

Article

Grasping Force Optimization and DDPG Impedance Control for Apple Picking Robot End-Effector

Xiaowei Yu ¹, Wei Ji ^{1,2,*} , Hongwei Zhang ¹ , Chengzhi Ruan ^{2,3}, Bo Xu ¹ and Kaiyang Wu ¹

¹ School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China; 18879209223@163.com (X.Y.); zhanghongwei991016@163.com (H.Z.); xubo@ujs.edu.cn (B.X.); 2222307061@stmail.ujs.edu.cn (K.W.)

² The Key Laboratory for Agricultural Machinery Intelligent Control and Manufacturing of Fujian Education Institutions, Wuyishan 354300, China; ranczhi@163.com

³ College of Mechanical and Electrical Engineering, Wuyi University, Wuyishan 354300, China

* Correspondence: jiwei@ujs.edu.cn; Tel.: +86-13862445908

Abstract: To minimize mechanical damage caused by an apple picking robot end-effector during the apple grasping process, and on the basis of optimizing the minimum stable grasping force of apple, a variable impedance control strategy based on a reinforcement learning deep deterministic policy gradient (DDPG) algorithm is proposed to achieve compliant grasping control for apples. Firstly, according to the apple contact force model, the gradient flow algorithm is adopted to optimize grasping force in terms of the friction cone, force balancing condition, and stability assessment index and to obtain a minimum stable grasping force for apples. Secondly, based on the analysis of the influence of impedance parameters on the control system, a variable impedance control based on the DDPG algorithm is designed, with the reward function adopted so as to improve the control performance. Then, the improved control strategy is used to train the optimized impedance control. Finally, simulation and experimental results indicate that the proposed variable impedance control outperforms the traditional impedance control by reducing the peak grasping force from 4.49 N to 4.18 N while achieving a 0.6 s faster adjustment time and a 0.24 N narrower grasping force fluctuation range. The improved impedance control successfully tracks desired grasping forces for apples of varying sizes and significantly reduces mechanical damage during apple harvesting.



Academic Editor: José Lima

Received: 28 March 2025

Revised: 3 May 2025

Accepted: 6 May 2025

Published: 8 May 2025

Citation: Yu, X.; Ji, W.; Zhang, H.; Ruan, C.; Xu, B.; Wu, K. Grasping Force Optimization and DDPG Impedance Control for Apple Picking Robot End-Effector. *Agriculture* **2025**, *15*, 1018. <https://doi.org/10.3390/agriculture15101018>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: apple picking robot; end-effector; reinforcement learning; impedance control; grasping force; DDPG

1. Introduction

The global agriculture sector is facing acute labor shortages in fruit harvesting, with the FAO's 2024 statistical yearbook showing that the agricultural workforce proportion has declined from 40% in 2000 to 26% in 2022, a trend that has both accelerated demand for automated harvesting solutions and spurred significant advancements in intelligent apple picking robot development [1,2]. An apple picking robot generally consists of a walking device, robot arm, end-effector, vision system and controller [3–5]. As an important part of an apple picking robot, the end-effector needs to be further studied for its efficient control technology [6,7].

The first part of the apple picking process is the selection of an appropriate grasping force to avoid the apple grasping failure that is caused by a grasping force that is too small or too large [8,9]. In terms of grasping force optimization, Jia et al. [10] reduced and

linearized grasping system contact limitations and created an optimization model with joint torque as the objective function. Here, a crucial constraint-set-based real-time effective method optimized joint torque and promoted dexterous hand control. Zhang et al. [11] presented an inexact multiblock alternating direction method for the contact-point friction model of the force optimization problem (FOP). A parallel inexact multiblock alternating direction method was used to solve the FOP and ultimately demonstrated a global convergence toward the proposed method. Zhang et al. [12] used obstacle functions to create regularization optimization issues, represented objective functions with multiple dimensions as compromises, and added a penalty element to create an enhanced optimization objective function. Adjusting the penalty factor might create a more compact, steady, relaxed, or flexible grasping scheme for a certain operating job. Mu et al. [13] suggested grasping force optimization with projection and contraction. The projection and contraction methods were globally convergent to the optimal grasping force, making them ideal for hot start technology. Li et al. [14] examined a dexterous hand as it grasped diverse objects and planned a robot hand grasping position and finger force that were able to accomplish steady grasping with little force.

The above studies are related to the optimization of grasping force, and can provide a basis for the acquisition of the grasping force of an apple. To achieve compliant control of the end-effector, it is essential to address the specific challenges in force control for fruits, including sensitivity to mechanical damage, high size variability and surface irregularities, while employing suitable force control algorithms. To prevent mechanical damage during fragile fruit picking, Wang et al. [15] employed a tactile-based closed-loop grasping force control method and demonstrated its effectiveness in reducing early-stage bruising through spatial frequency domain imaging (SFDI) experiments. Lama et al. [16] achieved adaptive force control through a spring-based rotating mechanism, enabling the stable grasping of objects ranging from 5 to 9 cm in diameter with varying sizes and shapes, without needing a closed-loop control system. Rabenoroso et al. [17] presented a hybrid force/position control strategy that treated the robot position and force free spaces as complementary orthogonal subspaces and conducted position control in one and force control in another. For drone–environment control issues, a force/position hybrid control strategy-based drone gliding grasp analysis approach was presented to improve a drone’s dynamic glide grasp control [18]. Kumar et al. [19] suggested a neural network-based force/position hybrid control strategy for limited reconfigurable manipulators. The boundedness of position and force tracking errors was achieved and the tracking error of the joints converged to zero asymptotically. The switching force/position hybrid control scheme can easily generate system instability by switching the end-effector force and position control. For compliant grasping control with an imperfect environmental model, this strategy is unsuitable. Hogan [20] suggested impedance control, which was intended to convert end-effector–environment interaction to inertia, damping, and stiffness. Adjusting impedance settings provided an optimal dynamic link between robot end force and position. Xie et al. [21] have suggested an integrated contact force control system using expected force feedforward and adaptive variable impedance control. The nonlinear tracking differentiator smoothed the expected force input step signal to reduce force overshoot. Then, modeling the force tracking error created an adaptive damping rule that was used to compensate for the disturbance.

While the above studies have considered grasping force optimization and force control algorithms to a certain extent, the application of agricultural picking robots is relatively lacking. This study will conduct further research in order to optimize the minimum stable grasping force for apples under comprehensive valuation index constraints, and will apply the idea of reinforcement learning to the robot apple picking control in order to achieve compliant grasping. Therefore, the main contributions of this paper are as follows:

- (1) According to the apple contact force model, the gradient flow algorithm is used to optimize the contact force under friction cone, force balance condition, and stability evaluation index constraints. The minimum stable apple grasping force is calculated to provide the desired grasping force for the control strategy.
- (2) A deep deterministic policy gradient (DDPG) optimized variable impedance control is designed. The position-based impedance control is improved by using reinforcement learning DDPG algorithm, and the reward function is designed to correct the apple grasping damping coefficient and enhance the control performance during continuous grasping.
- (3) An end-effector grasping experimental platform is setup to verify the effectiveness of the proposed method. The simulation and experimental results indicate that the improved impedance control has a smoother grasping force and better dynamic performance for varying sizes apples.

2. Materials and Methods

2.1. Apple Grasping Contact Force Optimization

When the picking robot end-effector grasps the apple, the force directly affects the success or failure of the grasping task. According to the apple contact force model [22], the grasping force could be computed from fruit deformation, but this is only considered the maximum grasping force. The minimum stable contact force can be obtained by optimization, and the following will introduce apple picking robot grasping force optimization.

2.1.1. Constraint Conditions

The picking robot end-effector developed by our research consists of two parallel aluminum hemispherical grippers (80 mm diameter, 24 mm depth) driven by a stepper motor and a stainless-steel cutting blade driven by a DC motor (shown in Figure 1a). Additionally, the classic Coulomb friction model [23] is selected in the friction point contact model, as shown in Figure 1b. The diagram shows one of the two fingers, where f_x and f_y are the grasping force components along the local coordinate system x and y at the contact point tangent direction, respectively, and f_z is the grasping force component along the contact surface normal direction.

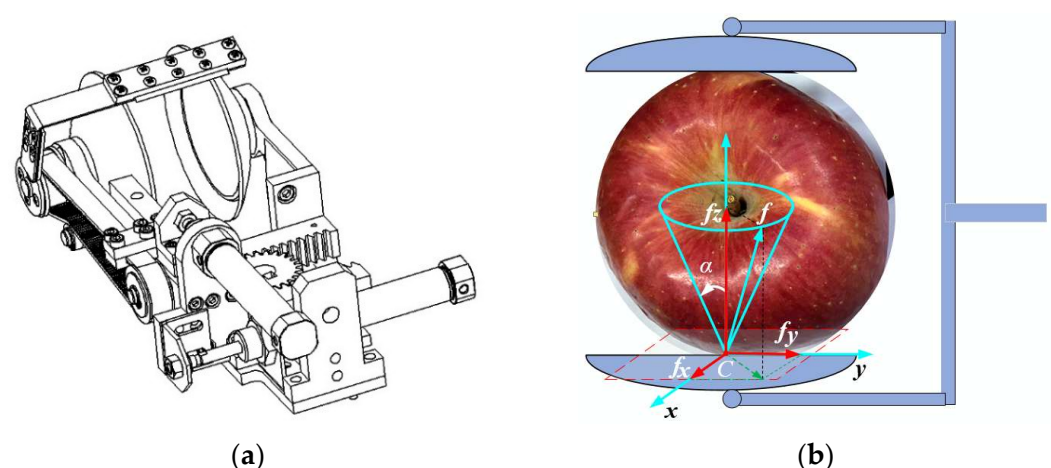


Figure 1. Friction cone with a friction point contact schematic diagram of a two-finger end-effector. (a) Two-finger end-effector and (b) friction cone with friction point contact.

In three-dimensional space, when the direction of the object changes its motion trend, the direction of the grasping force that provided static friction also changes. The continuously changing grasping force would then form a friction cone FC in space with the contact

point C as the vertex and the normal grasping force as the axis. The angle α between the grasping force and the z-axis is called the half-apex angle of the friction cone, as shown in Equation (1).

$$\alpha = \arctan \mu \quad (1)$$

where μ is the static friction coefficient and the friction coefficient between the end-effector and the captured apple is 0.412 [24].

