Detection method for the guidance directrix of daylily in the field environment

Junhui Feng[1,2*], Zilin Zhang[1], Daoqi Li[1], Gang'ao Li[1], Zongbao Huang[3], Xiuhan He[1], Zhiwei Li[3*]

(1. *College of Agricultural Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China*;
2. *Shanxi Key Laboratory of Key Technologies and Equipment for Agricultural Machinery, Taigu 030801, Shanxi, China*;
3. *College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China*)

Abstract: The daylily (*Hemerocallis citrina* Baroni) is an herbaceous perennial whose flowers are rich in nutritional and functional components. It is typically cultivated in rows but harvested manually, a labor-intensive process that this study aims to automate by developing a robotic harvesting system. A critical component of such a system is autonomous navigation, which poses significant challenges in unstructured field environments due to changing natural light, randomly distributed weeds, and varying inter-row density. To address these challenges, this study adopted vision-based navigation technology and proposed a guidance directrix detection algorithm. The proposed approach begins with converting the color model from RGB to HSV to decouple the brightness, thereby mitigating the impact of natural light variations. Morphological dilation is applied to suppress noise from weeds in the inter-row regions. Furthermore, an innovative coarse segmentation strategy based on hue and saturation is introduced to handle the problem of different sparsity in the inter-row. Finally, the inter-row axis is accurately extracted by employing concepts from physics, namely, the center of mass and moment of inertia. Experimental results demonstrate an accuracy of 98.25%, a recall of 100%, an average navigation processing time of 8.7521 ms, and a compact model size of only 18 KB. These findings empirically confirm that the proposed approach achieves high precision and real-time performance under unstructured, uncertain, and dynamically changing field conditions. Additionally, the algorithm operates with high computational efficiency and requires neither expensive hardware nor large-scale training datasets.

Keywords: visual navigation, robotic daylily harvester, guidance directrix, inter-row axis
DOI: 10.25165/j.ijabe.20261902.9706

Citation: Feng J H, Zhang Z L, Li D Q, Li G A, Huang Z B, He X H, et al. Detection method for the guidance directrix of daylily in the field environment. Int J Agric & Biol Eng, 2026; 19(2): 262–271.

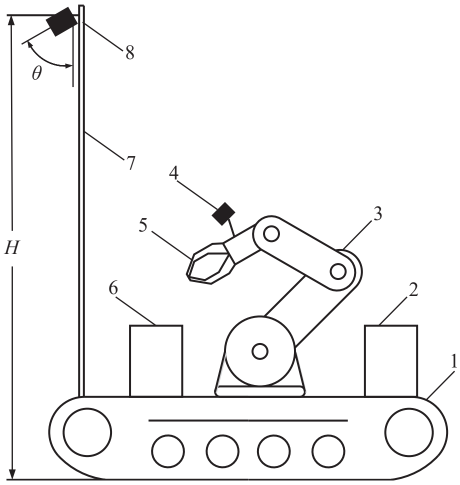
1 Introduction

The daylily (*Hemerocallis citrina* Baroni) is rich in nutrition the human body needs and is a good accompaniment to meals. It is extensively cultivated in China, with smaller-scale cultivation occurring in countries such as Malaysia, Japan, Indonesia, and Madagascar. Harvesting daylilies is highly labor-intensive. The harvest season spans approximately 40-50 d, during which peak harvesting periods require 10-12 h of work daily. To alleviate this labor burden, the development of a robotic daylily harvester was proposed by the team of authors of this study, the schematic of which is shown in Figure 1. The robotic system is designed to identify mature daylilies and locate picking points using a detection camera, grasp the flower with a harvesting manipulator and end-effector, and guide a motion vehicle with a navigation camera. To enable autonomous navigation, the image captured by the navigation camera was processed to generate a guidance directrix (or course). This directrix, along with other tractor state parameters such as velocity and orientation, was used by the navigation planner to produce steering commands. This study focused on the extraction of the guidance directrix to facilitate autonomous navigation for the robotic daylily harvester.

As the daylily is cultivated in rows, it is effective to explore a row-recognition system, which will allow the robotic daylily harvester to follow a row of daylily plants accurately. During the last several decades, many innovations in vision-based navigation technology and row-recognition algorithms have been explored[1], which can be divided into two kinds. The first is to extract some discriminative points of the crop and then cluster or fit lines to these points. The clustering algorithm, such as the Hough transform, was reported in several studies[2-5]. The fitting algorithms, such as the Thei-Sen estimator[6], least squares[7-10], and other robust linear regression methods[11,12] have been investigated. The second is to segment the rows and/or inter-rows, calculate the center lines of these regions, and further gain the guidance directrix[13]. Han et al. achieved row segmentation by the *K*-means clustering algorithm, row detection by a moment algorithm, and guidance line selection by a cost function[14]. Our research team segmented the crops based on color features, extracted the axes using a rotary inertia method, and selected navigation baselines by angle tolerance and distance tolerance[15]. With the development of deep learning in recent years, semantic segmentation and instance segmentation have been used to

Received date: 2025-01-23 **Accepted date:** 2026-03-16
Biographies: Zilin Zhang, Master candidate, research interest: image processing, Email: 18635368522@163.com; **Daoqi Li**, Master candidate, research interest: computer vision, Email: 2790884065@qq.com; **Gang'ao Li**, Master candidate, research interest: intelligent agricultural equipment, Email: 17866613116@163.com; **Zongbao Huang**, Master candidate, research interest: agricultural information monitoring and control technology, Email: 17702517264@163.com; **Xiuhan He**, Master candidate, research interest: agricultural electrification and automation, Email: hxh1234560824@163.com.
***Corresponding author: Junhui Feng**, PhD, Associate Professor, research interest: agricultural information detection and control. College of Agricultural Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China. Tel: +86-354-6289384, Email: fengjunhui@sxau.edu.cn; **Zhiwei Li**, PhD, Professor, research interest: precision agriculture and intelligent agricultural equipment. College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, Shanxi, China. Tel: +86-354-6289384, Email: lizhiweitong@163.com.

extract crop rows or inter-rows, from which guidance directives are subsequently derived based on either boundary contours or center lines[16-20].



1. Motion vehicle 2. Control system 3. Harvesting manipulator 4. Camera for detection 5. End-effector 6. Container for daylilies 7. Fixing rod located on the horizontal midline of the motion vehicle 8. Camera for navigation

Note: H is the height of the navigation camera, m, and θ is the angle between the camera's optical axis and the fixing rod, (°).

