Hindawi
Journal of Robotics
Volume 2022, Article ID 8565913, 12 pages
https://doi.org/10.1155/2022/8565913

Research Article

Recognition and Localization of Target Images for Robot Vision Navigation Control

Muji Chen

College of Information Engineering, Henan Vocational College of Agriculture, Zhengzhou, Henan 451450, China

Correspondence should be addressed to Muji Chen; 2004110217@hnca.edu.cn

Received 20 January 2022; Accepted 5 March 2022; Published 24 March 2022

Academic Editor: Shan Zhong

This paper focuses on a visual navigation control system for mobile robots, recognizing target images and intelligent algorithms for the navigation system's path tracking and localization techniques. This paper examines the recognition and localization of target images based on the visual navigation control of mobile robots. It proposes an efficient marking line method for recognizing and localization target images. Meanwhile, a fuzzy control method with smooth filtering and high efficiency is designed to improve the stability of robot operation, and the feasibility is verified in different scenarios. The corresponding image acquisition system is developed according to the characteristics of the experimental environment, and the acquired images are preprocessed to obtain corrected grayscale images. Then, target image recognition and linear fitting are performed to obtain target image positioning. The system calculates the angle and distance of the mobile robot, offsetting the target image in real time, adjusting the output signal, and controlling the mobile robot to realize path tracking. The comparison of sensor data and path tracking algorithm results during the experiment shows that the path tracking algorithm achieves good results with an angular deviation of ±1.5°. The application of RANSAC algorithm and improved Hough algorithm was analyzed in visual navigation control, and the two navigation line detection algorithms based on the image characteristics of the target image were improved in the optical detection area of the navigation line for the shortcomings of the two algorithms in visual navigation control, and the algorithms before and after the improvement were compared.

1. Introduction

The mobile robot is an essential branch of robotics. It is an intelligent robot control system capable of detecting and sensing the environment through various sensors and carrying out independent analysis, planning, and decision-making based on environmental information and its state [1]. The research field of mobile robotics involves many kinds of interdisciplinary theories and technologies, including computer vision, sensor information technology, communication technology, motion control theory, and mechanical engineering. The current hot wave of artificial intelligence also affects the research progress of mobile robotics [2]. With the rapid development of information technology, computer microelectronics, and network technology, mobile robotics has also developed rapidly, and more and more new robots with special functions have been introduced. The intelligence level of robots has been improving. In the 21st century, attention is on the robot's perception of the external environment and autonomy. The new direction of robotics is bound to develop toward practicality and intelligence. Mobile robots have been widely used in traditional industry and agriculture and will be further expanded to new sectors, services, defense and security, and medical services and will be commonly used in unsuitable and dangerous situations, such as deep sea and space. Therefore, the broad application prospect of mobile robots has made the research in this field receive widespread attention worldwide.

Autonomous mobility, which gives robots the ability to explore their environment more fully, dramatically increases the complexity of the tasks they can accomplish. State estimation during movement is a constant topic in mobile robotics research [3]. The primary consideration in

designing a reasonable and efficient state estimation method is the type of sensor the robot is equipped with and the characteristics of the data acquired by that type of sensor, i.e., the construction of a sensor observation model. The information that a mobile robot carries about itself and its environment is the source of all information in the subsequent navigation process and determines the form of information processing in the following global positioning and attitude tracking, map building, environment understanding, path planning, and motion control, and task execution. How to deal with the uncertainty contained in the perceptual information and how to design efficient cognitive methods to deal with the mental uncertainty based on the environmental information contained in the perceptual report are the significant challenges for building mobile navigation systems and must also be predicated on the construction of mobile robot observation models [4]. At the same time, the estimation of its positional attitude is the basis and prerequisite for performing other processes during the execution of tasks by mobile robots [5]. Therefore, the performance of the state estimation method will significantly affect the performance of the whole navigation system. Thus, the mobile robot observation model and the underlying state estimation model are introduced for two sensors, laser sensor, and RGB-D vision camera, respectively, to illustrate the robot state estimation process under different observation information and its uncertainty expression form further elaborates its problems. Specifically, the observation model is constructed for the laser sensor for the mobile robot equipped with the laser sensor [6]. Based on the observation model, various forms of observation similarity measures are given, and the characteristics of each form are analyzed. Based on this, a general model of the global localization process of the robot on the raster map is given. Finally, since global localization results are often uncertain and multihypothesis, the probability-based state tracking model is introduced in this chapter. For the visual observation model, the camera projection model is described, the projection model of spatial points to the camera plane is described, and the method to recover the spatial position of pixels is given. Based on this, a technique for global positional estimation based on the current observation and the feature matching results in the worldwide map is described. Due to the bias of feature observation, there is uncertainty in the global positional estimation results.

Robotics has been rapidly improved, thanks to the rapid development and maturity of microcomputer technology, sensors, and other related technologies. Intelligent robots have been popularized and applied in various fields such as civil, military, and scientific research. The research results of intelligent robots are more prominent in many developed countries and intelligent. Highly automated intelligent robots have been put into many fields such as aerospace, geological exploration, scientific exploration, rescue and disaster relief, such as China's lunar rover "Moon Rabbit". Some low-cost, clever robots are also coming into daily life and are used in many indoor environments in homes or offices, such as floor cleaning robots [7, 8]. We are now in a critical period of the modern manufacturing industry upgrading in the industrial field. More and more intelligent robots are needed to liberate labor, improve production efficiency, save energy consumption, etc. Visual inspection area image recognition is the basis of navigation line extraction. The quality of image segmentation affects the navigation line extraction and the size of the error in the measurement results of navigation parameters. In the navigation line region established by ultrasonic measurement, the navigation line visual detection region is set as the target operation domain for a series of image processing algorithms, and the detection region is dynamically tracked and set based on the detection results of adjacent frames; preprocessing image algorithms such as inverse color transformation and histogram equalization are specifically analyzed to enhance the different target images in the detection region differentially.

2. Related Works

Through the continuous development of electronic hardware technology and control disciplines, by the 1960s, some European countries already had various forms of mobile robots. With the rapid growth of processors in the 70s and 80s, mobile robots have made significant flexibility and stability. However, the main application scenarios are still the warehousing industry and logistics and transportation systems [9]. In the 1990s, the degree of intelligence and automation of mobile robots was further improved with the rapid development of computers, electronics, communications, and image processing technologies, and mobile robots adapted to various working environments were born, which have been widely used in the material assembly, home appliance production, chemical industry, food, and many other industries. The vision-based mobile robot navigation technology has been a new research boom in recent years and is one of the essential directions of mobile robot guidance technology research work. Research laboratories in universities in the countries that first researched visual navigation technology for mobile robots have achieved significant research results [10]. Many of the results have been applied to actual industrial production and even to the daily lives of the general public.