The apple picking robot end-effector should meet the friction coagulation and force–balance constraint conditions, as shown in Equations (2) and (3).

$$FC = \left\{ f \in R^3 \mid \sqrt{(f_x^2 + f_y^2)} \leq \mu f_z, f_z \geq 0 \right\} \quad (2)$$

$$F_e = -F = -G \cdot f_c \quad (3)$$

where $F_e \in R^6$ is the external force spiral, F is the grasping force spiral mapping at the apple mass center, $G = [G_1 G_2 \cdots G_n] \in R^{6 \times 3n}$ is the grasping matrix, and $f_c = [f_1 f_2 \cdots f_n]^T \in R^{3n}$ is the grasping vector with each grasping force.

2.1.2. Stability Evaluation Index Grasping Force Optimization Model

Assuming that the apple external force spiral is known, the general solution for end grasping force expression can be obtained according to Equation (3):

$$f_c = -G^+ F_e + N(G) \lambda \quad (4)$$

where G^+ is the generalized inverse grasping matrix, $N(G)$ is the grasping matrix null space vector, and $\lambda \in R^{3n}$ is an arbitrary vector in the real number set.

In order to avoid the friction cone boundary influence on the grasping stability, this paper proposes a grasping stability evaluation index to measure the end-effector ability to resist external force disturbance during grasping of the object. The grasping stability evaluation index φ_i at the contact point i is defined as follows:

$$\varphi_i = \frac{\beta}{\alpha} \quad (5)$$

where α is the friction cone half-apex angle, and β is the inner normal vector angle.

The inner normal vector angle is the angle between the grasping force f_i and the tangent plane normal line f_{iz} , as shown in Figure 2. Where f_i is the grasping force at the contact point C_i , $f_{it} = \sqrt{f_{ix}^2 + f_{iy}^2}$ is the resultant force of the tangential force components f_{ix} and f_{iy} , and f_{iz} is the normal force of f_i along the internal normal vector direction. To make the end-effector stably grasp the apple, the grasping force at the contact point C_i must be satisfied in the friction cone, that is, $0 \leq \beta \leq \alpha < 90^\circ$. The angle of the inner normal vector is expressed as follows:

$$\beta = \arctan \frac{f_{it}}{f_{iz}} \quad (6)$$

The grasping stability evaluation index describes the deviation between the grasping force and the friction cone constraint boundary degree so that the possible solution can avoid instability in the critical stage.

To determine the minimum stable grasping force, the norm ω of the grasping force is optimized as the objective function. The expression for the objective function is as follows:

$$\omega(f_c) = f_c^T Q f_c \quad (7)$$

where Q is the weight matrix [25], usually the unit matrix. Therefore, the minimum stable grasping force optimization model for the end-effector, based on the stability evaluation index, is as follows:

$$\begin{aligned} \text{obj. min } \omega(f_c) &= f_c^T Q f_c \\ \text{s.t. } \begin{cases} G f_c + F_e = 0 \\ FC \\ \varphi \end{cases} \end{aligned} \quad (8)$$

where $\omega(f_c)$ is the grasping force minimum value under the constraint condition of stable grasping and where s.t. represents the three constraints for the objective function, which are the force balance equation, friction cone and grasping stability evaluation index. Each variable content has been described.

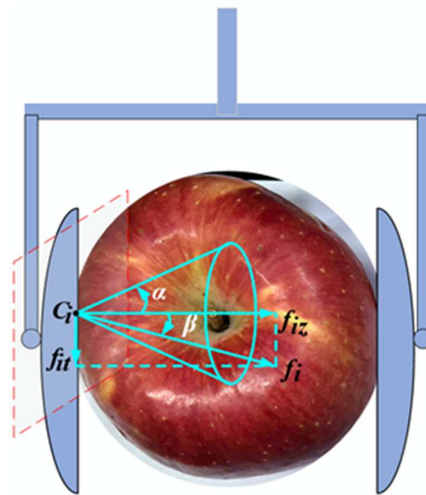


Figure 2. Inner normal vector angle diagram.

2.2. Solution of Contact Force Optimization Model

Equation (8) is a quadratic programming problem with equality constraints and inequality constraints. In the study, the force balance constraints and friction cone constraints can be linearized, and the gradient flow is used to solve the problem [26]. The optimal grasping internal force f_n under the constraint of grasping stability evaluation index is obtained and then substituted into Equation (4) to obtain the minimum stable grasping force f_c . The minimum stable grasping force is used as the desired input force of the control system to guide the end-effector to complete the apple stable grasping operation.

In Lei et al. [27], the nonlinear friction cone constraint condition is turned into a positive, definite linear constraint. Grasping force linear programming was made possible after linearizing all stable grasping constraints. Gradient flow solves optimization problems with affine constraints [28]. Thus, for a given external force spiral, the initial iteration value is first calculated by Equation (4). Then, the iterative operation judgment error a is selected as the iterative process-stopping criterion, and the linearized optimization model is solved by using the gradient flow algorithm shown in Figure 3.

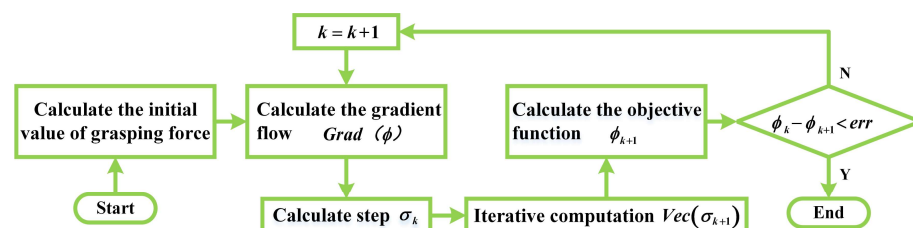


Figure 3. Optimization algorithm flow chart.

Here, $Grad(\phi)$ is the gradient operator, $\phi = tr(w_1 \cdot I \cdot P) + tr(w_2 \cdot I \cdot P^{-1})$ is the optimization objective function, tr represents the trace operator of the matrix, w_1 and w_2 are two weight parameters, I is a unit matrix, and err is the error threshold to stop the iteration, which is set at 0.03. $tr(w_1 \cdot I \cdot P)$ represents the magnitude of the grasping force generated by the end-effector and $tr(w_2 \cdot I \cdot P^{-1})$ indicates the relative distance between the grasping force and the friction cone boundary. P is the description matrix, $P = diag(P_1, P_2, \dots, P_n)$, as shown in Equation (9); σ_k is the iteration step size, as shown in Equation (10); $Vec(P_{k+1}) = Vec(P_k) + \sigma_k \cdot Vec(\dot{P}_k)$ is the iteration formula; $Vec(\dot{P}) = -Grad(\phi) = (I - A^+ \cdot A) \cdot Vec(P^{-1} \cdot w_2 \cdot I \cdot P^{-1} - w_1 \cdot I)$ is the linear gradient flow; A is a linear constraint matrix, $A = diag(A_1, A_2, \dots, A_n)$, as shown in Equation (11); A^+ is the Moore–Penrose generalized inverse matrix; and k is the iteration step number.

$$P_i = \begin{bmatrix} \mu f_{iz} & 0 & f_{ix} \\ 0 & \mu f_{iz} & f_{iy} \\ f_{ix} & f_{iy} & \mu f_{iz} \end{bmatrix} \geq 0 \quad (9)$$

$$\sigma_k = \frac{0.02}{\max |Vec(\dot{P}_k)|} \quad (10)$$

$$A_i = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (11)$$

2.3. DDPG-Optimized Variable Impedance Control for Robot End-Effector

In a complex environment, the traditional impedance control cannot adjust the impedance parameters in time for different sizes of apples, resulting in the failure to track the different forces accurately. Reinforcement learning can learn in the process of interacting with the environment to adapt to the needs of that environment, and can be used to solve the problems of traditional impedance control [29]. Therefore, a variable impedance control system based on reinforcement learning DDPG algorithm optimization is presented as shown in Figure 4.

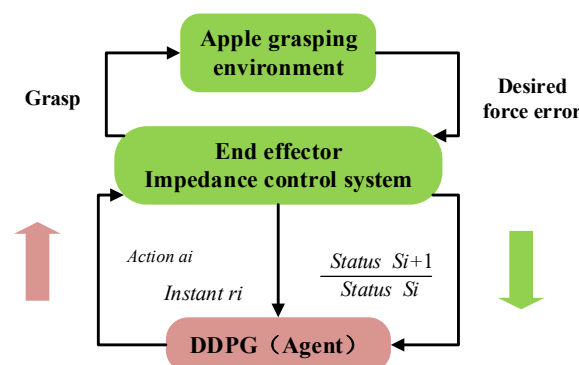


Figure 4. End-effector grasping force impedance control system based on DDPG.

The DDPG can select the optimal strategy to compensate for the impedance parameters according to the amount of reward, so as to correct the impedance characteristics in different grasping stages, track the grasping force of the target, avoid overshoot, and achieve compliant grasping.

2.3.1. Grasping Contact Process Analysis

Impedance control establishes a dynamic link between robot end position and grasping force. Synchronous position and force control is achieved by dynamically altering the robot end-to-work environment contact form [30]. Impedance control in apple grasping relies on the damping spring mass system and first-order spring system models. Apples with complex biology are simulated using the first-order spring model. The grasping contact process between the end-effector and environment (target fruit) can be divided into three stages based on their relative positions, as shown in Figure 5. In Figure 5, K_e represents the stiffness coefficient of the external environment, X_e represents the environment location, and X represents the actual position of the end-effector. Stage I $X < X_e$ is the no-load feeding stage; Stage II $X = X_e$ is the contact deceleration stage; and Stage III $X > X_e$ is the stable clamping stage.

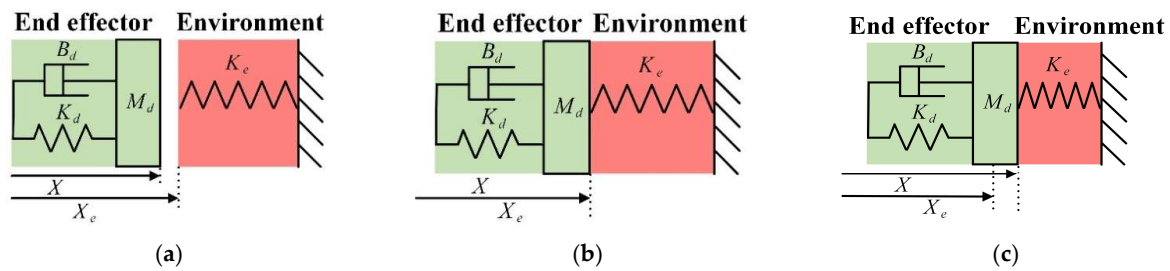


Figure 5. The grasping process of the end-effector and environment (target fruit): (a) Stage I; (b) Stage II; (c) Stage III.

