

Article

Deep Vision Servo Hand-Eye Coordination Planning Study for Sorting Robots

Tao Ning ^{1,2,*}, Changcheng Wang ² and Yumeng Han ²

¹ Institute of Computer Science and Engineering, Dalian Minzu University, Dalian 116000, China

² Big Data Application Technology Key Laboratory of State Ethnic Affairs Commission, Dalian Minzu University, Dalian 116000, China; dalianjh@djtu.edu.cn (C.W.); daliannt@dlmu.edu.cn (Y.H.)

* Correspondence: ningtao@dlmu.edu.cn

Abstract: Within the context of large-scale symmetry, a study on deep vision servo hand-eye coordination planning for sorting robots was conducted according to the problems of low recognition-sorting accuracy and efficiency in existing sorting robots. In order to maintain the symmetry of the picking robot, a small telescopic sorting robot with RealSense depth vision servo embedded in the manipulator was developed. The workspace and posture of picking parcels were analyzed, and the coordinate transformation model of hand-eye coordination was established for the “Eye-in-hand” mode. The hand-eye coordinated sorting test shows that the average positioning accuracy of the end in the X, Y and Z directions is 3.49 mm, 2.76 mm and 3.32 mm respectively, and the average time is 19.19 s. Among them, the average time for the mechanical arm to pick up the package from the initial position is 12.02 s, the average time for intermediate identification and calculation is 3.79 s, and the average time for placing the package is 6.9 s. The time consumed by robot arm’s action accounts for 79.8% of the whole cycle. The robot structure and the hand-eye coordination strategy with RealSense depth vision servo embedded in the robot can meet picking operation requirements, and the design of a picking robot proposed in this paper can greatly improve the coordination symmetry of fruit target recognition, detection, and picking.



Citation: Ning, T.; Wang, C.; Han, Y. Deep Vision Servo Hand-Eye Coordination Planning Study for Sorting Robots. *Symmetry* **2022**, *14*, 152. <https://doi.org/10.3390/sym14010152>

Academic Editor: Jan Awrejcewicz

Received: 12 November 2021

Accepted: 2 December 2021

Published: 13 January 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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: sorting robot; deep vision servo; hand-eye coordination; action planning

1. Introduction

In order to maintain the symmetry of the picking robot, “hand-eye” coordination is the main concern of an express parcel sorting robot and the key to completing an autonomous picking operation. The complexity of parcel types, colors, postures, etc., as well as the randomness and variability of one stacking, poses a great challenge to the wisdom of “perception-action” of the robot hand-eye system. The “Eye-to-Hand” structure has been first widely researched in the field of sorting robots and has strongly promoted the progress of picking robot technology. However, external camera telemetry resulted in excessive redundant information in the image, a complicated search process for target parcels, a prominent problem of mutual occlusion between hand-eye coordination and parcels, large recognition reliability, and visual localization errors [1]. This “see first, move later” static fragmentation approach lacks the ability to adapt to unstructured and complex environments and dynamic changes in logistics; at the same time, it relies too much on accurate hand-eye coordinate system conversion, robot kinematic modeling, and inverse problem solving, resulting in the accumulation of robotic errors and affecting positioning accuracy [2]. In contrast, the “Eye-in-Hand” (seeing and moving) mode of hand-eye coordination enables continuous visual feedback and correction of target positioning at the end so as to achieve real-time visual servo control, thus gaining importance in the field of sorting robots. In Japan, the United States, Belgium, and other countries, logistics sorting robot systems for different colored cartons and bags have adopted the “Eye-in-Hand”

configuration with palm or wrist mounted cameras. For the “Eye-in-Hand” hand-eye coordination control problem, LUO et al. designed a picking robot based on the greedy algorithm using CCD camera in hand mode, and laboratory detection and picking success rate reached 97% and 79%, but the robotic pick and place parcel’s path was too long, and too much time in motion planning and action execution was consumed. Moreover, a single operation time required up to 106 s [3]. Wang et al. designed a visual servo positioning accuracy of the end at about 15 mm, which can only meet the picking and positioning requirements of medium and large parcels [4]. Meanwhile, the existence of parcel prongs in the path will affect the servo’s control performance or even cause damage to the end-effector, and the measurement noise has a greater impact on the effect of feedback control based on high gain. Alipour et al. developed a humanoid dual-arm carton parcel picking robot with Xtion and PrimeSense sensors installed in the head and wrist, respectively [5]. However, experimental accuracy was insufficient, and the autonomous hand-eye coordination control was not fully realized. The logistics picking robot studied by Kumar et al. [6] required establishing a three-dimensional point cloud model by scanning the parcels vertically and horizontally with RealSense on the hand; however, the experiment only achieved two-dimensional image pixel position approximation but not depth (distance) localization. Ning et al. [7] constructed a hand-eye system based on the RealSense RR300 sensor that was first positioned in front of the parcel to take a picture in order to determine the relative angle between the manipulator and the parcel, but the control study of hand-eye coordination has not been conducted. The disadvantages of former studies are shown in Table 1.