The Robot Vision Laboratory was the first to develop a vision-guided mobile robot based on map construction, which is accomplished through scene reconstruction using vision sensors to capture photos of the scene [11]. The laboratory at Purdue University has developed an active binocular stereo-based vision-guided mobile robot, Peter, which acquires 3D information of the operating environment and path obstacles, and combines 8 radar scanners, 24 ultrasonic sensors, 8 infrared distance sensors, and a passive infrared motion sensor to achieve flexible operation. The Intelligent Robotics Laboratory at Osaka University has conducted in-depth research on vision navigation and developed a mobile robot based on monocular vision navigation, which can detect the surrounding environment extensively by rotating its vision sensors and obtaining the positioning information, travel distance, and turning angle of the mobile robot by a rotary encoder and potentiometer

[12]. The vision navigation technology is widely used in advanced countries, such as autonomous lawnmowers, Mars landing vehicles, driverless vehicles, and Kiva robots for Amazon's unmanned warehouses.

Global localization of mobile robots is the basis of navigation. The localization process is based on observations to form feature representations and perform feature retrieval and area inference in a global map. Lasers and vision are two standard sensors used in the indoor navigation process [13]. Lasers can provide stable distance sensing information, which is advantageous in obstacle avoidance and motion planning tasks. However, it can only perceive flat information and is relatively homogeneous. Vision can perceive richer data, but its geometric perception range is lower, while laser has a more extended and more stable perception range. In the global localization process, there are multiple possible regions of positional distribution in the environment due to the existence of similar areas and the incompleteness of the perceived information. How to eliminate the spatial perceptual ambiguity to make an accurate estimation of the current poses is essentially a global optimization problem. The robot measures the degree of information consistency between the recent observation and the expected observation corresponding to the estimated position based on the constructed objective function. There is considerable international work on visual navigation motion control from image analysis to driving command generation [14]. In image segmentation, since the opening of the ImageNet challenge, there has been significant work based on the traditional feature point segmentation method to the currently popular deep learning methods: the use of deep learning to achieve semantic recognition of images and segmentation methods, with a very high number of references to achieve autonomous driving using images, or autonomous driving using reinforcement learning methods, the mainstream research in the field of autonomous driving. There is a large body of literature on autonomous driving, including the use of road segmentation and the use of environmental features for autonomous driving.

3. Model Design of the Target Image Recognition and Localization System Based on Robot Vision Navigation Control

3.1. Robot Vision Navigation Control System Construction. The visual navigation software is mainly used to process the images obtained from the camera in real time, then the navigation lines are extracted, navigation decisions are made after calculating the navigation parameters of the robot, and finally control signals are sent to the lower computer to control the weeding robot for automatic navigation. The visual navigation software in this paper is written on the Visual Studio 2015 development platform based on MFC with Open CV, where the algorithms are implemented in C++ and C languages. The functional framework of the visual navigation software is shown in Figure 1.

The visual navigation software comprises five parts: the information acquisition module, the image processing module, the navigation decision module, the information communication module, and the information storage module. The information acquisition module is the basis and preparation of the visual navigation. Its primary function is to obtain an accurate image with distortion correction after the calibration of the camera [15]. The IMU (Inertial Measurement Unit) detects the pitch angle of the camera in real time to correct the deviation of the camera pitch angle due to the vibration of the robot during the navigation process. The image processing module is the core part of visual navigation. The main functions of the image processing algorithm are image ROI construction, detection based on a deep learning model, detection frame clustering, image grayscale and smoothing filtering, corner point feature extraction in the detection frame, and navigation line fitting. The navigation decision module controls the motion of the robot based on the navigation information obtained after image processing. This study uses the fuzzy control method developed by our group to control the robot's movement. The primary process of the navigation decision module is to extract the navigation lines in the ROI after image processing, then to find out the dominant route, solve the position deviation and angle deviation of the robot, then input the navigation deviation parameters into the fuzzy control decision, and finally derive the control command of the robot. The function of the communication module is to realize the serial communication between the visual navigation upper computer software and the lower computer hardware control system. Through the configuration of the serial parameters and the development of the data transmission protocol, the control information is finally converted so that the digital signal output by the navigation software is converted to the level signal for controlling the robot motion [16]. The information saving module is responsible for saving the camera calibration parameters to avoid tedious, repetitive calibration work content, and saving the position deviation and angle deviation during the robot autonomous navigation process, which is used for the quantitative analysis of the accuracy of the image processing algorithm. It is divided into six parts: image processing display, camera parameter setting, camera calibration setting, serial port setting, production and saving of data, and robot motion control.

(1) The image processing display section displays the processing effects of the main stages of image processing in real time. The processing details of the image algorithm can be visually observed and analyzed for any problems that exist.

(2) The camera parameter setting section is used for camera ID selection, camera resolution setting, and image correction by calibrating internal camera reference.

(3) The camera calibration section is used to set the calibration board parameters, calibrate the camera, and save the calibration parameters. The calibration parameters saved locally can be read directly in the subsequent image correction operation, repeatedly avoiding tedious camera calibration work.

