ITC 2/54
Information Technology and Control
Vol. 54 / No. 2 / 2025
pp. 396-412
DOI 10.5755/j01.itc.54.2.37912

Improved Agricultural Machinery Navigation Algorithm Based on Machine Learning and Machine Vision Technology

Received 2024/07/05 | Accepted after revision 2025/11/19

HOW TO CITE: Zhu, S., Zhang, W., Zhao, Q., Meng, X., Zhao, C., Feng, W. (2025). Improved Agricultural Machinery Navigation Algorithm Based on Machine Learning and Machine Vision Technology. *Information Technology and Control*, 54(2), 396-412. https://doi.org/10.5755/j01.itc.54.2.37912

Improved Agricultural Machinery Navigation Algorithm Based on Machine Learning and Machine Vision Technology

Fengwu Zhu, Weijian Zhang, Qinglai Zhao, Xianzhang Meng, Weizhi Feng*
College of Engineering and Technology, Jilin Agricultural University, Changchun 130118, China

Chunkai Zhao
Research Institute of Straw Returning Application Technology, Jilin Academy of Agricultural Machinery, Changchun, 130022, China

Corresponding author: weizhifeng12@outlook.com

The automatic navigation of agricultural machinery is one of the important directions in intelligent agriculture research. To realize the automatic production of agricultural machinery, the automatic planning of the navigation route for agricultural machinery is the key. Considering the complexity of the agricultural production environment, the agricultural machinery navigation model is constructed based on binocular vision technology, and the optimized BP network is used to calibrate the binocular vision model. Considering the difficulty in crop identification by traditional machine vision technology, RGB space technology is used to complete image segmentation and noise processing. The optimized S-RANSAC algorithm is used to extract image features. The experimental results showed that in the multi-algorithm agricultural rice field image feature matching test, the S-RANSAC algorithm accurately identified the color difference, shape difference, and hydrological environment difference of seedlings. In contrast, other algorithms were unable to identify complex environmental features. At the same time, in the complex agricultural environment positioning test, the maximum error of the S-RANSAC algorithm was 4.16m, which was better than 5.17m of SURF and had the best positioning performance. It can be seen that the proposed technology has excellent application effects in practical scenarios, providing important technical references for the intelligent development of agriculture and the innovation of visual navigation technology.

KEYWORDS: Machine learning; Machine vision; Feature matching; Binocular vision.

1. Introduction

China is a typical agricultural production country. Agricultural economy is not only related to the economic development of farmers, but also affects the economic lifeline of China's social development. The food problem has already affected the development of the world's population. However, with the increase in the world population, the environment has been destroyed, and resources have been gradually depleted. In this context, the traditional agricultural production model needs to be improved, and automated agricultural production technology will be a new direction for the development of modern agriculture.

Agricultural machinery can be used to perform various tasks on a farm, including planting and harvesting crops, tilling the soil, and applying fertilizers and pesticides. Some common types of agricultural machinery include tractors, combines, plows, harrows, and sprayers. To navigate agricultural machinery safely and effectively, following the manufacturer's instructions and paying attention to the surrounding environment are important.

However, the current visual navigation technology applied to agricultural machinery operations still faces problems, such as complex environments, diverse crop appearances, and low recognition efficiency of visual vision technology, leading to insufficient automation in agricultural machinery. In order to address the current issue of insufficient navigation in agricultural machinery, a machine vision technology for agricultural machinery navigation is proposed [16]. This study constructs an agricultural machinery navigation model based on binocular vision technology, and uses an optimized Genetic Algorithm (GA) to optimize BP network for binocular vision model calibration. Compared with traditional manual calibration, gray matching, and feature point matching, the improved BP model can adapt to more complex agricultural scenes, which has better processing accuracy for nonlinear data [6, 14, 22]. In addition, RGB spatial technology is also used for image segmentation and noise processing, such as using Accelerated Up Robust Features (SURF) to extract image feature points [1, 23]. Considering the insufficient extraction of agricultural features by SURF, the Random Sample Consensus (RANSAC) is used to optimize the feature matching process. The Speeded Up Robust Features Random Sample Consensus (S-RANSAC) algorithm is constructed to extract image features. There are two innovations in the research. Firstly, compared to traditional visual systems, the study introduces an agricultural machinery navigation model based on binocular vision technology, which can dynamically plan in real time and has higher accuracy. The second aspect is the introduction of various visual algorithm techniques to optimize the process of image feature extraction, including the use of SURF algorithm and S-RANSAC algorithm, which enhances the extraction of agricultural image information and further improves the effectiveness of machine planning. There are two research contributions: (1) This study uses a binocular vision-based agricultural machinery navigation model and optimized AG for model calibration, improving the accuracy and efficiency of agricultural production; (2) The research technology provides technical support for the development of agricultural machinery automation, further promoting the development of agricultural economy.

2. Related Work

Machine vision technology provides important ideas for developing modern machine automation, which has been effectively deployed in industrial manufacturing, automotive, aerospace, and other fields [26, 4]. Foreign and domestic experts have carried out relevant research on this. Rehman et al. [17] found that machine vision technology in agricultural production increased crop yield. Machine vision was combined with machine learning technology to process and analyze the color, shape, and texture. Experimental tests showed that the two machine learning techniques had good application value in agricultural production [17].

Mennel et al. [11] found that the development of machine vision technology drove innovation in the smart industry. A new data preprocessing technique was developed to improve the signal processing effect of neural networks. In both supervised and unsupervised learning training, the trained sensor could effectively encode and process images optically projected onto the chip. However, this technology incurs significant costs and still requires improvement and optimization [11].

Chao et al. [2] found that foreign body detection had a more significant impact on tobacco quality. Based on machine vision technology, a new tobacco foreign body detection scheme was proposed, and three cameras were used to identify and extract foreign body color on the characteristic surface. Experimental simulation tests showed that the proposed scheme could remove various packaging foreign bodies in production. For complex tobacco scenes, the research technique still improves the effect on dark details [2]. Taherimakhsousi et al. [19] proposed an efficient non-destructive testing method that quantified the uniformity of large-area perovskite thin films through machine vision. The machine vision technology used was a sample experiment with a pixel resolution of 25. In the experimental test, the proposed technical scheme accelerated the multi-objective optimization of the thin film deposition process and improved the final work effect. However, this technology is targeted at specific substances and still needs further optimization in the future to adapt to more materials [19].

Fan et al. [7] found that machine vision technology has been effectively applied in many fields. A welding state detection method combined with machine vision technology was proposed. The optimization and improvement of the welding process were realized through the visual feature recognition of the welding process. The experimental results showed that this method was feasible and improved the welding quality [7]. Wang et al. [21] found that deep learning has a good application effect in the field of detection and analysis. Hence, a new visual detection method based on deep technology was proposed. The method classified product defects without losing image accuracy. The performance test showed that the proposed method could realize the classification and detection of products well and had high recognition accuracy. However, complex feature extraction and recognition still face difficulties. In the future, feature preprocessing is needed to improve recognition performance [21].

The combination of machine vision and machine language technology in the navigation field has promoted the intelligent development of related industries. Devo et al. [3] found that goal driven visual navigation is the focus of modern society and has been effectively validated in practical scenarios. However, current state-of-the-art methods need to implement adjustments to the parametric model of the environment and objects. A new structural model consisting of two networks is proposed to optimize the generalization problem of traditional vision models. Through specific experimental tests, the proposed scheme had good generalization ability. However, the cost process of network training still requires further optimization of parameters [3]. Zhu et al. [29] found that identifying the landmarks on the planet's surface and using visual navigation technology to achieve a precise landing is an arduous task in the future exploration of Mars and asteroids. Therefore, the detection of the six-degree-of-freedom state was achieved according to the pixel values of the navigation landmarks. Taking the angle invariance characteristic of optical images, the sight vector of landmarks was selected as the observation value. The results showed that the proposed scheme had good feasibility. Although the technology has good application effects, it is still difficult to recognize certain specific angles and needs further improvement [29].