In the grasping target fruit process by the end-effector, the above three stages of grasping force changes are shown in Figure 6. In the figure, $0 \sim t_1$ corresponds to Stage I. At this time, the end-effector moves in free space. As the end-effector does not contact the target fruit, the stage does not produce force, so $F = 0$. In the period of $t_1 \sim t_2$, the end-effector contacts the fruit, and the grasping process enters Stage II. The fruit deforms in the constraint space formed by the end finger, and the interaction force F is generated in the contact area between the two. After an adjustment period, the desired grasping force is finally achieved in the $t_2 \sim t_3$ period, and the stable clamping state shown in Stage III is entered.

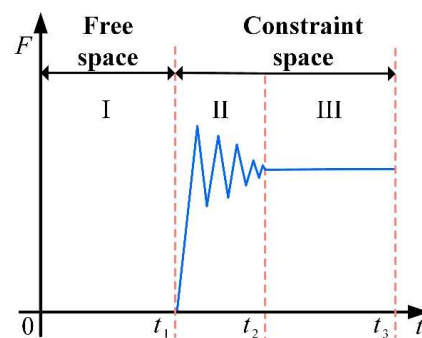


Figure 6. Schematic diagram of contact force change.

According to the above analysis, when the robot end-effector moves in free and constrained space, the actual grasping force between the end-effector and the environment can be expressed as follows:

$$F = \begin{cases} 0 & X < X_e \\ K_e(X - X_e) & X \geq X_e \end{cases} \quad (12)$$

2.3.2. Impedance Control Based on DDPG Algorithm Optimization

Traditional Impedance Control

To analyze the end-effector grasping force tracking problem, the desired grasping force is supplied, and the force deviation is employed as the impedance control input signal. In addition, the impedance equation includes velocity and acceleration terms. With Equation (12), the ambient contact impedance model is as follows:

$$M_d(\ddot{X} - \ddot{X}_r) + B_d(\dot{X} - \dot{X}_r) + K_d(X - X_r) = F_r - K_e(X - X_e) = \Delta F \quad (13)$$

where M_d , B_d and K_d are inertia parameter, damping parameter and stiffness parameter, respectively; X_r , $\dot{X}_r$ and $\ddot{X}_r$ are desired position, velocity and acceleration, respectively; F_r is desired grasping force; and F is actual grasping force.

When the end-effector works in free space $F = 0$, the control task is to ensure that $\Delta X = X - X_r = 0$ in order to achieve the output control of the end position. When the end-effector works in constraint space, we find $F \neq 0$. The goal of the control task is to realize the contact between the end-effector and the environment to meet the desired target impedance relationship, so as to accurately control the grasping force.

For different task environments, the key to grasping force control is to select appropriate parameter values for M_d , B_d and K_d in the impedance model. Figure 7 is the grasping force curve when the different impedance parameters change. It can be seen that the inertia and stiffness parameters easily cause oscillation in the system, while the damping parameter fluctuates but can significantly reduce the overshoot [22].

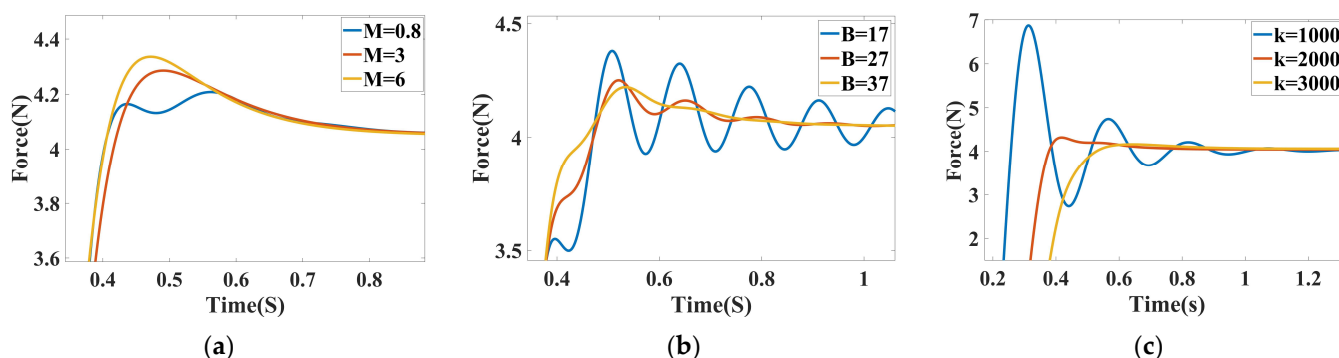


Figure 7. Grasping force curve when impedance control parameters change. (a) Inertia parameter, (b) damping parameter, and (c) stiffness parameter.

Traditional impedance control can track the grasping force stably under certain conditions, but if the impedance parameters cannot be altered in time, the apple may be irreparably damaged. The above analysis is used to create a reinforcement learning DDPG-optimized impedance control strategy to handle this challenge.

DDPG-Optimized Impedance Control

The problem of reinforcement learning can be described as a Markov decision process, which can be represented by tuples (S, A, P, R, γ) [29]. Here, S is the state space, expressed by current grasping force; A is the action space, expressed by impedance compensation parameters; P is the state transition probability; R is the corresponding reward given according to the state; and γ is the discount factor (ranging from 0 to 1). The reinforcement learning goal is to find an optimal strategy π that meets the task requirements to maximize future rewards, as shown in Equation (14) [31,32].

$$R_t = \sum_{i=0}^{\infty} \gamma^i r(s_t, a_t) \quad (14)$$

where T is the moment at the end round, $r(s_t, a_t)$ represents the reward when the state is s_t and the action is a_t .

The value function is used to evaluate the strategy quality, which is expressed by the Bellman equation shown in Equation (15):

$$Q^\pi(s, a) = r(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) \sum_{a' \in A} \pi(s', a') Q^\pi(s', a') \quad (15)$$

where $\pi(s', a')$ represents the strategy under s', a' , $p(s'|s, a)$ represents the state transition probability, and $Q^\pi(s, a)$ represents the state action pair s, a , which is the cumulative expected reward obtained by strategy π .

The optimization goal divides reinforcement learning into value-based and policy-based algorithms. Policy-based optimization algorithms use random and deterministic strategies. Deterministic strategies are consistent, but random strategies produce diverse actions under the same state S [33]. The same strategy gradient is also separated into random and deterministic strategy gradients. This study applies the deep deterministic policy gradient algorithm (DDPG) to stabilize, eliminate overshoot and speed up the impedance control system. Strategy networks in DDPG develop behavioral strategies, while value networks evaluate actions [34]. Figure 8 shows the algorithm structure, where the DDPG algorithm employs a dual-network structure for both the policy-based network and value network.

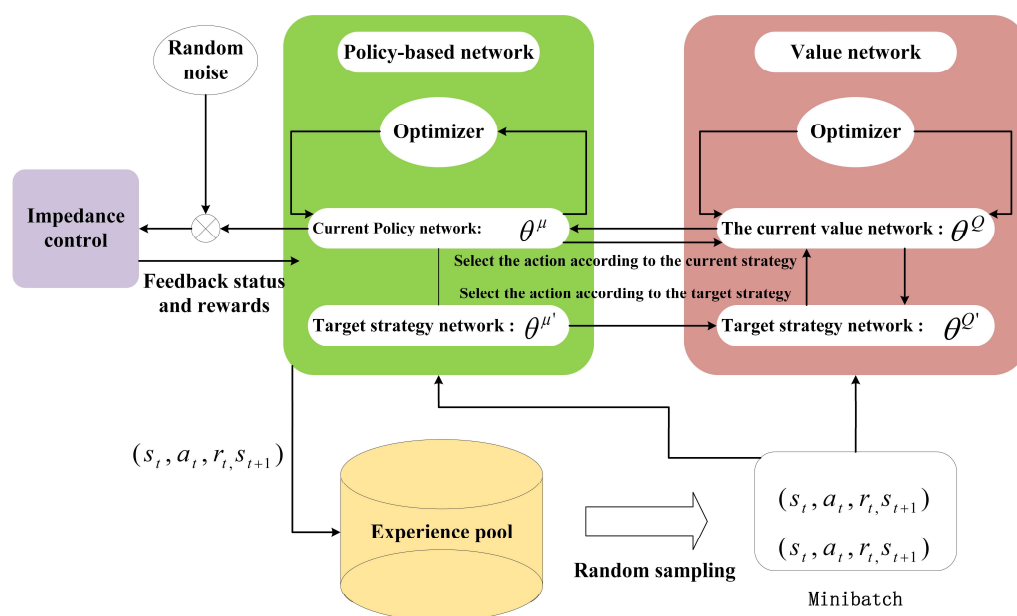


Figure 8. DDPG algorithm structure diagram.

In Figure 8, the strategy network simultaneously outputs the action value to the value network. The value network chooses the best action value and adjusts network settings to minimize loss. Equations (16) and (17) show the value network loss function and target action value.

$$J(\theta^Q) = \frac{1}{n} \sum_{k=1}^n \omega_k (y_i - (s_t, a_t / \theta^Q))^2 \quad (16)$$

$$y_i = r_{t+1} + \gamma Q(s_t, a_t / \theta^{Q'}) \quad (17)$$

where θ^Q is the value network parameter, n is the samples number, ω_k is the sample weight, $Q(s_t, a_t / \theta^Q)$ is the sample reward at time t , and $\theta^{Q'}$ is the updated value network

parameter, which is shown in Equations (18) and (19) with the updated policy network parameter θ^μ .

$$\theta^{Q'} = \tau\theta^Q + (1 - \tau)\theta^{Q'} \quad (18)$$

$$\theta^{\mu'} = \tau\theta^\mu + (1 - \tau)\theta^{\mu'} \quad (19)$$

where τ is the learning rate.