Table 1. Differences between the different previous methods and the proposed method.

Methods	Characteristic
Greedy algorithm [3]	path is too long, too much time spent
Visual servo robot [4]	only meet medium and large parcels
Humanoid dual-arm robot [5]	experimental accuracy is insufficient, and hand-eye coordination control is not fully realized
RealSense picking robot [6]	the experiment only achieved two-dimensional image pixel position
Hand-eye system based on RSRR300 [7]	the control study of hand-eye coordination has not been conducted
Proposed method	lifting picking robot and “far and near” hand-eye coordination control

In order to achieve efficient picking robot operation under replicated real-world conditions, “Eye-in-Hand” hand-eye coordination control based on RGB-D sensors has become an objective need and research consensus [8]. In this paper, a RealSense depth servo picking robot design and far and near scene motion planning were carried out and verified by experiments based on the RealSense depth sensor’s reliable identification of 160–1200 mm depth of field and 160 mm close range package, combined with the structural design of lifting picking robots, the new hand-eye configuration mode, and the combination of “far and near” hand-eye coordination control.

2. Picking Robot Structure

The structure of the RealSense depth servo small telescopic picking robot is shown in Figure 1, which is composed of an autonomous mobile chassis, a scissor lift mechanism, a three-turn robot arm, and a wrist RealSense SR300 depth sensor.

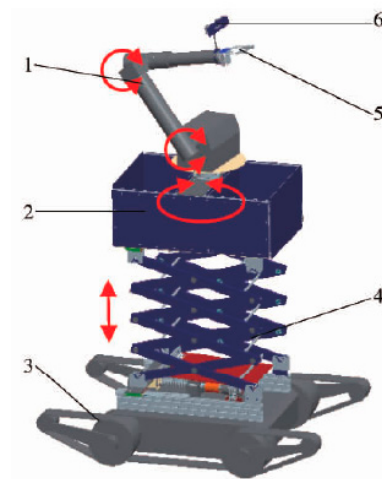


Figure 1. Structure diagram of small lifting picking robot. 1. 3 slewing manipulator; 2. Pull type pick-up box; 3. Autonomous mobile chassis; 4. Picking lifting mechanism; 5. Effector; 6. RealSense SR300 Depth sensor.

- (1) A multi-stage picking and lifting mechanism driven by a flush DC motor was designed, and a flexible auxiliary start-stop device was designed to improve starting performance and achieve a buffer for retraction, thus effectively solving the contradiction between the miniaturization of the chassis for narrow space passage and the larger telescoping needs for mobile operation.
- (2) On the basis of the two-joint robot arm, the horizontal servo slewing mechanism of the base of the picking and telescoping platform was designed, and the combination of the large and small arms, horizontal slewing, and space telescoping becomes a four-degree of freedom PRRR robot arm. Moreover, picking and parcel placement can be completed in a specific height area only by the three-swing action planning of the robot arm, which greatly reduces path length, motion planning complexity, and time consumption of the cycle operation and satisfies the standardized logistics picking requirements.
- (3) For the “pick-and-place” cycle operation with a few degrees of freedom and short path, the “Eye-in-Hand” system was formed by the Real Sense depth sensor installed on the wrist of the robot arm, using its depth of field of 160–1200 mm and a large field of view of 1280×720 depth detection characteristics and in-hand RealSense depth servo control, thus ensuring hand-eye coordination and accurate positioning of the combination of near and far views.

3. Hand-Eye Coordination of Posture and Coordinates

3.1. Picking Workspace and Posture Analysis

The picking robot achieves a wide range of picking operations by vertical lifting and lowering and 3 degrees of freedom rotation of the robotic arm, while only 3 degrees of freedom rotation of the robotic arm is required to achieve the placement of parcels [9–11]. The end work space and posture with respect to lifting and turning should meet the need of accurately placing parcels in the reserved area.