Palossi et al. [15] found that vision-capable robots still face challenges in navigation. Therefore, a visual navigation scheme was proposed based on neural learning technology. The software mapping technology was discussed and the application of the technology was realized. This scheme had good application advantages in the test and promoted the further development of modern nano-UAVs. The results indicated that high-tech had good feasibility, but the recognition of this technology was susceptible to the influence of azimuth angle and made further improvement in the future [15]. Mousavi et al. [13] analyze a robot localization problem with sparse visual feature selection. Based on the problem information matrix analysis, a scalable random sampling algorithm was proposed to obtain the estimation quality by extracting more informative features. Experimental performance tests showed that the proposed scheme outperformed existing greedy methods and performed more efficiently. However, it is susceptible to parameterization issues and further improvements are necessary in the future to address parameter problems [13]. Yasuda et al. [25] analyzed the positioning method of autonomous robots and found that the robot moved in the environment to plan the path effectively. The technology and related literature were reviewed to solve the problem. A systematic review of vision systems suitable for dynamic indoor environments was carried out. The review results showed that the proposed method still faced problems in testing and detection processes, and the relevant standards should be standardized [25].

Zhao et al. [27] found that traditional machine navigation methods could be more computationally efficient and expensive. A mobile robot path planning scheme based on panoramic vision was proposed. It used the environment to explore and map. The performance test showed that the proposed scheme had a good effect on path navigation planning, the length of path planning was shorter, and the time-consuming was also lower. However, the overall computational cost of this technology is relatively high, and future optimization of computational load is necessary [27]. Zheng and Jiang [28] compared the performance of Convolutional Neural Networks (CNN) and Vision Transformers for traffic sign recognition tasks. They found that CNN was more competitive than Transformers in traffic sign classification, pointing out the performance gap of Transformers in this task, and proposed suggestions for improvement [28]. Tummala et al. [20] explored the application of Vision Transformer (ViT) models in brain tumor classification. The pre-trained and fine-tuned ViT models were used for the tri-classification task of MRI images. The ViT model performed excellently in diagnosing brain tumors, with an overall test accuracy of up to 98.7%, demonstrating the potential of ViT model ensembles in computer-aided diagnosis. However, the technology has not been used for the diagnosis of other diseases, and further expansion of data training is needed in the future [20].

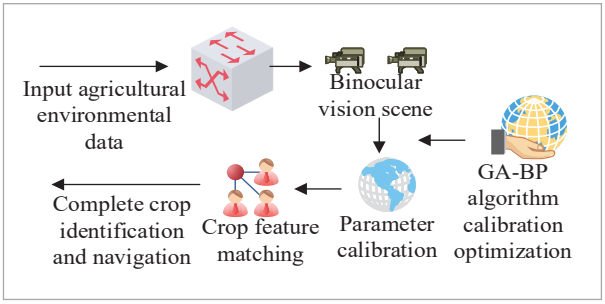
Through the research on the latest navigation technology, it can be found that machine vision and machine learning have important applications in the navigation field and have achieved good application results. However, it still faces low recognition rate in agricultural production. Therefore, a technology based on mechanical vision improvement is proposed for agricultural machinery navigation to accelerate agricultural development.

3. Construction of Agricultural Machinery Navigation Model Based on Machine Vision

3.1 Construction of Agricultural Machinery Navigation Model Based on Binocular Vision

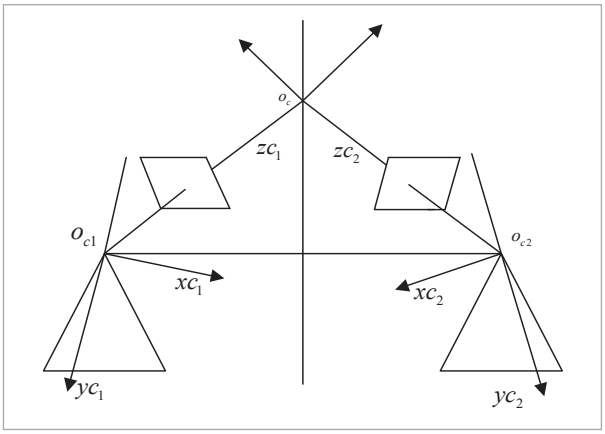
Currently, path planning and navigation are becoming increasingly important in mechanical design [10]. Binocular vision mainly realizes the recognition and collection of image data by two cameras, which can effectively simulate the effect of human eye recognition [9]. It has high recognition accuracy, low technical cost, and high agricultural automation. Therefore, binocular vision and machine algorithm are used to construct an agricultural machinery navigation model. The workflow is shown in Figure 1.

Figure 1
Technical workflow process



Binocular vision is the key to agricultural machinery navigation, which uses binocular vision technology to determine environmental positioning parameters. The principle of binocular vision stereo vision is shown in Figure 2.

Figure 2
Schematic diagram of binocular vision



The binocular camera tilt angle is defined as $\beta=\beta_1=\beta_2$. The angle between cameras is $a = a_1 = a_2$. The focal length is f, and the baseline distance is b. Then, the relationship between the camera world coordinates and the camera position is shown in equation (1):

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ O^T & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{1}$$

In Equation (1), t represents the translation of the space vector. R represents the rotation matrix. $O^T = [0, 0, 0]$. Among them, the new matrix M calculated by the rotation matrix $M = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$ is shown in Equation (2):

$$M = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha & 0 \\ \sin\alpha\sin\beta & \cos\beta & \cos\alpha\sin\beta & 0 \\ \sin\alpha\cos\beta & -\sin\beta & \cos\alpha\cos\beta & k \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

In Equation (2), K represents the position parameter, which is the balance between the image plane and the optical axis under ideal conditions. The transformation relationship between image and camera coordinates is shown in Equation (3):

$$Z_c \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \tag{3}$$

In Equation (3), f is the conversion expression between the world coordinates and the image obtained from Equations (2)-(3), as shown in equation (4):

$$Z_c \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} M \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{4}$$

Equation (2) is combined with equation (4) to obtain the Equation (5):

$$M' = \begin{bmatrix} f\cos\alpha & 0 & -f\sin\alpha & 0 \\ f\sin\alpha\sin\beta & f\cos\beta & f\cos\alpha\sin\beta & 0 \\ \sin\alpha\cos\beta & -\sin\beta & \cos\alpha\cos\beta & K \end{bmatrix} \tag{5}$$

Then, the transformation relationship between the world coordinates and the image is expressed, as shown in Equation (6):

$$Z_c \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = M' \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \tag{6}$$

From Equations (5)-(6), the expression of the binocular camera axis convergence coordinate is obtained, as shown in Equation (7):

$$\begin{cases} X_W = -(K - Z_c)\cos\beta\sin\alpha + \dfrac{(X\cos\alpha\sin\beta)Z_C}{f} \\ Y_W = (K - Z_c)\sin\beta + \dfrac{Y\cos\beta Z_c}{f} \\ Z_W = -(K - Z_c)\cos\alpha\cos\beta + \dfrac{(-X\sin a + Y\sin\beta\cos\alpha)Z_c}{f} \end{cases} \tag{7}$$

The transformation expression between physical coordinates and image pixel coordinates is shown in Equation (8):

$$\begin{cases} u = C_x + \dfrac{X}{dx} \\ v = c_y + \dfrac{Y}{dy} \end{cases} \tag{8}$$

In Equation (8), the coordinates (c_x, c_y) of the of the image center are expressed. dx and dy are the physical dimensions of the X axis and the Y axis expressed by the triangular relationship. The relationship between the camera and the image when the axis is parallel is obtained, as shown in Equation (9):

$$\begin{cases} X_c = \dfrac{u_r, b}{d} \\ Y_c = \dfrac{vb}{d} \\ Z_c = \dfrac{fb}{d} \end{cases} \tag{9}$$

In Equation (9), (u_l, v) is left image coordinates, and (u_r, v) is the right image coordinates. The parallax is $d = u_r - u_l$. In the binocular system, the camera position calibration is the key, then the pose relationship of the camera in the world coordinates is shown in equation (10):

$$M_{ci} = [R\ t] Mc2 \tag{10}$$

In equation (10), M_{ci} is camera parameter matrix. $[R\ t]$ is the external parameters. The relevant expressions are shown in equation (11):

$$\begin{cases} M_{c1} = t_1 + R_1 M_{W1} \\ M_{c2} = t_2 + R_3 M_{W2} \end{cases} \tag{11}$$

In Equation (11), M_{c1} and M_{c2} represent the two camera parameter matrices. $[R_1\ t_1]$ and $[R_1\ t_1]$ both represent the external parameters of the two-ray camera. Considering that the two cameras of agricultural machinery are fixed on a calibration board, there exists $M_{W1} = M_{W2}$. The external parameters of the camera are expressed as shown in Equation (12):

$$\begin{cases} R = R_1 R_2^{-1} \\ t = t_1 - R t_2 \end{cases} \tag{12}$$

Through the analysis of BP neural network, it is better to use BP neural network to calibrate the camera than the traditional calibration method. However, BP neural network is easy to fall into local convergence, so GA is used to optimize PB to avoid BP network falling into local convergence [8].