Figure 9 shows the DDPG-optimized variable impedance control strategy structure diagram. The DDPG algorithm obtains sample data $(s_t, a_t, r_{t+1}, s_{t+1})$ from the observed impedance control model, where s_t is the difference between the actual grasping force and the desired grasping force in the impedance control model at time t , which is denoted by ΔF ; r_{t+1} is the system reward value to perform the action in the previous state; and s_{t+1} is the ΔF of the system after adding the output action value B_p and the initial damping parameter of impedance model, so as to realize the DDPG-optimized variable impedance control and softly grasp different sized apples. The DDPG execution action is shown in Equation (20), and the variable impedance control rate is shown in Equation (21).

$$a_t = \mu(s_t|\theta^\mu) + N_t \quad (20)$$

$$B_d(t) = B_r + B_p \quad (21)$$

where N_t is random noise, function $\mu(s_t|\theta^\mu)$ is the optimal behavior strategy, B_r is the initial damping parameter, and B_p is the action value output by DDPG algorithm to compensate the damping parameter. The improved variable impedance control model is as follows:

$$M_d\ddot{\Delta X} + (B_r + B_p)\dot{\Delta X} + K_d\Delta X = \Delta F \quad (22)$$

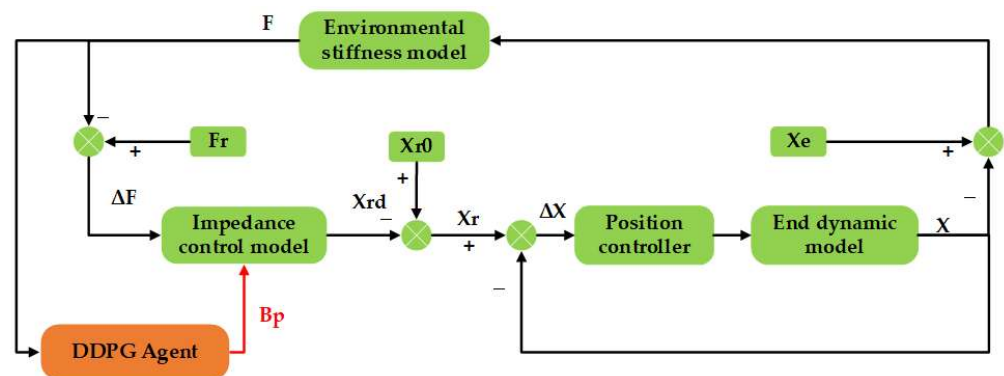


Figure 9. Block diagram of a variable impedance control system based on DDPG optimization.

2.3.3. Stability Analysis and Reward Function Design of Control System

Control system stability is especially important during apple picking. This work studies the compensation parameter value range after DDPG training using the Routh criterion to adjust the reward function, in order to ensure that the control system matches grasping needs as much as feasible while avoiding system instability. The Laplace transform of Equation (22) yields the system transfer function, as follows:

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{M_d s^2 + (B_r + B_p)s + K_d} \quad (23)$$

The characteristic equation of the system is as follows:

$$D(s) = M_d s^2 + (B_r + B_p)s + K_d \quad (24)$$

It is not difficult to obtain the following Routh table:

$$\begin{array}{c} S^2 \quad M_d \quad K_d \\ S^1 \quad B_r + B_p \quad 0 \\ S^0 \quad K_d \quad 0 \end{array} \quad (25)$$

Therefore, according to the Routh criterion, the controller parameters should be in the following range:

$$\begin{cases} M_d > 0 \\ B_p > -B_r \\ K_d > 0 \end{cases} \quad (26)$$

The reward function of DDPG designed is shown in Equation (27).

$$reward = 0.05 \times \left(\frac{1}{|e|} \times 0.05 \right) - 10 \times (|e| \geq 0.2) - 100 \times (|e| \geq 1) - 1000 \times (B_p \leq -B_r) \quad (27)$$

where e is the difference between the actual grasping force and the desired force; $0.05 \times (\frac{1}{|e|} \times 0.05)$ indicates that the reward value is given according to the grasping error size, in which the smaller the error, the greater the reward value. The formula $-10 \times (|e| \geq 0.2)$ indicates that the reward is -10 when the grasping force is greater than or equal to the desired value of 0.2 N; $-100 \times (|e| \geq 1)$ indicates that the reward is -100 when the grasping force is greater than or equal to the desired value of 1 N; and $-1000 \times (B_p \leq -B_r)$ indicates that the reward value is -1000 when the system is unstable.

3. Results and Discussion

According to the theoretical analysis above, this section solves the contact force optimization of the two-finger end-effector, inputs the result as the desired grasping force into the DDPG optimized variable impedance control, and verifies the effect of the control strategy through simulation and actual experiments.

3.1. Optimization Solution of Contact Force of Two-Finger End-Effector

In the experiment, an apple with a radius of 45 mm and a mass of 0.22 kg, as shown in Figure 10a, is used as the grasping sample, and the minimum grasping force will be solved using the above two finger contact force optimization model. The mass distribution of an apple is uniform, and its center of mass basically coincides with the center of the sphere of the apple. The schematic diagram of the end-effector grasping the apple is shown in Figure 10b. The contact positions between the end-effector and the apple are C_1 and C_2 , respectively. The two contact positions for the grasping models are the friction point contacts. The apple centroid is selected as the origin, the global coordinate system $O - xyz$ is established, and the local coordinate system $C_i - x_i y_i z_i$, $i = 1, 2$ is established at each contact point. The x_i axis and y_i axis of each local coordinate system are located on the contact surface and are perpendicular to each other. The z_i axis is perpendicular to the contact surface, and the direction is the same as the internal normal vector of the contact point. The specific coordinate information of two contact points C_1 and C_2 , based on the global coordinate system $O - xyz$, is shown in Table 1.

Table 1. Position coordinates of contact points.

Contact Point	x (mm)	y (mm)	z (mm)
C_1	44.98	−1.17	0
C_2	−44.99	0.78	0

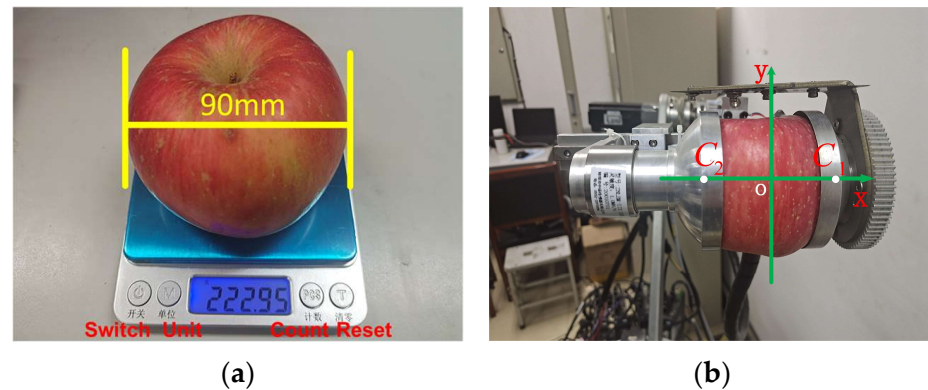


Figure 10. End-effector grasp apple schematic diagram. (a) Apple weight measurement and (b) apple grasping schematic.

The foregoing criteria establish the optimization model for apple picking grasping force, which is optimized using the gradient flow algorithm in Figure 3. Stability evaluation indexes φ_1 and φ_2 are set at 0.8 to avoid critical circumstances. According to the minimum stable grasping force optimization model, the gradient flow algorithm is employed for iterative computation, yielding a set of optimal grasping forces as follows:

$$f_c = [-0.9539, -0.1536, 3.8947, -1.2488, -0.1196, 3.8959]^T$$

In the simulation process, with the increase in the number of iterations, each contact point grasping force component also changes, and the change trend is shown in Figure 11, which shows that optimization variables vary constantly and converge to the optimal solution during iteration. The normal component of the grasping force changes most, and an increasing normal component increased positive pressure on the apple. Positive pressure increases friction at the contact point between the end-effector and the apple, countering fruit gravity and achieving stable grasping.

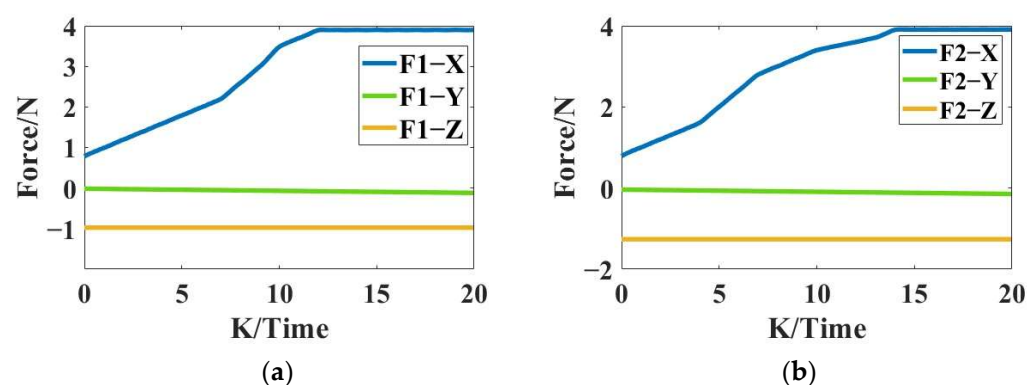


Figure 11. The convergence curve of grasping force under the action of external spiral force. (a) Contact force component at C_1 and (b) contact force component at C_2 .

With the grasping force component optimal solution, each contact point grasping force is calculated as follows:

$$f_{c1} = 4.0127, f_{c2} = 4.0929$$

The viable grasping force solution is obtained, and the solution of two fingers are almost the same, as the forces act on each other. To simplify the control, the grasping force is rounded as the desired force for the variable impedance control to evaluate controller performance.

3.2. DDPG-Optimized Impedance Control Simulation

The DDPG optimization variable impedance control strategy is tested using a simulation model as shown in Figure 12. First, the variable damping control system for the desired grasping force for apples is based on the grasping force optimization results. The impedance controller input is the difference between real and desired grasping force. To dynamically change the damping coefficient, the starting damping coefficient is added to the dynamic compensatory damping coefficient after DDPG training, and the desired position adjustment X_{rd} is output. In the new predicted position, the position controller quickly moves the end-effector finger to track the grasping force accurately. Among these, the position controller consists of a PID controller and apple picking robot end-effector transfer function [19]. Combined with the influence of the impedance parameters on the control system [35], the specific parameters of the impedance controller are set as in Table 2.

