuring desirable target landing locations and over-net heights. A virtual environment was built for effective pre-training, and experimental results show superior performance compared to traditional aerodynamics-based methods, with an average landing error around 80 mm and a landing-within-table probability over 70%, though it is sensitive to position estimation errors [17]. An augmented reality system (avaTTAR) with on-body and detached visual cues was developed for table tennis stroke training, enabling real-time comparison with expert strokes via 3D pose estimation and IMU (Inertial Measurement Unit) sensors, and user studies confirmed its effectiveness in improving training results [18]. An improved multi-modal table tennis robot system with four frame-based cameras, two event-based cameras, and a KUKA robot arm was developed, featuring novel calibration for multi-modal perception, accurate spin estimation, and SNN (Spiking Neural Network)-based ball detection using event cameras, with high accuracy and fast reaction; however, spin estimation may fail for extremely high or low spins, and event-based camera calibration has higher reprojection errors compared to frame-based ones [19]. To address simultaneous input dead-zone and field-of-view (FoV) constraints in visual servoing, an adaptive neural network controller, featuring a novel zone barrier Lyapunov function (zBLF), was developed in [20]. The key advantages of this method are its ability to strictly enforce state constraints while significantly reducing control energy consumption by activating the controller only when necessary. However, these advanced systems share a common dependency on precise camera calibration. This calibration process is not

only time-consuming and requires technical expertise, but it must be meticulously repeated whenever the camera is moved or replaced. This dependency poses a significant practical barrier to developing low-cost, flexible, and truly "plug-and-play" robotic training systems, creating a clear need for robust calibration-free alternatives.

In response to the challenge about camera calibration, numerous scholars have conducted relevant research on the rapid calibration of cameras and direct control under uncalibrated camera conditions [21, 22]. A transformer-based method for single-image camera calibration was proposed, integrating image and line segment features with an auxiliary line classification task to estimate camera parameters, outperforming existing methods in accuracy for vertical orientation, pitch, roll, FoV, and horizon line AUC (Area Under Curve); however, it relies on detected line segments, which may limit performance in scenes with few lines [23]. In [24], a new method for single-image camera calibration was proposed by introducing Perspective Fields, a per-pixel representation containing up-vectors and latitude values, which enables robust estimation of camera parameters via neural networks and outperforms existing methods on cropped images. A method using referring expression segmentation was proposed for uncalibrated image-based visual servoing (UIBVS), with CLIPUNetr network leveraging CLIP's vision-language representations and U-shaped architecture to generate high-quality segmentations, improving boundary and structure measurements by 120% and enabling real-world robot control [25]. In [26], a novel uncalibrated model-free visual servo control scheme for robotic-assisted minimally invasive surgery was proposed, featuring gradient neural network-based Jacobian and interaction matrix estimation, a quadratic programming framework incorporating multiple constraints, and a predefined-time convergent solver, with simulations showing effectiveness in feature regulation and tracking. As an engineering example, in the research [27], to address the challenges of jamming and computational complexity in multi-pin assembly, an uncalibrated visual servoing framework was developed which employs a projective homography-based task function for computational efficiency, a virtual image plane for spatial path constraint, and an LSTM-compensated Kalman filter for robust Jacobian matrix estimation. Further advancing this trend, the work in [28] presents a position-based visual servoing (PBVS) scheme that operates without prior hand-eye calibration. Instead of estimating the unknown parameters online, the authors approach the problem from a robust control perspective. The unknown camera-to-flange transformation was treated as a bounded uncertainty, allowing the corresponding Jacobian to be embedded within a convex polytope. This framework enables the design of a controller with proven convergence, demonstrating a valuable alternative to adaptive methods for uncalibrated control. While these studies provide a valuable foundation, there remains a need for an integrated, dynamics-aware framework specifically tailored to the high-speed requirements of a table tennis training robot.

Motivated by these challenges, this paper proposes a visual control scheme for a table tennis robot that uses an uncalibrated pinhole camera. Our control system employs an uncalibrated monocular vision camera with unknown intrinsic and extrinsic parameters for visual feedback. The camera only needs to provide images obtained through its projection. Through feature recognition technology, we can calculate the error between the table tennis racket and the ball. Then, by constructing a depth-independent visual feedback controller and a parameter update law for adaptively updating controller parameters, we establish a calibration-free control framework for the manipulator of the table tennis training robot. Furthermore, the stability of the proposed scheme is proven using the Lyapunov method. Finally, its feasibility is verified through simulation experiments. In summary, the contributions of this

study can be concluded as follows:

1) A practical "plug-and-play" visual control architecture is proposed, which eliminates the complex and time-consuming camera calibration process. This significantly lowers the technical barrier and cost for deploying table tennis robots, allowing the system to operate effectively even if the camera is moved or replaced.
2) A lightweight online parameter estimation scheme is equipped for the uncalibrated camera of the table tennis robot. By constructing a regression matrix to directly estimate the task-relevant image Jacobian, the method avoids the need for explicit reconstruction of camera parameters, making it computationally efficient and suitable for real-time applications.
3) In contrast to purely kinematic-based control approaches, the proposed scheme incorporates the robotic dynamics, thereby exhibiting enhanced disturbance rejection and greater robustness.

2. System models

The system under consideration is a table tennis training robot equipped with a fixed robotic arm, with a racket mounted on its end-effector. A fixed monocular vision camera is utilized to provide feedback on the positional error between the racket and the table tennis ball. This visual feedback enables the designed controller to drive the robotic manipulator to complete the ball-intercepting task.