The number of nodes in the three layers of the neural network is defined as respectively n_i, L_n, n_o. The weight parameter and the threshold parameter of the unified coding are used for calculation. The coding length is shown in Equation (13):

$$Length = n_j \times L_n + L_n \times n_o + L_n + n_o \tag{13}$$

The image coordinates are used as the input of the network, and the result is the predicted world coordinates. Then, the relationship between the world coordinate value and the expected value is used as the fitness function, as shown in Equation (14):

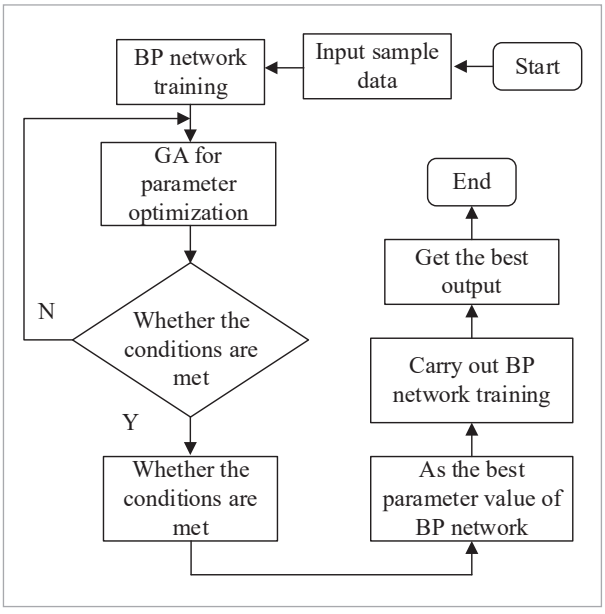
$$f(x) = \frac{1}{SSE(Value - newValue)} \tag{14}$$

In Equation (14), *Value* represents the expected output value. *newValue* represents the actual world coordinate value. To get the best calibration parameter value, it is necessary to select the best individual in the population for reproduction [12]. The roulette method is used to determine the probability of the next generation through arithmetic cross-calculation, and the expression is shown in Equation (15):

$$\begin{cases} c_1 = p_1 a + p_2 (1 - a) \\ c_2 = p_1 (1 - a) + p_2 a \end{cases} \tag{15}$$

In Equation (15), a is a random number in the range (0,1). c_1, c_2 all represent cross individuals. The population parameters are adjusted by non-uniform probability distribution to realize the variation of the population and achieve the optimal solution to obtain the optimal output result. The optimized GA-BP workflow is shown in Figure 3.

Figure 3
GA-BP work flow chart



3.2 Algorithm Construction Based on Crop Feature Matching

There are many common image feature extraction algorithms. The three-dimensional visual map of the agricultural environment is constructed through three-dimensional feature matching, and the automatic navigation of agricultural machinery

is achieved. In the binocular stereo vision system of agricultural machinery, the same target can be presented from different perspectives. Stereoscopic feature matching is the process of matching identical feature points in two camera images. However, the agricultural environment is complex and easily affected by dust, low light, and different types of crop morphology. In order to ensure effective navigation and planning of agricultural machinery, it is necessary to uniformly identify and analyze the characteristics of farmland. Considering that different crops are mainly green plants, RGB space is used for image segmentation processing. The Speeded Up Robust Feature (SURF) method is used to extract image feature points, which can adapt well to the influence of complex environments [24]. SURF algorithm has good robustness in image feature extraction and is a common local feature extraction method, which describes feature points in the form of a Hessia-Laplace matrix [5]. The detection feature point is defined as $H(x, \sigma)$, as shown in Equation (16):

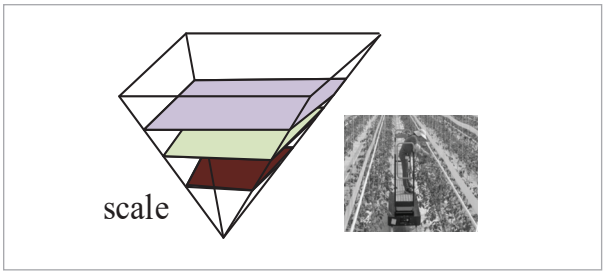
$$H(x,\sigma)=\begin{bmatrix} L_{xx}(X,\sigma)\ L_{xy}(X,\sigma) \\ L_{xy}(x,\sigma)\ \ L_{yy}(X,\sigma) \end{bmatrix} \tag{16}$$

In Equation (16), σ represents the scale parameter. $L_{xx}(X, \sigma)$, $L_{xy}(X, \sigma)$ and $L_{yy}(X\sigma)$ represent image convolutions. In the process of image processing, the accelerated calculation and integration method are used for data processing. They are described in the form of a matrix, corresponding to convolution and filtering D_{xx}, D_{xy}, D_{yy}, respectively, and the expression is shown in Equation (17):

$$\det(H_{app}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \tag{17}$$

In Equation (17), ω represents the proportional coefficient, which is 0.9. In graphic data processing, Gaussian smoothing filter is used. The determinant in Equation (17) $\det(H_{app})$ takes a positive value, indicating that the point is an extreme point. In the extraction of image data, the SURF algorithm uses the box-type filter to achieve scale transformation and improve the processing effect of the algorithm [18]. The schematic diagram of the SURF spatial scale is shown in Figure 4.

Figure 4
Schematic diagram of SURF spatial scale

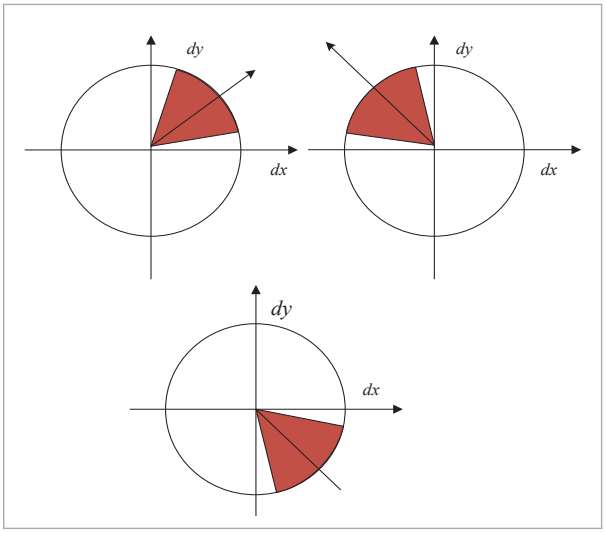


The SURF spatial scale mainly comes from Octaves and divides the pyramids into groups. Layering is performed on the basis of groups, and there is a relationship between each pyramid's upper and lower layers. Different filtering templates are used to ensure that the scales between each group will not overlap.

After the extraction process of the algorithm is clarified, the Harr wavelet features in the feature points are counted. Generally, the entire area is traversed by radians, and the end with the longest vector is used as the main direction. The schematic diagram of the main direction is shown in Figure 5.