3.1.1. Workspace Analysis

The workspace of the manipulator was solved according to the Monte Carlo method [12]. The method involves drawing a point cloud diagram of the manipulator by using the functions of Rand and Plot in MATLAB robot toolbox [13].

On the basis of the lifting mechanism, the rotary robot arm can pick parcels within 278 mm of the vertical range. When the chassis is 160 mm away from packages, the maximum picking depth in the horizontal direction is 753 mm, thus meeting the bilateral picking of parcels in the logistics center within 1.5 m.

3.1.2. Posture Analysis

Combined with the robot's working space and parcel size, the posture of the picking robot was analyzed. According to the parcel location and end accessibility, robot picking presents three different postures: top picking, flat picking, and back picking. At the same time, according to the relative installation relationship between parcel placement position and the robot arm, the posture of the robot arm end when placing the parcel is 70–110 degrees between the forearm and the horizontal angle.

3.2. Coordinate Transformation of Hand-Eye Coordination

For the “Eye-in-Hand” mode, parcel positioning needs to obtain the position of the parcel relative to RealSense [14], and then it obtains the position coordinates of the parcel relative to the robot coordinate system according to the relative installation position of RealSense and the robot arm and kinematic equations.

Assuming that the coordinates of the parcel position obtained by RealSense SR300 is $S_0 = (x_0, y_0, z_0)$, the position S_1 of the target parcel relative to the robot coordinate system is $S_1 = S_0 M$ [14].

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & z_0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta_1 & -\sin \theta_1 & 0 \\ 0 & \sin \theta_1 & \cos \theta_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & L_1 \\ 0 & 0 & 1 & L_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta_2 & 0 & \sin \theta_2 & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_2 & 0 & \cos \theta_2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & L_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & L_4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta_3 & 0 & \sin \theta_3 & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_3 & 0 & \cos \theta_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & L_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & L_6 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

In Equation (1), M is the transformation matrix of wrapped camera coordinates and robot coordinates.

z_0 —Picking lift mechanism vertical linear displacement, mm.

1—Horizontal rotational joint angular displacement, rad.

2—Rotational angular displacement of the robot arm's large arm, rad.

3—Angular displacement of small arm rotation, rad.

L_1 – L_6 —The length of each bar of the robot arm and the relative mounting dimensions of RealSense and the robot arm.

L_1 – L_6 are 168 mm, 8 mm, 313 mm, 100 mm, 285 mm, and 125 mm [14].

4. Robotic Arm Motion Planning Based on RealSense Combined near and Far View

4.1. Hand-Eye Coordination Strategy for Near and Far View Based on RealSense Servo

Compared with CCD passive vision, active depth vision has less interference information and is less affected by light [15]. However, consumer-grade RGB-D cameras are affected by resolution and accuracy and cannot obtain a clear depth image of the parcel individual. Moreover, they may have large positioning errors occurring after only one vision identification and positioning [16,17]. Using RealSense SR300 as a vision sensor, the depth and coordinate information of package pairs can be obtained in the range of 160–1200 mm depth of field. Meanwhile, the “large and clear” image acquisition and reliable identification and positioning of a limited number of package boxes can be achieved in ultra-close ranges and small-view fields of 160 mm. Based on RealSense, the extensive remote detection of the top of the package pile, sub-area division, and positioning and the close range precise identification and positioning of the package are effectively combined to realize the coarse and fine guidance and precise positioning of the robot arm and can improve the real-time and stability of the “Eye-in-Hand” image servo control through segmental path planning and a combination of key points in the path.

The RealSense servo-based far-near view hand-eye coordination action flow is shown in Figure 2. $OABCD$ is the initial picking path; $DBCD$ is the cycle operation path.

