Virtual Reality-Integrated Solutions for Extracting Central Crop Row Skeleton Lines Using the Maximum Diamond Principle Algorithm

Beibei Wu

1Sanquan College of Xinxiang, Medical University Henan, Xinxiang 453003, China

Corresponding author: Beibei Wu, hncjwbb@163.com

Abstract. Crop line skeleton extraction technology is the basis of realizing the automatic navigation path of machine vision in the process of precision agricultural pesticide application. It plays a vital role in the automatic walking of agricultural machinery and improving the accuracy of pesticide application. In order to overcome the problems of many redundant branches, poor connectivity, and long-running time in common skeleton extraction algorithms, a crop row skeleton extraction algorithm based on the maximum diamond is proposed in this paper. Firstly, the improved over green grayscale (1.65G-R-B) algorithm is used to grayscale the collected field crop row images. Then, the binary image is obtained by filtering and using the Otsu threshold algorithm to segment the crop line and background. Because crop rows have the characteristics of a linear structure, this paper adopts a $3 \times 1$ linear structure and $3 \times 3$. The method of combining square structure elements to carry out the morphological closed operation on the binary image. Finally, the crop row skeleton is obtained by the proposed maximum diamond skeleton extraction algorithm, and the central crop row skeleton is fitted by random Hough transform so as to realize the accurate positioning of crop rows. The experimental results show that the algorithm can maintain the unity and connectivity of the skeleton and has a strong anti-interference ability to the noise of crop rows. Compared with the commonly used topology thinning and maximum disk skeleton extraction algorithms, the algorithm extracts fewer redundant branches of the skeleton and has a short running time.

Keywords: Machine vision, Skeleton extraction, Maximum diamond, Navigation line; Virtual Reality-Integrated Solutions.
DOI: https://doi.org/10.14733/cadaps.2024.S17.37-48

1 INTRODUCTION

With the continuous development of modern agricultural technology, smart agriculture has become a major trend in current agricultural development. Precisely targeted spraying is an important...
component of agricultural intelligence, which can promote the large-scale development and progress of smart agriculture. The accuracy of agricultural machinery navigation path recognition determines the accuracy and effectiveness of spraying, while the accurate extraction of crop skeleton lines is the key to the extraction of navigation lines for precise targeted spraying systems and the basis for achieving precise targeted spraying technology. It plays an important role in improving the accuracy of field operations and reducing operating costs. In recent years, domestic and foreign scholars have done a lot of research on crop skeleton extraction. Huang Chenglong et al. [11] used a parallel thinning algorithm to extract plant skeletons from complete binary images and removed small burrs. They used the Hough transform to detect vertical lines in the maize skeleton image and identify the stem part of the skeleton. Zhang Weizheng et al. [29] proposed a maize plant image stem-leaf segmentation method based on skeleton extraction and binary tree analysis. They established the skeleton model of the plant through skeleton thinning algorithm. Brichet et al. [1] first used the mean shift and HSV threshold method to extract maize plants, then used the axis algorithm to extract the plant skeleton, and used the shortest path algorithm to detect the main stem, thereby achieving stem-leaf segmentation. Cabrera et al. [2] used the thinning algorithm to extract plant skeletons from binary images, and used the plugin Analyze Skeleton in Imagej to further obtain the intermediate nodes of the main stem and the endpoint of leaf branches, thereby achieving stem-leaf segmentation. Choudhury et al. [5] used the fast marching distance transformation method to skeletonize the binary image, then used threshold-based skeleton pruning to remove false branches, and then detected the nodes and endpoints in the plant skeleton to complete stem-leaf segmentation based on prior knowledge of the plant. Through the research of the above scholars, it was found that the skeletons extracted by them mostly have problems such as poor anti-interference ability, poor connectivity, and long running time. This study introduces an innovative integration of Virtual Reality solutions with the Maximum Diamond Principle Algorithm to extract central crop row skeleton lines. By combining the immersive capabilities of VR with the efficiency of the algorithm, this approach aims to revolutionize the visualization and analysis of agricultural fields, empowering farmers with accurate insights for informed decision-making and enhanced crop management.

In order to accurately obtain the crop line skeleton line and accurately apply drugs to the target, this paper studies the corn crop line skeleton extraction algorithm based on Visual C++6.0 software system. Based on the improved image preprocessing method, a practical crop line skeleton extraction algorithm based on the maximum diamond principle is proposed. The algorithm can maintain the connectivity of crop row discontinuity area, has strong anti-interference ability in complex background environments, has short running time, and has good adaptability to other grain crop rows with similar linear structures.