In the process of image feature extraction in the agricultural environment, the image contains various complex scenes, including weeds, straws, stones, different light environments, etc. These factors will

Figure 5
Schematic diagram of main direction extraction



cause a great interference to image feature matching. To improve the actual matching effect of the image, it is necessary to optimize the extracted data. The preprocessing includes three steps: binarization, filtering, and image segmentation. In the agricultural production environment, it is only necessary to ensure the accuracy of mechanical visual recognition. In order to avoid interference from complex environments on the recognition of agricultural machinery, binary image segmentation is performed using the Red green blue (RGB) method on the basis of the original color image. Its function is to preserve the most important information features in the image to the greatest extent possible. The specific operating principle is that the green component is greater than the red and blue components as the reserved parts for image binarization, as shown in Equation (18):

$$f(x,y)=\begin{cases}1, G>R\ AND\ G>B\\ 0,\ otherwise\end{cases} \tag{18}$$

In order to remove the influence of lighting and other factors on the binarization operation, statistical filtering is used to set a threshold to preserve the corresponding image details and perform denoising. The principle is to dynamically adjust the binarization threshold by setting a window to calculate the statistical characteristics of local regions in the image. This method can effectively reduce the impact of uneven lighting, preserve image details, and improve the binarization effect. Its mathematical expression is shown in Equation (19):

$$G(x,y)=T\left[f(x,y)\right]_{k\times l}=\begin{cases}1 & if\left(\frac{n}{k\times 1}\right)>Threshold\\ 0, & otherwise\end{cases} \tag{19}$$

In Equation (19), T represents the template processing method. k, l represent the length and width of the template, respectively. n represents the quantity value. $\frac{n}{k\times 1}$ represents the set threshold parameter. When the ratio $\frac{n}{k\times 1}$ is lower than the $\frac{n}{k\times 1}$ value, $G(x, y)$ is set to 0. Otherwise, its value is set to 1. The machinery in the agricultural environment also needs to adapt to the proportion of green between different crops. The adaptive statistical filtering method is used for calculation based on the binarized image. The proportion of each color pixel in the image is set as the threshold weight to obtain the threshold between different images. Then, the effect of image filtering processing can be realized, and the expression is as shown in Equation (20):

$$G(x,y)=T\left[f(x,y)\right]_{k\times l}=\begin{cases}1 & if\left(\frac{n}{k\times 1}\right)>k\times 1\times\frac{N}{\mathrm{Re}\,solution}\\ 0, & otherwise\end{cases} \tag{20}$$

In Equation (20), Re *solution* represents the resolution of the image. N represents the number of white pixels in the binarized image. By performing statistical filtering on the image, more detailed features of the image are extracted. However, the image features extracted by traditional SURF algorithms still face interference from complex factors in farmland. In order to extract more ideal feature points, the RANSAC is used to optimize the matching results of the SURF algorithm and construct the S-RANSAC algorithm. The RANSAC algorithm has robustness and randomness characteristics. In order to improve the processing performance of the RANSAC algorithm, research has increased the number of iterations to enhance robustness, while adjusting the error threshold to accurately control the distinction between inliers and outliers. Image segmentation reduces the number of features in data processing, which optimizes computational efficiency and results. The sampling set is defined as A to obtain the correct mathematical calculation model results M, and samples n. The RANSAC algorithm can well screen the abnormal data of the sampling set, and ensure the training effect of the model. When there are too many left and right image features of agricultural machinery, it is necessary to perform matching screening. The random interior point probability is shown in Equation (21):

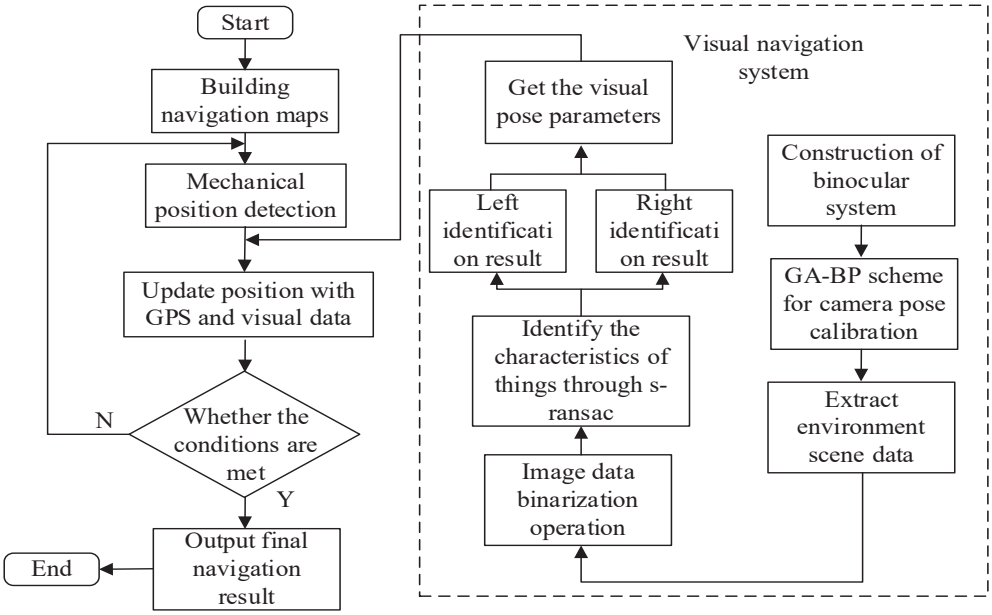
$$1-p=(1-W^n)^k \tag{21}$$

In Equation (21), P represents the probability value of randomly selected points. W represents the probability of points in a specific sample set, and the iteration times table is shown in Equation (22):

$$k=\frac{\log(1-p)}{\log(1-W^n)} \tag{22}$$

When the agricultural machinery binocular system extracts feature information, the left, and right camera pose, perspectives and images are different, and the camera pose parameters are also different, which will reduce the matching effect of image feature points. Therefore, the original matching method is optimized, as shown in the agricultural machinery positioning process in Figure 6.

Figure 6
Positioning flow chart of agricultural machinery



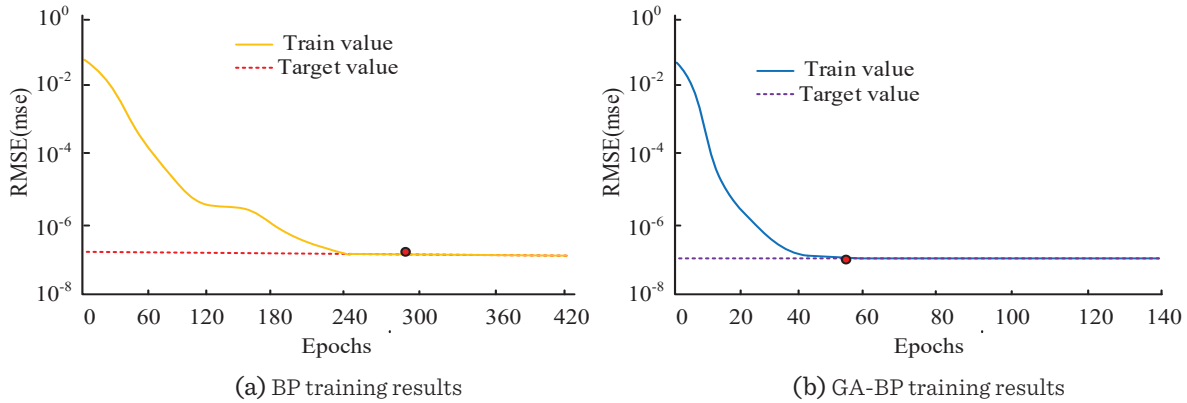
4. Experimental Testing and Analysis

4.1 Binocular Camera Calibration Performance Test

The agricultural machinery binocular camera positioning experiment was completed using MATLAB tools, with a testing system of Win7 and a running memory of 16GB. The study used GA-BP network to obtain calibration parameters, where the number of neural network nodes was 2-5-2. The learning rate of GA algorithm was set to 0.01, which could balance convergence speed and stability. The number of iterations was set to 500, and the mutation rate was set to 0.05 to prevent premature convergence. A total of 40 experimental tests were conducted to evaluate the performance accuracy of the neural network in calibration using Root Mean Square Error (RMSE). Figure 7 shows the results of neural network training.