- (1) The wrist of the robot arm moves from the initial position *O* to the perspective position *A*. RealSense SR300 obtains a depth map and the corresponding coordinate information of the larger parcel pile area from the perspective position *A*. According to the vertical and horizontal view range of RealSense ($55^\circ \times 71^\circ$) and the target distance, the field of view of different distances can be obtained. Then, according to the optimal close-range view range of RealSense SR300, the perspective depth area can be divided into several close-range subareas, and the location of the subareas is completed (Figure 3); the recognition symmetry of fruit tree information scanning area plays a guiding role in the selection of subsequent design strategies.
- (2) According to the coordinates of a sub-region, the robot arm rapidly approaches from the distant view position *A* to the subregions near view position *B* and completes the precise identification and positioning of parcels in the subregion's near view at the near view position.
- (3) According to the parcel identification positioning results in the close view position, the robotic arm sends the end-effector to parcel *C* and completes the picking of the parcel.
- (4) The end moves from parcel position *C* to placement position *D* of the target placement area, returns to the near view position *B* of this subregion after completing the picking and placing operation, picks the remaining parcels in this subregion in turn, and sends them to the parcel placement position *D* after completing picking at picking point *C*.
- (5) After picking the parcels in this subregion, the picking and placing operation continues for the next subregion according to the above process until parcel picking is completed, and the picking robot moves to the next vision position for the next cycle of the picking operation.

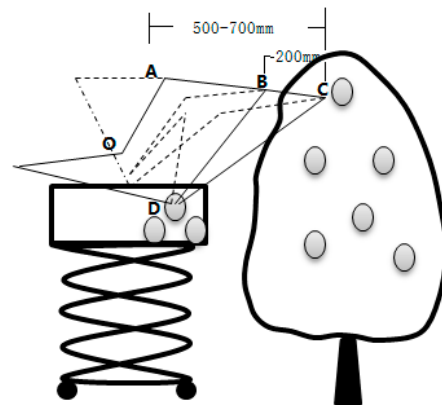


Figure 2. Action flow of far-to-close hand-eye motion.

In this strategy, RealSense returns different depth information of “coarse” and “fine” at the distant view position and the close view position, respectively, and guides the robot arm to approach the target from the far and near point and completes the selection. The method is based on a depth vision servo with limited key points to achieve high precision robot arm guidance and continuous operation with less computation and effective reliability and stability, thus providing assurance of fast and high success rate of robot sorting operation.

4.2. Critical Path Points for Hand-Eye Coordination

RealSense vision sensors are installed on the wrist of the robotic arm, which feeds back target depth information at the far and near points, respectively, by using the far and near point coordination strategy to guide robotic arm movement in order to target parcel picking.

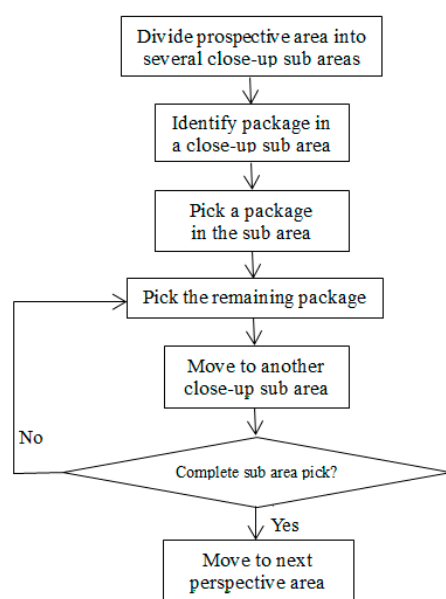


Figure 3. Flow chart of picking process.

According to the operating range of the robot arm, it is known that the best vision detection range of RealSense SR300 is 500–700 mm [17], and the best close view recognition distance is 200 mm. The settings of the key position are critical to the precise positioning of the robot arm, and the hand-eye tasks at each key position are as follows:

- (1) Far-field position: The robot arm is placed 500–700 mm away from the top of the parcel pile to scan the top of the parcel pile at a horizontal altitude and can smoothly obtain complete depth data of the top of the pile at a certain height far field for subregional division. The 700 mm detection area is 1008 mm × 729 mm, and the 500 mm detection area is 720 mm × 520 mm [17].
- (2) Close view position: After determining the subregion, the robot arm moves to the center of the targeted subregion, which is the best close view position at 200 mm from the subregion, and still uses the horizontal altitude as the ideal altitude. The vision sensor detection area is 288 mm × 208 mm [18].
- (3) Picking position: According to the coordinates of the targeted parcel, the robot arm reaches the picking position through altitude planning and picks the parcels up in a horizontal or upward or downward altitude.

The flowchart of the picking method is shown in Figure 3.