2 PRINCIPLES OF MAXIMUM DIAMOND SKELETON EXTRACTION

2.1 Definition of Maximum Diamond

In order to better explain the subsequent skeleton extraction process, the relevant definitions of the algorithm are described as follows.

1. Target pixel: refers to the point with a pixel value of 1 in the binary image after crop row and background segmentation after preprocessing, that is, the black pixel in the binary image.

2. Neighborhood pixel: as shown in Fig. 1 (a), if pixel O is the target pixel, O1O3O5O7 is defined as the first-order neighborhood pixel of target pixel O; Similarly, as shown in Fig. 1 (b), O1O2O4O0180602208 is the second-order neighborhood pixel of the target pixel O; And so on, the target pixel O has n-order neighborhood pixels.
3. Number of neighborhood pixels: the set of points with a pixel value of 1 in the n-order neighborhood pixels of the target pixel O, which is recorded as Nn, that is, Nn = \{S|N1\cup N2\cup \cdots \cup Nn\}.

4. Maximum Diamond: if a point with a pixel value of 1 appears in the first-order neighborhood pixel of the target pixel O, the diamond composed of the first-order neighborhood pixel is the maximum diamond of the target pixel O; If there are only 0 points in the first-order neighborhood pixels of the target pixel O, continue to search for the second-order neighborhood pixels until the searched n-order neighborhood pixels contain points with a pixel value of 1, then the diamond composed of the n-order neighborhood pixels is the largest diamond of the target pixel O, as shown in Fig. 1 (a) (b).

![First Order Neighborhood Pixel](image1.png) ![Second Order Neighborhood Pixel](image2.png)

**Figure 1:** Definition of maximum diamond.

### 2.2 Target Skeleton Point Extraction

As shown in Fig. 2 (a), the maximum diamond skeleton point extraction method is described with a specific example. In this paper, the second layer pixel point in Figure 2 is selected as an example for description. According to the above definition, the second layer pixel point contains three target pixels. According to the definition of the maximum diamond, find the maximum diamond of the three target pixels in turn, and then calculate and compare the number of neighborhood pixels contained in the three maximum diamonds. If the maximum diamond of one target pixel contains the largest number of neighborhood pixels, then the target pixel is the target skeleton point to be found; If the number of neighborhood pixels is equal in the maximum diamond of the target pixel of this layer, the target pixel close to the centerline is specified as the target skeleton point to be found.

According to the above extraction method, the number of neighborhood pixels of the first target pixel in Fig. 2 is N1 = 1; The number of neighborhood pixels of the second target pixel N2 = 3; Similarly, N3 = 2; By comparing N1, N2 and N3, it can be seen that the second target pixel is the target skeleton point of the layer, as shown in Fig. 2 (b), which is the target skeleton point of the second layer.
3 CROP LINE IMAGE PREPROCESSING

3.1 Image Acquisition and Experimental Development Platform

In this paper, the northern corn crop behavior is studied. The image acquisition equipment is the CCD industrial camera mv-vdo030sm / SC produced by Weishi Digital Technology Co., Ltd. the height from the ground during image acquisition is 1.6m, the shooting time is July 2021, and the output image is an 8-bit RGB color image, as shown in figure 3 (a).

A computer with Windows 7 operating system and Intel (R) Pentium (R) CPU g630 @ 2.70 GHz and 4G memory is used for this experimental research, and the skeleton extraction algorithm is studied and developed under the MFC application platform under Microsoft Visual C + + 6.0.

3.2 Grayscale Processing

In field environments, factors such as weeds, shadows, and non-target areas can interfere with crop row recognition, making direct recognition difficult to achieve ideal results. Grayscale conversion can increase the color differentiation between target and non-target areas. Converting RGB color images to HIS[21], YCrCb[16], Lab[3], and other color models can eliminate some shadow interference. Reasonable setting of the region of interest (ROI)[15],[30]can reduce interference from non-target rows and reduce computational load. The maize crop images collected in this study have large leaves and overlapping leaf phenomena, and the background is complex, with interference from wheat stubble, soil, weeds, and other complex backgrounds. In order to eliminate more complex background information, this paper uses the green characteristic of crop rows to adopt a super-green grayscale algorithm to grayscale the obtained crop row images, in order to reduce background noise interference. However, the traditional super-green grayscale algorithm still has significant background noise in the processed crop rows. Therefore, this paper improves this super-green grayscale algorithm. The improved algorithm is:

\[
G(x,y) = \begin{cases} 
1.65G-R-B & G \geq R \oplus G \geq B \\
\frac{G}{255} & \text{else}
\end{cases}
\]

