DOI: 10.12398/j.issn.2096-7217.2024.04.004

Visual navigation in orchard based on multiple images at different shooting angles

MA Zenghong[1, 2, 3], YUE Jiawen[1, 2, 3], YIN Cheng[1, 2, 3], ZHAO Runmao[1, 2, 3], CHANDA Mulongoti[1, 2, 3], DU Xiaoqiang[1, 2, 3*]

(1. *School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou* 310018, *China*; 2. *Key Laboratory of Transplanting Equipment and Technology of Zhejiang Province, Hangzhou* 310018, *China*; 3. *Key Laboratory of Agricultural Equipment for Hilly and Mountainous Areas in Southeastern China (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, Hangzhou* 310018, *China*)

Abstract: The orchards usually have rough terrain, dense tree canopy and weeds. It is hard to use GNSS for autonomous navigation in orchard due to signal occlusion, multipath effect, and radio frequency interference. To achieve autonomous navigation in orchard, a visual navigation method based on multiple images at different shooting angles is proposed in this paper. A dynamic image capturing device is designed for camera installation and multiple images can be shot at different angles. Firstly, the obtained orchard images are classified into sky and soil detection stage. Each image is transformed to HSV space and initially segmented into sky, canopy and soil regions by median filtering and morphological processing. Secondly, the sky and soil regions are extracted by the maximum connected region algorithm, and the region edges are detected and filtered by the Canny operator. Thirdly, the navigation line in the current frame is extracted by fitting the region coordinate points. Then the dynamic weighted filtering algorithm is used to extract the navigation line for the soil and sky detection stage, respectively, and the navigation line for the sky detection stage is mirrored to the soil region. Finally, the Kalman filter algorithm is used to fuse and extract the final navigation path. The test results on 200 images show that the accuracy of visual navigation path fitting is 95.5%, and single frame image processing costs 60 ms, which meets the real-time and robustness requirements of navigation. The visual navigation experiments in *Camellia oleifera* orchard show that when the driving speed is 0.6 m/s, the maximum tracking offset of visual navigation in weed-free and weedy environments is 0.14 m and 0.24 m, respectively, and the *RMSE* is 30 mm and 55 mm, respectively.

Keywords: orchard; visual navigation; multiple shooting angles; region segmentation; Kalman filter

CLC number: S24　　**Documents code:** A　　**Article ID:** 2096-7217 (2024) 04-0051-15

MA Zenghong, YUE Jiawen, YIN Cheng, ZHAO Runmao, CHANDA Mulongoti, DU Xiaoqiang. Visual navigation in orchard based on multiple images at different shooting angles [J]. Journal of Intelligent Agricultural Mechanization, 2024, 5(4): 51−65

马锃宏, 岳佳文, 殷程, 赵润茂, CHANDA Mulongoti, 杜小强. 基于不同拍摄角度多幅图像的果园视觉导航[J]. 智能化农业装备学报(中英文), 2024, 5(4): 51−65

0　Introduction

Tree fruit industry plays an important role in agricultural production, and has become a pillar industry in rural economies. With the continuous expansion of tree fruit planting scale, the contradiction between

Received date: 2023-10-25　　**Revised date:** 2023-12-29

Foundation item: National Key Research and Development Program of China (2022YFD2202103); National Natural Science Foundation of China (31971798); Zhejiang Provincial Key Research & Development Plan (2023C02049、2023C02053); SNJF Science and Technology Collaborative Program of Zhejiang Province (2022SNJF017); Hangzhou Agricultural and Social Development Research Project (202203A03)

First Author: MA Zenghong, PhD, Associate Professor, research interests: agricultural machinery navigation and unmanned driving. E-mail: mzh2018@zstu.edu.cn

* **Correspondence Author:** DU Xiaoqiang, PhD, Professor, research interests: machine vision and agricultural robot, agricultural machinery equipment design. E-mail: xqiangdu@zstu.edu.cn

heavy labor demand and labor shortage has become increasingly prominent. It is urgent to guide and develop mechanized and intelligent production of agricultural equipment[1-2].

Compared with traditional manual operation mode, intelligent agricultural equipment can significantly reduce labor intensity, improve operation efficiency and production quality, and alleviate labor shortage[3-4]. Autonomous navigation is an essential technology for intelligent agricultural equipment. The research on autonomous navigation technology of agricultural robots in orchard environment has gradually become focus and hotspot[5].

GNSS plays an important role in absolute position acquisition and absolute path navigation[6-7], and is mostly applied in field[8]. However, fruit trees are mostly planted in mountainous and hilly areas where the terrain is rough, the tree canopy and the weeds are thick. GNSS can be hardly applied due to the problems such as signal occlusion, multipath effects, and radio frequency interference developed orchard robot navigation methods by using Lidar. But the Lidar sensor has high cost and poor real-time performance due to high algorithm complexity and large computational power[9-10].

Visual sensors are less expensive than radars and commonly used in inter-row navigation, which play a key role in extracting relative navigation path[11]. The image segmentation technology is mainly used in visual navigation to obtain relative path, and the deviation acquisition method is used to calculate the pose to achieve steering control of agricultural robots. OPIYO et al. [12] developed a method to eliminate the sky region using $L \times a \times b$ space, and the image was divided into canopy region and soil region. Then a central line was used to fit the navigation path. Since the fitted navigation path is mainly straight line, the method is not generally applicable.

SHARIFI and CHEN[13] divided the original image into ground and background areas according to the orchard environment. Then the linear detection method was used to extract ground navigation path, which was suitable for the uniformly planted orchards on even ground. If there are dense weeds on the ground, the accuracy of path extraction will deteriorate when only the ground information is used to extract path.

RADCLIFFE et al. [14] considered the tree canopy and sky together, and used the sky region to fit navigation path, which has high accuracy in path extraction. But the method relies on the tree canopy condition. If the canopy is sparse, its robustness is poor. ZHOU et al. [15-17] also used sky region for the navigation path extraction, which has a good extraction effect. However, the influence of ground is not considered on path fitting, which results in poor adaptability to unstructured orchards.

These orchard navigation methods can be applied to specific planting scenarios due to the limitations, which are: (1) low adaptability, poor robustness to dense and non-dense planting conditions; (2) fixed sensing angle that cannot collect enough effective information, especially in mountainous and hilly orchard.

Deep learning and neural network have achievements in complex environment background removal and scene segmentation. HAN et al. [18] used U-Net network to segment soil area in image, and then realized path recognition through boundary points. This method shows the good segmentation effect of semantic segmentation. BI and WANG [19] used Mask R-CNN model to identify roads and trunks, and solved the relative pose of fruit trees according to road geometry model.

YANG et al. [20] established a corn rhizome recognition algorithm based on convolutional neural network to realize inter-row navigation. Although the semantic segmentation and deep learning methods have a good effect on extracting navigation paths, the training period is long and the calibration effect of the training set directly affects the segmentation effect.

In order to achieve the autonomous navigation in orchard, especially in mountainous and hilly terrain, a dynamic image capturing device was designed to trigger image capturing at different capturing angles. The sky and soil regions in each image were accurately segmented by scene segmentation algorithm. The visual navigation path was extracted subsequently by Kalman filter algorithm. A set of orchard visual navigation control system was developed and installed on a crawler-type fruit collector. Experimental validation in a *Camellia oleifera* orchard was conducted

to test the accuracy of the proposed method.

1　Materials and methods

1.1　Orchard navigation path extraction

The orchard navigation path extraction mainly includes four steps: image scene segmentation, sky and soil navigation line extraction, dynamic weighted filtering on the navigation line and fusion of the navigation path based on Kalman filter. The specific algorithm flowchart is shown in Figure 1.