4.3. Segmented Planning of Picking Manipulator's Working Trajectory

4.3.1. Trajectory Planning Method

In order to achieve control of the robot arm, the poses of the end of the robot arm at the initial and target points need to be provided before planning. The robot's poses in Cartesian space are converted to joint angles, and the trajectory is described by fitting the joint angles with a function. The commonly used trajectory interpolation methods in joint space include a cubic polynomial and quintuple polynomial. The cubic polynomial function only constrains the position and velocity on the trajectory. In order to ensure smoothness of the robot arm's motion, it also needs to constrain acceleration at the starting and stopping points of the joint. Trajectory interpolation is used for trajectory interpolation. The quintuple polynomial interpolation function equation [19] is described as follows.

$$\theta(t) = a_0 + a_1t + a_2t^2 + a_3t^3 + a_4t^4 + a_5t^5 \quad (2)$$

By taking the first and second order derivatives of Equation (2), the time functions of angular velocity and acceleration of the joint of the manipulator can be obtained. The

linear equations can be determined by the constraints of each point chosen directly from the start and end points, thus solving for each coefficient of the equation.

After setting each key point and determining the linear trajectory, certain interpolation points need to be added to each segment of the trajectory, and the interpolation of these difference points is calculated by a 5-times polynomial to derive the rate of change of the positional coordinates of the different points with time, i.e., the relationship between the positional changes with time at the end of the process from the starting point to the end point is described by interpolation points.

4.3.2. Segmented Trajectory Planning

The trajectory of robotic arm operation can be divided into two trajectories: picking and placing. Considering the operation efficiency and control difficulty, the linear trajectory is adopted as the shortest path planning scheme in this paper for the two sections from the initial position to the prospective position and moving to the placing point after picking.

The process from the far view position to the near view range and then to the parcel picking position can be understood as the operation path of the robot arm approaching the parcel from the far view position. After completing parcel picking, the robot arm then moves the parcel from the picking position to the placement position and then returns to the vision position. Joint angle, angular velocity, and acceleration are generated from the picking position and then combined with the pose of the telepresence position for the inverse solution, and the joint angle, angular velocity, and acceleration are generated after trajectory planning in the joint space so that the robot arm arrives at the given positioning pose.

4.3.3. Track Articulation

From the initial position to the prospective position, the velocity at the end of the manipulator is 0 at both ends of the path. The process from the prospective point to the near scenic spot is a fast approaching process, and the velocity is not 0. However, the velocity at both ends of the trajectory from the picking position to the placement position is 0. According to the picking process and each key point, it is necessary to plan three straight line trajectories from the initial position to the prospective position, from the prospective position to the picking position and from the picking position to the placing position. Firstly, the coordinates of key points in Cartesian space need to be converted into the angles of revolute joints and the positions of straight joints (up and down) in joint space, as shown in Table 2.

Table 2. *D* Coordinates of key points and joint parameters.

Key Point	Coordinates Value/mm	Joint Parameters
Initial position P_0	(698, −57.3, 0)	(0 mm, 0°, 0°, 0°)
Prospective position P_1	(618.207, 339.391, 451.683)	(163 mm, 22.1°, 51.8°, −41.8°)
Picking position P_2	(648.107, 194.283, 468.602)	(163 mm, 31.8°, 13.9°, −25.6°)
Placing position P_3	(451.687, −342.238, 129.343)	(163 mm, −11.8°, 44.1°, −108.2°)

In order to ensure the accuracy of the linear trajectory of each fitting point, an interpolation point is selected every 15 mm, and the key parameters of the manipulator corresponding to each interpolation point are obtained by inverse solution. The linear interpolation trajectory of each segment of P_0 to P_1 , P_1 to P_2 , and P_2 to P_3 was obtained with MATLAB simulation.

5. Prototype and Test Verification

5.1. Prototype Design

The structure diagram of the small lifting picking robot includes four main picking parts: an autonomous mobile chassis, a scissor lift mechanism, a three-turn robot arm, and a wrist RealSense SR300 depth sensor.

The main technical parameters and performance indicators of this prototype are shown in Table 3.

Table 3. Prototype parameters.

Key Point	Coordinates Value/mm
Whole machine size/(mm × mm × mm)	380 × 570 × 1200
Whole machine quality/kg	≤45
Rated operating power/W	900
Working height/m	0.368~2.708
RealSense Detection distance/m	0.2~0.7
Navigate mode	GPS + SINS
Control mode	CAN Bus Control
Drive method	Servo DC motor drive

5.2. Verification Test

5.2.1. Test Material

As shown in Figure 4, two customized transparent scale plates of 1 m × 1 m and 2 mm interval size were built into a coordinate detection device with a right-angle coordinate system through aluminum profiles in order to facilitate the detection of the picking robot accuracy test, and a laser pointer was installed on the wrist of the robot arm (Section 3.1.1. Workspace analysis).