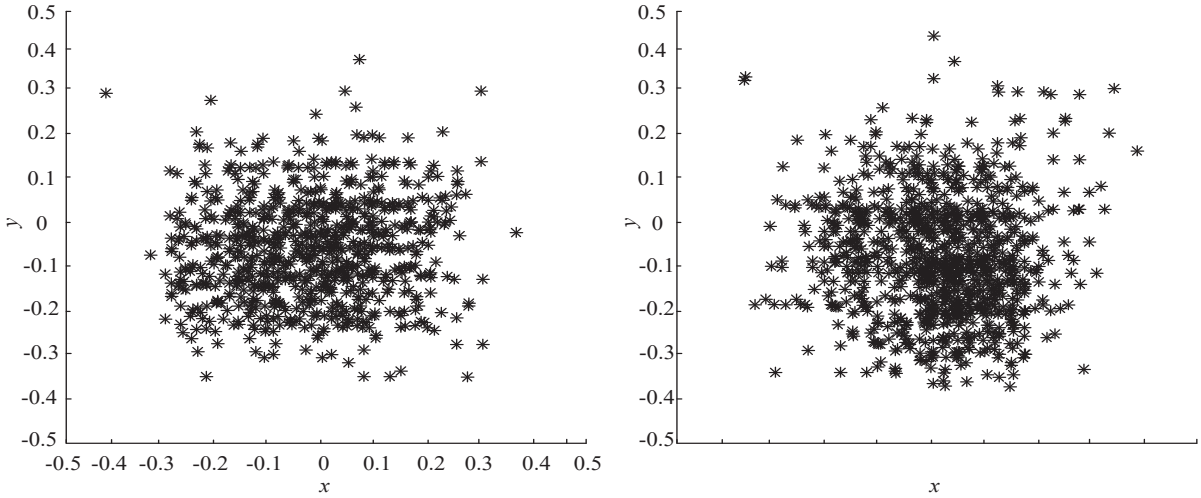
Figure 7 shows the training performance of the neural network in calibration. From Figure 7(a), the traditional BP algorithm tended to converge after 240 iterations and reached the target value at this moment, and the RMSE error was 10^{-7}. Figure 7(b) is the calibration training result of the GA-BP algorithm. Compared with the training effect of the BP algorithm, the GA-BP algorithm had obvious advantages. It tended to converge after 60 iterations, and the training efficiency was better. Compared with the BP algorithm, RMSE improved by 12%. To sum up, the GA-BP algorithm prevents the BP algorithm from falling into the local optimal solution problem. At the same time, it guarantees training efficiency and training accuracy. Figure 8 shows the calibration error distribution of the left and right cameras.

Figure 7
Neural network calibration test results



(a) BP training results (b) GA-BP training results

Figure 8
Calibration distribution of agricultural machinery cameras



(a) Left projection error (b) Right projection error

Figure 8 shows the calibration distribution of agricultural machinery cameras. The camera calibration parameters are optimized through GA-BP algorithm training, and the final calibration results are obtained. The calibration error of the left camera was -0.4-0.5 in the x direction and -0.4-0.4 in the y direction. The calibration error of the right camera was -0.4-0.4 in the x direction and -0.4-0.5 in the y direction. The pixel error ranges of the left and right cameras are close, and the error points are both close to the center, indicating that the cameras have achieved better calibration results. The final calibration parameters of the agricultural machinery binocular system are shown in Table 1.

Table 1
Final calibration parameters of the agricultural machinery binocular system

Distance / cm	Left focal length (horizontal and vertical coordinates)		Right focal length (horizontal and vertical coordinates)		Left error (horizontal and vertical coordinates)		Right error (horizontal and vertical coordinates)	
70	817.25	819.54	817.19	819.45	0.175	0.156	0.157	0.153
100	814.75	817.64	814.65	825.64	0.137	0.152	0.129	0.141
130	812.15	813.56	807.65	810.54	0.134	0.122	0.134	0.129
160	813.86	815.16	833.54	834.22	0.121	0.124	0.123	0.121
190	768.35	771.36	829.54	832.54	0.116	0.107	0.122	0.115
220	802.65	805.56	835.64	840.35	0.136	0.122	0.123	0.115
250	836.75	857.48	785.25	786.58	0.133	0.135	0.338	0.137

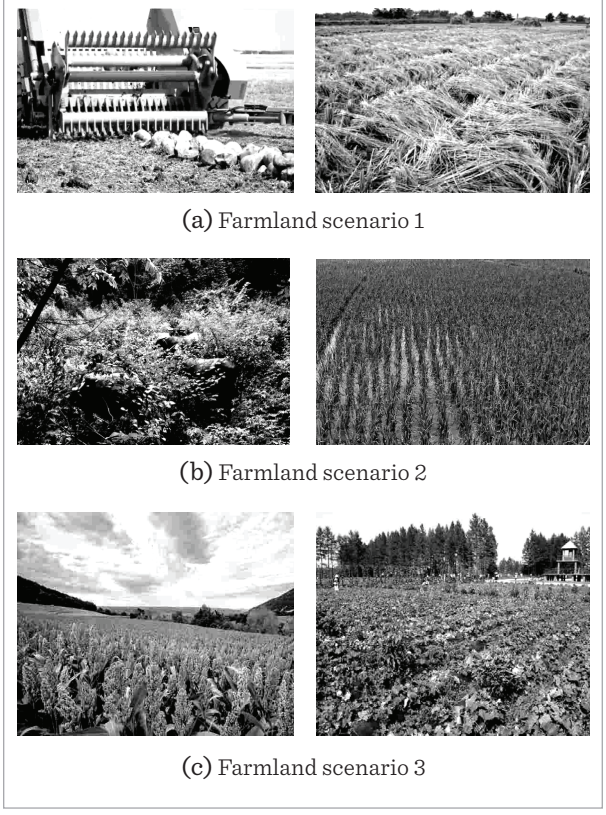
Different calibration object distances have different error results. Table 1 shows the calibration imaging error results at different object distances. As the calibration distance increased from 70cm to 150cm, the overall pixel error showed a downward trend. The error gradually increased when the calibration distance reached 190cm. Therefore, when calibrating the camera, it is necessary to fully consider the relationship between the distance and the error. At the same time, a longer distance will lead to a decrease in image clarity. Therefore, considering the above factors, the calibration distance is selected to be 190cm to ensure the imaging effect of the agricultural machinery navigation process.

4.2 Performance Test of Agricultural Machinery Navigation Algorithm

In the image feature matching test, Visual Studio is used to complete image processing. Related tests are completed on the OpenCV platform. The testing platform is Win7 with 16GB of running memory. The complex characteristics of agricultural environments, including weeds, stones, straw, etc., can interfere with experiments and affect the final matching performance. Therefore, the binarization operation and statistical filtering method are used to preprocess the image. Figure 9 displays the preprocessed test sample data.

Figure 9 is the original data map of the processed farmland, including special scenes such as straw, stones, green crops, agricultural machinery, flowers, lighting, etc. Features from Accelerated Segment Test (FAST) and Scale-invariant Feature (SIFT) are selected. Transform, SIFT, and SURF are three standard matching algorithms involved in the matching test, and the matching effect of each algorithm in the agricultural paddy field is shown in Figure 10.

Figure 9
Original data of farmland after processing

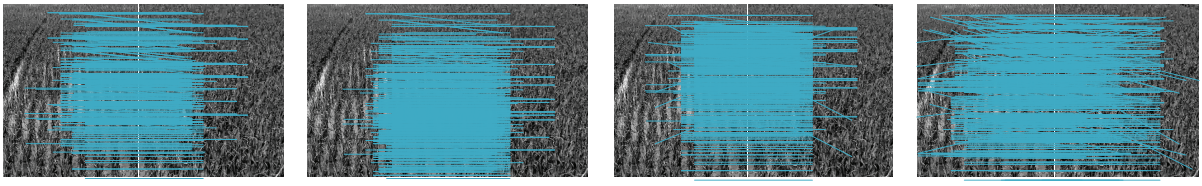


(a) Farmland scenario 1

(b) Farmland scenario 2

(c) Farmland scenario 3

Figure 10
The test results of four different matching schemes

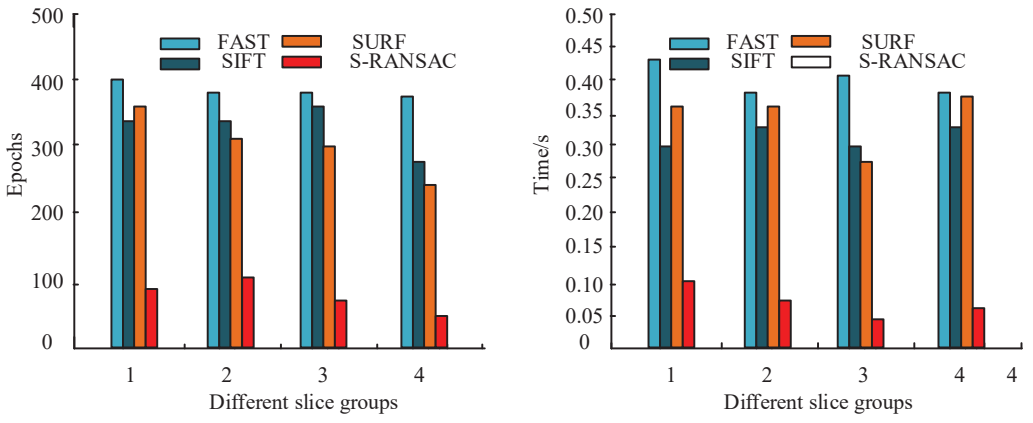


(a) FAST matching result (b) SIFT matching result (c) SURF matching result (d) S-RANSAC matching result