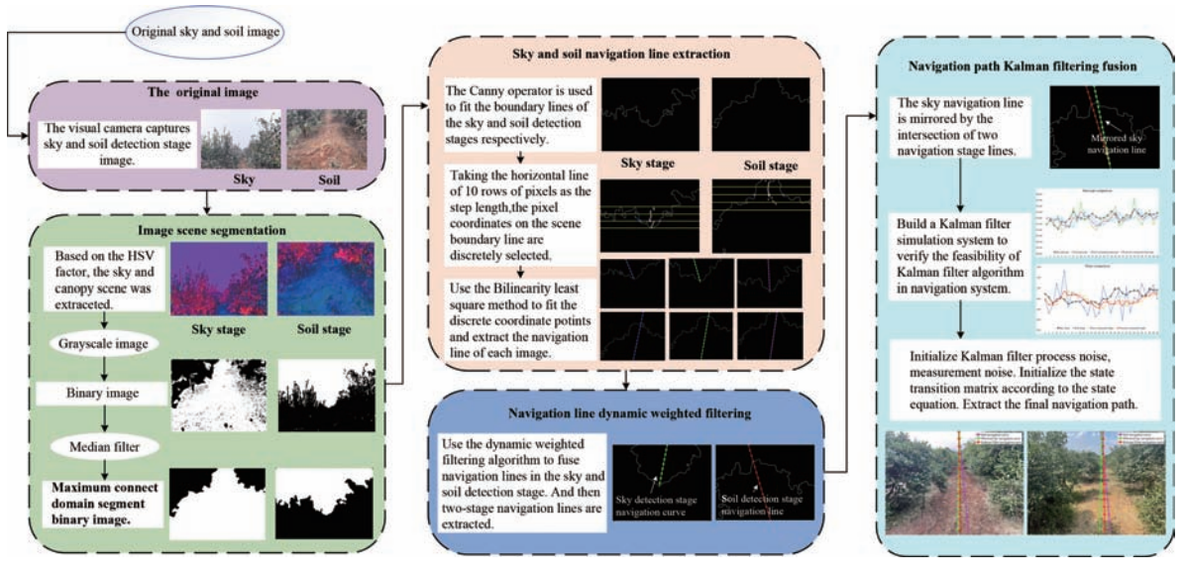


Figure 1　Algorithm flow chart of the orchard navigation path extraction

1.1.1　Orchard image acquisition

The original images of orchard were shot in the *Camellia* oleifera orchard of Dongfanghong Forest Farm, Jinhua City, Zhejiang Province, China. The row spacing is about 4 m, and the plant spacing is about 2 m.

As shown in Figure 2, a dynamic image capturing device was designed and installed 1.3 m above the ground. The device is driven by a stepper motor to pitch up and down from −60° to +60° with the horizon as 0°. A monocular industrial camera (MV-CA013-20GC) is installed on the device and set to trigger mode to capture the images. The main parameters of the camera are shown in Table 1.

By using an angle sensor, the device can collect one image every 15°. The image will not be collected when the camera is horizontal. Therefore, a total of 8 images are collected in a cycle.

As shown in Figure 3, when the pitch angle is positive (the center axis of the camera is above horizontal line), the collected 4 images are classified as the images at sky detection stage, otherwise they are the images at soil detection stage. Finally, a total of 200 inter-row images were collected as the image data set. Figure 4 shows 8 example images in a cycle.

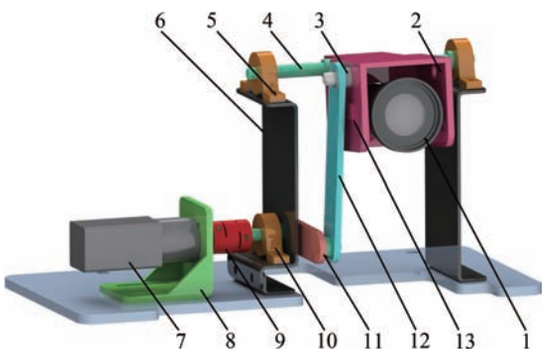


Figure 2　Dynamic image capturing device

1. Monocular industrial camera　2. Camera bracket　3. Rod　4. Shaft　5. Bearing seat　6. Support frame　7. Electric motor　8. Electric motor bracket　9. Coupling　10. Bearing seat　11. Crank　12. Link　13. Angle sensor

Table 1　Technical parameters of MV-CA013-20GC industrial camera

Parameter	Numerical or description
Sensor type	CMOS, global shutter
Pixel size/(μm × μm)	4.8 × 4.8
Target surface size	1/3.6″
Resolution	1 280×1 024
Maximum frame rate/fps	116
Data interface	GigE
Power/VDC	12
Dimensions/(mm×mm×mm)	29×29×42

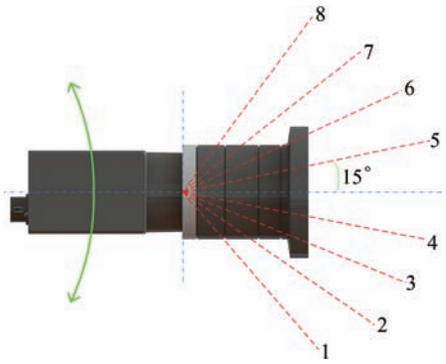
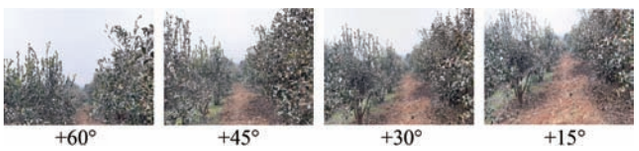
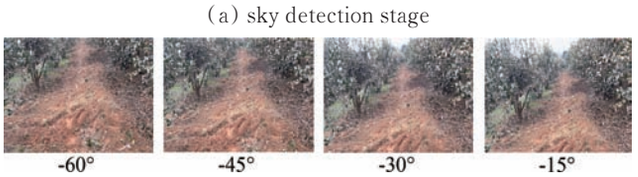


Figure 3 Image acquisition at different pitch angles of the camera



(a) sky detection stage



(b) soil detection stage

Figure 4 Original image

1.1.2 Image scene segmentation

The orchard image usually includes three scenes: sky, canopy and soil.

After analyzing the color characteristics of sky and canopy, HSV color transform is used to preprocess the images. *S* and *V* channels are selected to preliminarily extract canopy scene, with *S* channel thresholds of 0-26, and *V* channel thresholds of 217-255. Similarly, the canopy scene is preliminarily extracted, with *H* channel thresholds of 35-77, *S* channel thresholds of 43-255, and *V* channel thresholds of 46-255.

The preprocessing is performed on 8 images acquired in a cycle. The preprocessing results of the images at sky detection stage and at soil detection stage are shown in Figure 5 and Figure 6, respectively. After preprocessing, the image is smoothed by median filtering and morphological processing to accurately segment the sky and canopy scenes. When the sky and canopy scenes are extracted, the left is the soil scene. Figure 7 shows the example of image scene segmentation results.