2.1. Kinematics model

In our design, the world frame and base frame are coincident and marked on the table. The transformational relationship from the base frame to the camera frame can be described in the following form:

$$x^C = Tx, \tag{2.1}$$

where x_C and x are 4×1 position vectors, represent the spatial coordinates of the racket in the camera frame and base frame, respectively, and T is the transform matrix. Consider a rigid-link robotic manipulator, based on the kinematics, where the joint angle is denoted as an $n \times 1$ vector q, where n is the number of links. From the principles of kinematics, we have the following Jacobian relationship:

$$\dot{x} = J\dot{q}, \tag{2.2}$$

where J is the robotic Jacobian matrix. Consider the camera visual feedback image, where the feedback image coordinates are denoted as $y = (y_1, y_2, 1)$, and the pinhole projection model is given as

$$\begin{aligned} y &= \frac{1}{z} OTx \\ &= \frac{1}{z} Px, \end{aligned} \tag{2.3}$$

where O is the intrinsic parameter matrix of the camera and $P = OT$ is the projection matrix, whose rank is 3 and $z > 0$ is the depth value of the racket with respect to the camera, which is expressed as

$$z = p_3 x, \tag{2.4}$$

where p_j is the jth row of P. Furthermore, through differential operations, we have

$$\dot{y} = \frac{1}{z}\Gamma\dot{x}, \tag{2.5}$$

$$\Gamma = \left(P - y^T p_3\right), \tag{2.6}$$

where the matrix Γ is called the depth-independent image Jacobian matrix, since its elements are not related to the camera depth information. In the following proposition it holds that:

Property 1. *For any matrix-vector product of Γ and a vector ϖ, it holds that*

$$\Gamma\varpi = \Sigma\vartheta, \tag{2.7}$$

where Σ is a regression matrix, which excludes camera parameters. ϑ is a vector, which includes all unknown camera parameters.

2.2. Dynamics model

For a rigid robotic manipulator, we use the Lagrange equation to model its dynamic characteristics as follows:

$$M(q)\ddot{q} + \frac{1}{2}\dot{M}(q)\dot{q} + C(q,\dot{q})\dot{q} + g = \tau, \tag{2.8}$$

where $M(q)$ is the Inertia matrix, $C(q,\dot{q})$ is Coriolis force matrix, g represents the gravity effect, which is related to the robot state q, and τ is the joint torque acting on the robot joint, which is the system control input. Further, the robotic rigid dynamics equation holds for the following proposition:

Property 2. *For any vector a, we have*

$$a^T C(q,\dot{q})a = 0. \tag{2.9}$$

3. Controller design

The primary control objective is to drive the racket, mounted on the end-effector, to intercept the table tennis ball. This task is accomplished using visual feedback from a single, uncalibrated camera. The physical platform framework is as show in Figure 1.

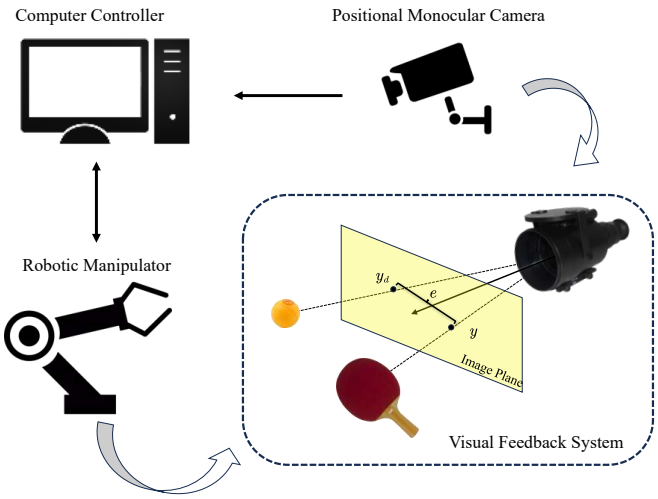

Figure 1. The physical platform framework employed.

3.1. Visual controller

To accomplish this task, we identify the positions of the table tennis racket and the table tennis ball in real time via visual feature recognition technology. Herein, the image feature position of the table tennis racket is denoted as y, and the image feature position of the table tennis ball is denoted as y_d. Therefore, the control error for this study is set in the following form:

$$e = y - y_d. \tag{3.1}$$

Applying the adaptive technology, the adaptive visual controller is designed in the following form:

$$\tau = g - K_1\dot{q} - J^T\left(\hat{\Gamma} + \frac{1}{2}e^T\hat{p}_3\right)K_2 e, \tag{3.2}$$

where K_1, K_2 are the adjustable control gain matrices, and $\hat{\Gamma}, \hat{p}_3$ are the estimation parameters of Γ, p_3, which are updated by an adaptive algorithm. In addition, $J \in \mathbb{R}^{3\times n}$ is the robotic linear velocity Jacobian matrix. It should be noted that this Jacobian maps the joint velocities to the time derivative of the 3D homogeneous coordinates, which is specifically required for our visual servoing formulation.

Combining with the robotic dynamics in Eq (2.8), we have the following closed-loop dynamics equation:

$$\begin{aligned} M(q)\ddot{q} + \frac{1}{2}\dot{M}(q)\dot{q} + C(q,\dot{q})\dot{q} &= -K_1\dot{q} - J^T\left(\hat{\Gamma} + \frac{1}{2}e^T\hat{p}_3\right)K_2 e \\ &= -K_1\dot{q} - J^T\left(\Gamma + \frac{1}{2}e^T p_3\right)K_2 e - J^T\left(\tilde{\Gamma} + \frac{1}{2}e^T\tilde{p}_3\right)K_2 e, \end{aligned} \tag{3.3}$$

where $\tilde{\Gamma} = \hat{\Gamma} - \Gamma, \tilde{p}_3 = \hat{p}_3 - p_3$ are the estimation errors of the adaptive parameters. It then follows from Property 1 that:

$$-J^T\left(\tilde{\Gamma} + \frac{1}{2}e^T\tilde{p}_3\right)K_2 e = Q(q,y)\tilde{\vartheta}, \tag{3.4}$$

where $Q(q,y)$ is a regression matrix, which is independent of the unknown camera parameters, and $\tilde{\vartheta} = \hat{\vartheta} - \vartheta$ is the estimation error.

3.2. Adaptive algorithm

Since in the actual practice environment the camera is usually hard to calibrate, the following adaptive rule is designed to update the controller parameters $\hat{\Gamma}, \hat{p}_3$. The adaptive law is designed as follows:

$$\dot{\hat{\vartheta}} = -K_3^{-1}Q^T(q,y)\dot{q}, \tag{3.5}$$

where K_3 is the positive updating gain matrix. It should be noted that the parameter adaptive update law designed in this paper is only applicable to updating the parameters of the image Jacobian matrix in the controller, and does not necessarily converge to its actual physical true value. The flowchart of the proposed calibration-free visual control algorithm is shown in Algorithm 1. The relevant convergence proof will be elaborated in the next section.