Figure 1 Schematic of the robotic daylily harvester

In a standardized organic daylily planting base, the density of inter-rows varies due to differences in the growth and developmental stages of the plant during the harvest season. There are kinds of weeds arbitrarily growing between rows, which complicates the segmentation of inter-rows. Moreover, the performance of visual navigation is adversely affected by changing natural light because of the sun's direction, clouds, and other conditions. The row-recognition algorithms based on discriminative points of the crop are not accessible undoubtedly; therefore, we intend to segment the inter-rows and then measure the center lines to obtain the guidance directrix. While semantic segmentation and instance segmentation could be considered for inter-row segmentation, their high computational demands make them unsuitable for hardware-constrained applications. As presented above, the image processor embedded in the control system of the robotic daylily harvester in Figure 1 is responsible for multiple tasks in real time, including recognizing mature daylilies, locating picking points, and detecting the guidance directrix. It is therefore critical to minimize the computational resources allocated to guidance directrix detection.

To address these challenges, we aim to develop a robust and efficient algorithm that integrates a suite of traditional image processing techniques. This approach is designed to accurately and reliably extract the guidance directrix under challenging field conditions, such as variable light, random weeds, and varying density of the inter-rows, while operating with minimal computational load to ensure real-time performance on resource-limited hardware.

2 Materials and methods

2.1 Image acquisition and downsampling

Images were taken during the 2023 and 2024 harvest seasons at the Hemerocallis resource garden of Shanxi Agricultural University. Two color cameras were used in the experiment: Canon EOS 550D with a resolution of 2592×1728 pixels and HUAWEI HLK-AL10 with a resolution of 4000×3000 pixels. The height ($H = 2$ m) and angle ($\theta = 50°$) were set based on the height of daylily plants and the shooting distance. Image processing was performed in MATLAB (R2018a) on an AMD Ryzen 9 5900HX @3.3 GHz, NVIDIA GeForce RTX 3070 GPU, 32.0 GB RAM laptop. Figure 2 presents representative field-captured examples of daylily plants with the following characteristic measurements: row spacing ranging from 1.2 m to 1.4 m, plant canopy spread between 1.5 m and 2.8 m, and plant height varying from 1.05 m to 1.25 m. The left image was taken under cloudy conditions with low illumination, whereas the right image was acquired on a sunny day with significantly better lighting. Additionally, the left image exhibits dense inter-row vegetation, in contrast to the right image, which shows a sparser inter-row distribution.

a. Cloudy, dense inter-row

b. Sunny, sparse inter-row

Figure 2 Illustrations of images captured in the daylily field

The original captured image is of substantially high resolution, which can result in a heavy computational load for visual navigation. To improve computational efficiency, the image is downsampled while maintaining acceptable quality. The downsampling process involves selecting every tenth row and every tenth column of pixels from the original image. As shown in Figure 3, the downsampled versions of the images from Figure 2 retain sufficient clarity for visual navigation tasks.

2.2 Coarse segmentation

The inter-row area serves as the operational zone for the robotic daylily harvester. During the harvest season, leaves from adjacent daylily rows often come into contact or even overlap. The density of the inter-row vegetation varies considerably depending on the growth stage of the plants. Furthermore, the inter-row space contains randomly distributed weeds, and the images captured under real-field conditions exhibit varying illumination levels.


a. Downsampled version of Figure 2a

b. Downsampled version of Figure 2b

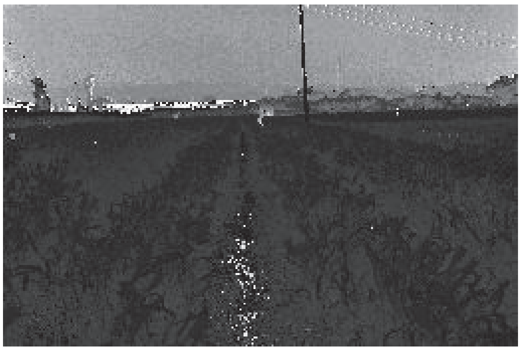
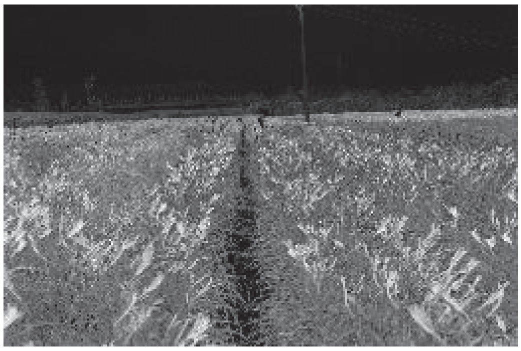
Figure 3 Downsampled versions of the images in Figure 2

To achieve more effective inter-row segmentation, the HSV (hue, saturation, value) color model is employed in place of the RGB (red, green, blue) model. Hue is a color attribute associated with the dominant wavelength in a mixture of light waves, saturation represents the relative purity or the amount of white light mixed with a hue, and value is the brightness and is used to describe the color sensation. Hue and saturation taken together are called chromaticity, and a color is characterized by its brightness and chromaticity[21]. In this study, images are captured under varying lighting conditions, resulting in inconsistent brightness levels. The HSV color model is particularly suitable in such scenarios because it separates the brightness component from the chromatic information. This separation allows hue and saturation to be used for the coarse segmentation of daylily inter-row regions. For a given image in the RGB color model, the hue H and saturation S components of each RGB pixel can be derived using the following equations:

$$H=\begin{cases}\arccos\left\{\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}, & B\le G\\ 2\pi-\arccos\left\{\dfrac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}, & B>G\end{cases}\tag{1}$$

$$S=\frac{\max(R,G,B)-\min(R,G,B)}{\max(R,G,B)}\tag{2}$$

According to the above calculation, the hue and saturation images can be obtained. Figure 4a and Figure 4b show the hue and saturation images corresponding to the left image in Figure 3, respectively, while Figure 4c and Figure 4d correspond to the hue and saturation images of the right image in Figure 3.

a. Hue image of Figure 3a

b. Saturation image of Figure 3a


c. Hue image of Figure 3b

d. Saturation image of Figure 3b

Figure 4 Hue images and saturation images of the images in Figure 3

In the hue images (Figure 4a and Figure 4c), the soil and dark plastic mulch between the daylily rows exhibit relatively higher hue values compared to the daylily plants. In contrast, the saturation images (Figure 4b and Figure 4d) show lower saturation levels in these areas, as they are often shaded by the plants. These characteristics enable the segmentation of the inter-row regions based on hue and saturation.

For images with dense inter-row vegetation (as in Figure 4a and Figure 4b), the region of soil and mulch is limited in the hue image but more distinguishable in the saturation image. Conversely, for images with sparse inter-row growth (Figure 4c and Figure 4d), the dark plastic mulch is clearly separable from plants in the hue image, while less evident in the saturation image. Accordingly, dense inter-row regions are segmented primarily using saturation, whereas sparse inter-row regions are segmented based on hue.