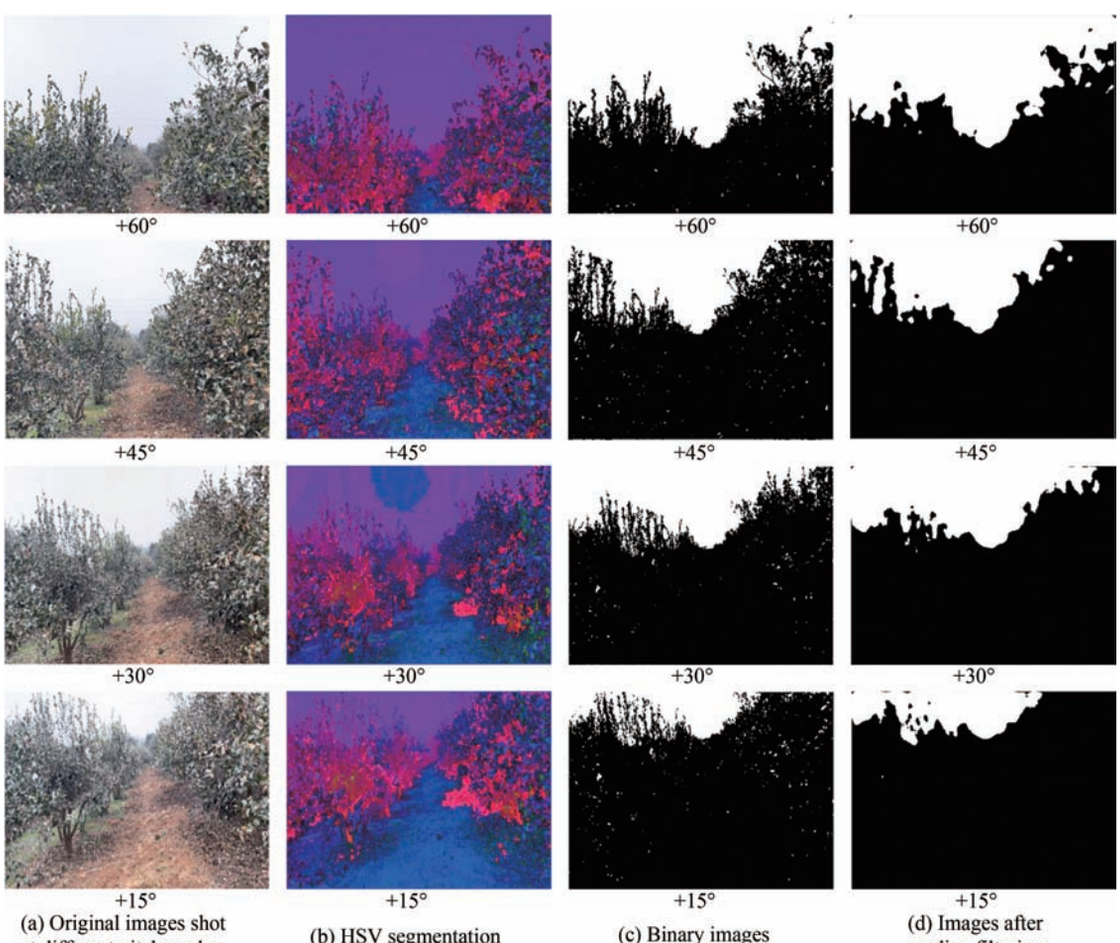


Figure 5 Preliminary scene extraction at sky detection stage

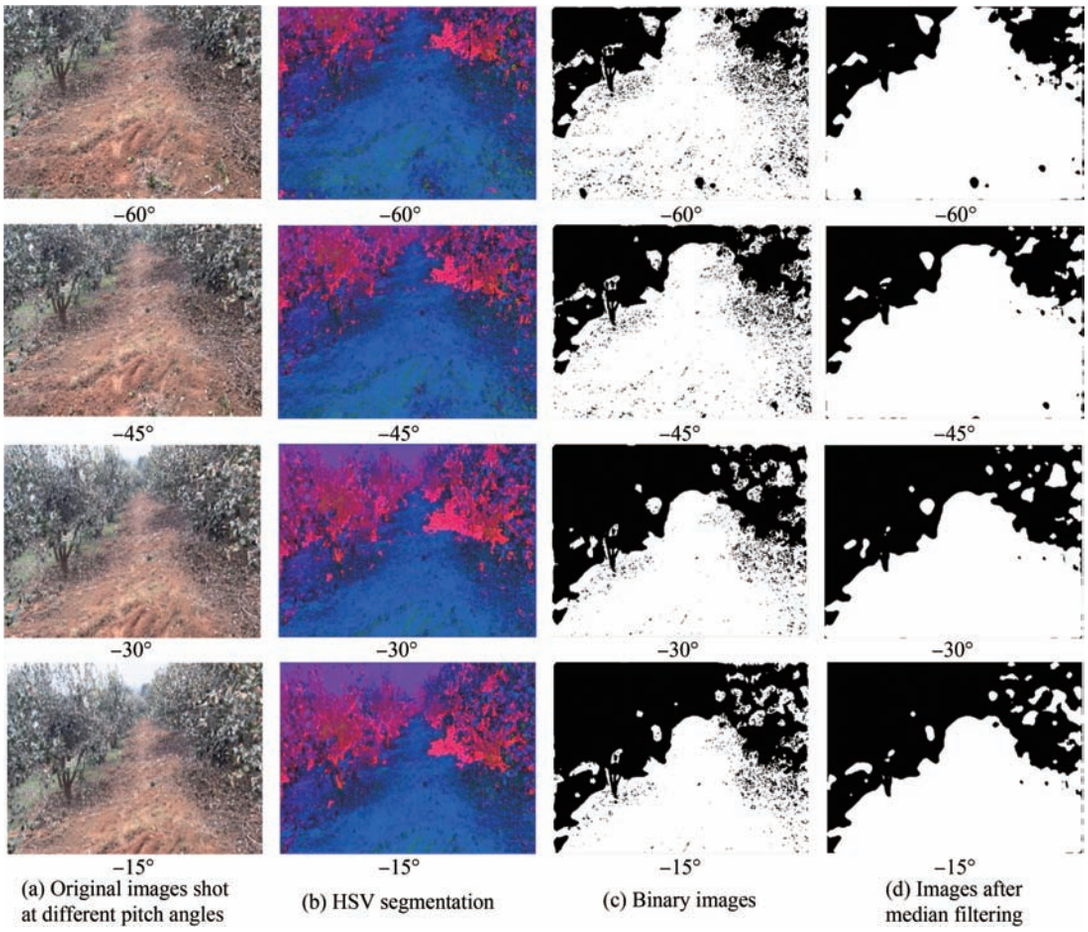


Figure 6　Preliminary image extraction at soil scene detection stage

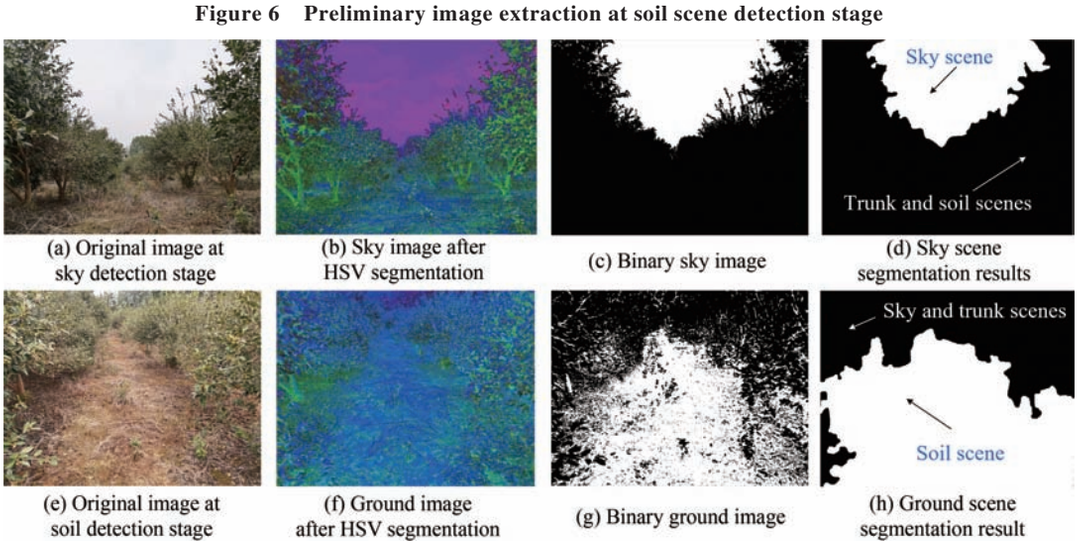


Figure 7　Image scene segmentation

1.1.3　Extraction of navigation line at sky and soil detection stages

After median filtering of binary image, there is still a lot of white noise. In order to avoid increasing the boundary fitting error, the binary image needs to be further processed.

In this study, the maximum connected region algorithm is used to sort the region area in image. The small region is classified as background to reduce impact on navigation line fitting.

The remaining white region is the sky or soil region. Then, the sky or soil region can be accurately extracted and the Canny operator is used to fit the region boundary.

Figure 8 shows the region boundary fitting results of the two images with pitch angles of +60° and −60°.

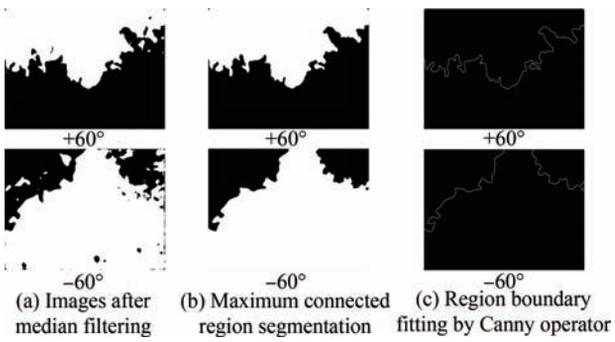


Figure 8 Region boundary fitting of the images with pitch angles of +60° and −60°

Pixels on the region boundary are used to extract the navigation line. However, not all pixels are useful because some pixels without reference value will deteriorate the accuracy of the navigation line. Therefore, it is necessary to filter the region boundary pixels.