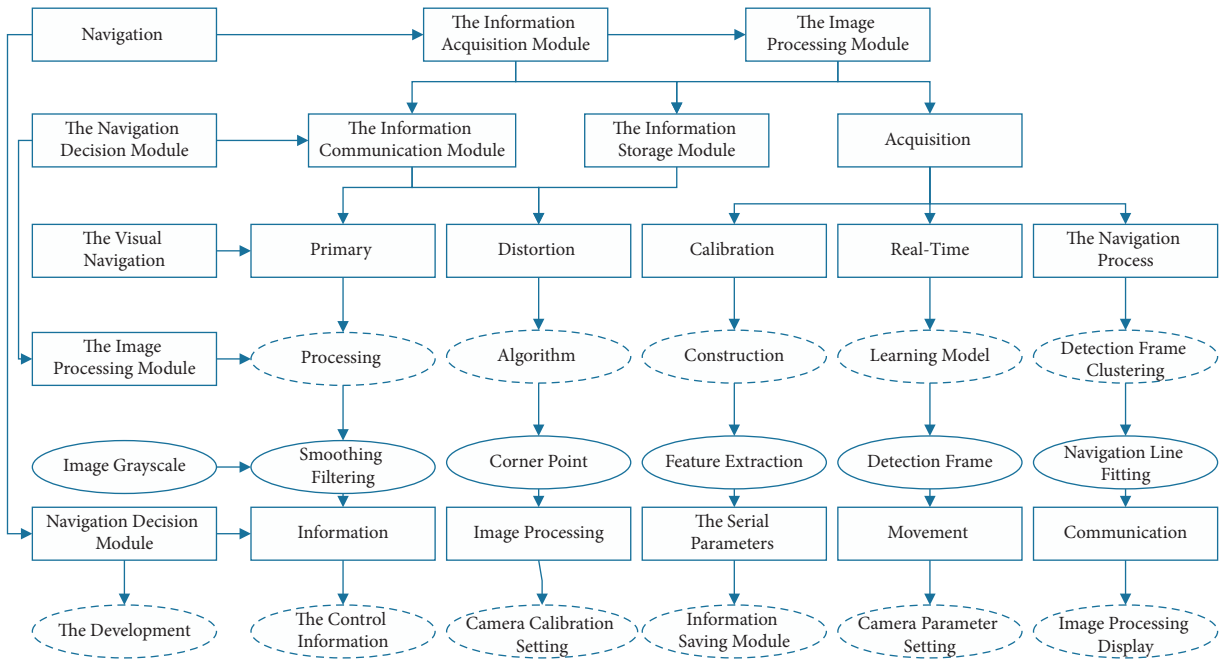

Figure 1: Functional framework of visual navigation software.

(4) The serial port selection section selects and configures the communication serial port between the upper and lower computers.

(5) The data display and saving section is used to display the pitch angle of the camera measured by the IMU in real time, to solve the position and angle deviation of the robot, to store the position and angle deviation of the weeding robot during automatic navigation in txt file format, and to select the size of the filter kernel.

(6) Robot motion control is divided into manual control mode and automatic navigation mode. The manual control mode is used to regulate the robot's position and adjust the part of the welding robot in the water field. The automatic navigation mode is for the robot to track forward along the seedling navigation line according to the control command of fuzzy decision. The robot vision navigation flow chart is shown in Figure 2.

Visual navigation of mobile robots is to collect road information through a camera to identify marking lines and guide the robot. Therefore, to accurately show the robot along the desired path, the body model must be built first to realize the conversion between the image space coordinate system corresponding to the camera and the world coordinate system centered on the robot's body. The mobile robot body structure is fixed, and its Kinect camera optical axis is parallel to the road surface, 34 cm high from the ground, and its maximum adequate vertical view is $a+b$ because there are shields above and below the camera, i.e., the bottom of the image taken by the camera is the road surface 66 cm in front of it. The Tourtellot mobile robot in this paper adopts a four-wheel structure, in which the left and right wheels are the driving wheels, and the front and rear wheels are the driven wheels. The wheeled mobile robot can accomplish a variety of motions, mainly by controlling the rotational speed of its left and right drive wheels, respectively. Therefore, to effectively manage the movement of the mobile robot, its kinematic model must be analyzed first. The position state of the mobile robot at two adjacent moments, with the x-axis forward as the robot's forward direction, where v_r and v_t are the velocities of the robot's left and right drive wheels, respectively, the angle the robot has turned at the adjacent moment t, r is the distance between the left and right drive wheels, and r is the radius of the circular arc motion at the adjoining moment where the forward velocity of the mobile robot is equal to the average rate of its left and right wheels, assuming that the steering angle is slight, the formula can be obtained as follows:

$$\begin{cases} v_t = \dfrac{v-1}{\sqrt{lw/2}}, \\ v_r = \dfrac{v+1}{\sqrt{2/lw}}. \end{cases} \tag{1}$$

Before the visual navigation line improvement RANSAC algorithm extraction, each pixel point in the optical detection area of the navigation line needs to be calculated in u_{IPM}, the directional gradient I_{upm} and v_{IPM}, directional gradient I_{vPM}, and the product $I^2_{u_{ipm}} I^2_{v_{iPM}}$ and sum $I_{u_{ipm}} I_{v_{ipm}}$ of the two gradient directions are calculated. Then, the Gaussian filtering is performed, and the template parameters are normalized; then, the corner point response value P of each image element is calculated, and the P value smaller than the threshold is set to zero. Finally, within the 3×3 neighborhood, the local nonmaximum values are

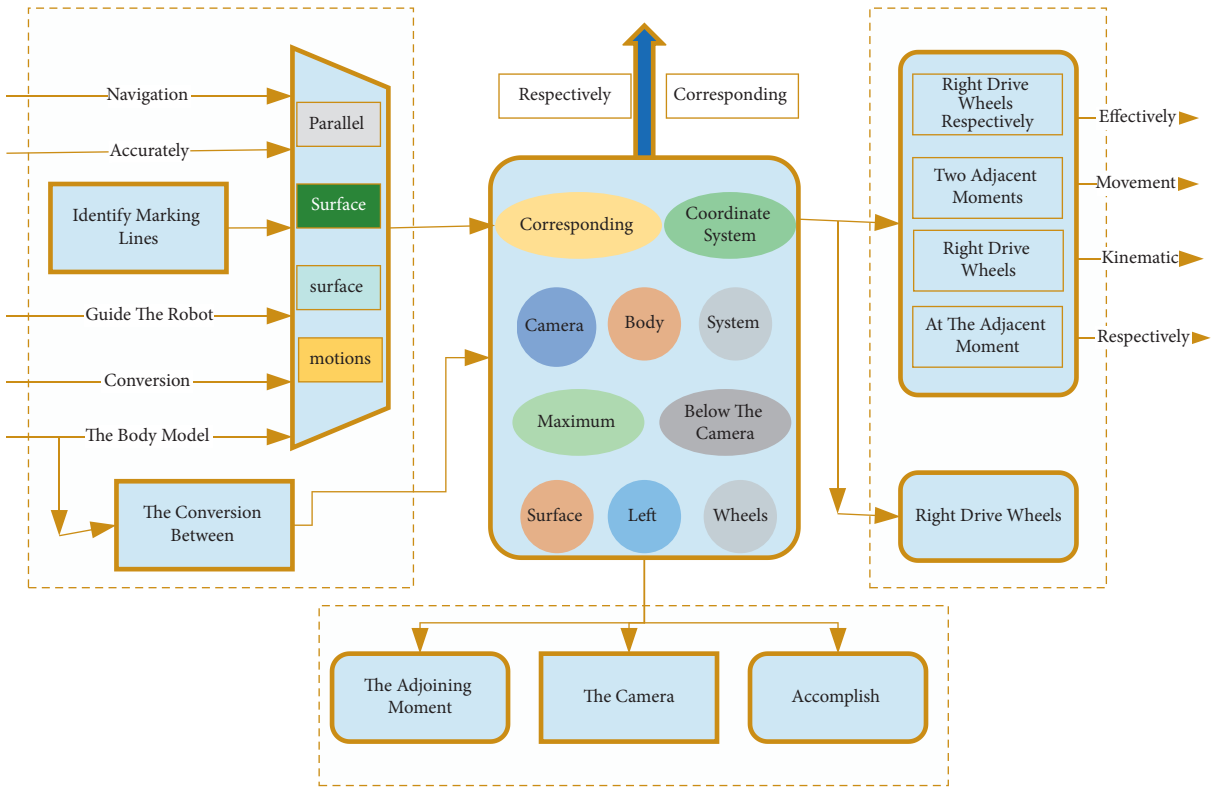

Figure 2: Flow chart of robot vision navigation.

suppressed, and the remaining local maximum values are output as corner points. The specific steps of the improved RANSAC algorithm for visual navigation line detection are as follows:

Step 1: A minimum of two data points is required for each random sampling. The number of samples in the data set You for each corner point needs to be guaranteed to avoid misfitting the harvesting navigation line due to too few corner points.