Algorithm 1 Adaptive visual control for a table tennis robot

Notation:

$q, \dot{q}$: Robot joint angle and velocity vectors.
y, y_d: Image coordinates of the racket and the ball, respectively.
e: Image error vector $(y - y_d)$.
$\hat{\vartheta}$: Vector of estimated unknown camera parameters.
τ: Joint input torque vector.
K_1, K_2, K_3: Positive-definite gain matrices.

Initialization:
Set gain matrices K_1, K_2, K_3.
Initialize the parameter estimation vector $\hat{\vartheta}(0)$.

while $e \neq \mathbf{0}$ **do**
    Step 1: Acquire real-time system state
    Read racket image position $y(t)$ and ball image position $y_d(t)$ from the camera.
    Read robot joint angle $q(t)$ and velocity $\dot{q}(t)$ from encoders.
    Step 2: Calculate image error
    $e(t) \leftarrow y(t) - y_d(t)$.
    Step 3: Calculate control torque using current estimates (3.2)
    Get robot's gravity term $g(q(t))$ and Jacobian $J(q(t))$.
    Construct estimated matrices $\hat{\Gamma}(t)$ and $\hat{p}_3(t)$ from the current parameter vector $\hat{\vartheta}(t)$.
    $F_{visual} \leftarrow J(q(t))^T \left(\hat{\Gamma}(t) + \frac{1}{2} e(t)^T \hat{p}_3(t)\right) K_2 e(t)$.
    $\tau(t) \leftarrow g(q(t)) - K_1 \dot{q}(t) - F_{visual}$.
    Apply torque $\tau(t)$ to the robot's joints.
    Step 4: Update parameter estimates for the next cycle (3.5)
    Construct the regressor matrix $Q(q(t), y(t))$ as defined in (3.4).
    Calculate the rate of change of parameters: $\dot{\hat{\vartheta}}(t) \leftarrow -K_3^{-1} Q(q(t), y(t)) \dot{q}(t)$.
    Update parameters via numerical integration: $\hat{\vartheta}(t + \Delta t) \leftarrow \hat{\vartheta}(t) + \dot{\hat{\vartheta}}(t) \cdot \Delta t$.
end while

4. Stability

In this section, we prove the stability of the proposed control scheme using the Lyapunov method. We assume that the robotic manipulator operates within a singularity-free workspace.

Theorem 1. *The robotic manipulator equipped with the controller (3.2) with the adaptive parameter updating rule (3.5) can catch the table tennis ball through visual feedback using an uncalibrated camera and the image error converges to zero.*

Proof. For the visual control system, the positive Lyapunov function is given in the following form:

$$L = \frac{1}{2} \left(\dot{q}^T M \dot{q} + z e^T K_2 e + \tilde{\vartheta}^T K_3 \tilde{\vartheta} \right). \tag{4.1}$$

Remark 1. *It should be noted that the term $ze^T K_2 e$ is positive, since the z with respect to the camera frame keeps a positive value. z is only used for the stability analysis, and is not required for the controller's implementation. Hence, this ensures that the overall function L remains positive definite, thereby satisfying the conditions required by Lyapunov's second method.*

Taking the time derivative of L yields:

$$\dot{L} = \dot{q}^T M\ddot{q} + \frac{1}{2}\dot{q}^T \dot{M}\dot{q} + \frac{1}{2}\dot{z}e^T K_2 e + e^T K_2 \dot{e} + \tilde{\vartheta}^T K_3 \dot{\tilde{\vartheta}}. \tag{4.2}$$

By pre-multiplying $\dot{q}$ to the closed-loop Eqs (3.3) and (3.4), we hold

$$\begin{aligned} \dot{q}^T M(q)\ddot{q} + \frac{1}{2}\dot{q}^T \dot{M}(q)\dot{q} + \dot{q}^T C(q,\dot{q})\dot{q} = & \\ - \dot{q}^T K_1 \dot{q} - \dot{q}^T J^T \left(\Gamma + \frac{1}{2}e^T p_3\right) K_2 e + \dot{q}^T Q(q,y)\tilde{\vartheta}. & \end{aligned} \tag{4.3}$$

Using the Proposition 2 and Eqs (2.2), (2.4), and (2.5), we can get

$$\dot{q}^T M(q)\ddot{q} + \frac{1}{2}\dot{q}^T \dot{M}(q)\dot{q} = -\dot{q}^T K_1 \dot{q} - z\dot{y}K_2 e - \frac{1}{2}\dot{z}e^T K_2 e + \dot{q}^T Q(q,y)\tilde{\vartheta}. \tag{4.4}$$

Substituting Eq (4.4) into Eq (4.2), we obtain the expanded form of $\dot{L}$:

$$\dot{L} = \left(-\dot{q}^T K_1 \dot{q} - z\dot{y}^T K_2 e + \frac{1}{2}\dot{z}e^T K_2 e + \dot{q}^T Q(q,y)\tilde{\vartheta}\right) + \frac{1}{2}\dot{z}e^T K_2 e + ze^T K_2 \dot{e} + \tilde{\vartheta}^T K_3 \dot{\tilde{\vartheta}}. \tag{4.5}$$

We now simplify this expression. Furthermore, as the gain matrix K_2 is positive-definite and symmetric ($K_2 = K_2^T$), the scalar term $ze^T K_2 \dot{e}$ is equal to its transpose, $z\dot{e}^T K_2^T e = ze^T K_2 \dot{e}$. Therefore, the terms $-z\dot{y}^T K_2 e$ and $+ze^T K_2 \dot{e}$ cancel each other out. The expression for $\dot{L}$ simplifies to:

$$\dot{L} = -\dot{q}^T K_1 \dot{q} + \dot{q}^T Q(q,y)\tilde{\vartheta} + \tilde{\vartheta}^T K_3 \dot{\tilde{\vartheta}}. \tag{4.6}$$

Next, we substitute the corrected adaptive law from Eq (3.5), $\dot{\tilde{\vartheta}} = -K_3^{-1} Q^T(q,y)\dot{q}$. Noting that ϑ is a constant vector, we have $\dot{\tilde{\vartheta}} = \dot{\hat{\vartheta}}$. The last term in Eq (4.6) becomes:

$$\tilde{\vartheta}^T K_3 \dot{\tilde{\vartheta}} = \tilde{\vartheta}^T K_3 \left(-K_3^{-1} Q(q,y)^T \dot{q}\right) = -\tilde{\vartheta}^T Q(q,y)^T \dot{q}. \tag{4.7}$$