2.2.1 Coarse segmentation based on saturation

In the saturation image, the soil and dark plastic mulch exhibit lower values compared to the daylily plants. Therefore, a saturation threshold can be defined to classify the pixels into two dominant categories: foreground and background, thereby achieving coarse inter-row segmentation. The binarization is performed using the following equation:

$$I_{\mathrm{BW}}(x,y)=\begin{cases}1, & S(x,y)<T\_\mathrm{saturation}\\0, & \mathrm{otherwise}\end{cases} \tag{3}$$

where, (x, y) is any point of the image, I_{BW} is the binary image after coarse segmentation, and $T\_\mathrm{saturation}$ is the saturation threshold.

2.2.2 Coarse segmentation based on hue

In the hue image, the soil and dark plastic mulch are higher than those of the daylily plants. Accordingly, a hue threshold can be set to achieve image binarization. The corresponding equation is as follows:

$$I_{\mathrm{BW}}(x,y)=\begin{cases}1, & H(x,y)>T\_\mathrm{hue}\\0, & \mathrm{otherwise}\end{cases} \tag{4}$$

where, T_hue is the hue threshold.

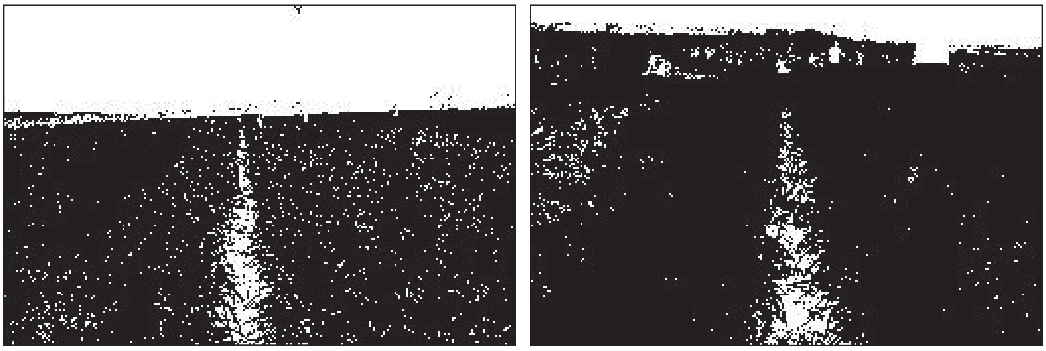
The binary images of Figure 3 are shown in Figure 5. The left image is segmented based on saturation, and the right is based on hue.

2.3 Noise reduction

In the binary image after coarse segmentation based on hue or saturation, a significant number of discrete foreground points appear within the background regions. These points, which can be regarded as noise, may adversely affect the segmentation accuracy of the inter-row region and should therefore be removed. Image denoising can be achieved through spatial filtering, frequency-domain filtering, or morphological filtering. The latter, which includes opening and closing operations, is also an effective method for noise reduction[22]. Specifically, the opening operation typically removes thin protrusions and isolated foreground points from the background, while the closing operation, conversely, fills narrow gaps and small holes within the foreground. Accordingly, a morphological opening operation is applied here to eliminate noise points in the binary image, which can be expressed as follows:

$$I_{\mathrm{denoised}}=I_{\mathrm{BW}}\circ SE=(I_{\mathrm{BW}}\Theta SE)\oplus SE \tag{5}$$

where, SE is a structuring element, I_{denoised} is the denoised image based on the morphological opening operation, $\circ$ is the morphological opening operator, Θ is the morphological erosion operator, and $\oplus$ is the morphological dilation operator. The structuring element used in our noise reduction method is shown in Figure 6a, and the denoised images of Figure 5 are shown in Figure 6b and Figure 6c. Figure 6b is the denoised result of the left image in Figure 5, and Figure 6c corresponds to the right one in Figure 5.



a. Binary image of Figure 3a b. Binary image of Figure 3b

Figure 5 Binary images of the images in Figure 3

0	1	0
1	1	1
0	1	0

a. Structuring element b. Denoised result of Figure 5a c. Denoised result of Figure 5b

Figure 6 Structuring element used in noise reduction and denoised results of the images in Figure 5

2.4 Region of interest extraction

In the denoised image, the foreground includes not only the daylily inter-row region but also portions of the sky. It is therefore necessary to extract the region of interest (ROI) to ensure that the foreground corresponds exclusively to the inter-row region. Since the sky typically occupies the upper portion of the image, ROI extraction is primarily conducted along the longitudinal direction. Accordingly, the longitudinal foreground density distribution is first analyzed by computing the following statistical measure:

$$sum\_fore(i)=\sum_{j=1}^{n}I_{\mathrm{denoised}}(i,j),\quad i=1,2,\ldots,m \tag{6}$$

where, I_{denoised} is the denoised image, whose height is m, and width is n, while sum_fore is the longitudinal foreground density distribution. The statistical measure described above was conducted on Figures 6b and 6c, and the resulting curves are plotted in Figure 7.

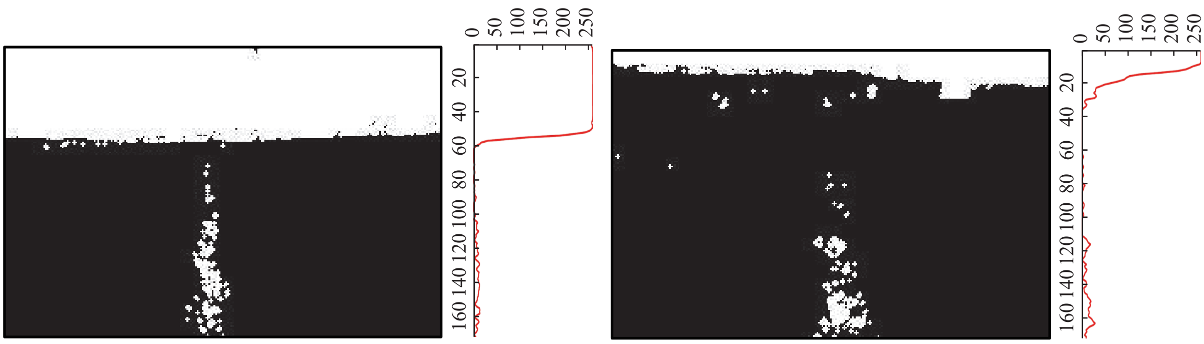
In the longitudinal foreground density distribution diagrams shown in Figure 7, the upper portion of the image corresponds to the sky, which is classified as foreground and thus exhibits high density. The region between the sky and the inter-row is identified as background, resulting in a significantly lower density. Due to linear distortion introduced by the camera perspective, the number of points and corresponding density value within the inter-row region increase in areas nearer to the camera. In the image, these areas correspond to the lower sections. In the area closest to the camera, if the plants within a row are sparse, the intra-row region may be misclassified as part of the inter-row region. To accurately extract the region of interest (ROI) along the longitudinal direction, it is necessary to determine the starting and ending rows, which are defined by the following expressions:

$$start\_row = \max(\max\{i \mid sum\_fore(i) > T\_start\_\max \;\&\; i < m/2\},\ \max\{j \mid sum\_fore(j) < T\_start\_\min \;\&\; j < m/2\}) \tag{7}$$

$$end\_row = \max\{k \mid sum\_fore(k) < T\_end\_\min \;\&\; k > m/2\} \tag{8}$$

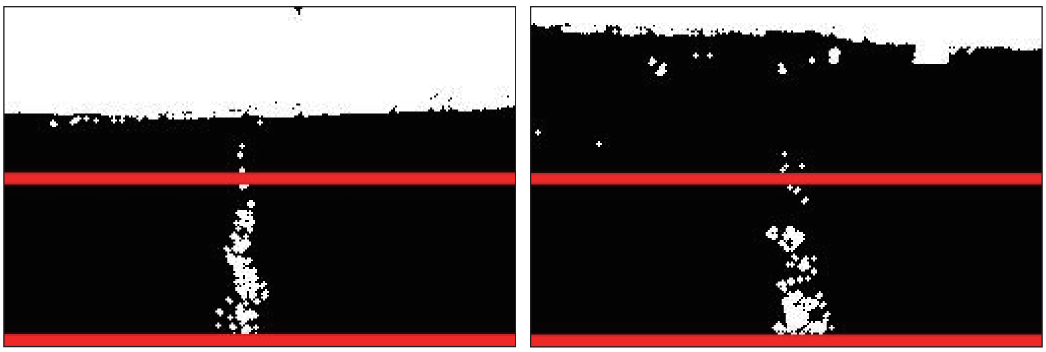
$$I_{\mathrm{ROI}} = I_{\mathrm{denoised}}(start\_row : end\_row, 1 : n) \tag{9}$$

where, $T\_start\_\max$ is the threshold which is the upper limit of the starting row's longitudinal foreground density, $T\_start\_\min$ is the threshold which is the lower limit of the starting row's longitudinal foreground density, $T\_end\_\min$ is the threshold which is the lower limit of the ending row's longitudinal foreground density, and I_{ROI} is the result of ROI extraction, namely, the ROI image. As shown in Figure 8, red lines mark the starting (top) and ending (bottom) rows determined for Figure 6b and 6c. The ROI image is then defined as the sub-image enclosed between these two rows in each respective figure.



a. Statistics of Figure 6b b. Statistics of Figure 6c

Figure 7 Longitudinal foreground density distribution statistics



a. Extracting ROI result of Figure 6b b. Extracting ROI result of Figure 6c

Figure 8 Extracting ROI results of the images in Figure 6

2.5 Morphological dilation

In the ROI image, numerous gaps are present within the inter-row regions due to the distribution of weeds and leaves from adjacent rows. To address this, the morphological dilation operation is applied to expand the boundaries, bridge the gaps, and restore broken foreground areas. The morphological dilation operation can be expressed as follows:

$$I_{dilated} = I_{\mathrm{ROI}} \oplus SE_{dilated} = \left\{ z \,\middle|\, \left[(\widehat{SE_{dilated}})_z \cap I_{\mathrm{ROI}} \right] \subseteq I_{\mathrm{ROI}} \right\} \tag{10}$$

where, $SE_{dilated}$ is the structuring element used in the morphological dilation, $(\widehat{SE_{dilated}})_z$ is the reflection of $SE_{dilated}$ about its origin, and shifting this reflection by z, $I_{dilated}$ is the dilated image. In this study, a 10×9 quasi-square structuring element is employed based on the slightly larger vertical extent of the gaps compared to the horizontal extent. The dilation results of the ROI images derived from Figure 8 are shown in Figure 9.

2.6 Connected component labeling

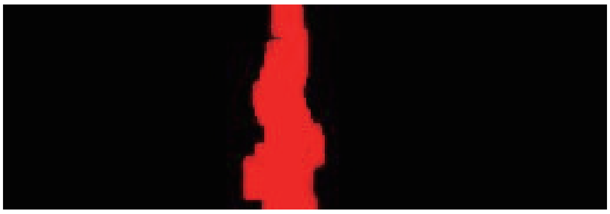
In the dilated ROI image, the connectivity of the inter-row regions is significantly enhanced. However, certain segments of these regions remain fragmented into multiple sections due to weed distribution. To facilitate guidance directrix detection, connected component labeling may be applied to the ROI image. This allows for further analysis of each connected component, enabling accurate identification of inter-row areas and ultimately supporting the detection of guidance directrix within the daylily inter-row regions.

A variety of classical connected component labeling algorithms are available. In this study, a space-efficient two-pass algorithm was employed, as described in Computer and Robot Vision[23]. Following the labeling process, each foreground region is assigned a distinct numeric label. To facilitate visual interpretation, different connected components are filled with different colors. Figure 10 shows the connected component labeling results corresponding to the dilated ROI images shown in Figure 9.

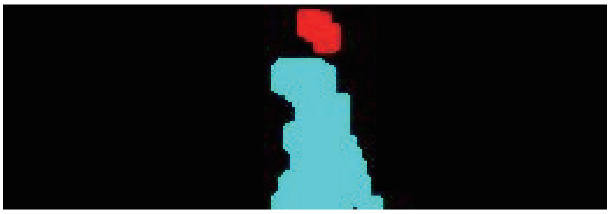
a. Dilated result of the ROI image derived from Figure 8a

b. Dilated result of the ROI image derived from Figure 8b

Figure 9 Dilated results of the ROI images derived from Figure 8



a. Connected component labeling result of Figure 9a



b. Connected component labeling result of Figure 9b

Figure 10 Connected component labeling results of the images in Figure 9

2.7 Connected components of interest extraction

In the labeled ROI image, not every connected component corresponds to the dominant daylily inter-row region. Hence, it is necessary to extract the connected components of interest (COI). Since the navigation camera is mounted along the horizontal midline of the robotic daylily harvester, the inter-row region typically aligns with this midline in the image. Accordingly, the COI extraction algorithm is formulated as follows:

$$\mathrm{COI} = \{C_i | C_i \cap R_{\mathrm{mid}} \neq \varnothing\},\ \ i = 1, 2, \ldots, N \tag{11}$$

$$R_{\mathrm{mid}} = I_{dilated}(1: m_{\mathrm{ROI}}, n_{\mathrm{ROI}}/2 - R\_\mathrm{radius}: n_{\mathrm{ROI}}/2 + R\_\mathrm{radius}) \tag{12}$$

$$I_{\mathrm{COI}}(x, y) = \begin{cases} 1, & (x, y) \in \mathrm{COI} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where, C_i is the ith connected component, and there are N connected components altogether in the labeled ROI image; R_{mid} is the rectangle neighborhood of the horizontal midline in the dilated ROI image $I_{dilated}$, R_radius is the distance from the horizontal midline to the vertical boundary of R_{mid}, and the width of R_{mid} is $2\times R\_\mathrm{radius}+1$; the size of $I_{dilated}$ is $m_{\mathrm{ROI}}\times n_{\mathrm{ROI}}$; COI is the set of COI, I_{COI} is the result of COI extraction.