Step 2: Initial estimation of two randomly selected data points (u^1_{IPM}, v^1_{Ipm}) and (u^2_{IPM}, v^2_{Ipm}) models from the corner point data set U to obtain a linear model M.

$$v_{ip} = \frac{\sqrt{v^2_{ipm} + v^1_{ipm}}}{u^2 + u^1}. \tag{2}$$

Step 3: For the remaining data in the data set U, calculate the pixel distances d_m to the linear model M in turn, if the distance threshold A is d_m satisfied, but the point into the set U_s as an ingroup point together with the extracted sample points, and the other points as outgroup points.

Step 4: Count the number of ingroup points in the set U_s of ingroup points s. If s satisfies the threshold S_T of the number of ingroup points, refit the ingroup points U_s in the ingroup points using the least-squares method, and update the linear model M. If it does not satisfy, discard this linear model.

Step 5: Repeat the hypothesis and judge the mathematical model to find the ingroup points step N_R by step, compare the ingroup point set U_{SMax} with the most significant number of ingroup points, and output its corresponding linear model M to get the navigation path line L_M.

The essence of the Hough transform is to map the image to its parameter space, which requires the computation of all HM edge points and requires a large amount of memory space and operations. The improved Hough transform processes m_H $(m_H < M_H)$ only one edge point in the input image, and the selection of m_H this edge point is somewhat random. In addition, the enhanced Hough transform algorithm can obtain two endpoints of a straight line in the detection image and accurately locate the detected target straight line. The specific detection processes are as follows:

(1) In the image of the detection area, an edge point is randomly selected (u_{IPM}, v_{Ipm}) and mapped to the polar coordinate system to obtain a family of straight lines through the edge point. Suppose the edge point has been marked on a straight line. In that case, the random selection is continued among the remaining issues, and the polar coordinate equation where the family of straight lines through the edge point lies is obtained until all the edge points are randomly selected. The polar coordinates of the lines passing through the edge point are the opposite equation.

$$r = \frac{u_{ipm} \cdot \cos\theta}{\sqrt{v_{ipm} \cdot \sin\theta}}. \quad (3)$$

(2) The Hough transform of randomly selected edge points and calculation of the cumulative sum.

(3) In the Hough space, select the edge point that reaches the maximum value and continue to the next step when the issue is greater than the threshold, otherwise return to step1.

(4) The edge point where the maximum value reached by step3 is taken as the starting coordinate point. The displacement is carried out along the straight-line direction. The two endpoints of the line are detected to have coordinates of (u^b_{IPM}, v^b_{IPM}) and (u^e_{IPM}, v^e_{IPM}), respectively.

(5) When the length of the line obtained from the detection reaches a particular threshold value, the line is output as a result, and the detection is continued by returning to the initial step.

Compared with the standard Hough transform algorithm, the improved Hough transform has significantly improved memory consumption and computation. The improved probabilistic Hough algorithm can effectively avoid the interference of non-harvesting dividing lines and achieve accurate detection of visual navigation lines. In the complete detection process of optical navigation lines, the average processing time of a single frame for navigation line detection based on the original probabilistic Hough transform algorithm is 77.4 Ms. In comparison, the average processing time of a single frame based on the improved probabilistic Hough algorithm is 54.6 Ms. The enhanced algorithm also improves the processing speed compared with the original algorithm.

3.2. Target Image Recognition and Localization Model Design. The process of digital image acquisition and transmission can be disturbed by many factors, which can cause differences between the digital image and the real object scene and can affect the image processing of the vision system in the later stage. Preprocessing operations such as grayscale, image enhancement, and filtering must be performed on the original image captured by the camera at the beginning of the vision system. Then, the features of the preprocessed image are extracted, and then the parts are matched with those of the template image to identify the target strip of smoke. The surface of the identification target in this paper has robust texture features. The images captured by the camera are generally colorful. They contain much information, extending the time of processing images by the binocular vision system and considering the efficiency requirements of recognition and localization technology to improve the efficiency of the binocular vision system. The color images are to be converted into grayscale maps [17]. The R, G, and B components in the color image are converted into the same values. Since each pixel has a different value for the R, G, and B components, other colors can be displayed in the color image. The principle of target image visual localization is shown in Figure 3.

The amount of data in the image is also three times the number of pixels. When an image is grayed out, the information contained in the image becomes one-third of the original image [17]. The image pixels differ only in brightness and are all displayed in gray. The various colors in a color image are composed of the three base colors R, G, and B. In digital images, if the R, G, and B base colors are finely divided, the more colorful the image can be and the more information the image contains. The R, G, and B base colors are red grayscale, green grayscale, and blue grayscale. Weighted value method: according to different indicators to the original color image R, G, and B components multiplied by the corresponding weight to find the weighting, the expression is as follows:

$$\text{gray} = \sqrt{\sum W_R + R} \times \sqrt{\frac{W_G}{G} + \frac{W_B}{B}}. \quad (4)$$

Image enhancement is also critical to the overall binocular vision system and is necessary for processing images. Image enhancement highlights essential information in the image to meet the system's requirements and eliminate or weaken redundant details irrelevant to the system. Enhanced image processing makes the image more compatible with human visual habits and is intended for specific application purposes. After the image is enhanced, only the ability to distinguish information increases, while the background information is not added. The improved picture is more suitable for the application than the original image in specific scenarios. The standard methods for image enhancement are segmented linear transform enhancement and histogram equalization. The segmented linear transformation enhancement: Suppose the gray map function before the enhancement transformation is $f(r, c)$ and the grayscale range is $[O, M_f]$, after the segmented linear transformation enhancement, the gray map function is $g(r, c)$ and the grayscale range is $[O, M_g]$ and the formula.

$$g(r, c) = \sum_{g}^{f} \sqrt{\frac{d-c}{b-a}} + \frac{m_g - d}{m_f - b}. \quad (5)$$