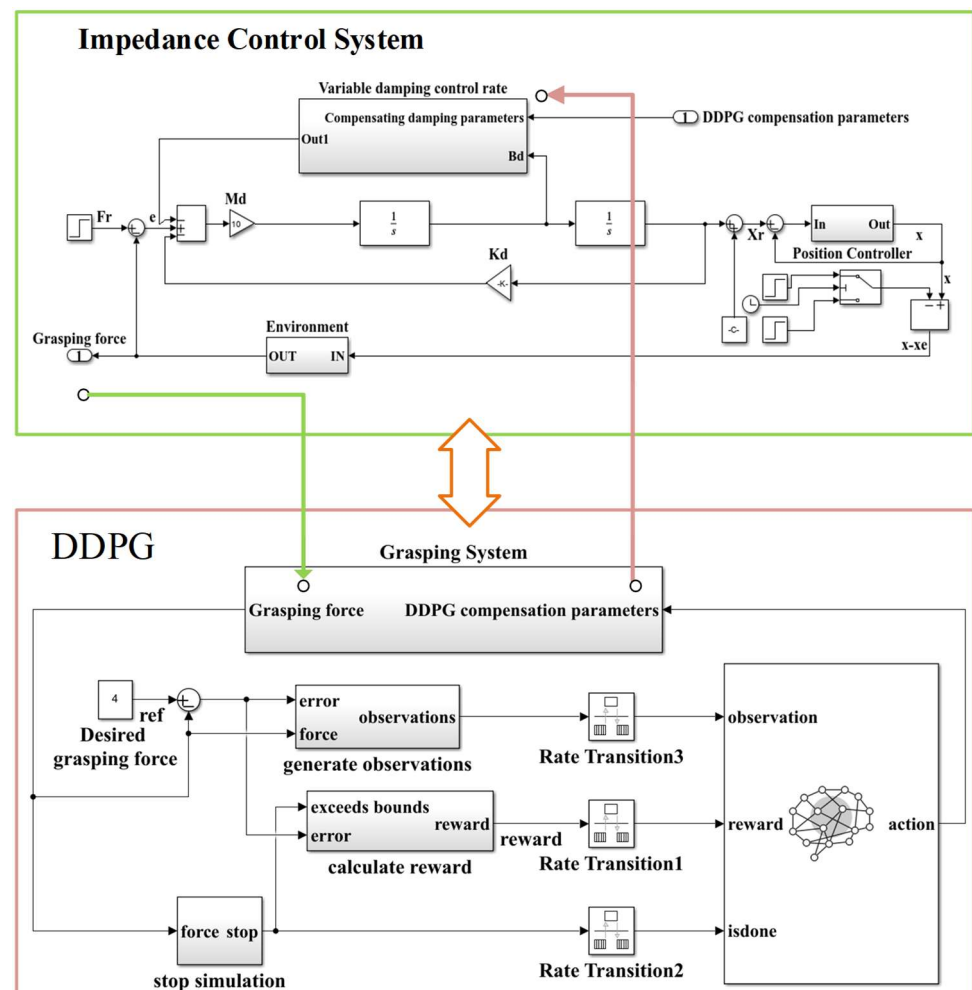


Figure 12. Simulation model of variable impedance control strategy based on DDPG optimization.

Table 2. Variable impedance controller parameters.

Parameters	M	B_0	K	Fr	Ke	Xe	Xr_0
Value	1	70	4000	4	7500	0.004	0.0045

3.2.1. Simulation Under Fixed Environment and Position

Firstly, the above-mentioned actual apple desired grasping force $F_r = 4$ N is trained. The agent sampling time is set to 0.1 s, the period is 3 s, and the training end condition is that

the average reward reaches 900. The DDPG reward function curve, which shows training speed and convergence, is shown in Figure 13. This reflects the learning condition when accurately tracking the 4 N desired grasping force. At each moment, the cumulative return after compensation is the reward value. The training result reaches the stop condition in 301 rounds, and the reward is 903. When the training stops, the grasping force tracking results of different controllers are shown in Figure 14, the variable impedance control peak value is 4.03 N, the overshoot is 0.75%, and the time to enter the steady state is 0.85 s. All control performances are greatly improved when compared with traditional impedance control.

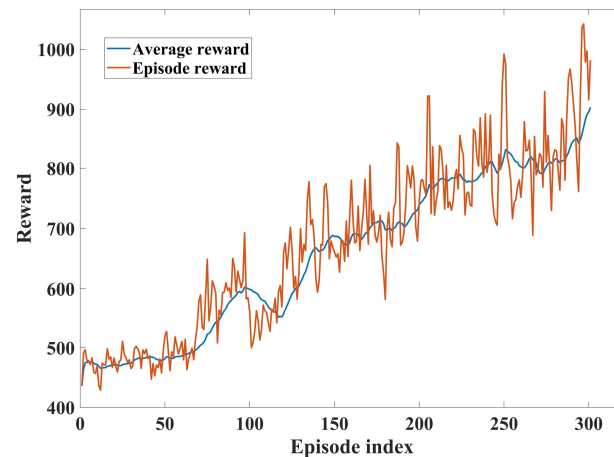


Figure 13. The DDPG reward curve.

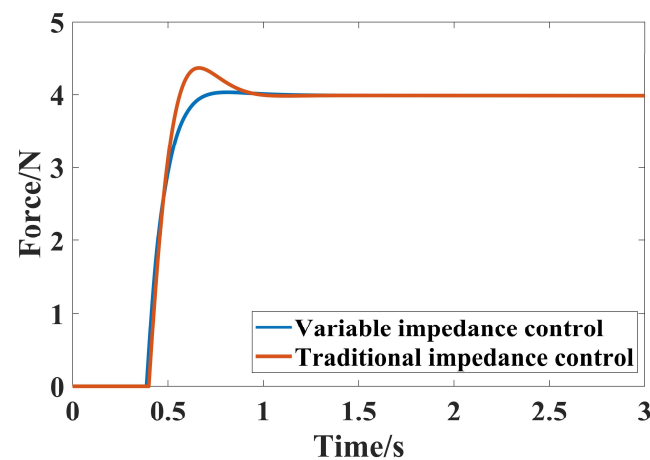


Figure 14. Fixed desired grasping force tracking experiment.

Simulation experimental results to verify the grasping effect of the proposed method for apples of different sizes, using desired grasping forces 5 N and 6 N, are shown in Figure 15, and the performance indicators under varied situations are in Table 3.

Table 3. Control performance index under variable desired grasping force.

Number	Variable Impedance Control			Traditional Impedance Control		
	Peak Value/N	Stable Time/s	Stable Value/N	Peak Value/N	Stable Time/s	Stable Value/N
5 N	5.02	0.92	5.006	5.38	1.09	5.007
6 N	6.03	0.80	6.002	6.37	1.05	6.005

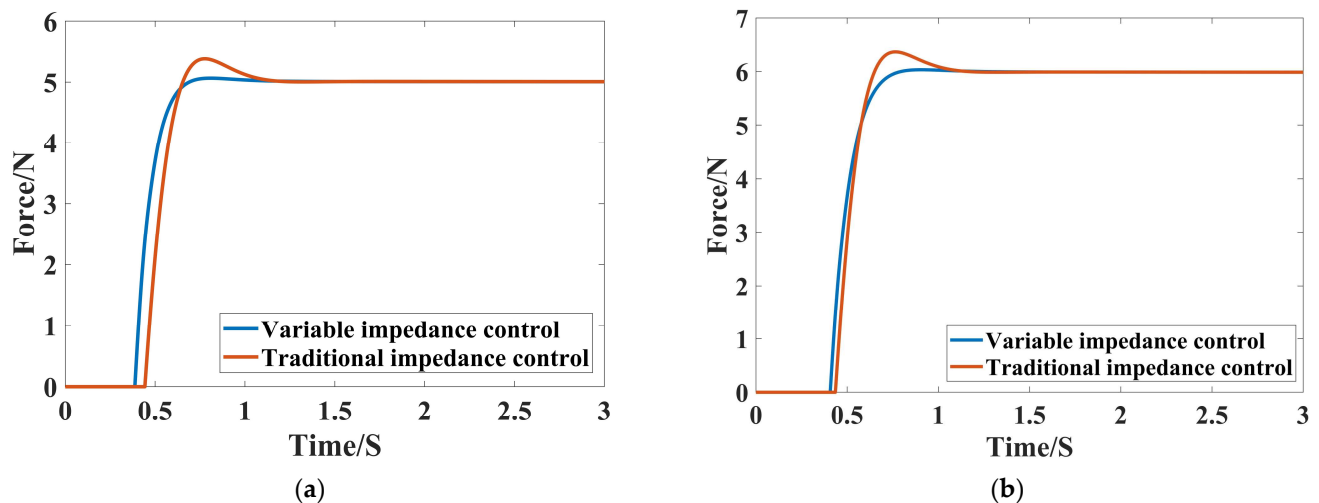


Figure 15. Desired grasping force tracking experiment under fixed parameters. (a) 5 N and (b) 6 N.

Figure 15a,b show that the DDPG-optimized variable impedance control overshoot is decreased and response speed is improved in various conditions. Table 3 shows that grasping force errors are within 0.1 N, and that the overshoot is less than 1%. The results indicate that the variable impedance control can carry out adaptive picking of apples under different desired forces. In order to further verify the stability of the system, the following simulation experiments are carried out on the control system under disturbed conditions.

3.2.2. Simulation Under Variable Environmental Stiffness Conditions

The following changing environmental stiffness studies verified the end-effector anti-interference ability. When the sampling duration is 2 s, the environmental stiffness increases by 1000. Figure 16 shows the grasping force tracking curve and Table 4 shows the performance index.