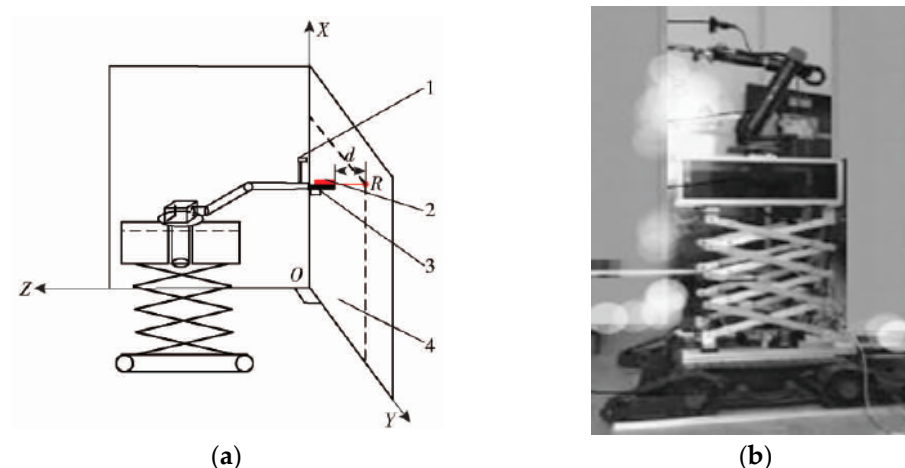


Figure 4. Schematic of measurement system. (a) Abstract diagram of measurement system, 1. RealSense SR300 Depth sensor; 2. Laser Pen; 3. Effector; 4. Coordinate detection device. (b) Prototype.

Adjust and determine the initial position of the manipulator. The depth direction of RealSense camera was used as the Z axis, and the initial position of the manipulator was measured in the height direction, i.e., X direction is 800 mm, and the center line of the waist rotating joint is the Y direction. After recording the reference point (800 mm, 0, 0) relative to the XOY plane on the grid scale plate, the X and Y coordinates of each measured end coordinate are read by the coordinate point relative to the reference point, and the Z direction can be measured by the vernier caliper, thus obtaining the actual coordinates of the picking robot end.

5.2.2. Test Method

- (1) The chassis of the picking robot is fixed, the mechanical arm is adjusted to the initial position to the initial position, and several express packages are fixed on the transparent scale plate in the Z direction. The center point of the package is taken as the target point, and the actual coordinates of the package are recorded according to the scale.
- (2) RealSense servo-based far and near field hand-eye coordination strategy and parcel picking robot arm operation trajectory segmentation planning are applied to perform

hand-eye tasks in different key positions and among key positions to complete single-cycle parcel picking.

- (3) The distance between the aluminum frame and the robot base coordinate system in the Z direction is adjusted, and the position of the parcels in different $Z_1, Z_2, \dots, Z_n$ planes is changed so that hand-eye coordination of parcel picking can be performed independently for different positions.
- (4) The coordinates of the projection of the end in the XOY plane are read from the mark point of the laser pointer on the scale plate. The distance between the end and the scale plate is measured by a vernier caliper, and the coordinate value of the end on the Z axis is obtained by combining the distance of the scale plate in the Z direction of the robot base coordinate system. At the same time, a DV camera was used to record the running process, and video editing software was used to capture the video clips of the manipulator's movement and to obtain its running time.

5.2.3. Analysis of the Test Results

The measured values of the coordinate points actually reached at the end of the picking robot and the corresponding elapsed time are shown in Table 4.

Table 4. Measurement results of end-effector position.