Take the image at sky detection stage as the example to explain the filtering process. As shown in Figure 9(a), by setting the top left corner of the image as the origin, the downward direction as the positive X-coordinate and the right direction as the positive Y-coordinate, the coordinates of the region boundary pixels are selected discretely. The point of maximum X-coordinate on the region boundary line in Figure 9(b) is taken as the start point *Ultimate_point* (X_u, Y_u).

If the region boundary intersects with the image low border, the midpoint between the left limit white pixel and the right limit white pixel on the image low border is taken as the start point. Then search upward every 10 rows of pixels, and count the white pixels in the current row.

If the number of white pixels is less than 2, the pixel coordinates of that row are discarded. If the number of white pixels is greater than or equal to 2, the pixel closest to the left side of the start point is selected as the left reference point *Left_point* (X_l, Y_l), and the pixel closest to the right side of the start point is selected as right reference point *Right_point* (X_r, Y_r).

If the $Y_{\text{Left_point}}$ and $Y_{\text{Right_point}}$ are both located on the left or right side of $Y_{\text{Ultimate_point}}$, their coordinates are discarded. The midpoint between the retained *Left_point* and *Right_point* is the central reference point, which is used for fitting and denoted as *Middle_point* (X_m, Y_m). The calculation equation of Y_m is

$$Y_m = (Y_l + Y_r)/2,\ Y_l \leqslant Y_u \leqslant Y_r \quad (1)$$

Finally, the least square method is used to fit all the retained *Middle_point* and extract the navigation line in the image. The searching of *Middle_point* in soil detection stage is opposite to the process in sky detection stage. The selection of *Middle_point* in sky and soil detection stage is shown in Figure 9(b) and Figure 9(d), respectively. The navigation line extraction results of 8 images in one cycle are shown in Figure 10.

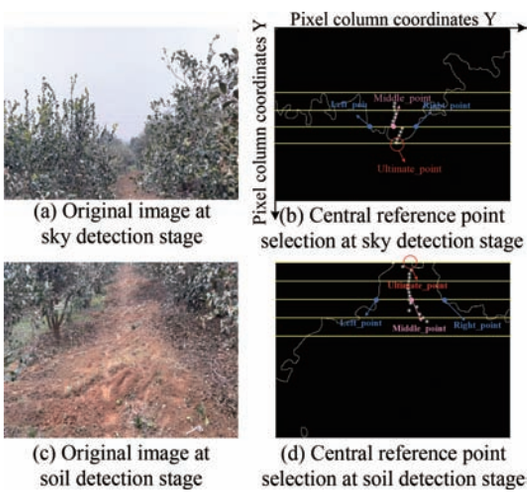


Figure 9 Discretely select the central reference points at sky and soil detection stages

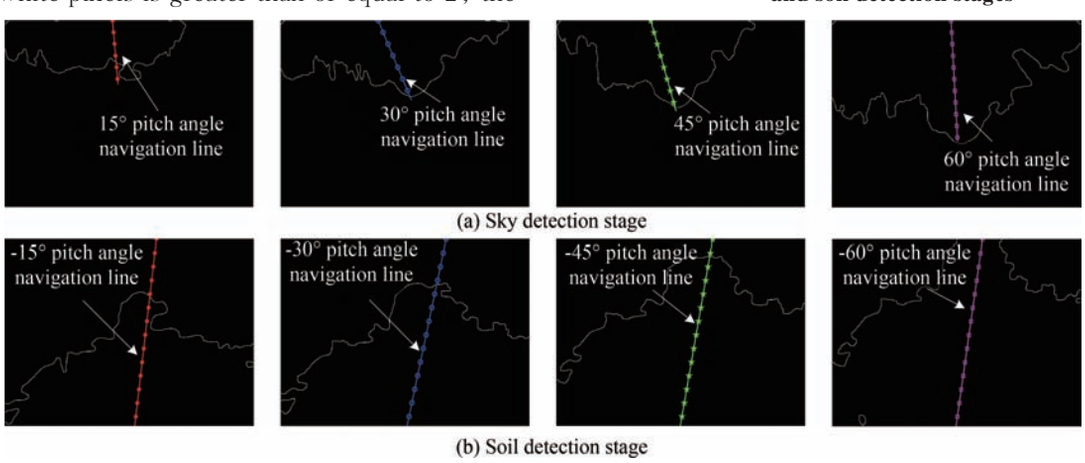


Figure 10 Navigation line extraction results in one cycle

1.1.4　Navigation line dynamic weighted filtering

Since there are four images in each detection stage, it is necessary to filter the navigation line parameters. In this study, the dynamic weighted filtering method is used to construct the deviation confidence in each detection stage. The weights of the four navigation line parameters in each detection stage are assigned based on corresponding deviation confidence. Finally, the slope and intercept of navigation line are weighted filtering, respectively.

Take the 4 images in sky detection stage as example to explain the process of weighted filtering. Generally, the larger the pitch angle is, the wider is the scene and the more accurate path fitting can be obtained. Therefore, the initial weights are defined as 0.1, 0.2, 0.3 and 0.4 for 4 images with pitch angle of 15°, 30°, 45° and 60°, respectively. The filtered navigation line equation of 4 images according to initial weights is shown in Equation (2).

$$\begin{cases} A_0 = \sum_{i=1}^{4} \dfrac{ia_i}{10} \\ B_0 = \sum_{i=1}^{4} \dfrac{ib_i}{10} \\ Y_0 = A_0 X + B_0 \end{cases} \tag{2}$$

where i represents the image index number (i=1, 2, 3, 4), a_i is the slope of the navigation line, b_i is the intercept of the navigation line.

According to parameter difference between the fitted navigation lines and the initial filtered navigation line, the proportion of the difference is calculated as the current filtering weight to be assigned. The weight is calculated in Equation (3).

$$\begin{cases} \eta_i = \dfrac{a_i - A_0}{\sum_{i=1}^{4}(a_i - A_0)} \\ \varepsilon_i = \dfrac{b_i - B_0}{\sum_{i=1}^{4}(b_i - B_0)} \end{cases} \tag{3}$$

where η_i is the slope weight of each navigation line after weight calculation, ε_i is the intercept weight of each navigation line after weight calculation.

The current filtering weight is sorted. The larger the parameter weight is, the smaller confidence the weight has, and the smaller the dynamic filtering weight is. Therefore, if the order of η_i and ε_i (i=1, 2, 3, 4) is m_i and n_i, respectively after sorting, the η_i and ε_i are replaced by Equation(4).

$$\begin{cases} \eta_i = \eta_{4+1-m_i} \\ \varepsilon_i = \varepsilon_{4+1-n_i} \end{cases} \tag{4}$$

The slope and intercept of the filtered navigation line of 4 images after dynamic weighted filtering are

$$\begin{cases} A = \sum_{i=1}^{4} \eta_i a_i \\ B = \sum_{i=1}^{4} \varepsilon_i b_i \end{cases} \tag{5}$$

The filtered navigation line equation of 4 images after dynamic weighted filtering is

$$Y = AX + B \tag{6}$$

1.1.5　Navigation line fusing using Kalman filter

The extracted navigation lines in sky and soil detection stage do not share the same coordinate system. In this study, the navigation line in sky detection stage is mirrored to make its coordinate system consistent with the soil detection stage. The pixel row where two navigation lines at sky and soil detection stage intersect is used as mirror axis.

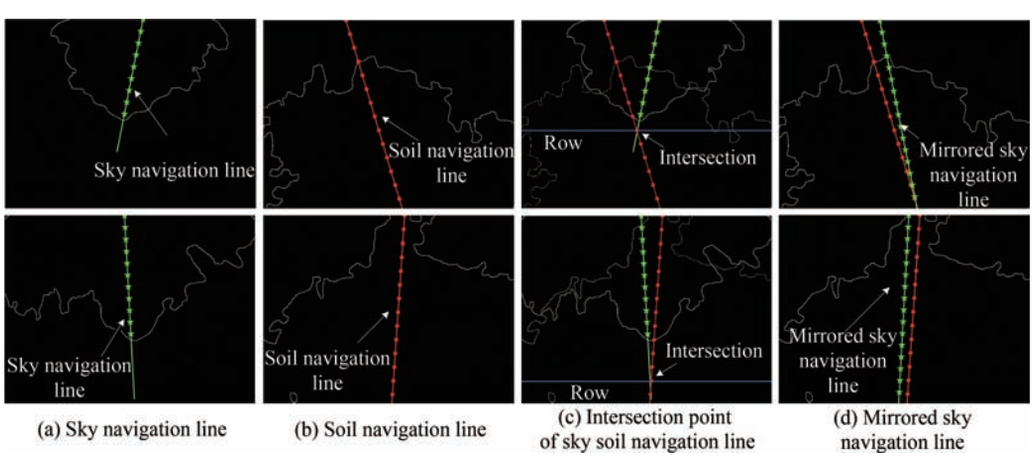


Figure 11　Mirror projection of the navigation line at sky detection stage

The unified navigation lines are shown in Figure 11.