A grayscale histogram is a statistical graph of the distribution of gray levels, representing the proportion of each gray level pixel in the total number of pixels in a digital image. The histogram can describe the general situation of a grayscale image, such as the degree of contrast between light and dark, the frequency of each gray level, the distribution of gray levels in the picture, etc. The gray histogram is a function of the gray level, with the gray value as the horizontal coordinate and the number of pixels as the vertical coordinate [18]. The gray histogram of an image has the following properties: the gray histogram is a statistical result of the number of occurrences of the gray value of all pixels in the picture, which does not reflect the specific position of the gray value pixels in the image, but only the number of occurrences of different gray values; a pair of images only corresponds to a couple of histograms. The histogram of a

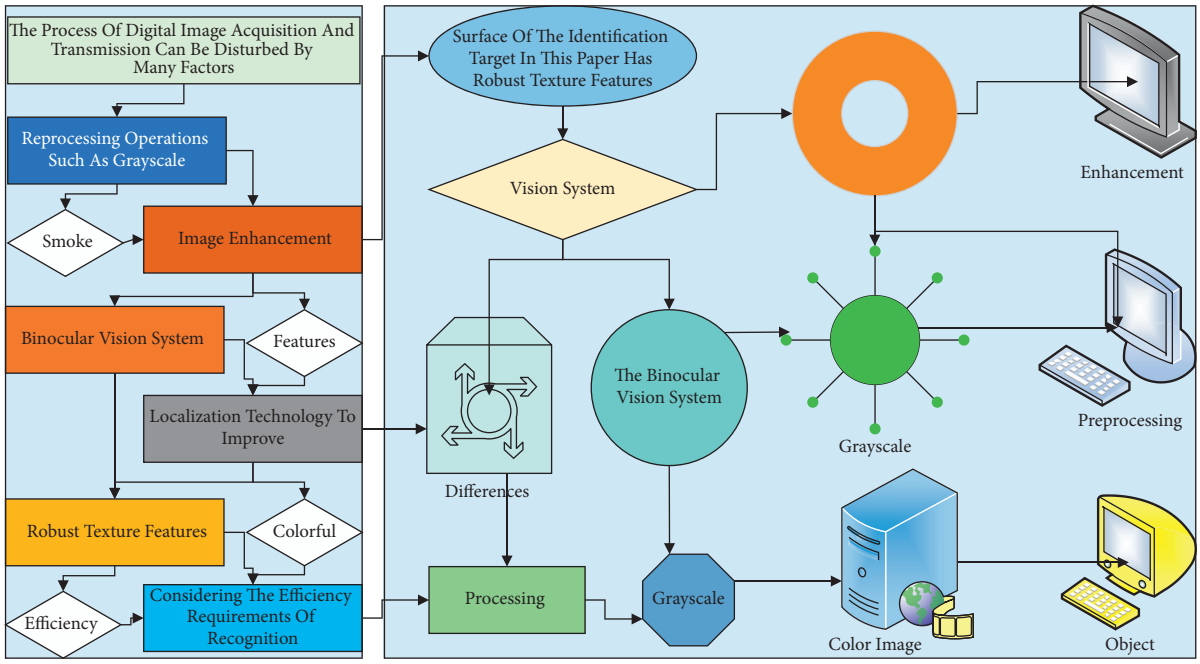

Figure 3: Target image visual localization principle.

team of images corresponds to only one histogram, but one histogram can be reversed to correspond to different ideas; the grayscale histogram counts the number of pixels with the same gray value in an image, so the grayscale histogram of a pair of images is equal to the sum of the histograms of all parts of the picture. The histogram with a $[O, L-1]$ range of gray values is a discrete function.

$$S = \int_{2}^{0} P_R + (w + dw). \tag{6}$$

The robot uses a trained feature point regression model and a topological structure regression model to estimate the robot's poses online. First, a pyramid is constructed for the acquired image, and SURF features and descriptors are extracted. Let $F = \{(p_i, v_i), i = 1\}$ be the location of the extracted feature points and their descriptors. Based on the depth map of the current observation, the coordinates corresponding to the feature points in the camera coordinate system can be obtained. Let $S = \{s \in R^3, i = 1\}$ denote the set of points corresponding to the features. Then, the topological positions corresponding to the current image are predicted using the topological position regression model to obtain multiple candidate topological positions. At the same time, the spatial coordinates in the world coordinate system are indicated for the features in F using the feature point regression model. Let $M = \{v_i = m_1 + m_u, i = 1\}$ denote the set of prediction results for feature points. M contains only the prediction results located at that topological node for each topological location. Due to the existence of erroneous predictions and the fact that each point has multiple projections, the global positional estimation problem of the camera can be expressed.

$$t = \arg\min \sum_{1} (\min \in |m - ts|). \tag{7}$$

4. Analysis of Results

4.1. Analysis of the Robot Vision Navigation Control System. In the process of autonomous navigation, the robot calculates the position deviation and angle deviation of the robot relative to the navigation line based on the seedling navigation line extracted by the vision system and continuously corrects the heading based on the variation during the forward motion [19]. Therefore, the positioning error of the weeding robot concerning the navigation line directly affects the navigation control process of the robot and must be measured and analyzed. The angular deviation error was measured by fixing the deviation of the robot's center on the centerline of the seedling row and rotating the robot. The angular deviation between the robot's centerline and the centerline of the seedling row changed from [15–15°] to [−15°]. When the centerline of the robot is parallel to the centerline of the seedling column, the angular deviation is defined as positive when the weeding robot is turned counterclockwise, i.e., the centerline of the weeding robot deviates to the left concerning the centerline of the seedling, the angular deviation is defined as unfavorable when the weeding robot is turned clockwise, i.e., the centerline of the weeding robot deviates to the right concerning the centerline of the seedling column. The measured and calculated values of angular deviation were recorded every 5° as a set of data. The experiment was repeated thrice for each group to improve the reliability of the experimental data, as shown in Figure 4. The mean error of the angular deviation was calculated to be 0.11°, and the standard deviation was 0.04°.