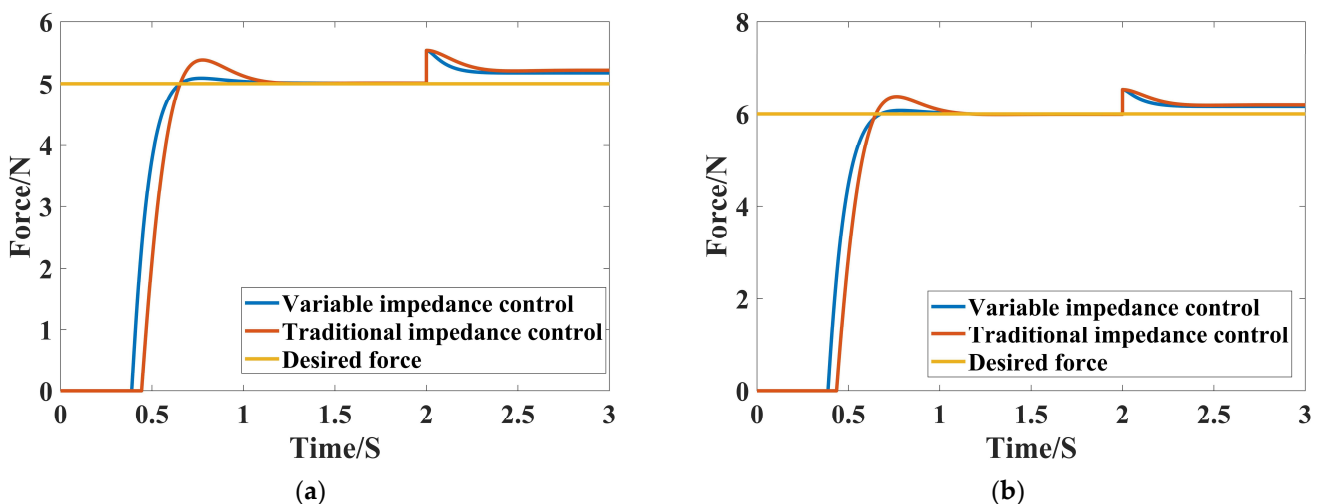


Figure 16. Desired grasping force tracking experiment under variable environmental stiffness: (a) 5 N; (b) 6 N.

Table 4. Impedance control performance index under different desired grasping force and variable environmental stiffness.

Number	Variable Impedance Control			Traditional Impedance Control		
	Mutation Force/N	Stable Time/s	Stable Value/N	Mutation Force/N	Stable Time/s	Stable Value/N
5 N	5.535	2.29	5.18	5.537	2.34	5.39
6 N	6.518	2.38	6.16	6.520	2.43	6.33

As illustrated in Figure 16a,b, the DDPG-optimized variable impedance control can better adjust to unexpected environmental stiffness changes than can the traditional impedance control. The figure shows that variable impedance control adjustment is smoother and faster. Table 4 shows that the adjustment time of variable impedance control is about 0.05 s faster than that of traditional impedance control, and that the stable error of the grasping force is within 0.2 N, which is about 0.2 N lower than that of a traditional impedance control, showing good resistance to the sudden change interference of environmental stiffness.

3.2.3. Simulation Under Variable Environmental Position Conditions

To verify the system anti-interference ability, the two set outcomes for the changeable environment location tracking of grasping force studies are as follows. The environmental position changes from 0.004 to 0.0042 at 2 s. The grasping force tracking curve is shown in Figure 17 and the performance metrics are shown in Table 5.

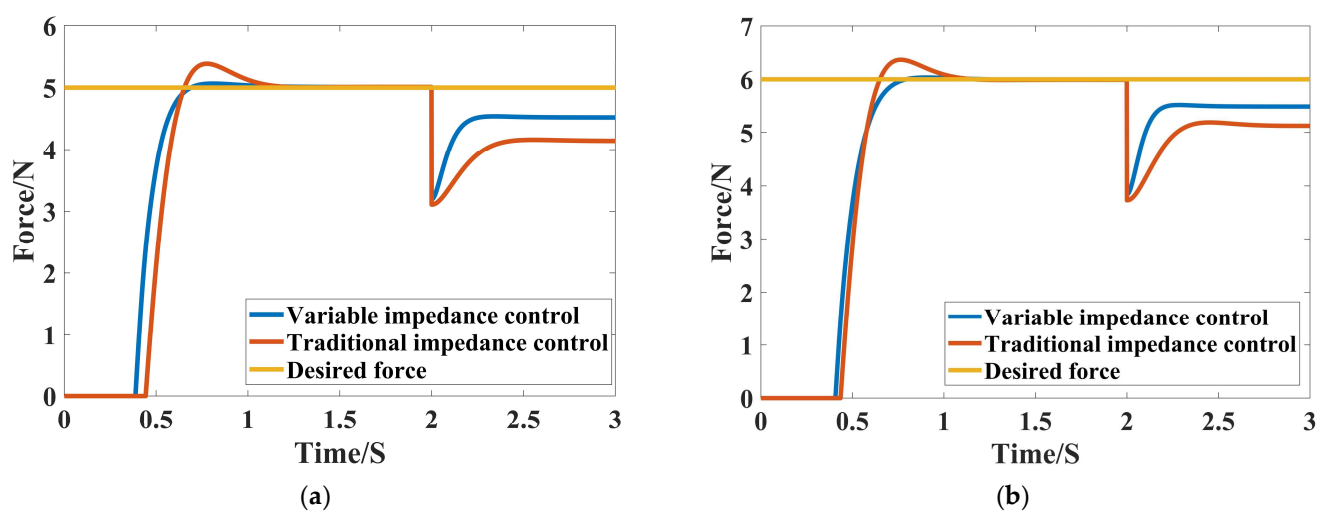


Figure 17. Desired grasping force tracking experiment under variable environment position. (a) 5 N and (b) 6 N.

Table 5. Impedance control performance index under different desired grasping force and variable environmental position.

Number	Variable Impedance Control			Traditional Impedance Control		
	Mutation Force/N	Stable Time/s	Stable Value/N	Mutation Force/N	Stable Time/s	Stable Value/N
5 N	3.201	2.26	4.63	3.181	2.76	4.14
6 N	3.947	2.43	5.69	3.930	2.60	5.15

As illustrated in Figure 17a,b, DDPG-optimized variable impedance control technique can better adjust to rapid ambient position changes than traditional impedance control. The graphic shows that variable impedance control tracks grasping force better, smoother, and faster. Table 5 shows that the adjustment time of variable impedance control is more than 0.1 s faster than that of the traditional impedance control, and the stable error of grasping force is kept within 0.4 N, which is about 0.5 N lower than that of a traditional impedance control, showing good resistance to environmental position mutation interference.

3.3. Actual Grasping Experiment

According to the simulation, the dynamic performance of the impedance control optimized by DDPG has been significantly improved when compared with the traditional

one, and it can resist environmental interference. The compensation parameters are applied to the actual apple grasping experiment for verification.

3.3.1. Experimental Equipment

The apple picking robot developed by our research, shown in Figure 18a, consists mainly of the mobile platform, the 4-degrees-of-freedom manipulator, the two-finger end-effector, the vision system with camera and the control system. The end-effector grasp control system is shown in Figure 18b. When the camera vision system finishes identifying and positioning the target apple, the control system gives instructions to control the manipulator to align with the target apple and bring the end-effector to a position where it can finish grasping it. Then, a stepper motor drives two hemispherical grippers to close symmetrically and secure the fruit, followed by a DC motor-actuated cutting blade severing the stem. The cutting blade, manipulator and grippers are reset in sequence to realize the picking for target apple.

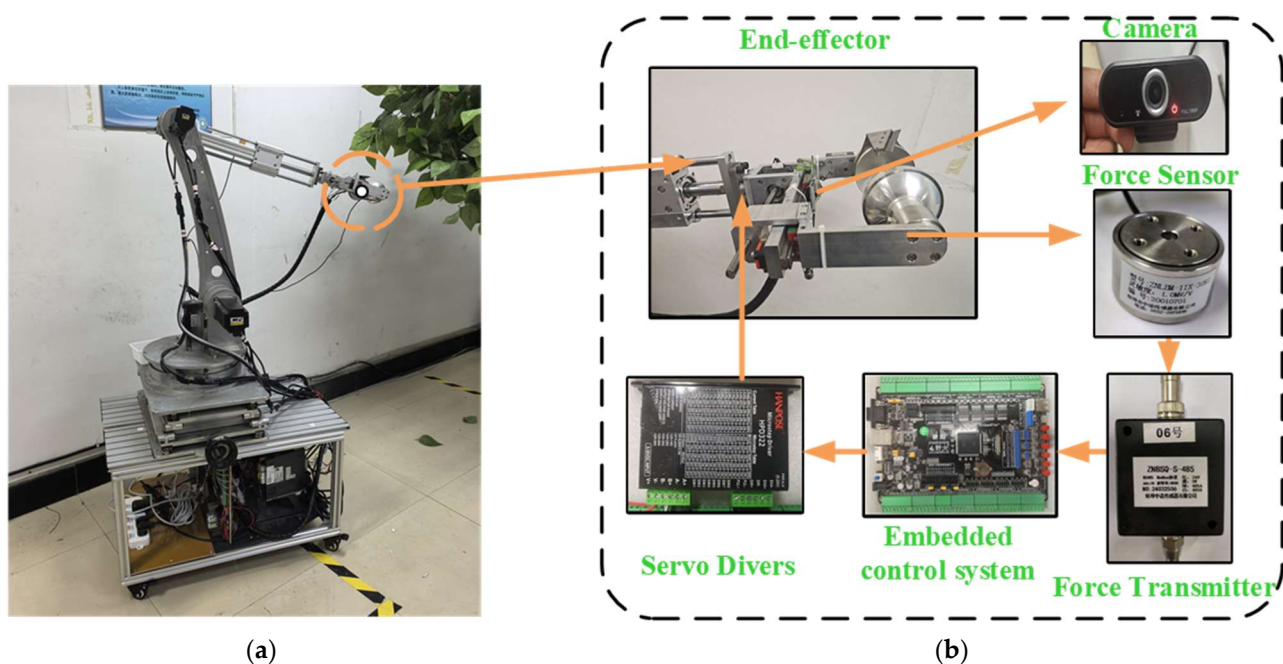


Figure 18. The experimental platform. (a) Apple picking robot and (b) end-effector grasp control system.

The end-effector grasp control system includes two-finger gripper with cutter, force sensor, force transmitter, embedded controller, servo driver and stepper motor. When the experimental platform works, the ZNLBM-IIX-30KG pressure sensor (manufactured by Bengbu Zhongnuo Sensor Co., Ltd., Bengbu, China) installed on the gripper sends grasping force signals to the STM32F407IGT6 (manufactured by STMicroelectronics, Switzerland/Italy/France) embedded controller through force transmitter, the embedded controller inbuilt impedance control algorithm moves the stepper motor based on the grasping force deviation from the intended force to grasp the apple. The pre-trained DDPG compensation amount is injected into the impedance control code in the development board to adjust damping parameters. Actual experiments will be used to verify that the proposed impedance control strategy works better than the old one. In the above contact force optimization, a 45 mm-radius apple with 0.22 kg is grasped.