Kalman filter (KF) method is used to fuse the navigation lines after mirror projection to obtain the inter-row navigation path[21-23]. KF is suitable for linear Gaussian system which has two variables, state variable and observation variable.

The observation variable is the actual value measured by the sensor, while the state variable is the theoretical value determined by the state equation of the navigation system. Compared with the soil detection stage, the navigation line obtained from the sky detection stage has more accurate reference significance because of no other disturbance such as weeds or uneven terrain on the ground[24-26].

Therefore, the slope and intercept of the mirrored sky navigation line are used as the state variables of KF. The slope and intercept of the soil navigation line are used as observation variables.

As a result, the parameters of the mirrored sky navigation line are used to correct the soil navigation line′s parameters. Therefore, the linear parameters at two stages are fused to obtain more accurate inter-row navigation path.

The state equation of KF in this study is shown in Equation (7) and Equation (8). The state variables of whole system follow the normal distribution of $P(\omega)\sim N(0,Q)$ and the observation variables follow the normal distribution of $P(\upsilon)\sim N(0,R)$.

$$X_t = AX_{t-1} + BU_{t-1} + \omega_{t-1} \tag{7}$$

$$Z_t = HX_t + \upsilon_t \tag{8}$$

where X_t is the state variable of the mirrored sky navigation line at time t, A is the state transition matrix, U_{t-1} is the control variable at time t, B is the control matrix, ω_{t-1} is the process noise, Z_t is the soil navigation line parameters at time t, H is the measure system parameter, υ_t is the measure noise.

In order to obtain the variables needed in state equation, it is necessary to analyse the parameter changes of the navigation line between time k and k-1, which are slope and intercept changes during Δt. Because the state observation has observation noise, the variables of the state equation are

$$X_t = X_{\text{observation}} = X_{\text{sky}} \tag{9}$$

$$Z_t = Z_{\text{measurement}} = Z_{\text{soil}} = HX_{\text{measurement}} \tag{10}$$

$$\begin{bmatrix} K_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} \triangle t \\ \triangle t \end{bmatrix} \begin{bmatrix} \triangle k \\ \triangle b \end{bmatrix} + \begin{bmatrix} \omega_{1,t-1} \\ \omega_{2,t-1} \end{bmatrix} \tag{11}$$

$$\begin{bmatrix} K_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \upsilon_{2,t} \end{bmatrix} \tag{12}$$

$$Z_t = Z_{\text{measurement}} = Z_{\text{soil}} = HX_{\text{measurement}} \tag{13}$$

$$\begin{bmatrix} K_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} \triangle t \\ \triangle t \end{bmatrix} \begin{bmatrix} \triangle k \\ \triangle b \end{bmatrix} + \begin{bmatrix} \omega_{1,t-1} \\ \omega_{2,t-1} \end{bmatrix} \tag{14}$$

$$\begin{bmatrix} K_2 \\ b_2 \end{bmatrix} = \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \upsilon_{2,t} \end{bmatrix} \tag{15}$$

where X_{sky} is the state variable of the sky navigation line parameter at time t; Z_{soil} is the soil navigation line parameters; K_1 and b_1 are the slope and intercept of the sky navigation line at time t, i. e., state variables; Δk and Δb are the slope and intercept changes at time t; $\omega_{1,t-1}$ and $\omega_{2,t-1}$ are the process noise of the slope and intercept; K_2 and b_2 are the slope and intercept of the soil navigation line at time t, i. e., observation variables; $\upsilon_{1,t}$ and $\upsilon_{2,t}$ are the measure noise of the slope and intercept.

The priori estimation and priori estimation of covariance equations of KF are

$$\hat{X}_t^- = F\hat{X}_{t-1} \tag{16}$$

$$P_t^- = FP_{t-1}F^T + Q \tag{17}$$

where $\hat{X}_t^-$ is the priori estimation of KF at time t, P_t^- is the priori estimated covariance of KF at time t, F is the state transition matrix, Q is the process noise variance. The Kalman gain at time t relative to time t-1 can be calculated according to the two parameters. Then the posteriori estimation at current time can be calculated. Finally, the posteriori estimation covariance is updated by the Kalman gain to achieve convergence in time dimension. The calculation process is

$$T_t = \frac{P_t^- H^T}{(HP_t^- H^T + R)} \tag{18}$$

$$\hat{X}_t = \hat{X}_t^- + T_t(Z_t - H\hat{X}_t^-) \tag{19}$$

$$P_t = \frac{I - T_t H}{P_t} \tag{20}$$

where T_t is the updated Kalman gain at time t; $\hat{X}_t$ is the posteriori state estimation, i. e., the optimal state at current time; I is the unit matrix; P_t is the posteriori estimation of covariance.

1.2　Orchard navigation system

As shown in Figure 12, a crawler-type fruit collector was used as the navigation control object. The dynamic image capturing device and the camera was installed in front of the collector, 1.3 m above the ground. The camera can capture and transmit the images to an upper computer controller. The upper computer controller is a Dell high-performance laptop (Win10 operation system, Intel i7 8700HQ processor, 16 GB DDR4 memory and 128 GB solid state drive) for information processing calculation, and realizes human-computer interaction through display. An STM32 lower computer mainly interacts with the controller, executes the upper computer control commands and analyses the angle information, and finally controls the steering solenoid valve to achieve navigation of the collector. The upper computer controller and the collector communicate through lower computer using RS232 serial port protocol. The communication mainly includes starting and stopping navigation, sending steering command, triggering image acquisition command and acquiring the status of the collector such as speed. The dynamic image capturing device is connected to the STM32 lower computer controlled by the upper computer. A USB serial port is used to communicate with the STM32 lower computer, and a network port is used to communicate with the industrial camera. The controller calculates the navigation deviation information by analysing and processing the image data obtained by the camera in real time, and uses the preview point tracking algorithm to make the fruit collector move along the central line of the row. The overall hardware framework of the navigation control system is shown in Figure 13.

A navigation control software with graphical user interface was developed by C++ and OpenCV based on QT platform. The graphical user interface is shown in Figure 14, which mainly includes image display, parameter setting, information display, serial communication, visual navigation and other functional modules.

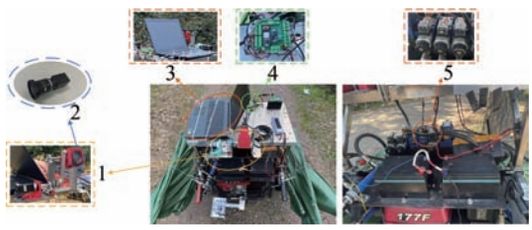


Figure 12　Crawler-type fruit collector

1. Dynamic image capturing device　2. Industrial camera　3. Upper computer controller　4. STM32 lower computer　5. Steering solenoid valves

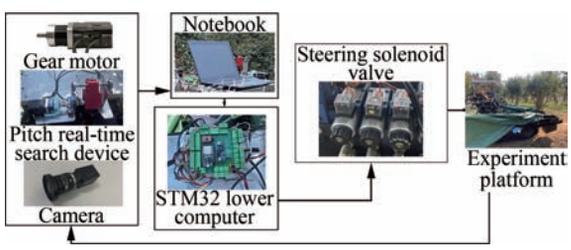


Figure 13　Overall framework of the navigation hardware system



Figure 14　Interface of the orchard navigation control software

2　Results and Discussion

2.1　Experiments of orchard navigation path extraction algorithm

2.1.1　Accuracy comparison of navigation line parameters after KF fusion

In order to verify the feasibility of KF algorithm, a KF fusion simulation system of sky and soil navigation lines is built. 30 original parameters of sky and soil navigation lines are used as state and observation variables of the simulation system. The parameters at time 0 are initialized as $P(\omega) \sim N(0, 16)$ and $P(\upsilon) \sim N(0, 36)$. The state noise variance is 16, the measure variance is 36 and unit matrix is 1. The initial

navigation line coincides with the image symmetrical line. So the posterior estimation slope is 0, the posterior estimation intercept is abscissa of the middle point of the image boundary. The image size is 640 pixel×480 pixel, then the posterior estimation intercept is 320. The posterior estimation parameters after KF fusion are compared with the original parameters of sky and soil navigation lines, as shown in Figure 15 and Figure 16.