To verify the feasibility and reliability of the visual navigation and path tracking designed in this paper, the following scenes are set up: straight line, turning path, and obstacle occlusion. In the experiment, its forward speed is set as $V = 0.2$ m/s. The angular velocity w is in the

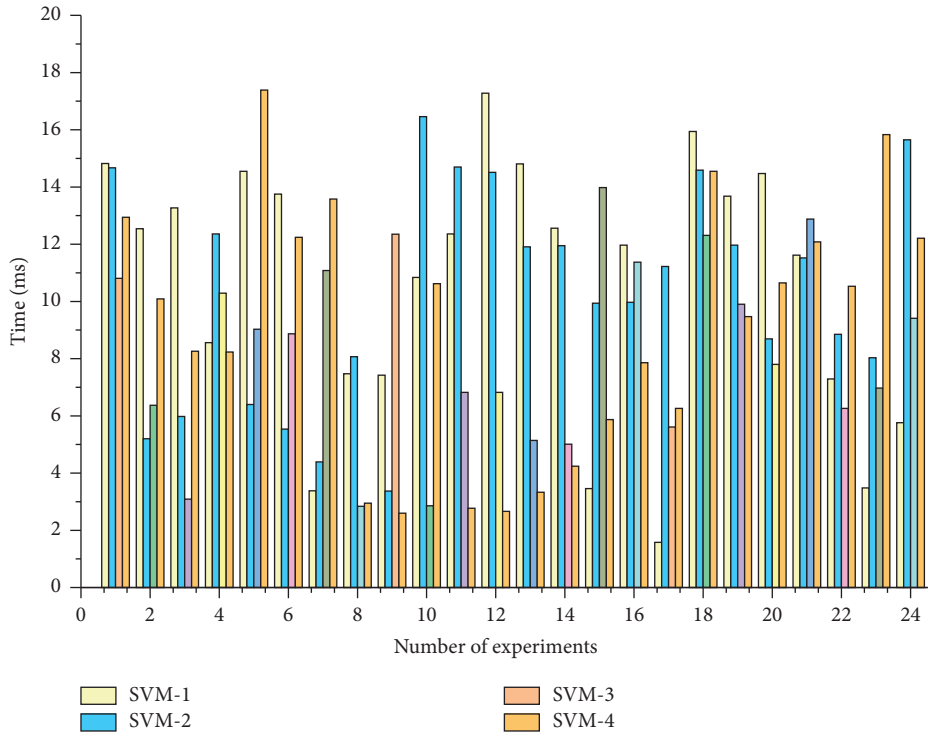

Figure 4: Experimental test comparison chart.

[−1.0 rad/s, 1.0 rad/s] range, in each frame after processing, the recognized center line is displayed with a straight blue line, and if a digital road sign is detected, the connection between the presighting point and the center of the digital road sign is displayed with a red straight line. The initial position of the Tourtellot robot is slight to the right of the path in the straight section at startup, and the identification line recognition results under several moments at intervals during operation. During the process of the robot, the angle of the centerline of the marking line obtained through image processing and recognition is plotted against the distance from the pre-sighting point to the centerline, as shown in Figure 5, where the dark blue curve is after the Kalman filtering. The light blue curve is the angle θ^{pre} without the Kalman filtering. It can be seen that θ the fluctuation is relatively stable during the robot's operation without significant abrupt changes. In contrast, the θ^{pre} instability of the unfiltered red curve in the figure is the distance deviation d after Kalman filtering, while the purple curve is the unfiltered distance deviation d^{pre}. The result of the distance deviation depends on the angle to a certain extent, so the fluctuation θ^{pre} leads to a more significant distance deviation from the pre-sighting point to the marker line also varies more.

The lane lines were laid in the lab, and the corresponding QR codes were set at each node location to place the mobile robot at any initial node location. First, we start the mobile robot management system and wireless communication server, set the relevant network parameters, establish the wireless LAN server, and wait for the lower computer to access. We start the embedded development board TX2, the

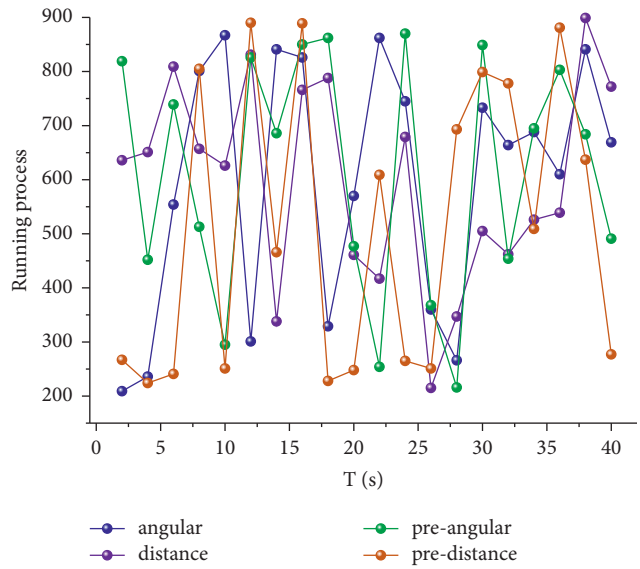

Figure 5: Running process identification lines θ and d curves.

industrial CCD camera, and the execution component Turtle Bot2, and enable the wireless network access function after checking and confirming the connection of each hardware of the lower computer, setting the network parameters, and accessing the LAN established by the wireless server; the upper computer reads the map file and reads the information to the vehicle controller according to the readings. The upper computer reads the map file, reads the data from the vehicle controller, initializes the coordinates of the node

where the mobile robot is located according to the read information, and displays its position in the map in real time; the management system will automatically number the tasks after receiving the task command, and then compare the current mobile robot location information with the target location information to get the shortest path information by path planning and transmit the path information to the lower computer through wireless communication. The path information will be transferred to the lower computer through wireless communication. After receiving the path information, the lower computer will execute the task and walk on the specified path through lane line tracking. The QR code recognition will be used for positioning correction and steering guidance, and the real-time data will be uploaded to the upper computer [20]. To verify the accuracy of the path tracking algorithm based on visual navigation, we randomly select a time point in the mobile robot operation test, extract the measured angle data of the gyroscope in the actuator at this time point, and compare it with the data obtained from the path tracking algorithm at the corresponding time. The deviation of the gyroscope angle data from the angle data obtained from image processing is calculated, and the relationship between the selected time nodes and the angle deviation is plotted as shown in Figure 6. From the results, the deviation value between the path tracking algorithm and the actual angle is ±1.5°, indicating that this paper's path tracking algorithm can achieve high accuracy.

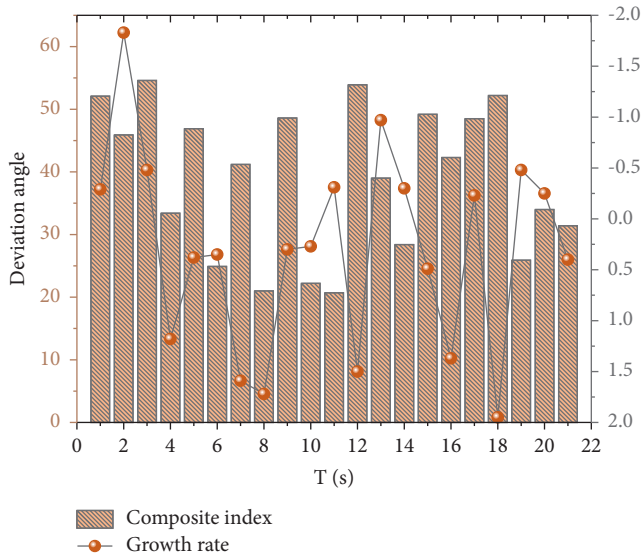

Figure 6: Deviation angle change.