As shown in Figure 10, the image feature matching results of four algorithms are compared in the same scene. The experimental site is an open green rice field scene, and the rice field environment is mainly affected by light factors, water surface refraction, and color factors. Based on the actual matching effect, FAST can accurately match seedlings in the rice field environment, but the shape and color of seedlings can easily affect the matching, and the effect of extracting visual edge features is poor. In the SIF image matching results, compared with the FAS matching results, this scheme can more accurately identify the shape of seedlings, but the overall processing is poor for areas with prominent light and shadows and locally darker areas. Finally, the SURF matching results are compared. This scheme has achieved recognition of seedling shape, hydrological environment, etc. However, there is a lack of detail processing for water surface edge details and image edge parts, and the overall matching effect is average. Finally, the S-RANSAC matching result shows that compared with the first three matching schemes, it can accurately identify the color differences, shape differences, and hydrological environment differences of seedlings. At the same time, it also has better ability to handle dark and edge details. The performance matching of each algorithm is shown in Figure 11.

Figure 11 shows the comparison effect of the matching performance of each algorithm. Four different types of agricultural scene images are selected for feature-matching tests, including various complex

Figure 11
Comparison of matching performance of various algorithms

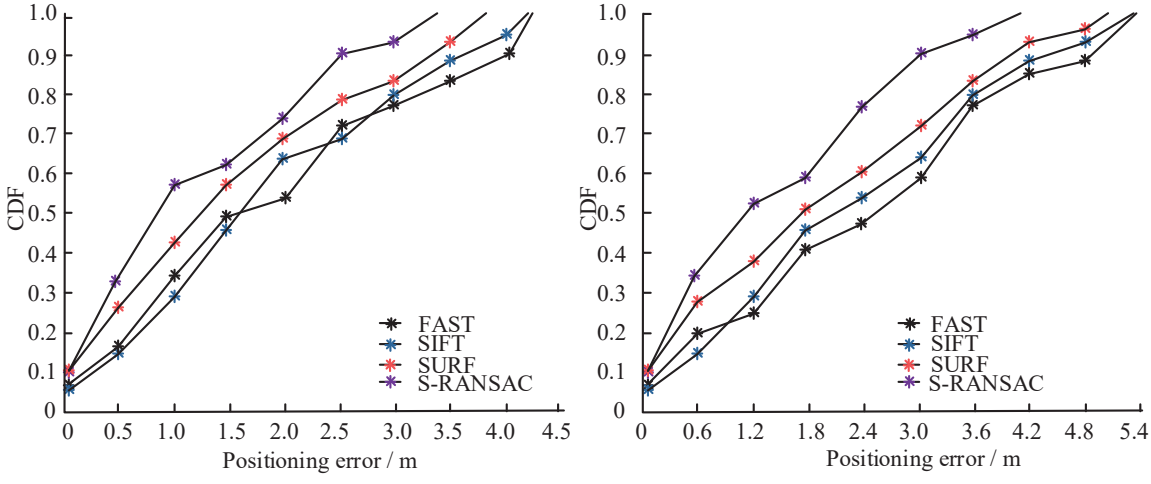


(a) Iteration relation result (b) Time relationship diagram

scenes such as stones, straws, seedlings, weeds, and rivers. Figure 11(a) shows the number of iterations for each scheme to complete image matching. The iterations of FAST, SIFT, SURF and S-RANSAC in Figure 11(a) were 400, 351, 362, and 98, respectively. In group 1, the iterations were 391, 356, 301, and 108. In group 3, the iterations were 390, 371, 299, and 94. In group 4, the iterations were 392, 285, 271, and 56, respectively. From the number of iterations, it can be seen that the number of iterations for image matching between different schemes in different agricultural scenarios is not the same. The S-RANSAC algorithm has excellent iterative performance in different scenarios. Figure 11(b) shows the matching time results of each scheme. From the image data in group 1, the matching time of FAST, SIFT, SURF, and S-RANSAC was 0.44s, 0.30s, 0.36s, and 0.10s, respectively. In group 4, the matching time of each scheme was 0.37s, 0.32s, 0.37s, and 0.07s, respectively. In different agricultural scenarios, compared with the other three algorithms, the S-RANSAC algorithm has the best performance in terms of matching iterations and matching time. Finally, the Cumulative Distribution Function (CDF) is used to test the navigation accuracy of different algorithms, and the results are shown in Figure 12.

Figure 12 shows the cumulative error results of the positioning system, which are tested in two different agricultural scenarios. Figure 12(a) is the location test result of the rice farmland environment. From the map location data, the maximum location errors under the FAST, SIFT, and SURF schemes were 4.25m, 4.15m, and 3.76m, respectively. Under the S-RANSAC algorithm, the maximum positioning error was 3.3m. Figure 12(b) displays the positioning test in a complex agricultural environment. The test environment is affected by complex environments such as light, color, and water waves. Compared with the rice positioning environment, the overall error is higher. The maximum positioning errors under the FAST, SIFT, SURF, and S-RANSAC schemes were 5.35m, 5.35m, 5.17m, and 4.16m, respectively. In different agricultural environments, the agricultural machinery under the S-RANSAC algorithm can obtain the best positioning accuracy and meet the requirements of automatic positioning of agricultural machinery. Finally, the proposed technology is compared with the latest navigation planning techniques, including Navigation Planning Techniques Based on the Coherent Structure (VGGNet) model, Regional Convolutional Neural Networks (R-CNN), Structure from Motion (SfM), and Multi View Stereo (MVS), as shown in Table 2.