Based on the COI extraction algorithm, the results corresponding to Figure 9 are shown in Figure 11, where all connected components labeled in Figure 10 are preserved. It should be noted, however, that not all labeled connected components necessarily represent COI in other experimental contexts.

a. Extracting COI result of Figure 10a

b. Extracting COI result of Figure 10b

Figure 11 Extracting COI results of the images in Figure 10

2.8 Axis measurement of COI

The COI can represent the main part of the inter-row region in the daylily field. By measuring the axis of the COI, the guidance directrix of the robotic daylily harvester can be detected. Assuming that each point within the COI has equal mass, the axis of the COI can be determined by calculating the moment of inertia of a rigid body with surface mass distribution[24].

In physics, the moment of inertia of a rigid body with surface mass distribution is defined as follows:

$$J = \int r^2 \sigma \mathrm{d}s \tag{14}$$

where, J is the moment of inertia, r is the perpendicular distance from each point-mass to the axis of rotation, σ is the areal density, and ds is the differential area element within the COI. Since the mass is uniformly distributed and the pixel coordinates of each point are known, σ can be set to 1. Thus, Equation (14) can be further expressed as:

$$J = \iint_D [(x-\bar{x})\sin\alpha - (y-\bar{y})\cos\alpha]^2 \mathrm{d}x\mathrm{d}y \tag{15}$$

where, D is the domain of the COI, α is the angle from the X-axis of the Cartesian coordinate system to the axis of rotation, and $(\bar{x}, \bar{y})$ is the center of mass of the COI. In the context of a two-dimensional digital image, Equation (15) can be expressed as follows:

$$J = \sum_{y=1}^{m_{\mathrm{ROI}}} \sum_{x=1}^{n_{\mathrm{ROI}}} [(x-\bar{x})\sin\alpha - (y-\bar{y})\cos\alpha]^2,\ if\ I_{\mathrm{COI}}(y, x) = 1 \tag{16}$$

$$\bar{x} = m_{10}/m_{00},\ \ \bar{y} = m_{01}/m_{00} \tag{17}$$

where, m_{pq} is the $(p+q)$th order moments, which are defined as:

$$m_{pq} = \sum_{y=1}^{m_{\mathrm{ROI}}} \sum_{x=1}^{n_{\mathrm{ROI}}} x^p y^q I_{\mathrm{COI}}(y, x) p,\ \ q = 0, 1, 2, \cdots \tag{18}$$

In Equation (16), the moment of inertia J is minimized when the axis at angle α passing through the center of mass $(\bar{x}, \bar{y})$ coincides with the axis of the COI. Through mathematical derivation, the angle α is given by the following expression:

$$\alpha = \frac{1}{2}\arctan\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right) \tag{19}$$

where, μ_{pq} are the central moments, which are defined as:

$$\mu_{pq} = \sum_{y=1}^{m_{\mathrm{ROI}}}\sum_{x=1}^{n_{\mathrm{ROI}}}(x-\bar{x})^p(y-\bar{y})^q I_{\mathrm{COI}}(y,x)p,\ \ q=0,1,2,\cdots \tag{20}$$

Since the angle α and the center of mass $(\bar{x},\bar{y})$ are known, the axis of the COI can be determined. According to the actual condition, the angle from the X-axis to the axis of COI may be 90°. Given that Equation (19) for the angle α involves the arctangent function, and for computational convenience, the axis of the COI can be expressed as follows:

$$x = k(y-\bar{y})+\bar{x} \tag{21}$$

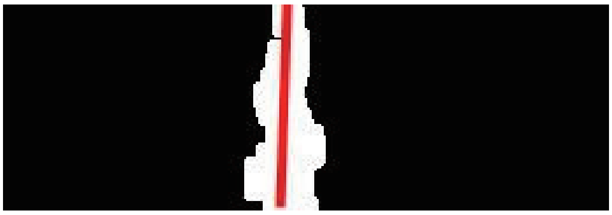
$$k = 1/\tan\alpha \tag{22}$$

Based on Equations (19)-(20), Equation (22) can be derived as:

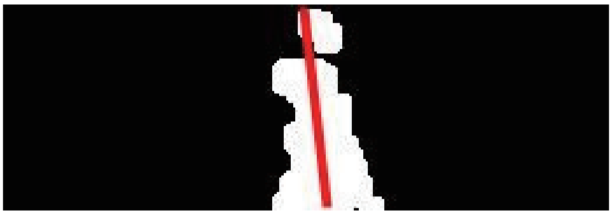
$$k = 1/(\lambda+\sqrt{1+\lambda^2}),\ \ \lambda = 2\mu_{11}/(\mu_{20}-\mu_{02}) \tag{23}$$

Thus, the calculation of the axis of the COI is completed. It should be noted that the Cartesian coordinate system used in this calculation differs from the image's pixel coordinate system, where the origin is located at the upper-left corner, the positive X-axis extends to the right, and the positive Y-axis extends downward.

Based on the axis measurement of the COI, the results for Figure 11 are shown in Figure 12, and the red line is the axis of COI.



a. Axis measurement of COI for Figure 11a

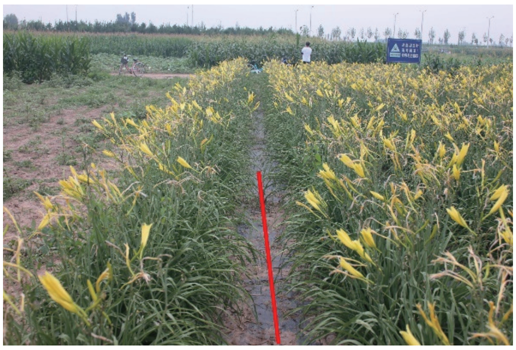


b. Axis measurement of COI for Figure 11b

Figure 12 Axis measurement of COI for the images in Figure 11

Based on the downsampling ratio and the starting row defined in Equation (7), the calculated axis of the COI is mapped back to the original image. Figure 13 shows the results of guidance directrix detection for Figure 2. It should be noted that the ending row specified in Equation (8) is not involved in this mapping process; instead, the guidance directrix is directly aligned with the bottom row of the original image.

a. Guidance directrix detection result for Figure 2a



b. Guidance directrix detection result for Figure 2b

Figure 13 Guidance directrix detection results for the images in Figure 2

2.9 Overall algorithm

Following a series of processing steps, including downsampling, coarse segmentation, noise reduction, ROI extraction, morphological dilation, connected component labeling, COI extraction, and axis measurement of COI, the guidance directrix has been successfully detected. However, the coarse segmentation step is designed to accommodate two distinct scenarios: dense and sparse inter-row regions. To enable adaptive selection of the coarse segmentation method, the foreground density, which is defined as the total number of foreground pixels, is computed for the binary image obtained from each method based on hue and saturation. The binary image with the higher foreground density is then selected and passed to the next step. Therefore, an overall workflow is proposed, as illustrated in Figure 14.