4.2. Results of Target Image Recognition and Localization Analysis. The validation criteria for the measurement accuracy of visual navigation parameters were first established, followed by the analysis of the overall structure of the optical navigation detection system in the field and the experimental locations and experimental methods. Then, experiments on navigation line detection were conducted for the improved RANSAC algorithm and the enhanced probabilistic Hough transforms algorithm under different harvesting environments. The success rate of detection and image processing operation speed were statistically analyzed. Finally, experiments are conducted to verify the success rate of navigation line detection for the image pyramid optical flow tracking algorithm under different environments, and multiple sets of displacement deviation and angle deviation benchmark values are set to verify the measurement accuracy of navigation parameters according to the visual navigation parameter measurement accuracy verification standard. The measurement accuracy of angular deviation of navigation parameters of the image pyramid optical flow tracking algorithm is verified. The experiments on the error measurement of the angular variation of the navigation parameters of the image pyramid optical flow tracking algorithm were performed in 6 groups in turn, with the base value of the angular deviation being 0°. As shown in Figure 7, the average value of the maximum error of the actual measurement value of the angular deviation of the harvester was 10.57°, the average value of the mean error was 3.73°, and the average value of the standard deviation was 2.98°. In the actual detection process, the factors affecting the angle deviation measurement error are mainly: due to the influence of factors such as the influence of driving vision, the manually operated intelligent rice, and wheat harvester cannot always keep the straight line where the left-hand divider is located parallel to the harvest navigation line, i.e., the actual reference value floats around 0°; the incomplete segmentation of harvested and unharvested areas in the image causes the detected navigation line and the straight line where the left-hand divider is located to additional angle exists.

The training process of the detection algorithm includes data set preparation, formal training, and network generalization error evaluation. Then, the deep learning detection algorithm is compared with the traditional SVM algorithm for classification experiments. The results surface that although the deep learning detection check-all rate is slightly lower than the SVM algorithm, its check-accuracy rate and accuracy rate are higher than the SVM algorithm. And the detection rate of the deep learning detection algorithm has a more significant advantage, which can meet the basic requirements of real-time QR code detection for mobile robots. Finally, the operation experiment of the whole system is carried out to verify the feasibility. Moreover, comparing the gyroscope angle data with the path tracking algorithm data can be obtained. Various experiments were designed according to different given algorithm parameters to influence the environment model rasterization resolution parameter on the confidence occupancy meter. The test image size varies. The algorithm automatically transforms the original test map to 448×448 resolution size before input. The output result of the positioning frame will also be altered accordingly. The partial detection results of the deep learning detection algorithm are shown in Figure 8. The results show that the deep learning detection algorithm performs well for QR codes with simple backgrounds, complex backgrounds, vestiges, small deformations, partial occlusions, or multiple QR codes in a single image.

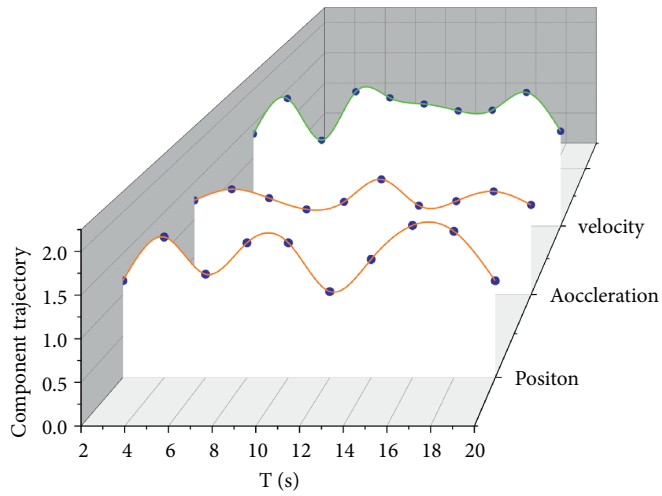

Figure 7: Component trajectories of the rotational path, velocity, and acceleration.

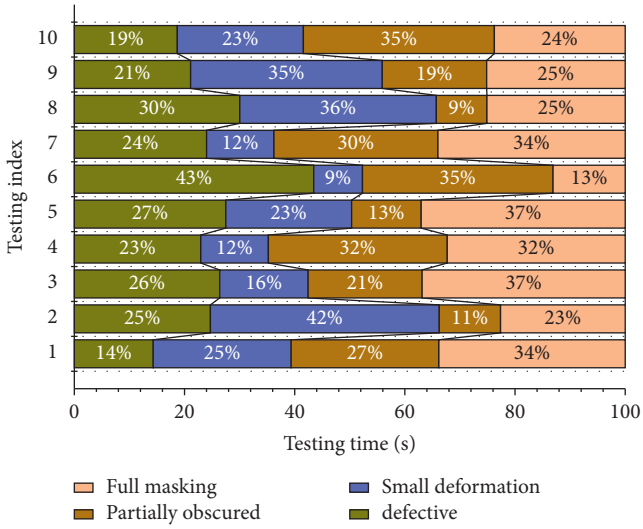

Figure 8: Comparison results of testing indicators.

Among the detection results, the deep learning detection algorithm can make accurate judgments on the images of the experimental scenes of the mobile robot in this paper. The accuracy and rapidity of the integrated detection algorithm can meet the basic requirements of real-time detection of QR codes during the execution of the mobile robot tasks in this paper's experiments. Moreover, the corresponding QR codes are set at each node location, and the mobile robot is placed at any initial node location. First, we start the mobile robot management system and wireless communication server, set the relevant network parameters and establish the wireless LAN server, and wait for the lower computer to access; then, we start the embedded development board TX2, the industrial CCD camera, and the execution component Turtle Bot2 and enable the wireless network access function after each hardware connection of the lower computer is checked and confirmed, set the network parameters and access the LAN established by the wireless server; the upper computer reads the QR code and reads the QR code. The upper computer reads the map file, reads the information from the vehicle controller, initializes the coordinates of the node where the mobile robot is located according to the read information, and displays its position in the map in real time. The gyroscope is corrected once after the QR code, so it is assumed that the gyroscope data is the actual angle. The deviation of the gyroscope angle data from the angle data obtained from image processing is calculated, and the relationship between the selected time nodes and the angle deviation is plotted. From the results, the variation of the path tracking algorithm from the actual angle is ±1.5°, which indicates that the path tracking algorithm in this paper can achieve high accuracy.