3.3.2. Verification of Minimum Stable Grasping Force

In Section 3.1, the optimization of the minimum stable grasping force is verified by the calculation of grasping examples, and the calculated desired force is 4 N. In order to

further verify its effect in actual grasping, three groups of apple grasp experiments with respective desired forces of 3 N, 4 N and 5 N, and without using any control algorithm, are carried out. The effectiveness of the optimization of minimum stable grasping force is demonstrated by observing whether the apple slips in the gripper. The experimental results are shown in Figure 19.

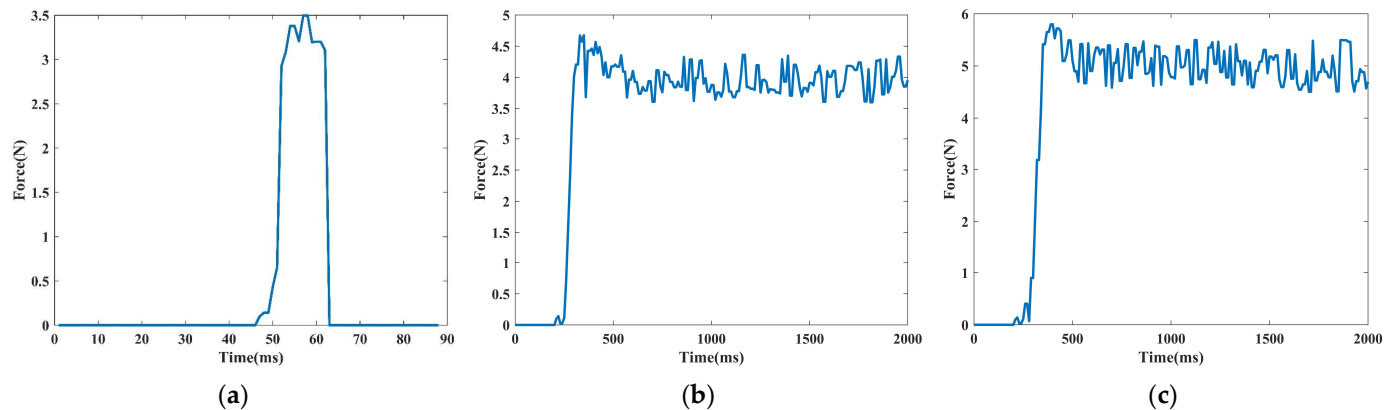


Figure 19. Grasping results under different desired stable grasping force. (a) 3 N, (b) 4 N, and (c) 5 N.

Figure 19a shows the apple grasping result under the condition of a 3 N stable grasping force. At the beginning, the grasping force of an apple increases continuously, and when it reaches the peak value, its grasping force no longer increases and shows a downward trend, which indicates that sliding occurs. The desired force needs to be increased. Figure 19b shows the apple grasping result under a 4 N desired stable grasping force, when the grasping force reaches its peak, the end effector adjusts the grasping force, tracks the desired force, and fluctuates around 4 N. There is no slip in the whole process, indicating that the 4 N desired force meets the actual grasping demand and can be used as the actual grasping force of the apple. The apple grasping result under a 5 N desired stable grasping force is shown in Figure 19c, the result proves that the apple can still be grasped stably under 5 N. The above results indicate that the optimization method of the minimum proposed stable grasping force can meet the desired force for the stable grasping of an apple. When the actual grasping force is greater than or equal to the calculated value, the apple will not slip in the gripper. When the force is less than the calculated value, the insufficient normal force will cause the apple to slip.

3.3.3. End-Effector Grasping Force Control Experiment Result

The actual grasping experiment is divided into three stages, as shown in Figure 20. In Figure 20a, before the end-effector is in contact with the apple, the contact force received by the force sensor is zero, and the controller drives the stepper motor to continuously approach the apple. Figure 20b is the second stage, where the end-effector simply touches the apple, the force sensor begins to feel the contact force, and the controller continues to drive the stepper motor to move forward. Under the adjustment of impedance control, the contact force overshoots, and the stepper motor backtracks to adjust the contact force, before continuing to advance and retreat in the process of tracking the contact force and then finally entering the third stage after the contact force has been stabilized, as shown in Figure 20c. Figure 20d shows the grasping force control curve under different controllers. The grasping force controlled by the DDPG-optimized variable impedance method is always smaller than the traditional one in the grasping process, which effectively reduces the overshoot. The peak value is 4.18 N, while the traditional impedance control peak is 4.49 N. The stabilization time is reduced by 0.6 s, and the fluctuation interval is reduced

by 0.24 N. Compared with those of Ji et al. [23] (0.31 s adjustment time reduction and 5.5% overshoot suppression), the results demonstrate superior performance, with a 0.6 s adjustment time reduction and 7.75% overshoot suppression.

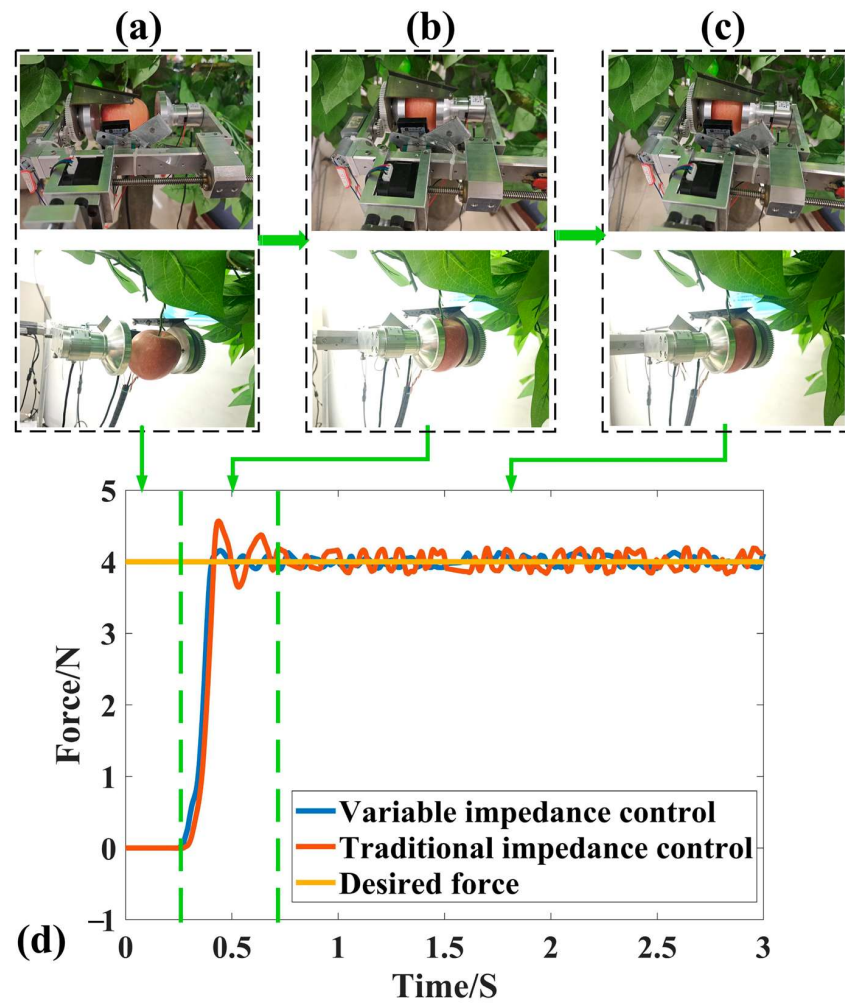


Figure 20. Experiment result of the end-effector grasping force control. (a) The end-effector with no contact with the apple, (b) the end-effector with contact with the apple, (c) the end-effector with stable contact with the apple, and (d) the grasping force control curve.

In the variable impedance control, the DDPG compensation value compensates the damping parameters of impedance control. The compensation value of DDPG is analyzed below, as shown in Figure 21.

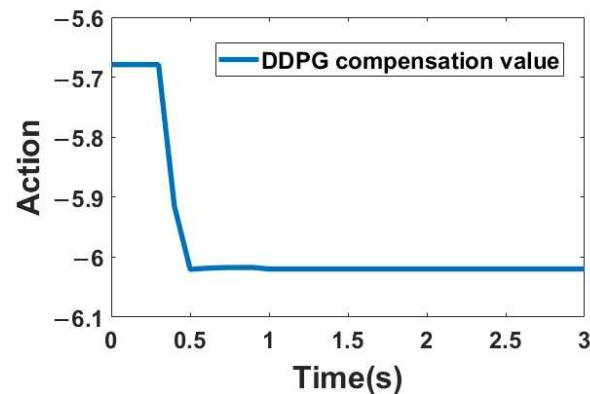


Figure 21. DDPG compensation value curve.

The compensation value output by DDPG is compensated according to the actual grasping force tracking curve trend, and the output is a smooth curve, which shows that the compensation of DDPG to impedance control is convergent. The damping compensation value of DDPG is maintained at -5.68 in the stage of $0-0.4$ s without grasping the apple, while the compensation value in the stage of $0.4-0.5$ s during the grasping process decreases, from -5.68 to -6.03 . The compensation value then changes, with the overshoot existing before the grasping stability. In the $1-3$ s stage, as the tracking in the grasping is stable near the desired grasping force of 4 N, the compensation value of DDPG is also stable, near -6.02 .

The grasping force tracking errors between two control strategies can be seen in Figure 22. The error of the variable impedance control optimized by DDPG is smaller and the stability speed is faster. The traditional impedance control maximum error before stabilization is 0.49 N, and the variable impedance control optimized by DDPG is 0.18 N. In the stable stage, the traditional impedance error control is within 0.21 N, and the variable impedance control error is within 0.15 N.

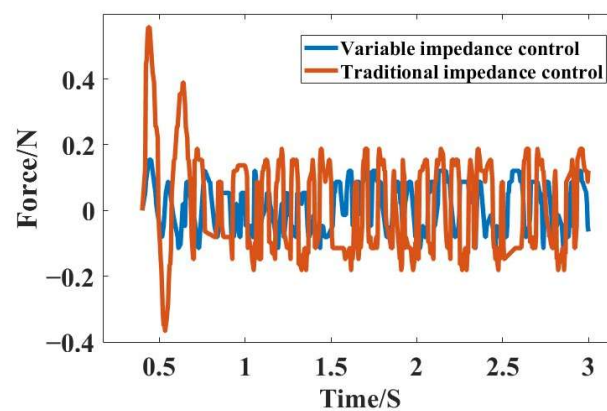


Figure 22. Grasping force tracking error.