The term $\dot{q}^T Q(q,y)\tilde{\vartheta}$ is a scalar and is therefore equal to its own transpose: $(\dot{q}^T Q\tilde{\vartheta})^T = \tilde{\vartheta}^T Q^T \dot{q}$. Substituting this back into Eq (4.6), we can see that the final cancellation occurs:

$$\begin{aligned} \dot{L} &= -\dot{q}^T K_1 \dot{q} + \left(\tilde{\vartheta}^T Q^T(q,y)\dot{q}\right) - \left(\tilde{\vartheta}^T Q^T(q,y)\dot{q}\right) \\ &= -\dot{q}^T K_1 \dot{q}. \end{aligned} \tag{4.8}$$

We have $L \geqslant 0$ and $\dot{L} \leqslant 0$. Hence, $L, e, \dot{q}$, and $\tilde{\vartheta}$ are bounded. Furthermore, $\ddot{q}$ is bounded and $\dot{q}$ is uniformly continuous. Apply Barbalet's lemma, and we conclude that

$$\lim_{t\to\infty} \dot{q} = 0. \tag{4.9}$$

Furthermore, at the equilibrium where $\lim_{t\to\infty} \dot{q} = 0$, the system dynamics reduce to:

$$J^T \left(\Gamma + \frac{1}{2} e^T p_3 \right) K_2 e = 0. \tag{4.10}$$

For the stability analysis to hold, it is required that the robot's trajectory $q(t)$ remains in a region where the manipulator Jacobian $J(q)$ is of full rank. Specifically, we define the singularity set as $\mathcal{S} = \{q \in \mathcal{Q} | \text{rank}(J(q)) < \min(m, n)\}$, where $\mathcal{Q}$ is the joint space and m, n are the dimensions of the task and joint spaces, respectively. The control objective is valid under the condition that the initial configuration $q(0)$ and the target y_d are chosen such that the resulting trajectory $q(t)$ does not enter $\mathcal{S}$. This implies that the term in the parentheses must be zero:

$$\left(\Gamma + \frac{1}{2} e^T p_3 \right) K_2 e = 0. \tag{4.11}$$

Since the gain matrix K_2 is positive-definite, $K_2 e = 0$ if and only if $e = 0$. Now, we prove by contradiction that $e = 0$ is the only possible solution to (4.11).

Assume that the system reaches equilibrium with a non-zero error, i.e., $e \neq 0$. Let $v = K_2 e$, so $v \neq 0$. The condition becomes:

$$\left(\Gamma + \frac{1}{2} e^T p_3 \right) v = 0. \tag{4.12}$$

Expanding $\Gamma = P - y^T p_3$, we get

$$\left(P - y^T p_3 + \frac{1}{2} e^T p_3 \right) v = 0. \tag{4.13}$$

$$Pv = \left(y^T p_3 - \frac{1}{2} e^T p_3 \right) v. \tag{4.14}$$

Recalling that $P = [p_1, p_2, p_3]^T$, we can write this as

$$[p_1, p_2, p_3]^T v = \alpha p_3, \tag{4.15}$$

$$v_1 p_1 + v_2 p_2 + v_3 p_3 = \alpha p_3, \tag{4.16}$$

$$v_1 p_1 + v_2 p_2 + (v_3 - \alpha) p_3 = 0, \tag{4.17}$$

where $\alpha = \left(y - \frac{1}{2} e \right) v$, v_i are the elements of the non-zero vector v. This equation states that the vectors p_1, p_2, p_3 are linearly dependent, a non-trivial linear combination of them equals zero. However, for a standard pinhole camera model, the projection matrix P must be of full rank, meaning its three row vectors (p_1, p_2, p_3) must be linearly independent. The result of linear dependence thus contradicts the fundamental properties of the camera projection model.

The contradiction arises from our initial assumption that $e \neq 0$. Therefore, the assumption must be false. This rigorously proves that the only possible equilibrium state is when the error is zero. Thus, we can conclude that:

$$\lim_{t\to\infty} e = 0. \tag{4.18}$$

□

Remark 2. *It is important to note that the proposed controller (3.2) does not include an explicit singularity avoidance mechanism. The stability proof relies on the Jacobian $J(q)$ remaining full-rank along the trajectory. In practical applications, this is often ensured through careful task planning, workspace definition, and selection of initial/target configurations far from known singular positions. For tasks where the robot must operate near singular configurations, the control law would need to be augmented with techniques such as the damped least-squares (DLS) method. The development of an explicit singularity-robust controller is an important direction for our future work.*

5. Simulation

In this section, a simulation using a PUMA robotic manipulator was conducted to verify the proposed control scheme. The simulation environment was constructed using MATLAB R2021a/Simulink. The physical characteristics of the robotic arm were described using the Denavit-Hartenberg (D-H) parameters and implemented with the Robotics Toolbox. The corresponding D-H parameters are presented in Table 1. The initial parameter configuration of the controller is as follows:
$K_1 = [30, 32, 18]$, $K_2 = diag([0.0035, 0.0045, 0.0025])$,
$K_3 = diag([0.0005, 0.0005, 0.0005, 0.0005, 0.0005, 0.0005, 0.0005, 0.0005, 50, 50, 50])$, and
$\hat{\vartheta}_0 = [759.1338, 364.8237, -352.3898, 478.9258, 6.3236, -162.8226, -890.7926, 511.7079, 0.9754, 10.2822, 8.9988]$,
where $diag(\cdot)$ is the diagonal matrix operator.

Table 1. Manipulator parameters.

Parameter Symbol	Joint 1	Joint 2	Joint 3
θ (rad)	q_1	q_2	q_3
d (m)	0	0	0.15005
a (m)	0	0.4318	0.0203
α (rad)	$\pi/2$	0	$-\pi/2$
m (kg)	0	17.4	4.8

*Note: θ = joint angle, d = offset, a = link length, α = twist angle, m = mass.