Test Number	Actual Coordinates of the Target Point/mm	Far and Near View Guide End Coordinates/mm	Far and Near View Coordinate Error/mm	Time Consuming/s
1	(860, 60, 700)	(858.2, 59.25, 698.23)	(−1.8, −0.75, −1.77)	20.86
2	(860, −60, 700)	(860.9, −62.27, 702)	(0.9, −2.27, −2)	19.98
3	(900, 200, 650)	(902.8, 202.18, 647.45)	(2.8, 2.18, −2.55)	20.31
4	(900, −200, 650)	(902, −195.30, 656.74)	(2, 4.7, 6.74)	19.74
5	(1100, 200, 600)	(1098.87, 202.09, 598)	(−1.13, 2.09, −2)	19.92
6	(1100, −200, 600)	(1097.36, −201.80, 601.05)	(−2.64, −1.8, 1.05)	19.30
7	(1100, 200, 550)	(1103.2, 201.11, 548.25)	(3.2, 1.11, −1.75)	19.02
8	(1100, −200, 550)	(1102.8, −198.06, 551.9)	(2.8, 1.94, 1.9)	18.03
9	(1200, −150, 500)	(1198.5, −151.08, 502.98)	(−1.5, −1.08, 2.98)	17.24
10	(1200, 150, 500)	(1201.95, 148.90, 498.11)	(1.95, −1.10, −1.89)	17.88
11	(1200, 200, 200)	(1203.25, 200.99, 198.06)	(3.25, 0.99, −1.94)	16.08
12	(1200, −200, 200)	(1196.33, −199.82, 198.34)	(−3.67, 0.18, −1.66)	15.99

As shown in Table 4, under the hyperopia and myopia coordination strategy involving RealSense information feedback, the average errors of the robot in X, Y and Z directions are 3.49 mm, 2.76 mm and 3.32 mm respectively. The time consumption of pick-up operation increases with the increase of package distance, and the average time consumption is 19.19 seconds. The average time consumption from the initial position of the manipulator to pick-up action is 12.02 s, the average time consumption of intermediate identification and calculation is 3.79 s, the average time consumption of package placement action is 6.9 s, and the time consumption of robot arm's action accounts for 79.8% of the whole process. The hand-eye coordination movement accuracy can meet the picking requirements, and achieves a good comprehensive effect of picking accuracy and efficiency. However, its accuracy and efficiency still have great room for improvement. There are some errors in each joint of the self-developed robot; In particular, there is a certain gap between the joints of the three rotating machinery. The maximum clearance of the manipulator under static test is (2.75, 2.1, 2.1, 2.65) mm, accounting for about 79% of the actual positioning error. Through optimization, the hand-eye coordination accuracy of the robot can be improved significantly. In the experiment, the path of the manipulator is calculated through interpolation with an interval of 15 mm. The actual arrival accuracy of the manipulator will improve with the reduction of the interval of interpolation points, and the improvement of the robot's computing ability will further reduce the recognition, positioning and interpolation calculation consuming.

6. Discussion

The results show that the average positioning accuracy of the end in the X, Y and Z directions are 3.49 mm, 2.76 mm and 3.32 mm respectively. The average time-consuming is 19.19 s, of which the average time-consuming of the robot arm from the initial position is 12.02 s, the average time-consuming of intermediate recognition and calculation is 3.79 s, the average time-consuming of package placement action is 6.9 s, and the action of the robot arm accounts for 79.8% of the whole course of treatment. Next, the construction of image feature vector of fruit tree picking object will be optimized to improve the accuracy of image recognition.

7. Conclusions

The design of the picking robot proposed in this paper can greatly improve the coordination symmetry of fruit target recognition, detection, and picking. In order to improve recognition sorting accuracy and efficiency of existing fruit tree picking robots, a small telescopic sorting robot with a RealSense depth visual servo embedded in the manipulator was designed in this paper. According to the parameters of RealSense and the manipulation of “eye-in-hand,” the hand-eye coordination strategy from far to near based on RealSense depth servo installed at the hand was proposed for the sorting robot, and the operation flow and motion planning based on depth vision and far-near coordination strategy were completed. The robot structure and in-hand RealSense depth servo hand-eye coordination measurement can meet picking operation requirements, while the improvement of joint accuracy and optimization of path interpolation calculation can further improve the accuracy and efficiency of hand-eye coordination.