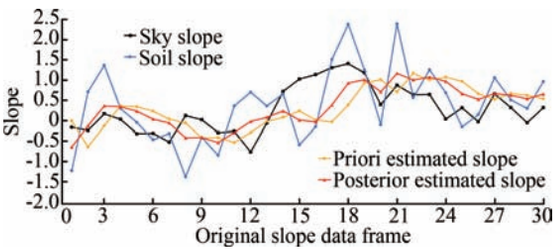


Figure 15 Comparison results of posteriori estimation slope and original slope parameters

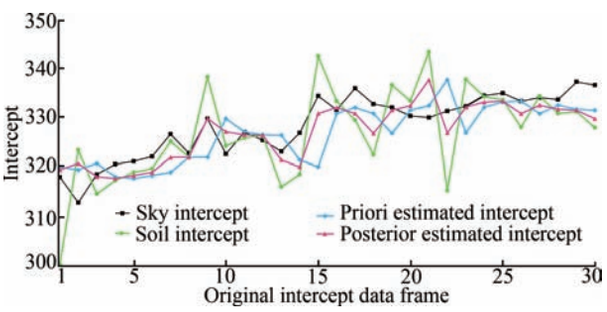


Figure 16 Comparison results of posteriori estimation intercept and original intercept parameters

The sky slope fluctuation range is between −1 and 1.5, the soil slope fluctuation range is between −1.5 and 2.5, and the priori estimated slope and the posterior estimated slope fluctuation range is between −1 and 1.25; the sky intercept fluctuation range is between 310 and 340, the soil intercept fluctuation range is between 310 and 345, and the priori estimated intercept and the posterior estimated intercept fluctuation range is between 315 and 340.

The comparison results show that the navigation line parameters obtained after KF fusion tend to be stable over time. The observation noise caused by system oscillation is reduced. Therefore, the accuracy of the obtained navigation line parameters can be effectively improved by using KF algorithm to fuse sky and soil navigation lines, which lay a foundation for improvement of subsequent navigation accuracy.

2.1.2 Impacts of environment light on KF fusion algorithm

By using the KF algorithm to fuse navigation lines at sky and soil detection stages, the final navigation path of orchard in four seasons is shown in Figure 17.

The path is obtained by fusing 8 navigation lines in two stages. The red line is the final navigation path, the green line is the mirrored sky navigation line and the purple line is the soil navigation line. It shows that the KF fusion algorithm can be used to extract the navigation path of different seasons, which verifies the robustness of the algorithm.

In this paper, the fitting of sky and soil navigation lines in different weather is compared to observe the robustness of the algorithm as well. As shown in Figure 18, the navigation line is extracted in sunny and cloudy weather. It can be seen from Figure 18 that in both sunny and cloudy weather, the fitting accuracy of sky and soil navigation line is high.

2.1.3 Accuracy of the navigation path extraction algorithm

In order to evaluate the reliability of the navigation path extraction algorithm, the inter-row path is determined manually by observing the orchard images, as shown in Figure 19(a), and compared with the extracted navigation line. 30-pixel offset from the inter-row path to left and right is set as the judgement range, as shown in Figure 19(b).

If the extracted navigation line is located within that range, the extracted path is correct, otherwise it is incorrect. Table 2 shows the accuracy of the navigation path extraction algorithm for 200 images. The accuracy rate is 95.5%, and the average processing time of each image is 60 ms.

2.2 Orchard navigation experiments

After the orchard navigation path is extracted, the orchard navigation experiments on the crawler-type fruit collector were conducted by using the preview point tracking algorithm to control the path tracking of the fruit collector. The preview point tracking algorithm has been deeply studied by us and successfully applied to the rice row tracking control of a crawler tractor[24].

The orchard navigation experiments were conducted at Dongfanghong Forest Farm, Jinhua City, Zhejiang Province, China. The test site is shown in Figure 20(a). The ground has small pits and slopes that

affect the running of the collector. In order to reduce the impact of weeds on visual recognition, weeds were removed. The rows after weed removal are shown in Figure 20(b).