Figure 12
Cumulative error result of the positioning system



(a) Agricultural scenario 1

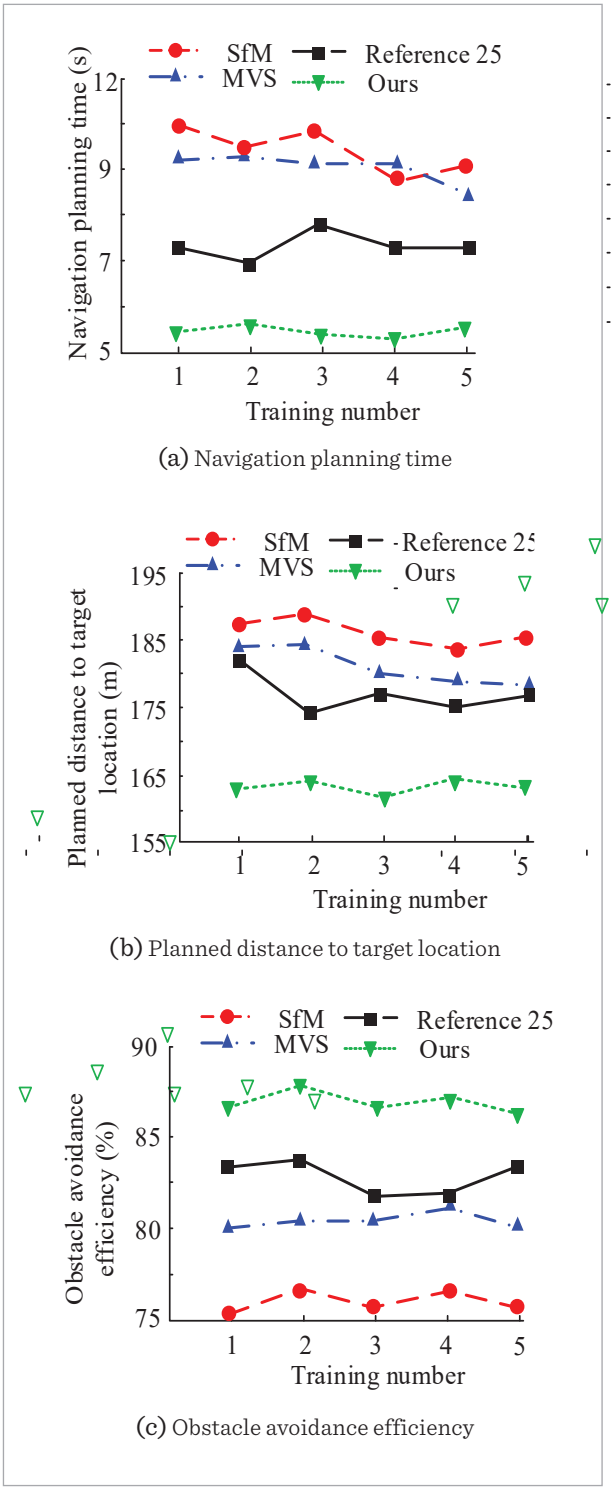
(b) Agricultural scenario 2

Table 2
Comparison with Latest Navigation Planning Technologies

Model type	Navigation planning time (s)	Planned distance to target location (m)	Obstacle avoidance efficiency (%)
VGGNet	15.6	198	75.65
R-CNN	16.6	201	78.65
SfM	9.6	187	73.65
MVS	8.2	185	79.35
The proposed model	5.7	161	86.56

Based on the data in Table 2, the navigation planning time, distance to target planning, and obstacle avoidance efficiency of different models are discussed and analyzed. Firstly, in terms of navigation planning time, the proposed model had the lowest navigation planning time of 5.7s, which was more efficient compared with other models. The VGGNet model had the highest navigation planning time, at 15.6s, followed by the R-CNN model at 16.6s. The navigation planning time for SfM and MVS models was 9.6ds and 8.2s, respectively. Secondly, in terms of planning distance to the target location, there was not much difference in distance between all models. The proposed model had the minimum planning distance of 161m, while the maximum planning distance was 198m of the VGGNet model. Finally, in terms of obstacle avoidance efficiency, the proposed model had the highest obstacle avoidance efficiency of 86.56%. The MVS model had a second highest obstacle avoidance efficiency of 79.35%, followed by the R-CNN model at 78.65%. The obstacle avoidance efficiency of the VGGNet model was the lowest, at 75.65%. In summary, the proposed model performs well in terms of navigation planning time, distance to target planning, and obstacle avoidance efficiency, with high efficiency and accuracy. It can be seen that the proposed technology has better application effects in the field of agricultural navigation Finally, the technology proposed in [8] is introduced to participate in the comparison, and corn is selected as the experimental object. In addition, considering the average performance of VGGNet and R-CNN, no comparison will be made in the following. The navigation effects of different agricultural machinery technologies are compared, as shown in Figure 13.

Figure 13
Comparison of different baseline agricultural machinery planning



(a) Navigation planning time

(b) Planned distance to target location

(c) Obstacle avoidance efficiency

Figure 13 (a)-(c) show the test results of navigation planning time, target planning distance, and obstacle avoidance efficiency, respectively. In terms of navigation planning time, the comprehensive effect of the research model is the best. For example, in 5 experiments, the average navigation planning time of the research model was 5.56s, while the average navigation planning time of the SfM model was 9.62s. The average navigation planning time of the MVS model was 8.62s, and the average navigation planning time of reference 25 was 7.43s. In the target planning distance, the research model had an average of 160.34m, which was better than similar models. In the comparison of obstacle avoidance efficiency, the average obstacle avoidance efficiency of the research model was 89.353%, which also performed the best.

5. Conclusion

Agricultural automation is an essential link in the development of modern agriculture. Problems such as low positioning accuracy of agricultural machinery in automated navigation need to be overcome, mainly because the complex environmental characteristics of agriculture affect the visual recognition effect of agricultural machinery. Therefore, an agricultural machinery navigation model was built based on the binocular vision platform. Considering the low calibration accuracy of the traditional camera calibration method, the BP neural network optimized by the GA was used to calibrate the binocular vision system. The complexity of agricultural environment affects machine vision recognition. Therefore, the RGB space method was used to achieve image segmentation and noise processing, and the optimized S-RANSAC algorithm was used to extract image feature information. The experimental test showed that in the camera calibration test, the GA-BP algorithm tended to converge after 60 iterations, and the BP algorithm tended to converge after 240 iterations. Compared with the BP algorithm, the calibration accuracy of the GA-BP algorithm was improved by 12%. In the agricultural machinery navigation test, the proposed S-RANSAC algorithm was superior to other algorithms in feature matching, time and accuracy. In the cumulative error test of the positioning system, the maximum positioning errors of FAST, SIFT, SURF, and S-RANSAC in complex environments were 5.35m, 5.35m, 5.17m, and 4.16m, respectively. It can be seen that S-RANSAC has excellent positioning accuracy performance. There are also shortcomings in the research. The navigation control mainly realizes the navigation effect by constructing the visual model, which lacks reliability for acquiring navigation information. In the later stage, ultrasonic and inertial navigation technology can be added to improve the navigation system. Finally, in harsh foggy, rainy, dusty, and extreme lighting conditions

The comparison results are shown in Table 3.

Table 3 shows the comparison results under harsh environments, selecting foggy, rainy, dusty, and extreme lighting environments to compare the planning effects of different technologies. According to the results, dust and extreme weather had the greatest impact on agricultural machinery navigation. For example, in dusty environments, SfM, MVS, reference 25, and research model navigation planning took 17.64s, 417.97s, 18.98s, and 12.55s, respectively. Overall, the research model performs the best. In the target planning distance, such as in extreme lighting environments, SfM, MVS, reference 25, and research model navigation planning took 259.4m, 256.76m, 229.95m, and 182.54m, respectively. The overall planning dis-

Table 3
Comparison under harsh environments

Environment	Navigation planning time (s))				Planned distance to target location (m))			
	SfM	MVS	Reference 25	Ours	SfM	MVS	Reference 25	Ours
Foggy days	15.97	14.59	13.48	10.26	256.4	245.46	224.45	175.46
rainy day	15.59	15.27	15.89	11.47	257.4	247.77	224.64	181.54
dust	17.64	17.97	18.98	12.55	265.5	257.68	231.85	183.54
Extreme lighting	18.69	18.98	17.56	11.95	259.43	256.76	229.95	182.54

tance of the research model is the shortest, which is suitable for foggy and rainy days. In addition, the memory usage and running time of different technologies are compared, as shown in Table 4.

Table 4
Technical memory usage and running time

Project	SfM	MVS	Reference 25	Ours
Memory usage rate (%)	63.32%	62.65%	60.25	43.25
Processing time (s)	1.23	1.16	1.02	0.86

The memory usage and running time of different technologies are displayed in Table 4. The research model showed better memory usage and processing time, with a running occupancy rate of 43.25% and a processing time of 0.86s. From this, it can be seen that research techniques have excellent shimmering effects under the actual Yangtze River. There are also shortcomings in the research. Navigation control mainly relies on visual model construction to achieve navigation effects, which lacks reliability in obtaining navigation information. In the later stage, ultrasonic and inertial navigation technologies can be added to improve the navigation system.

Funding

Projects of Jilin Provincial Department of Science and Technology: Research on Multi parameter Accurate Control Technology of Northern Pig House Environment Based on Support Vector Machine.