The control block diagram is shown in Figure 2. In the feedforward loop, the controller outputs the required torque for each joint, while the adaptive law is responsible for updating the variation of the adaptive parameters in the controller. In the feedback loop, the joint positions and velocities of the robotic arm are obtained through the encoders configured in the joints, while the image coordinates of the racket and table tennis ball are directly fed back by the vision camera. Considering the high-speed nature of table tennis rallies, the simulation time is set to 2 s in this study. The proposed algorithm is required to ensure that the error converges within 0.8 s. The following are the main experimental results of this simulation.

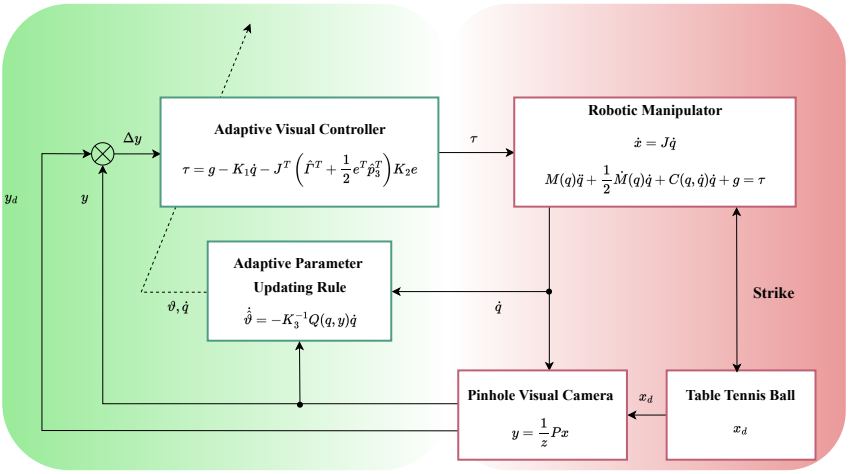

Figure 2. The framework of the proposed control scheme.

According to the simulation results, the variations of the joint velocities and joint angles of the robotic manipulator over time are shown in Figure 3. It can be observed that as time progresses, the system tends to stabilize, with the velocities decaying to zero and the joint angular positions approaching stability. This indicates that the proposed control scheme effectively stabilizes the system through visual feedback. The convergence of the image error over time is depicted in Figure 4. As shown, the image error converges to a value close to zero at approximately 0.6 s. By 0.8 s, the system error has completely converged to zero, further demonstrating the effectiveness and applicability of the proposed method. Furthermore, the trajectory of the racket is illustrated in Figure 5. We can see that after rapid movement, the racket reaches the desired position. In addition, the control torque output of the actuator is shown in Figure 6. It can be seen that there is no obvious chattering in the torque output in this simulation, suggesting that the system avoids singularities. This satisfies the condition that the Jacobian matrix J is of full rank as mentioned in Theorem 1. As can be seen from Figure 7, the adaptive parameters are updated in real time along with the convergence of the system, which further ensures the stability and convergence of the system.

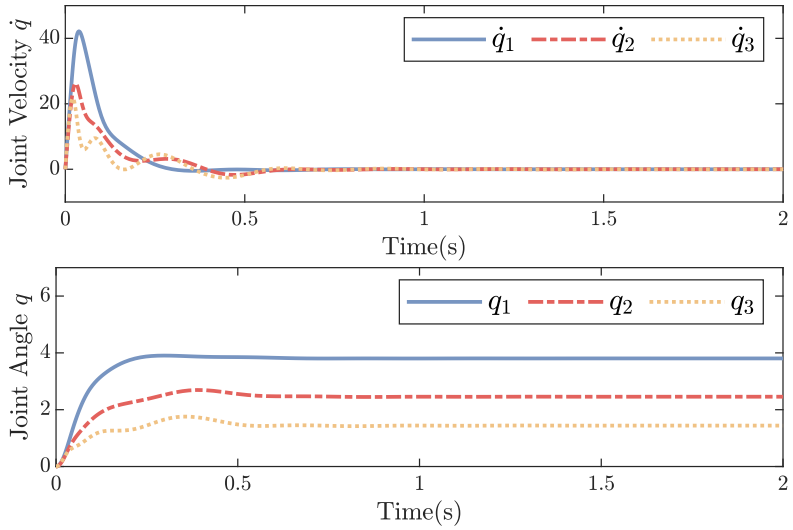

Figure 3. The variation of robot joint velocity $\dot{q}$ and angles q with time.

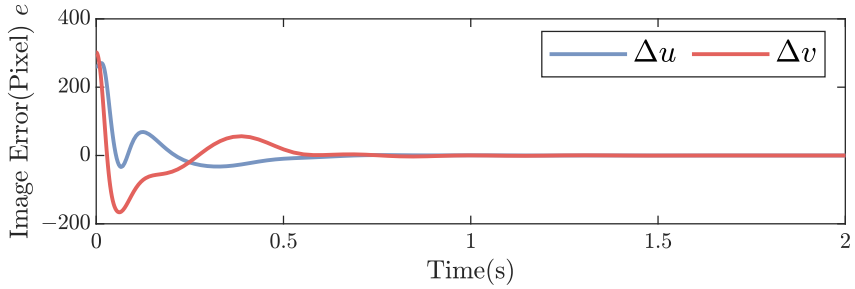

Figure 4. The variation of system error e with time.

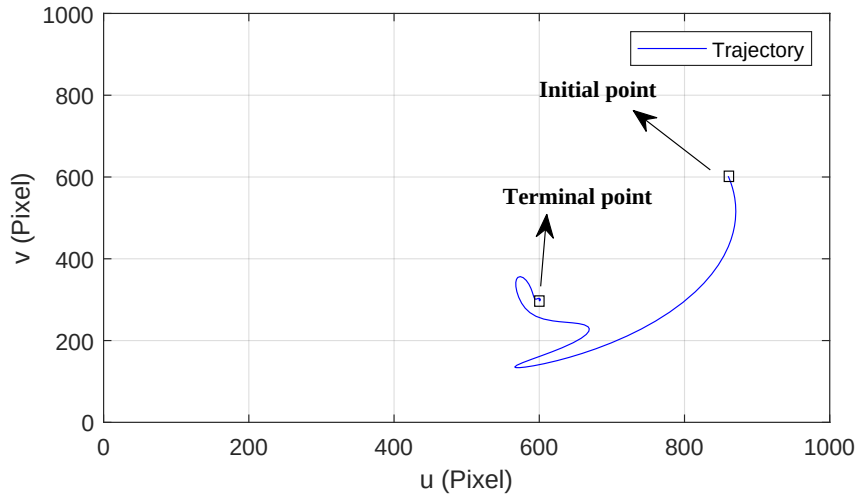

Figure 5. Image tracking trajectory of the table tennis racket.