3 Experiments and results

3.1 Experimental settings

In this study, the experimental parameters are listed in Table 1, where the parameters of $T\_start\_\max$, $T\_start\_\min$, and $T\_end\_\min$ are set based on the denoised image's width n.

Table 1 Parameters in the algorithm of detection guidance directrix

Parameters	Value	Description
T_saturation	0.3	The saturation threshold
T_hue	0.5	The hue threshold
$T\_start\_\max$	$0.4\times n$	The threshold, which is the upper limit of the starting row's longitudinal foreground density
$T\_start\_\min$	$0.01\times n$	The threshold, which is the lower limit of the starting row's longitudinal foreground density
$T\_end\_\min$	$0.15\times n$	The threshold, which is the lower limit of the ending row's longitudinal foreground density
R_radius	3	The distance from the horizontal center line to the vertical boundary of R_{mid}

3.2 Experimental results

Two cameras were used in this study, Canon EOS 550D and HUAWEI HLK-AL10, to capture 400 images during the 2023 and 2024 harvest seasons at the Hemerocallis resource garden of Shanxi Agricultural University. The pixel resolutions of these images are 2592×1728 and 4000×3000. The two illustrative images in the previous section are 2592×1728 pixels, and the following experiments contain two images with a resolution of 4000×3000

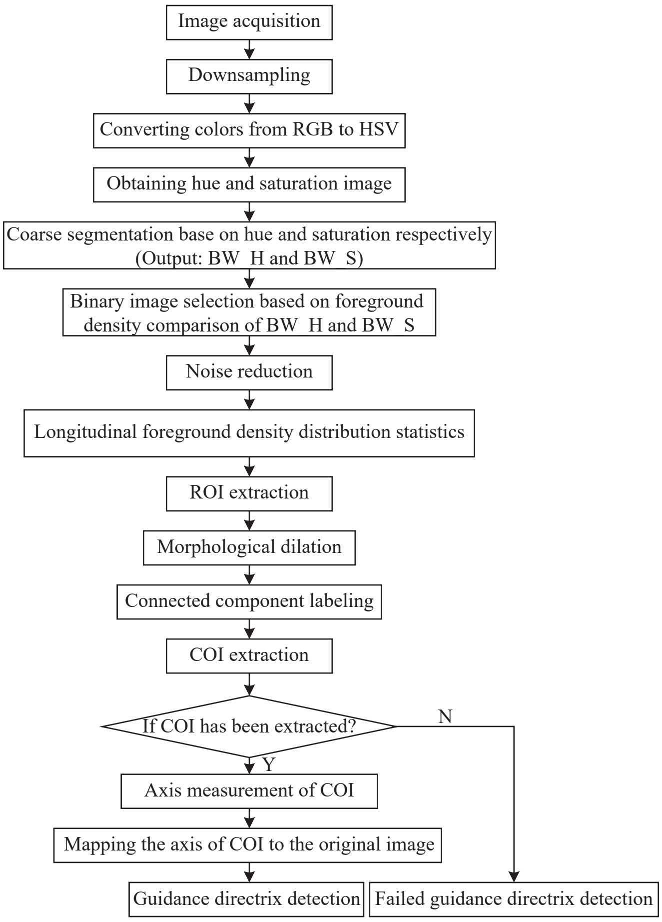


Figure 14 The overall workflow of our algorithm

pixels and one with a resolution of 2592×1728 pixels.

In the three experiments illustrated in Figure 15, the inter-row regions in Experiments 1 and 2 are relatively sparse, while that in Experiment 3 is denser. Accordingly, different coarse segmentation methods based on hue or saturation were applied, both of which yielded equally effective segmentation results. Experiment 1 exhibits noticeable weeds in the inter-row region, and Experiment 3 shows overlapping leaves from adjacent daylily rows within the inter-row region. Although these factors slightly affect inter-row segmentation, they do not impair visual navigation performance. The original images across the three experiments were captured using two different cameras with varying pixel resolutions, yet all achieved high accuracy.

Throughout the experiments, all captured images were successfully detected, with 393 out of 400 images accurately processed. This corresponds to an accuracy rate of 98.25% and a recall rate of 100%. Additionally, we recorded the navigation time, which averaged 8.7521 ms per image, with the longest execution time remaining under 13 ms.

3.3 Comparative experiments

This subsection presents a comparative evaluation of the proposed algorithm against other methods using various semantic segmentation models. Specifically, the semantic segmentation models are first employed to identify the inter-row regions, following which the inertia-based axis extraction method in Section 2.8 is applied to detect the guidance directrix. The 400 captured images were augmented by horizontal flipping, brightness enhancement, and brightness reduction, yielding 1600 images for the dataset. The experimental results are summarized in Table 2. The algorithm proposed in this study achieves slightly higher accuracy and recall rates compared to models such as YOLOv5n-seg-axis, YOLOv5s-seg-axis, YOLOv8n-seg-axis, YOLOv8s-seg-axis, and YOLOv11n-seg-axis, while demonstrating significantly lower average navigation time and a substantially reduced model size.

Table 2 Comparison of different methods

Method	Accuracy rate/%	Recall rate/%	Average navigation time/ms	Model size/B
YOLOv5n-seg-axis	96.32	96.57	12.6135	4.1 M
YOLOv5s-seg-axis	97.03	96.88	20.8432	15.2 M
YOLOv8n-seg-axis	96.81	96.67	19.1723	6.8 M
YOLOv8s-seg-axis	97.15	97.13	27.3667	23.9 M
YOLOv11n-seg-axis	97.88	98.84	18.1448	6.0 M
Our algorithm	98.25	100	8.7521	18.0 K

4 Discussion

Our algorithm is specifically tailored for the robotic daylily harvester. Some processes, such as coarse segmentation, ROI extraction, morphological dilation, and COI extraction, are closely related to the growth characteristics of daylily plants, the visual contrast between the crops and the field environment, as well as the practical demands of navigation. As a result, the algorithm is not directly applicable for detecting guidance directrices in other crop types, or even for daylilies at different developmental stages.