References

1. Abualigah, L., Almotairi, K. H. Dynamic Evolutionary Data and Text Document Clustering Approach Using Improved Aquila Optimizer Based Arithmetic Optimization Algorithm and Differential Evolution. Neural Computing and Applications, 2022, 34(23), 20939-20971. https://doi.org/10.1007/s00521-022-07571-0
2. Chao, M., Kai, C., Zhiwei, Z. Research on Tobacco Foreign Body Detection Device Based on Machine Vision. Transactions of the Institute of Measurement and Control, 2020, 42(15), 2857-2871. https://doi.org/10.1177/0142331220929816
3. Devo, A., Mezzetti, G., Costante, G., Fravolini, M. F., Valigi, P. Towards Generalization in Target-Driven Visual Navigation by Using Deep Reinforcement Learning. IEEE Transactions on Robotics, 2020, 36(5), 1546-1561. https://doi.org/10.1109/TRO.2020.2994002
4. Din, A. F. U., Mir, I., Gul, F., Mir, S., Alhady, S. S. N., Nasar, M. R. A., Alkhazaleh, H. A., Abualigah, L. Robust Flight Control System Design of a Fixed Wing UAV Using Optimal Dynamic Programming. Soft Computing, 2022, 1-12. https://doi.org/10.1007/s00500-022-07484-z
5. Ding, H., Zhang, B., Zhou, J., Yan, Y., Tian, G. Recent Developments and Applications of Simultaneous Localization and Mapping in Agriculture. Journal of Field Robotics, 2022, 39(6), 956-983. https://doi.org/10.1002/rob.22077
6. Espinoza-Fraire, A. T., Sáenz-Esqueda, A., Córtes-Martínez, F. Adjustment Mechanism with Sliding Mode for Adaptive PD Controller Applied to Unmanned Fixed-Wing MAV Altitude. International Journal of Robotics and Automation (IJRA), 2021, 10(4), 308-318. https://doi.org/10.11591/ijra.v10i4.pp308-318
7. Fan, X., Gao, X. D., Liu, G. Q., Ma, N., Zhang, Y. X. Research and Prospect of Welding Monitoring Technology Based on Machine Vision. The International Journal of Advanced Manufacturing Technology, 2021, 115(11), 3365-3391. https://doi.org/10.1007/s11119-022-09881-8
8. Fang, Q., Xu, X., Wang, X. T., Zeng, Y. J. Target- Driven Visual Navigation in Indoor Scenes Using Reinforcement Learning and Imitation Learning. CAAI Transactions on Intelligence Technology, 2022, 7(2), 167-176. https://doi.org/10.1007/s00170-021-07398-4
9. Hai, J., Hao, Y. T., Zou, F. Z., Lin, F., Han, S. C. A Visual Navigation System for UAV Under Diverse Illumination Conditions. Applied Artificial Intelligence, 2021, 35(15), 1529-1549. https://doi.org/10.1080/08839514.2021.1985799
10. Lu, W., Zeng, M., Qin, H. Intelligent Navigation Algorithm of Plant Phenotype Detection Robot Based on Dynamic Credibility Evaluation. International Journal of Agricultural and Biological Engineering, 2021, 14(6), 195-206. https://doi.org/10.25165/j.ijabe.20211406.6615
11. Mennel, L., Symonowicz, J., Wachter, S., Polyushkin, D. K., Molina-Mendoza, A. J., Mueller, T. Ultrafast Machine Vision with 2D Material Neural Network Image Sensors. Nature, 2020, 579, 62-66. https://doi.org/10.1038/s41586-020-2038-x

12. Morad, S. D., Mecca, R., Poudel, R. P. K., Liwicki, S., Cipolla, R. Embodied Visual Navigation with Automatic Curriculum Learning in Real Environments. IEEE Robotics and Automation Letters, 2021, 6(2), 683-690. https://doi.org/10.1109/LRA.2020.3048662

13. Mousavi, H. K., Motee, N. Estimation with Fast Feature Selection in Robot Visual Navigation. IEEE Robotics and Automation Letters, 2020, 5(2), 3572-3579. https://doi.org/10.1109/LRA.2020.2974654

14. Muneer, A., Dairabayev, Z. Design and Implementation of Automatic Painting Mobile Robot. IAES International Journal of Robotics and Automation (IJRA), 2021, 10(1), 68-74. https://doi.org/10.11591/ijra.v10i1.pp68-74

15. Palossi, D., Loquercio, A., Conti, F., Flamand, E., Scaramuzza, D., Benini, L. A 64-mW DNN-Based Visual Navigation Engine for Autonomous Nano-Drones. IEEE Internet of Things Journal, 2019, 6(5), 8357-8371. https://doi.org/10.1109/JIOT.2019.2917066

16. Pawlowski, A., Romaniuk, S., Kulesza, Z., Petrović, M. Trajectory and Optimization Using Learning from Demonstration with Meta-Heuristic Grey Wolf Algorithm. IAES International Journal of Robotics and Automation (IJRA), 2022, 11(4), 263-277. https://doi.org/10.11591/ijra.v11i4.pp263-277

17. Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., Shin, J. Current and Future Applications of Statistical Machine Learning Algorithms for Agricultural Machine Vision Systems. Computers and Electronics in Agriculture, 2019, 156, 585-605. https://doi.org/10.1016/j.compag.2018.12.006

18. Ruangurai, P., Dailey, M. N., Ekpanyapong, M., Soni, P. Optimal Vision-Based Guidance Row Locating for Autonomous Agricultural Machines. Precision Agriculture, 2022, 23(4), 1205-1225. https://doi.org/10.1007/s11119-022-09881-8

19. Taherimakhsousi, N., Fievez, M., MacLeod, B. P., Booker, E. P., Fayard, E., Matheron, M., Manceau, M., Cros, S., Berson, S., Berlinguette, C. P. A Machine Vision Tool for Facilitating the Optimization of Large-Area Perovskite Photovoltaics. NPJ Computational Materials, 2021, 7(1), 1-10. https://doi.org/10.1038/s41524-021-00657-8

20. Tummala, S., Kadry, S., Bukhari, S. A. C., Rauf, H. T. Classification of Brain Tumor from Magnetic Resonance Imaging Using Vision Transformers Ensembling. Current Oncology, 2022, 29(10), 7498-7511. https://doi.org/10.3390/curroncol29100590

21. Wang, J., Fu, P., Gao, R. X. Machine Vision Intelligence for Product Defect Inspection Based on Deep Learning and Hough Transform. Journal of Manufacturing Systems, 2019, 51, 52-60. https://doi.org/10.1016/j.jmsy.2019.03.002

22. Wang, S., Rao, H., Wen, C., Jia, H., Wu, D., Liu, Q., Abualigah, L. Improved Remora Optimization Algorithm with Mutualistic Strategy for Solving Constrained Engineering Optimization Problems. Processes, 2022, 10(12), 2606. https://doi.org/10.3390/pr10122606

23. Wu, D., Rao, H., Wen, C., Jia, H., Liu, Q., Abualigah, L. Modified Sand Cat Swarm Optimization Algorithm for Solving Constrained Engineering Optimization Problems. Mathematics, 2022, 10(22), 4350. https://doi.org/10.3390/math10224350

24. Yang, L., Zhu, W. Research on the Straight Line Driving and Visual Navigation Technology of Agricultural Machinery. International Core Journal of Engineering, 2020, 6(6), 304-309. https://doi.org/10.6919/ICJE.202006_6(6).0045

25. Yasuda, Y. D. V., Martins, L. E. G., Cappabianco, F. A. M. Autonomous Visual Navigation for Mobile Robots: A Systematic Literature Review. ACM Computing Surveys (CSUR), 2020, 53(1), 1-34. https://doi.org/10.1145/3368961

26. Zhang, X., Zhong, C., Abualigah, L. Foreign Exchange Forecasting and Portfolio Optimization Strategy Based on Hybrid-Molecular Differential Evolution Algorithms. Soft Computing, 2022, 1-19. https://doi.org/10.1007/s00500-022-07526-6

27. Zhao, Y., Gupta, R. K., Onyema, E. M. Robot Visual Navigation Estimation and Target Localization Based on Neural Network, Paladyn. Journal of Behavioral Robotics, 2022, 13(1), 76-83. https://doi.org/10.1515/pjbr-2022-0005

28. Zheng, Y., Jiang, W. Evaluation of Vision Transformers for Traffic Sign Classification. Wireless Communications and Mobile Computing, 2022, 2022(1), 304-311. https://doi.org/10.1155/2022/3041117

29. Zhu, S., Liu, D., Liu, Y., Cui, P. Observability-Based Visual Navigation Using Landmarks Measuring Angle for Pinpoint Landing. Acta Astronautica, 2019, 155, 313-324. https://doi.org/10.1016/j.actaastro.2018.10.011