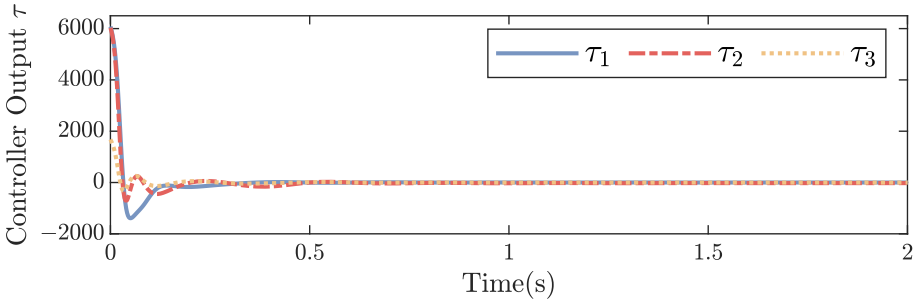

Figure 6. The variation of controller torque output τ with time.

To further validate the applicability of the proposed scheme, two additional experiments were conducted. First, based on the above experiment, the robotic manipulator was used to drive the racket to hit different targets, with the norm of the error adopted for evaluation. The experimental results are shown in Figure 8. It can be observed that the system exhibits excellent convergence performance in table tennis practice with different targets. Furthermore, we designed an experiment where the table tennis ball continuously hits different targets within 10 seconds to simulate the scenario of continuous training. The error convergence of this experiment is shown in Figure 9. It can be seen that the proposed scheme still maintains excellent control performance in the continuous hitting training environment.

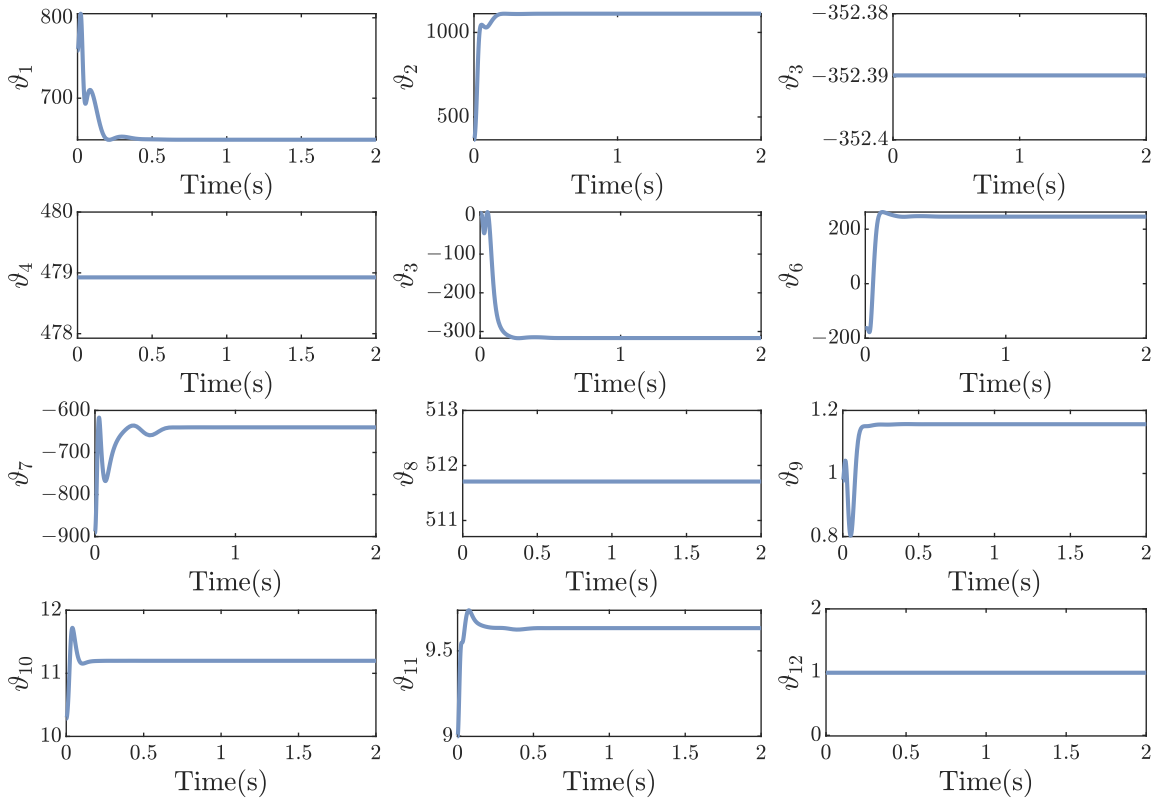

Figure 7. The variation of adaptive parameter $\hat{\vartheta}$ with time.

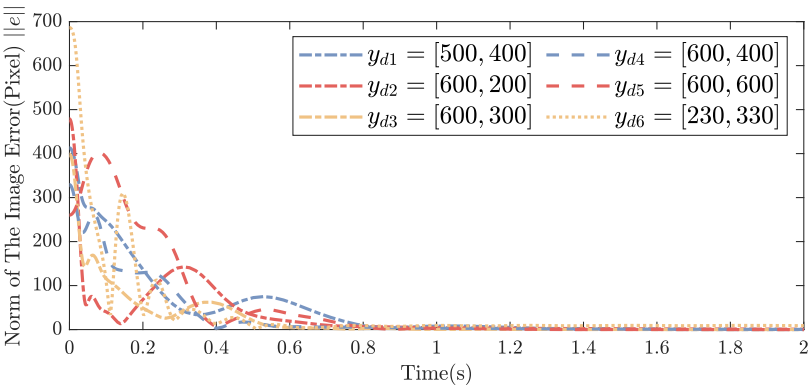

Figure 8. The convergence of the norm of image errors $\|e\|$ for different tracking endpoints.