The detected guidance directrix provides a visual representation of the navigation path, whereas the relative yaw angle deviation and lateral deviation of the robotic daylily harvester must be derived for motion control. Since the navigation camera is mounted along the horizontal midline of the vehicle, the yaw angle can be determined based on the value of α in Equation (19). More precisely, it equals to $90°-\alpha$. The lateral deviation is then given by $d\cdot\sin(90°-\alpha)$, where d represents the horizontal distance from the fixing rod (on which the navigation camera is mounted) to the center of mass of the vehicle.

The experimental images were captured using two different cameras at an earlier stage. However, during the development of the daylily harvesting robot, we selected the Z30C industrial camera as the navigation camera.

The experimental images were initially captured with two cameras: Canon EOS 550D and HUAWEI HLK-AL10. However, during the development of the daylily harvesting robot, we selected the Z30C (Daying digital industrial camera, made in Changsha, Hunan, China) industrial camera as the navigation camera. It features a resolution of 640×480 pixels, a maximum frame rate of 60 frames per second, and a 46° field of view. As presented in Section 2.1, the navigation camera is mounted at a height of 2 m, with its optical axis tilted at an angle of 50° relative to the fixing rod. Based on the schematic diagram shown in Figure 16, we analyze the effective navigation distance. Here, point C represents the navigation camera, OC corresponds to the installation height H (as introduced in Figure 1), θ denotes the angle between the optical axis and the fixing rod, and ω represents half of the field of view. In the captured image, the Y-coordinate range corresponds to the actual photographed distance between points A and B (i.e., AB). However, this photographed distance does not represent the actual navigation distance used in our algorithm. According to Equation (7), the starting row during ROI extraction is constrained to no more than $m/2$, where m is the image height. Consequently, the effective navigation distance spans from AD to AB.

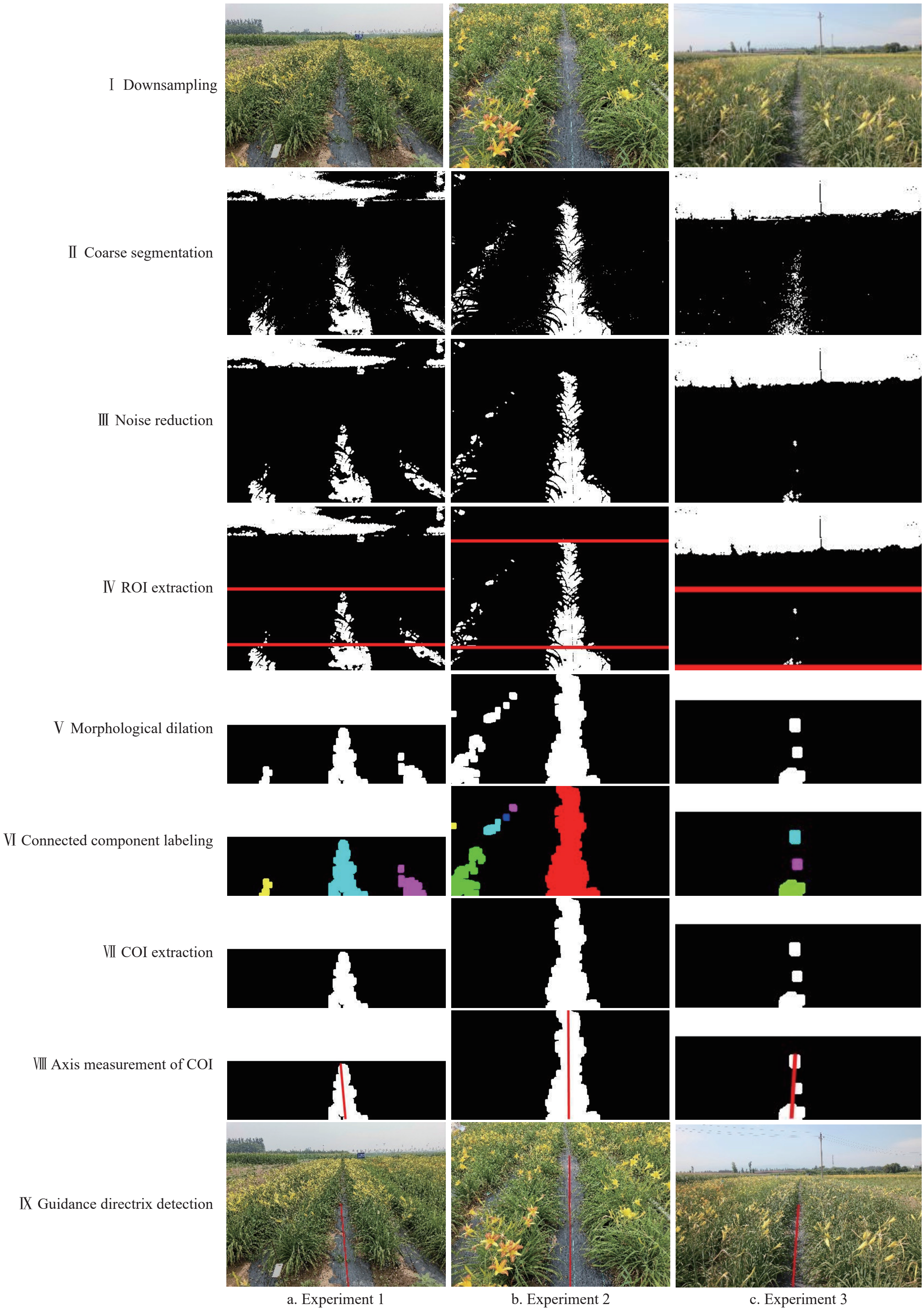


Figure 15 Experimental results at different pixel resolutions and inter-row densities

$$AD = H\cdot(\tan\theta - \tan(\theta-\omega)) \tag{24}$$

$$AB = H\cdot(\tan(\theta+\omega) - \tan(\theta-\omega)) \tag{25}$$

where, H is the height of the navigation camera, it is 2 m; θ is the angle between the camera's optical axis and the fixing rod, it is 50°; ω is half of the field of view, it is 23°, so the navigation distance ranges from 1.36 m to 5.52 m.

The pixel resolution of Canon EOS 550D is 2592×1728, HUAWEI HLK-AL10 is 4000×3000, and Z30C is 640×480. As described in subsection 2.1, images captured by the Canon EOS 550D and HUAWEI HLK-AL10 are downsampled at a factor of

10×10. To maintain a balance between image quality and computational efficiency, images from the Z30C can be downsampled at a factor of 2×2. As summarized in subsection 3.2, the longest navigation time remains below 13 ms, enabling the system to support image capture at the camera's maximum frame rate of 60 frames/s. This corresponds to a navigation rate exceeding 81 meters per second, which adequately meets the operational speed requirements of the robotic daylily harvester.