The above results indicate that the DDPG-optimized variable impedance control meets the design requirements in practical applications, can greatly reduce the overshoot when compared with the traditional impedance control, and can track more stably in the stable stage.

4. Conclusions

This paper realizes apple-compliant grasping by a picking robot end-effector. A minimum stable grasping force optimization method for apples based on grasping stability evaluation indicators is studied, and a DDPG-optimized variable impedance control is proposed, one which avoids the problem of a grasping force that is too large or too small and which therefore leads to apple grasping failure. By solving the contact force optimization model, the desired minimum stable grasping forces for different sizes of apples are obtained. After that, the DDPG-optimized variable impedance control is realized by introducing a DDPG algorithm to compensate for the damping parameters in traditional impedance controls.

For a real apple with a radius of 45 mm and a weight of 0.22 kg, the optimized grasping force is rounded to 4 N. The simulation results under fixed environment and position parameters, variable environment parameters, and variable position parameters show that the dynamic performance parameters of the variable impedance control are superior to traditional impedance control. In the actual experiment, the desired grasping force tracking is improved. The peak value is reduced by 0.31 N, the stabilization time is reduced by 0.6 s, and the fluctuation interval is reduced by 0.24 N, showing a better grasping effect that can achieve the desired goal and compliant grasping.

Author Contributions: Conceptualization, W.J. and X.Y.; methodology, B.X. and X.Y.; software, X.Y. and H.Z.; validation, H.Z., C.R. and K.W.; formal analysis, X.Y.; investigation, W.J.; resources, W.J.; data curation, X.Y.; writing—original draft preparation, X.Y.; writing—review and editing, W.J. and X.Y.; visualization, B.X.; supervision, W.J.; project administration, W.J.; funding acquisition, W.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (No. 61973141), a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (No. PAPD) and the Open Project Program of the Key Laboratory for Agricultural Machinery Intelligent Control and Manufacturing of Fujian Education Institutions [AMICM202401].

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhou, H.; Wang, X.; Au, W.; Kang, H.; Chen, C. Intelligent robots for fruit harvesting: Recent developments and future challenges. *Precis. Agric.* **2022**, *23*, 1856–1907. [\[CrossRef\]](#)
2. FAO. *World Food and Agriculture: Statistical Yearbook 2024*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2024.
3. Zhang, K.; Lammers, K.; Chu, P.; Li, Z.; Lu, R. System design and control of an apple harvesting robot. *Mechatronics* **2021**, *79*, 102644. [\[CrossRef\]](#)
4. Ji, W.; Zhang, T.; Xu, B.; He, G. Apple recognition and picking sequence planning for harvesting robot in a complex environment. *J. Agric. Eng.* **2024**, *55*, 1549.
5. Xu, B.; Cui, X.; Ji, W.; Yuan, H.; Wang, J. Apple grading method design and implementation for automatic grader based on Improved YOLOv5. *Agriculture* **2023**, *13*, 124. [\[CrossRef\]](#)
6. Chen, K.; Li, T.; Yan, T.; Xie, F.; Feng, Q.; Zhu, Q.; Zhao, C. A soft gripper design for apple harvesting with force feedback and fruit slip detection. *Agriculture* **2022**, *12*, 1802. [\[CrossRef\]](#)
7. Liu, J.; Liang, J.; Zhao, S.; Jiang, Y.; Wang, J.; Jin, Y. Design of a virtual multi-interaction operation system for hand-eye coordination of grape harvesting robots. *Agronomy* **2023**, *13*, 829. [\[CrossRef\]](#)
8. Ji, W.; He, G.; Xu, B.; Zhang, H.; Yu, X. A new picking pattern of a flexible three-fingered end-effector for apple harvesting robot. *Agriculture* **2024**, *14*, 102. [\[CrossRef\]](#)
9. Wang, X.; Kang, H.; Zhou, H.; Au, W.; Wang, M.; Chen, C. Development and evaluation of a robust soft robotic gripper for apple harvesting. *Comput. Electron. Agric.* **2023**, *204*, 107552. [\[CrossRef\]](#)
10. Jia, P.; Wu, L.; Wang, G.; Geng, W.; Yun, F.; Zhang, N. Grasping torque optimization for a dexterous robotic hand using the linearization of constraints. *Math. Probl. Eng.* **2019**, *2019*, 5235109. [\[CrossRef\]](#)
11. Zhang, Y.; Mu, X. An inexact multiblock alternating direction method for grasping-force optimization of multi-fingered robotic hands. *J. Inequal. Appl.* **2023**, *2013*, 30. [\[CrossRef\]](#)
12. Zhang, H.; Ji, W.; Xu, B.; Yu, X. Optimizing contact force on an apple picking robot end-effector. *Agriculture* **2024**, *14*, 996. [\[CrossRef\]](#)
13. Mu, X.; Zhang, Y. Grasping force optimization for multi-fingered robotic hands using projection and contraction methods. *J. Optim. Theory Appl.* **2019**, *183*, 592–608. [\[CrossRef\]](#)
14. Li, Y.; Cong, M.; Liu, D.; Du, Y.; Xu, X. Stable grasp planning based on minimum force for dexterous hands. *Intell. Serv. Robot.* **2020**, *13*, 251–262. [\[CrossRef\]](#)
15. Wang, Q.; Bai, K.; Zhang, L.; Sun, Z.; Jia, T.; Hu, D.; Li, Q.; Zhang, J.; Knoll, A.; Jiang, H.; et al. Towards reliable and damage-less robotic fragile fruit grasping: An enveloping gripper with multimodal strategy inspired by Asian elephant trunk. *Comput. Electron. Agric.* **2025**, *234*, 110198. [\[CrossRef\]](#)
16. Lama, S.; Deemyad, T. Using a rotary spring-driven gripper to manipulate objects of diverse sizes and shapes. *Appl. Sci.* **2023**, *13*, 8444. [\[CrossRef\]](#)
17. Rabenoroosa, K.; Clévy, C.; Lutz, P. Hybrid force/position control applied to automated guiding tasks at the microscale. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010.
18. Zhang, Z.; Chen, Y.; Wu, Y.; Lin, L.; He, B.; Miao, Z.; Wang, Y. Gliding grasping analysis and hybrid force/position control for unmanned aerial manipulator system. *ISA Trans.* **2022**, *126*, 377–387. [\[CrossRef\]](#)
19. Kumar, N.; Rani, M. Neural network-based hybrid force/position control of constrained reconfigurable manipulators. *Neurocomputing* **2021**, *420*, 1–14. [\[CrossRef\]](#)

20. Neville, H. Impedance control: An approach to manipulation: Part I—Theory. *J. Dyn. Syst. Meas. Control* **1985**, *107*, 1–7.
21. Xie, F.; Chong, Z.; Liu, X.; Zhao, H.; Wang, J. Precise and smooth contact force control for a hybrid mobile robot used in polishing. *Robot. Comput.-Integr. Manuf.* **2023**, *83*, 102573. [[CrossRef](#)]
22. Ji, W.; Tang, C.; Xu, B.; He, G. Contact force modeling and variable damping impedance control of apple harvesting robot. *Comput. Electron. Agric.* **2022**, *198*, 107026. [[CrossRef](#)]
23. Li, X.; Qian, Y.; Li, R.; Niu, X.; Qiao, H. Robust form-closure grasp planning for 4-pin gripper using learning-based attractive region in environment. *Neurocomputing* **2020**, *384*, 268–281. [[CrossRef](#)]
24. Zhang, X.; Yin, L.; Shang, S.; Chung, S. Mechanical properties and microstructure of Fuji apple peel and pulp. *Int. J. Food Prop.* **2022**, *25*, 1773–1791.
25. Bao, X.; Ren, M.; Ma, X.; Gao, S.; Bao, Y.; Li, S. Optimization of contact force for spherical fruit and vegetable picking dexterous hand based on minimum force. *J. Agric. Mach.* **2025**, *56*, 333–341.
26. Liu, Z.; Jiang, L.; Yang, B. Task-oriented real-time optimization method of dynamic force distribution for multi-fingered grasping. *Int. J. Humanoid Robot.* **2022**, *19*, 2250013. [[CrossRef](#)]
27. Lei, Q.; Wisse, M. Fast grasping of unknown objects using force balance optimization. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, 14–18 September 2014.
28. Deng, H.; Luo, H.; Wang, R.; Zhang, Y. Grasping force planning and control for tendon-driven anthropomorphic prosthetic hands. *J. Bionic Eng.* **2018**, *15*, 795–804. [[CrossRef](#)]
29. Wen, G.; Yang, T.; Zhou, J.; Fu, J.; Xu, L. Reinforcement learning and adaptive/approximate dynamic programming: A survey from theory to applications in multi-agent systems. *J. Control Decis.* **2023**, *38*, 1200–1230.
30. Chen, G.; Huang, Z.; Jiang, T.; Li, T.; You, H. Force distribution and compliance control strategy for stable grasping of multi-arm space robot. *J. Control Decis.* **2024**, *39*, 112–120.
31. Kitchat, K.; Lin, M.; Chen, H.; Sun, M.; Sakai, K.; Ku, W.; Surasak, T. A deep reinforcement learning system for the allocation of epidemic prevention materials based on DDPG. *Expert Syst. Appl.* **2024**, *242*, 122763. [[CrossRef](#)]
32. Zhu, M.; Wang, Y.; Pu, Z.; Hu, J.; Wang, X.; Ke, R. Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transp. Res. Part C Emerg. Technol.* **2020**, *117*, 102662. [[CrossRef](#)]
33. Sutton, R.; Barto, A. Reinforcement learning: An introduction. *IEEE Trans. Neural Netw.* **1998**, *9*, 1054. [[CrossRef](#)]
34. Wu, M.; Wang, X.; Jiang, Y.; Zhong, L.; Mo, F. Collaborative temperature control of deep deterministic policy gradient and fuzzy PID. *Control Theory Appl.* **2022**, *39*, 2358–2365.
35. Song, P.; Yu, Y.; Zhang, X. A Tutorial survey and comparison of impedance control on robotic manipulation. *Robotica* **2019**, *37*, 801–836. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.