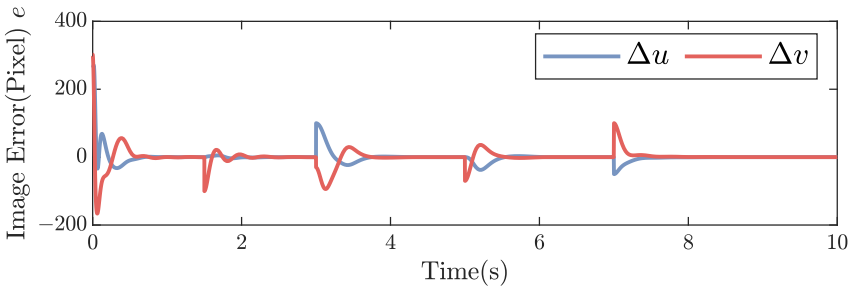

Figure 9. The image error convergence in continuous strokes tests.

To validate the controller's ability to handle the high-speed dynamics characteristic of a table tennis rally, a dynamic tracking experiment was designed. In this test, the desired target position in the image plane y_d was subjected to a sequence of high-frequency step changes at 0.07-second intervals (approximately 14.3 Hz). The results, presented in Figure 10, demonstrate the controller's excellent tracking performance. As can be seen, the tracking error consistently and rapidly converges to zero after each abrupt change in the target. A quantitative analysis of this response shows that the average settling time is 0.920 second.

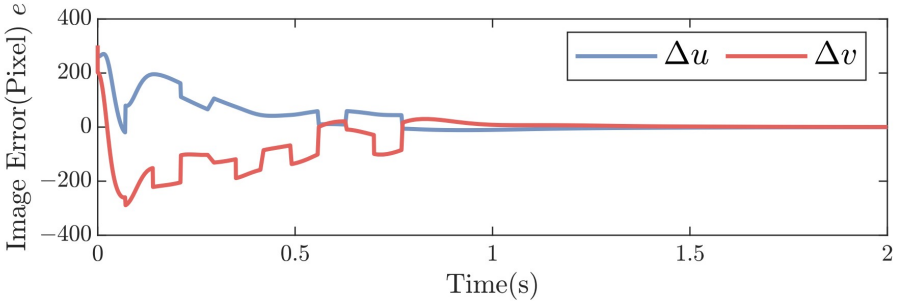

Figure 10. Convergence of the image error norm in response to rapid step changes in the target position.

Furthermore, to assess the algorithm's robustness under non-ideal conditions, a second experiment was established. In this scenario, the system was simultaneously subjected to several real-world factors: 1) Broadband white noise (Simulink's White Noise block, power = 0.1) was added to both the control signals and the image feature measurements; and 2) a random time delay, uniformly distributed between 0 and 50 ms, was introduced into the visual feedback loop. The results of this robustness test are shown in Figure 11. It is evident that even in the presence of these combined disruptive factors, the controller maintains system stability, confining the error to a small, bounded region around the origin. The steady-state root-mean-square (RMS) error is calculated to be 5.26 pixels, with the maximum error peak not exceeding 15.77 pixels, confirming the practical robustness of the proposed method.

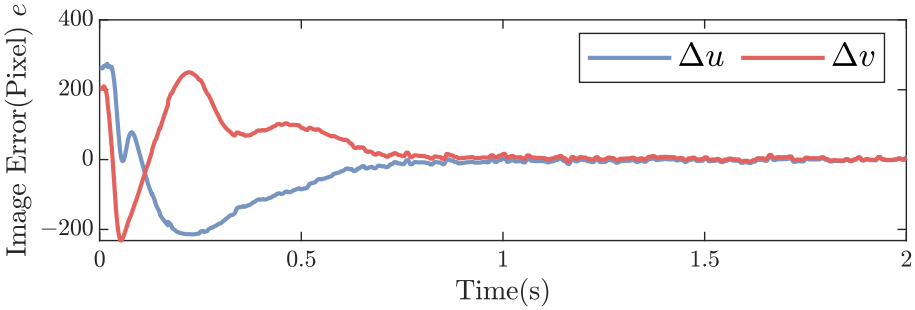

Figure 11. error convergence under dynamic tracking with simulated noise and delays.

6. Conclusions

This paper proposes a calibration-free visual control technique for table tennis training robots. The proposed control framework does not require precise camera parameters, but instead requires only the image information fed back by the camera. Within this framework, the unknown camera parameters

in the controller are updated in real time based on the system state through adaptive algorithms. The stability of the proposed algorithm was proven by the Lyapunov method, and the effectiveness and practicality of the proposed method were verified through simulation experiments.

Author contributions

Quanyu Song: Conceptualization, methodology, software, validation, formal analysis, visualization, writing-original draft; Rong Lu: Data curation, writing-review & editing; Yaobo Long: Investigation, supervision; Xiaobing Zheng: Conceptualization, formal analysis, funding acquisition, project administration, writing-review & editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research is jointly supported by the School-Level Fund Project of Guangdong University of Science and Technology (No. GKY-2024BSQDW-89), Special Innovation Program for General Colleges and Universities in Guangdong Province (2025KTSCX373), Artificial Intelligence Technology Application Research and Service Center of Dongguan Polytechnic, School-Level Fund Project of Dongguan Polytechnic (No. 2024a01), and SSL Sci-tech Commissioner Program (No. 20234400-01KCJ-G).

Conflict of interest

All authors declare no conflict of interest in this paper.