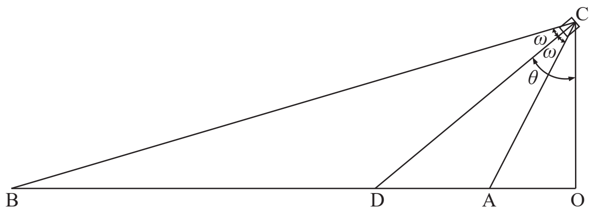


Figure 16 The schematic for the analysis of navigation distance

5 Conclusions

The proposed algorithm successfully detected guidance directrix for the robotic daylily harvester under challenging field conditions, including changing natural light, random weed distribution, and varying density of the inter-rows. Downsampling significantly reduced computational load without compromising navigation feasibility. Conversion from RGB to HSV color model effectively mitigated the impact of natural light variations. Coarse segmentation based on hue or saturation accommodated differences in inter-row density. Noise in binary images was suppressed through morphological opening operations. ROI extraction on the basis of longitudinal foreground density distribution excluded irrelevant areas such as the sky. Morphological dilation connected discontinuous regions caused by weeds and overlapping leaves from adjacent rows. COI were extracted to represent a valid inter-row region, and the axis was accurately determined via inertia-based calculation.

Experimental results demonstrated that the proposed algorithm achieves accurate and real-time detection of navigation lines in unstructured field environments, while also offering advantages such as low computational cost, no need for expensive hardware, and independence from large-scale training data. However, the algorithm has not yet been implemented on an actual vision-based robotic daylily harvester, representing a critical direction for future research and development.

Acknowledgments

This work was financially supported in part by the Natural Science Research Project of Shanxi Province (Grant No. 202303021221086), in part by Scientific Research and Initiating Engineering Project for Introducing Talents of Shanxi Agricultural University (Grant No. 2023BQ93), in part by the Sub-project of National Key Research and Development Program of China (Grant No. 2021YFD1600301-4), in part by Joint Research and Development Project for high-quality seeds of Shanxi Province under Grant YZGG-10, in part by Shanxi Modern Agro-industry Technology Research System (Grant No. 2025CYJSTX08-18), and in part by Central Government Guidance Fund for Local Science and Technology Development (Grant No. YDZJSX20231A042).

[References]

[1] Yao Z X, Zhao C J, Zhang T H. Agricultural machinery automatic navigation technology. iScience, 2023; 27(2): 108714.
[2] Marchant J A. Tracking of row structure in three crops using image analysis. Computers and Electronics in Agriculture, 1996; 15(2): 161–179.
[3] Åstrand B, Baerveldt A. A vision based row-following system for agricultural field machinery. Mechatronics, 2005; 15(2): 251–269.
[4] Choi K H, Han S K, Han S H, Park K H, Kim K-S, Kim S. Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields. Computers and Electronics in Agriculture, 2015; 113: 266–274.
[5] Li J B, Zhu R G, Chen B Q. Image detection and verification of visual navigation route during cotton field management period. Int J Agric & Biol Eng, 2018; 11(6): 159–165.
[6] Guerrero J M, Guijarro M, Montalvo M, Romeo J, Emmi L, Ribeiro A, et al. Automatic expert system based on images for accuracy crop row detection in maize fields. Expert Systems with Applications, 2013; 40(2): 656–664.
[7] Si Y S, Jiang G Q, Liu G, Gao R, Liu Z F. Early stage crop rows detection based on least square method. Transactions of the CSAM, 2010; 41(7): 163–167, 185.
[8] Song Y, Liu Y B, Liu L, Zhu D Q, Jiao J, Chen L Q. Extraction method of navigation baseline of corn roots based on machine vision. Transactions of the CSAM, 2017; 48(2): 38–44.
[9] Liang X H Z, Chen B Q, Wei C J, Zhang X C. Inter-row navigation line detection for cotton with broken rows. Plant Methods, 2022; 18: 90.
[10] Zhang X C, Chen B Q, Li J B, Fang X, Zhang C L, Peng S B. Novel method for the visual navigation path detection of the jujube harvester autopilot based on image processing. Int J Agric & Biol Eng, 2023; 16(5): 189–197.
[11] Billingsley J, Schoenfisch M. The successful development of a vision guidance system for agriculture. Computers and Electronics in Agriculture, 1997; 16(2): 147–163.
[12] Søgaard H T, Olsen H J. Determination of crop rows by image analysis without segmentation. Computers and Electronics in Agriculture, 2003; 38(2): 141–158.
[13] Liu W, Hu J P, Liu J X, Yue R C, Zhang T F, Yao M J, et al. Method for the navigation line recognition of the ridge without crops via machine vision. Int J Agric & Biol Eng, 2024; 17(2): 230–239.
[14] Han S, Zhang Q, Ni B, Reid J F. A guidance directrix approach to vision-based vehicle guidance systems. Computers and Electronics in Agriculture, 2004; 43(3): 179–195.
[15] Feng J H, Li Z W, Yang W, Han X P, Zhang X L. Detection navigation baseline in row-following operation of maize weeder based on axis extraction. Int J Agric and Biol Eng, 2020; 13(5): 181–186.
[16] Lin Y K, Chen S F. Development of navigation system for tea field machine using semantic segmentation. IFAC-PapersOnLine, 2019; 52(30): 108–113.
[17] Song Y, Xu F Y, Yao Q, Liu J L, Yang S. Navigation algorithm based on semantic segmentation in wheat fields using an RGB-D camera. Information Processing in Agriculture, 2023; 10(4): 475–490.
[18] Chen Z Y, Cai Y H, Liu Y X, Liang Z P, Chen H, Ma R J, et al. Towards end-to-end rice row detection in paddy fields exploiting two-pathway instance segmentation. Computers and Electronics in Agriculture, 2025; 231: 109963.
[19] Xie X L, Li Y C, Zhao L J, Jin X, Wang S S, Han X B. Method for the fruit tree recognition and navigation in complex environment of an agricultural robot. Int J Agric & Biol Eng, 2024; 17(2): 221–229.
[20] Lai X D, Huang J Z, Mo Y M, Li H W, Dong T Y, Wu T, He D Q. A lightweight deep learning model for synchronized crop stem detection and row segmentation at the seedling stage: Exploring their contribution to agricultural navigation line extraction. Computers and Electronics in Agriculture, 2026; 243: 111385.
[21] Gonzalez R, Woods R E. Digital image processing. Boston: Addison-Wesley Publishing Company. 1992; 793p.
[22] Maragos P, Schafer R. Morphological filters-part I: their set-theoretic analysis and relations to linear shift-invariant filters. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1987; 35(8): 1153–1169.
[23] Haralick R M, Shapiro L S. Computer and robot vision. New Jersey: Addison-Wesley Publishing Company. 1992; 672p.
[24] Ding S. Research on axis extraction methods for target's attitude measurement. Master of Engineering dissertation. Changsha: National University of Defense Technology. 2013; 70p.