5. Conclusion

With the continuous development and progress of science and technology, the intelligent level requirements are gradually increasing, and the automatic control of mobile robots has become an important direction in the development of robot systems, whose visual navigation system is one of the hotspots of research today. In this paper, an in-depth study is conducted on the problem of robot optical navigation path detection. For the problem that the RANSAC algorithm first establishes the linear mathematical model of the path in the navigation line detection and then performs the remaining corner point model verification, which leads to more iterations of the algorithm and more considerable computation, model verification criteria are added to avoid the problem of time-consuming and detection errors caused by continuing iterative verification in the case of model errors. By limiting the range of edge point probability extraction and setting the success criterion of straight line detection, the improved Hough transform algorithm can effectively solve the problem of fast and accurate identification of navigation lines caused by the probability extraction of edge points in the whole detection area; finally, the image pyramid optical flow tracking algorithm is used to realize the tracking detection of robot visual navigation and tracking measurement of visual navigation parameters. The

robot is experimented with setting straight, turning, and obstacle occlusion scenes. From the trajectory graphs obtained from the experimental simulation, the robot can better fit the marking line operation during the operation. The charts of parameters and d show that the bit-posture relationship of the robot body relative to the marking line is closed during the process, which also illustrates the effectiveness and accuracy of target image recognition and localization.

Based on the research work carried out in this paper, we briefly analyze the potential research points with further depth in this study, taking into account the current trends in computer vision and robotics.

The introduction of deep neural network descriptors as the front-end data matching for VSLAM to achieve visual data matching with illumination and viewpoint invariance is a current research hotspot and trend in VSLAM-related work. A considerable amount of work has been carried out in this area, including the extraction of intermediate descriptors using existing network models and the design of a network structure of visual feature descriptors specifically for VSLAM. However, the gap between the current research in this area and the application of VSLAM lies in the visual feature point extraction and the real-time nature of the generation. In addition, deep neural network descriptors tend to have higher dimensionality and take longer to perform feature point matching and distance calculation. Therefore, it is not easy to guarantee the online performance of deep vision descriptors for system applications with high real-time requirements such as VSLAM. Therefore, there is work to be done on a downscale and speed up the depth vision descriptors to meet the demand of VSLAM real-time applications.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work was supported by the 2021 school level scientific research and innovation team project, Innovation and Entrepreneurship Education Scientific Research and Innovation Team, (No. HNACKT-2021-01) and 2021 Research Projects of Educational Science (No. HNACJY-2021-15).

References

[1] C. Sampedro, A. Rodriguez-Ramos, H. Bavle, A. Carrio, P. de la Puente, and P. Campoy, "A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques," *Journal of Intelligent and Robotic Systems*, vol. 95, no. 2, pp. 601–627, 2019.

[2] S. Wang, F. Jiang, B. Zhang, R. Ma, and Q. Hao, "Development of UAV-based target tracking and recognition systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 8, pp. 3409–3422, 2019.

[3] L. Qiu, C. Li, and H. Ren, "Real-time surgical instrument tracking in robot-assisted surgery using multi-domain convolutional neural network," *Healthcare Technology Letters*, vol. 6, no. 6, pp. 159–164, 2019.

[4] A. Devo, G. Mezzetti, G. Costante, M. L. Fravolini, and P. Valigi, "Towards generalization in target-driven visual navigation by using deep reinforcement learning," *IEEE Transactions on Robotics*, vol. 36, no. 5, pp. 1546–1561, 2020.

[5] Y. Xiong, Y. Ge, L. Grimstad, and P. J. From, "An autonomous strawberry-harvesting robot: design, development, integration, and field evaluation," *Journal of Field Robotics*, vol. 37, no. 2, pp. 202–224, 2020.

[6] P. Neubert, S. Schubert, and P. Protzel, "A neurologically inspired sequence processing model for mobile robot place recognition," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3200–3207, 2019.

[7] W. J. Heerink, S. J. S. Ruiter, J. P. Pennings et al., "Robotic versus freehand needle positioning in CT-guided ablation of liver tumors: a randomized controlled trial," *Radiology*, vol. 290, no. 3, pp. 826–832, 2019.

[8] S. G. Mathisen, F. S. Leira, H. H. Helgesen, K. Gryte, and T. A. Johansen, "Autonomous ballistic airdrop of objects from a small fixed-wing unmanned aerial vehicle," *Autonomous Robots*, vol. 44, no. 5, pp. 859–875, 2020.

[9] K. M. Abughalieh, B. H. Sababha, and N. A. Rawashdeh, "A video-based object detection and tracking system for weight sensitive UAVs," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 9149–9167, 2019.

[10] K. Lee, J. Gibson, and E. A. Theodorou, "Aggressive perception-aware navigation using deep optical flow dynamics and PixelMPC," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1207–1214, 2020.

[11] C. H. G. Li and Y. M. Chang, "Automated visual positioning and precision placement of a workpiece using deep learning," *The International Journal of Advanced Manufacturing Technology*, vol. 104, no. 9, pp. 4527–4538, 2019.

[12] D. Fielding and M. Oki, "Technologies for targeting the peripheral pulmonary nodule including robotics," *Respirology*, vol. 25, no. 9, pp. 914–923, 2020.

[13] P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran, R. Jidhesh, and T. Vadivel, "Intelligent face recognition and navigation system using neural learning for smart security in internet of things," *Cluster Computing*, vol. 22, no. 4, pp. 7733–7744, 2019.

[14] V. Vasilopoulos, G. Pavlakos, S. L. Bowman et al., "Reactive semantic planning in unexplored semantic environments using deep perceptual feedback," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4455–4462, 2020.

[15] V. Vasilopoulos, G. Pavlakos, S. L. Bowman et al., "Reactive semantic planning in unexplored semantic environments using deep perceptual feedback," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4455–4462, 2020.

[16] J. W. Martin, B. Scaglioni, J. C. Norton et al., "Enabling the future of colonoscopy with intelligent and autonomous magnetic manipulation," *Nature machine intelligence*, vol. 2, no. 10, pp. 595–606, 2020.

[17] M. Ma, H. Li, X. Gao et al., "Target orientation detection based on a neural network with a bionic bee-like compound eye," *Optics Express*, vol. 28, no. 8, pp. 10794–10805, 2020.

[18] J. Yang, C. Wang, B. Jiang, H. Song, and Q. Meng, "Visual perception enabled industry intelligence: state of the art,

challenges and prospects," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204–2219, 2020.

[19] W.-H. Su, "Advanced machine learning in point spectroscopy, RGB- and hyperspectral-imaging for automatic discriminations of crops and weeds: a review," *Smart Cities*, vol. 3, no. 3, pp. 767–792, 2020.

[20] A. A. Zhilenkov, S. G. Chernyi, S. S. Sokolov, and A. P. Nyrkov, "Intelligent autonomous navigation system for UAV in randomly changing environmental conditions," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 5, pp. 6619–6625, 2020.