References

1. D. Seemiller, M. Holowchak, *Winning table tennis*, Human Kinetics, 1997.
2. R. McAfee, *Table tennis: Steps to success*, Human Kinetics, 2009. https://doi.org/10.5040/9781718219250
3. I. M. Lanzoni, R. D. Michele, F. Merni, Performance indicators in table tennis: A review of the literature, *Eur. J. Sport Sci.*, **14** (2014), 492–499. https://doi.org/10.1080/17461391.2013.819382
4. M. Pfeiffer, H. Zhang, A. Hohmann, A markov chain model of elite table tennis competition, *Int. J. Sports Sci. Coa.*, **5** (2010), 205–222. https://doi.org/10.1260/1747-9541.5.2.205
5. C. Ferrandez, T. Marsan, Y. Poulet, P. Rouch, P. Thoreux, C. Sauret, Physiology, biomechanics and injuries in table tennis: A systematic review, *Sci. Sport.*, **36** (2021), 95–104. https://doi.org/10.1016/j.scispo.2020.04.007

6. G. Martinent, E. Ansnes, A literature review on coach-athlete relationship in table tennis, *Int. J. Racket Sport. Sci.*, **2** (2020), 9–21. https://doi.org/10.30827/digibug.63717
7. Z. Zhang, D. Xu, M. Tan, Visual measurement and prediction of ball trajectory for table tennis robot, *IEEE T. Instrum. Meas.*, **59** (2010), 3195–3205. https://doi.org/10.1109/tim.2010.2047128
8. F. Basiri, A. Farsi, B. Abdoli, M. Kavyani, The effect of visual and tennis training on perceptual-motor skill and learning of forehand drive in table tennis players, *J. Modern Rehabilitation*, **14** (2020), 21–32. https://doi.org/10.32598/jmr.14.1.3
9. H. Ripoll, Uncertainty and visual strategies in table tennis, *Percept. Motor Skill.*, **68** (1989), 507–512. https://doi.org/10.2466/pms.1989.68.2.507
10. R. Nakazato, C. Aoyama, T. Komiyama, R. Himo, S. Shimegi, Table tennis players use superior saccadic eye movements to track moving visual targets, *Front. Sports Act. Liv.*, **6** (2024), 1289800. https://doi.org/10.3389/fspor.2024.1289800
11. X. Wang, J. Li, Towards a novel paradigm in brain-inspired computer vision, *Front. Neurorobotics*, **19** (2025). https://doi.org/10.3389/fnbot.2025.1592181
12. Y. Chen, *Intelligent tennis ball pick-up trolley based on visual recognition*, In: 2025 International Conference on Electrical Drives, Power Electronics & Engineering (EDPEE), IEEE, 2025, 628–633. https://doi.org/10.1109/edpee65754.2025.00114
13. H. I. Lin, Z. Yu, Y. C. Huang, Ball tracking and trajectory prediction for table-tennis robots, *Sensors*, **20** (2022), 333. https://doi.org/10.3390/s20020333
14. G. L. Cai, A method for prediction the trajectory of table tennis in multirotation state based on binocular vision, *Comput. Intel. Neurosc.*, **2022** (2022), 8274202. https://doi.org/10.1155/2022/8274202
15. A. Ziegler, T. Gossard, A. Glover, A. Zell, An event-based perception pipeline for a table tennis robot, *arXiv Preprint*, 2025. https://doi.org/10.48550/arXiv.2502.00749
16. Z. Zhao, H. Huang, S. Sun, J. Jin, W. Lu, *Reinforcement learning for dynamic task execution: A robotic arm for playing table tennis*, In: 2024 IEEE International Conference on Robotics and Biomimetics (ROBIO), IEEE, 2024, 608–613. https://doi.org/10.1109/robio64047.2024.10907471
17. L. Yang, H. Zhang, X. Zhu, X. Sheng, Ball motion control in the table tennis robot system using time-series deep reinforcement learning, *IEEE Access*, **9** (2021), 99816–99827. https://doi.org/10.1109/access.2021.3093340
18. D. Ma, X. Hu, J. Shi, M. Patel, R. Jain, Z. Liu, et al., *avattar: Table tennis stroke training with embodied and detached visualization in augmented reality*, In: Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology. UIST '24, Association for Computing Machinery, New York, USA, 2024. https://doi.org/10.1145/3654777.3676400
19. A. Ziegler, T. Gossard, K. Vetter, J. Tebbe, A. Zell, A multi-modal table tennis robot system, *arXiv Preprint*, 2023. https://doi.org/10.48550/arXiv.2310.19062
20. J. Jiao, Z. Li, G. Xia, Y. Chen, J. Xin, Adaptive neural network control of manipulators with input deadband and field-of-view constraints, *Appl. Math. Model.*, **150** (2026), 116311. https://doi.org/10.1016/j.apm.2025.116311

21. T. A. Clarke, J. G. Fryer, The development of camera calibration methods and models, *Photogramm. Rec.*, **16** (1998), 51–66. https://doi.org/10.1111/0031-868x.00113

22. X. Li, Y. Xiao, B. Wang, H. Ren, Y. Zhang, J. Ji, Automatic targetless lidar–camera calibration: A survey, *Artif. Intell. Rev.*, **56** (2023), 9949–9987. https://doi.org/10.1007/s10462-023-10404-0

23. J. Lee, H. Go, H. Lee, S. Cho, M. Sung, J. Kim, *Ctrl-c: Camera calibration transformer with line-classification*, In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2021, 16228–16237. https://doi.org/10.1109/iccv48922.2021.01592

24. L. Jin, J. Zhang, Y. H. Geoffroy, O. Wang, K. B. Matzen, M. Sticha, et al., *Perspective fields for single image camera calibration*, In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2023, 17307–17316. https://doi.org/10.1109/cvpr52729.2023.01660

25. C. Jiang, Y. Yang, M. Jagersand, *Clipunetr: Assisting human-robot interface for uncalibrated visual servoing control with clip-driven referring expression segmentation*, In: 2024 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2024, 6620–6626. https://doi.org/10.1109/icra57147.2024.10611647

26. M. Cao, L. Xiao, Q. Zuo, X. Yan, L. Li, X. Gao, Uncalibrated model-free visual servo control for robotic endoscopic with rcm constraint using neural networks, *IEEE T. Cybernetics*, **55** (2025), 4541–4553. https://doi.org/10.1109/TCYB.2025.3582866

27. J. Jiao, Z. Li, G. Xia, J. Xin, G. Wang, Y. Chen, An uncalibrated visual servo control method of manipulator for multiple peg-in-hole assembly based on projective homography, *J. Franklin I.*, **362** (2025), 107572. https://doi.org/10.1016/j.jfranklin.2025.107572

28. E. Salvato, F. Blanchini, G. Fenu, G. Giordano, F. A. Pellegrino, Position-based visual servo control without hand-eye calibration, *Robot. Auton. Syst.*, **193** (2025), 105045. https://doi.org/10.1016/j.robot.2025.105045

AIMS Press