Author Contributions: Conceptualization, T.N.; methodology, Y.H.; software, C.W.; writing-original draft preparation, T.N.; writing-review and editing, Y.H.; supervision, C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a project supported by Research Project on Economic and Social Development of Liaoning Province (2022lsjdybkt-014), Science and technology innovation fund program of Dalian (2021JJ13SN81), and Education Department Project of Liaoning Province (JDL2019022, L2020006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: None. This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Sassi, P.; Tripicchio, P.; Avizzano, C.A. A Smart Monitoring System for Automatic Welding Defect Detection. *IEEE Trans. Ind. Electron.* **2019**, *66*, 9450–9461. [\[CrossRef\]](#)
2. Natarajan, V.; Hung, T.Y.; Vaikundam, S.; Chia, L.T. Convolutional networks for voting-based anomaly classification in metal surface inspection. In Proceedings of the IEEE International Conference on Industrial Technology, Toronto, ON, Canada, 22–25 March 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 986–991.
3. Luo, Q.; Fang, X.; Liu, L.; Yang, C.; Sun, Y. Automated visual defect detection for flat steel surface: A survey. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 626–644. [\[CrossRef\]](#)
4. Wang, J.; Li, Q.; Gan, J.; Yu, H.; Yang, X. Surface defect detection via entity sparsity pursuit with intrinsic priors. *IEEE Trans. Ind. Inform.* **2020**, *16*, 141–150. [\[CrossRef\]](#)
5. Alipour, M.; Harris, D.K.; Miller, G.R. Robust pixel-level crack detection using deep fully convolutional neural networks. *J. Comput. Civ. Eng.* **2019**, *33*, 4–19. [\[CrossRef\]](#)
6. Kumar, S.S.; Abraham, D.M.; Jahanshahi, M.R.; Iseley, T.; Starr, J. Automated defect classification in sewer closed circuit television inspections using deep convolutional neural networks. *Autom. Constr.* **2018**, *91*, 273–283. [\[CrossRef\]](#)
7. Ning, T.; An, L.; Duan, X.D. Optimization of cold chain distribution path of fresh agriculture products under carbon tax mechanism: A case study in China. *J. Intell. Fuzzy Syst.* **2021**, *40*, 10549–10558. [\[CrossRef\]](#)

8. Silvestre-Blanes, J.; Albero Albero, T.; Miralles, I.; Pérez-Llorens, R.; Moreno, J. A public fabric database for defect detection methods and results. *Autex Res. J.* **2019**, *19*, 363–374. [[CrossRef](#)]
9. Tabernik, D.; Šela, S.; Skvarč, J.; Skočaj, D. Segmentation-based deep-learning approach for surface-defect detection. *J. Intell. Manuf.* **2020**, *31*, 759–776. [[CrossRef](#)]
10. Cheng, J.C.; Wang, M. Automated detection of sewer pipe defects in closed-circuit television images using deep learning technique. *Autom. Constr.* **2018**, *95*, 155–171. [[CrossRef](#)]
11. Phan, T.V.; Wang, G.; Liu, L.; Austin, R.H. Bootstrapped Motion of an Agent on an Adaptive Resource Landscape. *Symmetry* **2021**, *13*, 225. [[CrossRef](#)]
12. Ning, T.; Wang, X.P.; Hu, X.P. Disruption Management Decision Model for Vehicle Scheduling under Disruption Delay. *Syst. Eng.-Theory Pract.* **2019**, *39*, 1236–1245.
13. Li, Y.T.; Xie, Q.S.; Huang, H.S.; Yao, L.G.; Wei, Q. Surface defect detection based on fast regions with convolutional neural network. *Comput. Integr. Manuf. Syst.* **2019**, *25*, 1897–1907.
14. Noreen, I. Collision Free Smooth Path for Mobile Robots in Cluttered Environment Using an Economical Clamped Cubic B-Spline. *Symmetry* **2020**, *12*, 1567. [[CrossRef](#)]
15. Ning, T.; Shi, S.S.; Zhang, P.; Gou, T. Distribution management decision model for VRPDP under changes of customer distribution demand. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *2*, 2053–2063.
16. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiya, T.; Omata, H. Road damage detection and classification using deep neural networks with smartphone images. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1127–1141. [[CrossRef](#)]
17. Marino, S.; Beausery, P.; Smolarz, A. Weakly-supervised learning approach for potato defects segmentation. *Eng. Appl. Artif. Intell.* **2019**, *85*, 337–346. [[CrossRef](#)]
18. Ning, T.; Gou, T.; Liu, X.D. Simulation of Cold Chain Distribution Path of Fresh Agricultural Products under Carbon Tax Mechanism. *J. Syst. Simul.* **2021**, *10*, 1–11. Available online: <http://kns.cnki.net/kcms/detail/11.3092.V.20210223.1338.002.html> (accessed on 29 November 2021).
19. Ciszak, O.; Juskiewicz, J.; Suszyński, M. Programming of Industrial Robots Using the Recognition of Geometric Signs in Flexible Welding Process. *Symmetry* **2020**, *12*, 1429. [[CrossRef](#)]