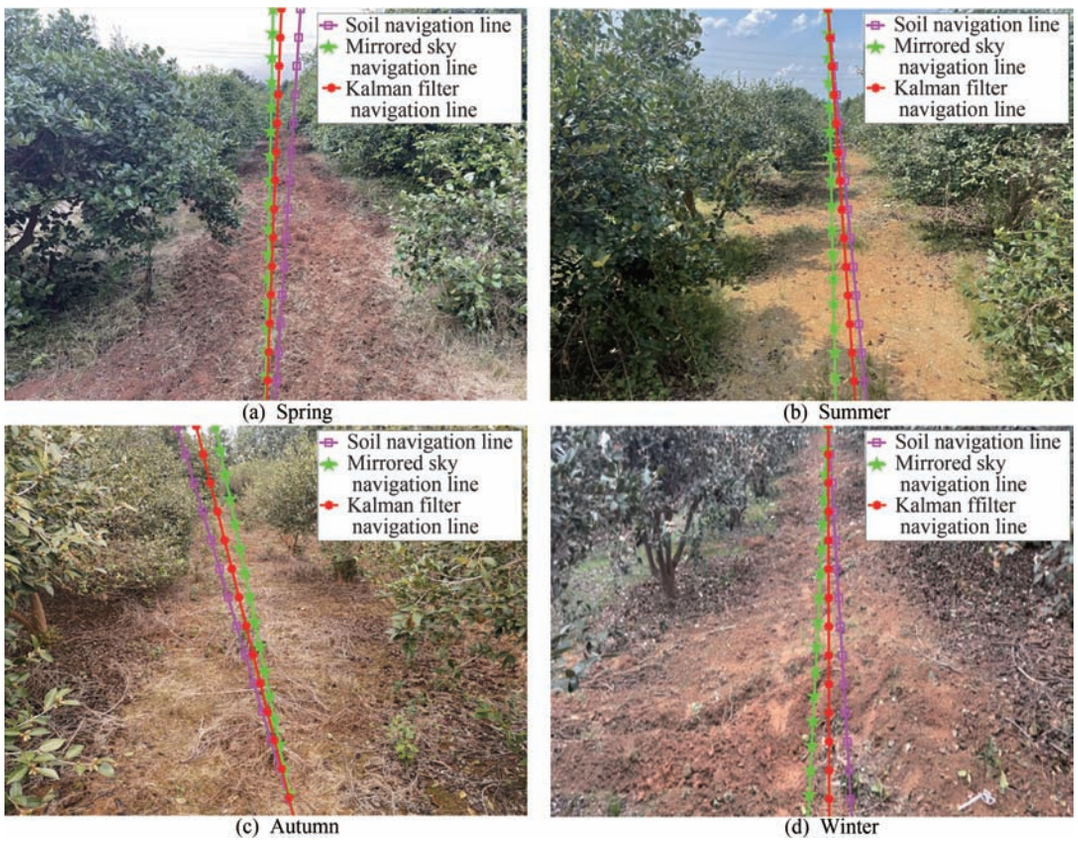


Figure 17　Navigation path extraction after KF fusion

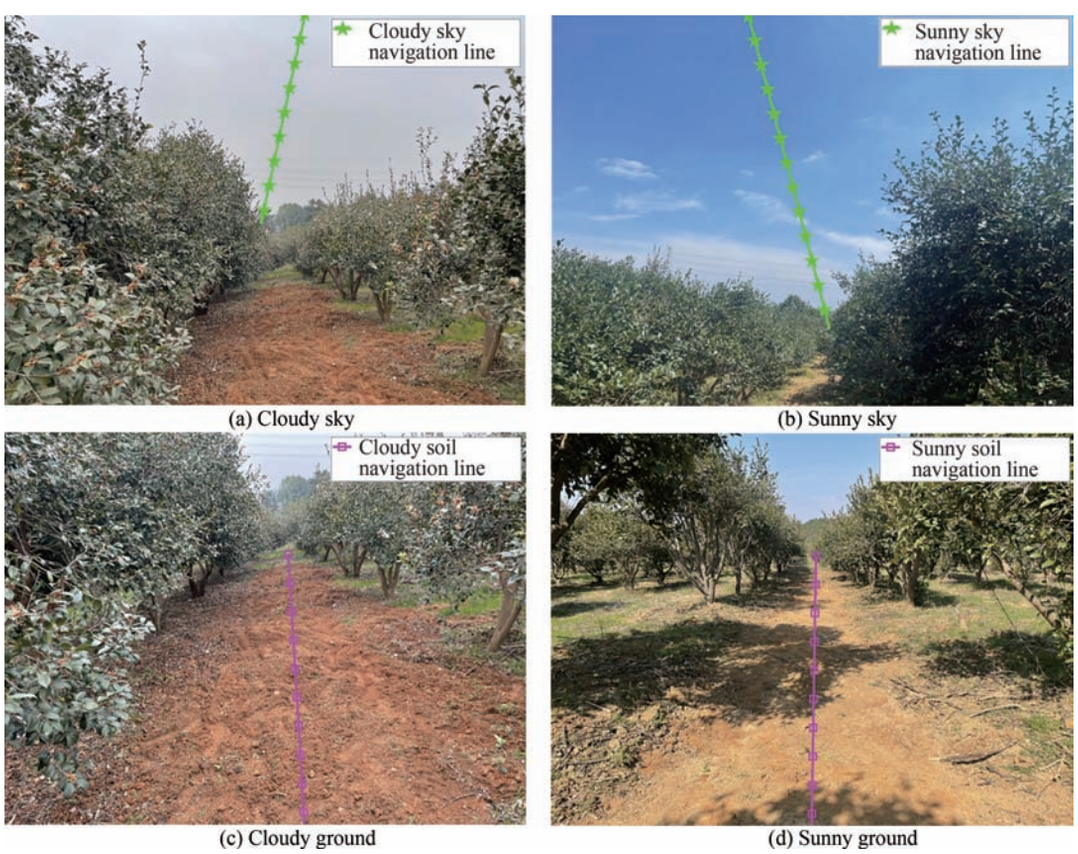


Figure 18　Navigation line fitting in different weather

The collector runs at 0.6 m/s. A container filled with white sand is installed on the collector to show the actual navigation path, as shown in Figure 21(a). Therefore, the actual navigation path can be observed directly as the white sand line in Figure 21(b). The manually determined inter-row path based on the position of tree on both rows is shown as the yellow line in Figure 21(b). The navigation accuracy is recorded by evenly measuring the lateral distance between two lines.

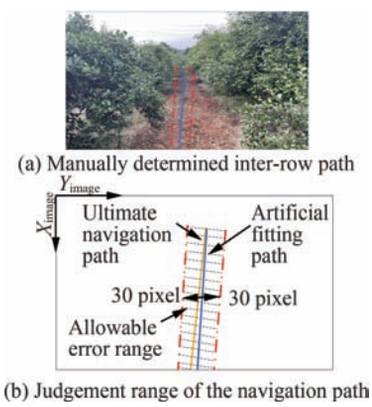


Figure 19 Reliability test of the navigation path extraction algorithm

Table 2 Results of orchard navigation path extraction

Image group	Total number of navigation path	Number of incorrect navigation path	Accuracy /%
Orchard original image	200	9	95.5



(a) Initial test site (b) Test site after weed removal

Figure 20 Navigation test scenario

The total effective path length of inter-row navigation is about 25 m. The accuracy statistics are carried out at an interval of about 5 cm. There are 440 sets of data, and the overall change of inter-row navigation accuracy is shown in Figure 22.

As shown in Figure 22, the overall inter-row navigation accuracy fluctuates around 0.08 m, and the maximum lateral offset is no more than 0.14 m. Because there is a distance between the position of container and camera, the path recorded by container has delay when the collector steers during navigation. This situation also has some effective influences on the accuracy of navigation.



(a) Installation of a sand container



(b) White sand line showing the actual navigation path

Figure 21 Actual navigation path in the orchard

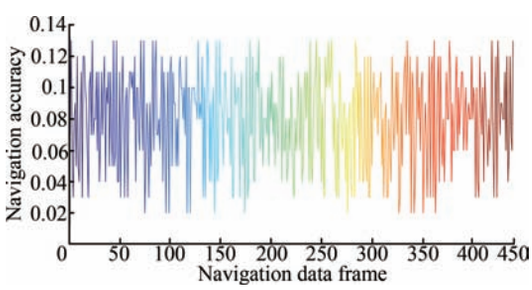


Figure 22 Inter-row navigation data graph

The variation of inter-row navigation accuracy is presented by the box plot 20 sets of data are included into a data block to draw box plot. There are 22 data blocks in total. The navigation accuracy box plot is shown in Figure 23.

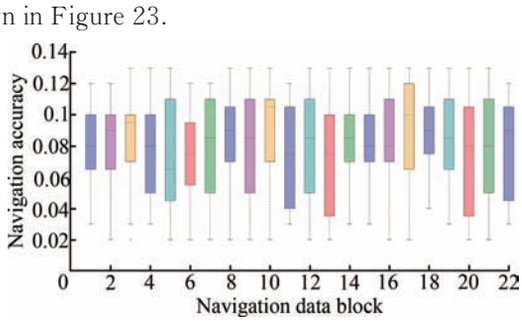


Figure 23 Navigation accuracy box plot

As shown in Figure 23, each block of box plot represents 80% of the data set. The upper and lower poles are represented by dash lines. The median line of plot block is about 0.08 m. The root mean square error (*RMSE*) of the inter-row navigation accuracy is 30 mm. The variation range is stable and reliable with respect to the entire inter-row navigation system. The navigation information acquisition algorithm can obtain

the deviation information correctly. The final test data also verifies that the inter-row navigation information acquisition method in this study fully meets the practical requirements.

The validation experiments are also conducted in the orchard with weed on the ground to verify the robustness of the inter-row navigation system. The experiment process is consistent with the operation process in weed free environment. The inter-row navigation is started immediately after the system is stabilized. The total length of effective path is about 20 m, and the accuracy statistics is implemented at an interval of about 5 cm. The overall change box plot of inter-row navigation accuracy in weedy environment is shown in Figure 24.

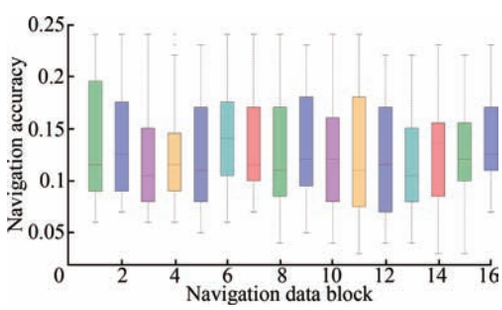


Figure 24　Navigation accuracy box plot in weedy environment

As shown in Figure 24, the weedy environment brings challenges to the algorithm. The median navigation accuracy of plot block is about 0.13 m, and the maximum deviation is no more than 0.24 m. The variation range of navigation accuracy in this environment is statistically analyzed. The *RMSE* of inter-row navigation accuracy in weedy environment is 55 mm. It indicates that in weedy environment, the collector can track along the inter-row navigation path stably.

3　Conclusions

In order to realize autonomous operations in orchard, a visual navigation method based on multiple images at different shooting angles is proposed in this paper, including real-time acquisition method of orchard image under multiple pitch angles, extraction of navigation line at sky and soil detection stages, navigation line dynamic weighted filtering and navigation line fusing using Kalman filter. A dynamic image capturing device is designed for camera installation and multiple images can be shot at different angles. On this basis, the orchard navigation control system of a crawler-type fruit collector was developed, and the navigation accuracy comparison test and analysis were carried out in weed-free environment and weedy orchard.

1) According to the terrain characteristics of mountainous and hilly orchards, a real-time searching device with pitch motion was designed to collect inter-row images of orchards from multiple angles, and the distribution characteristics of images in different colour spaces were analysed. The best colour space threshold combination was selected to distinguish the soil, canopy and sky scene information. The maximum connected region algorithm was used to accurately extract the soil and sky scene areas. Because there are multiple frames of images in soil and sky detection stage, a dynamic weighted filter was proposed to obtain the navigation line. The mirror projection was carried out based on intersection of two-stage navigation line to realize the unification of coordinate system. Finally, the Kalman filter algorithm was used to obtain final navigation path, which meets the real-time requirements of orchard navigation and provides technical support for the achievement of precision agriculture.

2) The hardware system of orchard navigation control was installed on the crawler fruit collector, and the corresponding navigation control software system was developed. The navigation test results show that when the collector travelled at a speed of 0.6 m/s in weed-free environment, the maximum inter-row navigation accuracy did not exceed 0.14 m, and the root mean square error *RMSE* was 30 mm. The inter-row navigation was stable and reliable during the test. The navigation accuracy of collector in weedy environment is no more than 0.25 m, and the root mean square error *RMSE* is 55 mm. The orchard experiments verify the robustness of the navigation system. Therefore, the orchard navigation algorithm described in this paper improves the navigation accuracy and robustness of agricultural robots in orchard environment, and can fully achieve the autonomous operation of agricultural robots in orchard.

References

[1] VERBIEST R, RUYSEN K, VANWALLEGHEM T, et al. Automation and robotics in the cultivation of pome fruit: Where do we stand today? [J]. Journal of Field Robotics, 2020.

[2] MARINOUDI V, CLAUS G. SØRENSEN, PEARSON S, et al. Robotics and labour in agriculture. A context consideration [J]. Biosystems Engineering, 2019, 184: 111－121.

[3] FUMIOMI T, WEI Y, CHANGYING L, et al. Applying new technologies to transform blueberry harvesting [J]. Agronomy, 2017, 7(2): 33.

[4] NAVAS E, ROEMI FERNÁNDEZ, DELIA SEPÚLVEDA, et al. Soft grippers for automatic crop harvesting: A review [J]. Sensors, 2021, 21(8): 2689.

[5] YANG Q, DU X, WANG Z, et al. A review of core agricultural robot technologies for crop productions [J]. Computers and Electronics in Agriculture, 2023, 206: 107701.

[6] LECHNER W, BAUMANN S. Global navigation satellite systems [J]. Computers and Electronics in Agriculture, 2000, 25: 67－85.

[7] GAN-MOR S, CLARK R L, UPCHURCH B L. Implement lateral position accuracy under RTK-GPS tractor guidance [J]. Computers and Electronics in Agriculture, 2007, 59: 31－38.

[8] LIAO J, WANG Y, YIN J, et al. An integrated navigation method based on an adaptive federal Kalman filter for a rice transplanter [J]. Transactions of the ASABE, 2021, 64 (2): 389－399.

[9] HELLSTRÖM T. Autonomous navigation for forest machines [J]. Pre-Study Report, Department of Computing Science, Umea University, Sweden, 2002.

[10] LI M, IMOU K, WAKABAYASHI K, et al. Review of research on agricultural vehicle autonomous guidance [J]. International Journal of Agricultural and Biological Engineering, 2009, 2(3): 1－16.

[11] BENGOCHEA-GUEVARA J M, CONESA-MUÑOZ J, ANDÚJAR D, et al. Merge fuzzy visual servoing and GPS-based planning to obtain a proper navigation behavior for a small crop-inspection robot [J]. Sensors, 2016, 16 (3): 276.

[12] OPIYO S, OKINDA C, ZHOU J, et al. Medial axis-based machine-vision system for orchard robot navigation [J]. Computers and Electronics in Agriculture, 2021, 185: 106153.

[13] SHARIFI M, CHEN X Q. A novel vision based row guidance approach for navigation of agricultural mobile robots in orchards [C]// 2015 6th International Conference on Automation, Robotics and Applications (ICARA). IEEE, 2015: 251－255.

[14] RADCLIFFE J, COX J, BULANON D M. Machine vision for orchard navigation [J]. Computers in Industry, 2018, 98: 165－171.

[15] ZHOU M, XIA J, YANG F, et al. Design and experiment of visual navigated UGV for orchard based on Hough matrix and RANSAC [J]. International Journal of Agricultural and Biological Engineering, 2021, 14 (6): 176－184.

[16] MALAVAZI F B P, GUYONNEAU R, FASQUEL J B, et al. LiDAR-only based navigation algorithm for an autonomous agricultural robot [J]. Computers and Electronics in Agriculture, 2018, 154: 71－79.

[17] BLOK P M, SUH H K, VAN BOHEEMEN K, et al. Autonomous in-row navigation of an orchard robot with a 2D LIDAR scanner and particle filter with a laser-beam model [J]. Journal of Institute of Control, Robotics and Systems, 2018, 24(8): 726－735.

[18] HAN Z H, LI J, YUAN Y W, et al. Path recognition of orchard visual navigation based on U-Net [J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52 (1): 30－39.

[19] BI S, WANG Y H. Inter-line pose estimation and fruit tree location method for orchard robot [J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52(8): 1－26, 39.

[20] YANG Y, ZHANG Y L, MIAO W, et al. Accurate identification and location of corn rhizome based on faster R-CNN [J]. Transactions of the Chinese Society of Agricultural Machinery, 2018, 49(10): 46－53.

[21] BLOK P M, VAN BOHEEMEN K, VAN EVERT F K, et al. Robot navigation in orchards with localization based on Particle filter and Kalman filter [J]. Computers and Electronics in Agriculture, 2019, 157: 261－269.

[22] BERGERMAN M, MAETA S M, ZHANG J, et al. Robot farmers: Autonomous orchard vehicles help tree fruit production [J]. IEEE Robotics & Automation Magazine, 2015, 22(1): 54－63.

[23] SHALAL N, LOW T, MCCARTHY C, et al. Orchard mapping and mobile robot localisation using on-board camera and laser scanner data fusion-Part B: Mapping and localisation [J]. Computers and Electronics in Agriculture, 2015, 119: 267－278.

[24] MA Z, YIN C, DU X, et al. Rice row tracking control of crawler tractor based on the satellite and visual integrated navigation [J]. Computers and Electronics in Agriculture, 2022, 197: 106935.

[25] HAN Leng, HE Xiongkui, WANG Changling, et al. Key technologies and equipment for smart orchard construction and prospects [J]. Smart Agriculture, 2022, 4(3): 1−11.

[26] LIU Limin, HE Xiongkui, LIU Weihong, et al. Autonomous navigation and automatic target spraying robot for orchards [J]. Smart Agriculture, 2022, 4(3): 63−74.

基于不同拍摄角度多幅图像的果园视觉导航

马锃宏[1, 2, 3]，岳佳文[1, 2, 3]，殷程[1, 2, 3]，赵润茂[1, 2, 3]，

CHANDA Mulongoti[1, 2, 3]，杜小强[1, 2, 3*]

（1. 浙江理工大学机械工程学院，浙江杭州，310018；2. 浙江省种植装备技术重点实验室，浙江杭州，310018；3. 农业农村部东南丘陵山区农业装备重点实验室（部省共建），浙江杭州，310018）

摘要：果园通常地形崎岖，树冠茂密，杂草丛生。由于信号遮挡、多径效应和射频干扰，GNSS难以在果园中进行自主导航。为了实现果园的自主导航，本文提出了一种基于不同拍摄角度下多幅图像的视觉导航方法。一种动态图像捕捉设备是为安装相机而设计的，可以在不同角度拍摄多个图像。首先，将获得的果园图像分为天空和土壤检测阶段。每个图像被转换到HSV空间，并通过中值滤波和形态学处理初步分割为天空、树冠和土壤区域。其次，利用最大连通区域算法提取天空和土壤区域，并利用Canny算子对区域边缘进行检测和滤波。第三，通过对区域坐标点的拟合，提取当前帧数中的导航线。然后使用动态加权滤波算法分别提取土壤和天空检测阶段的导航线，并将天空检测阶段导航线镜像到土壤区域。最后，利用卡尔曼滤波算法对最终导航路径进行融合提取。对200幅图像的测试结果表明，视觉导航路径拟合的准确率为95.5%，单帧图像处理耗时60 ms，满足导航的实时性和鲁棒性要求。在油茶园进行的视觉导航实验表明，当行驶速度为0.6 m/s时，无杂草和杂草环境下视觉导航的最大跟踪偏移量分别为0.14 m和0.24 m，均方根误差分别为30 mm和55 mm。

关键词：果园；视觉导航；多拍摄角度；区域分割；卡尔曼滤波器