In this paper, the traditional super green grayscale algorithm [10],[20],[27] and the improved super green grayscale algorithm are used to grayscale the collected 8-bit RGB color corn crop line image. The processed results are shown in Fig. 3 (b) (c). From the processing results, it can be seen that the improved algorithm can greatly reduce the interference of background noise and save time for subsequent filtering, The accuracy of navigation line fitting is improved.
3.3 Filtering Processing

After grayscale processing of the collected maize crop row images, there are still a few salt and pepper noises between rows. In order to improve the accuracy of crop row skeleton extraction, this paper adds a filtering step to the grayscale crop row image. Since median filtering has a good filtering effect on salt and pepper noise[4], this paper uses an improved median filtering algorithm to remove noise. However, for a 3x3 array window (9 elements), the original median filtering algorithm needs to sort all of them and then use the obtained median value to replace the grayscale value of the matrix center point. To obtain the median value of 9 pixels, 36 data comparisons are required, which will consume a lot of time and is not conducive to real-time image processing. Therefore, this paper improves the method of obtaining the median value of median filtering. The process of obtaining the median value is illustrated below using the example of a 3x3 pixel array A.

\[
A = \begin{bmatrix}
a & b & c \\
d & e & f \\
g & h & i \\
\end{bmatrix}
\]

First, sort each row vector of the matrix in descending order. Assuming \(a > b > c, d > e > f, g > h > i, e > b > h\), the maximum value among \(a, d, g\) is the maximum value of matrix A. Assuming \(d > a > g\); the minimum value among \(c, f, i\) is the maximum value of matrix A. Assuming \(i < f < c\), we have:

\[
\begin{align*}
a > b > c & \Rightarrow d > a > g > h > i, a > b > c \\
d > e > f & \Rightarrow i < f < c < b < a, b < e < d \\
g > h > i & \Rightarrow a > b > c, a > g > h > i \\
e > b > h & \Rightarrow f < c < b < a < d, b < e \\
d > a > g & \Rightarrow e > b > h, b > c > f > i \\
i > f > c & \Rightarrow h < b < e < d, b < a, h < g
\end{align*}
\]

Since the median of nine elements must be greater than four elements and less than four elements, it is unlikely that \(D, I, a, F, E, H\) are intermediate values, the minimum \(g\) of the maximum, the median \(B\) of the median, and the maximum \(C\) of the minimum in the rank comparison window matrix, then take the median of the three, which is the median of nine pixels, from the improved process, the improved algorithm only needs 21 sorting to find the median, the speed of the algorithm has been improved. Because the ultra-green gray algorithm has less noise, so it can filter salt-and-pepper noise by setting one filtering time. The result of the improved median algorithm is shown in Figure 4(a).

---

(a) Corn crop row image
(b) Traditional grayscale algorithm
(c) Improved super green grayscale algorithm

**Figure 3:** Grayscale processing.
3.4 Binarization Treatment

After median filtering, the corn crop row image has almost no noise interference. In order to extract the target area using morphological operations in the subsequent processing, the filtered grayscale image needs to be binarized. Binarization methods commonly used include Otsu's method based on maximum inter-class variance (OTSU) [6], iterative thresholding method [24], fixed thresholding method [7], mixed thresholding method [14], etc. The thresholding method is mainly determined based on experiments and experience, and is not suitable for practical applications due to large application errors for different images or images with different brightness. The Otsu method, on the other hand, is simple and fast to calculate, is not interfered by information such as image brightness and contrast, and can automatically binarize the image based on clustering [25]. Therefore, in this study, the Otsu algorithm was selected to binarize the filtered image, and the resulting binary image is shown in Fig. 4(b).

3.5 Connectivity Processing

From the binarized crop row image, it can be observed that there are gaps and interruptions between the corn plants. In order to maintain the connectivity of skeleton extraction and improve the accuracy of navigation lines, this study applied the morphological algorithm's closing operation to correct the binarized crop row image [9],[28],[8]. The method selected a combination of $3 \times 1$ linear structure and $3 \times 3$ square structure as the structural elements for morphological closing operation on the binarized image. Through experimentation, it was found that setting the $3 \times 1$ linear structure for 3 times dilation operation and the $3 \times 3$ square structure for 4 times erosion operation achieved the best connectivity preservation effect for the crop row. As shown in Fig. 4(c), the resulting image after connectivity processing shows good connectivity between the discontinuous plant points.

4 CENTRAL CROP ROW SKELETON EXTRACTION

Since the spacing between rows is symmetrical when planting corn crops in the north, when accurately applying drugs to the crop rows, just make the central target aim at the central crop row, and the other targets will automatically aim at the corresponding crop rows for accurate application. In this paper, the extraction steps of central crop row skeleton line are as follows:

1. Firstly, the binary image with good connectivity is obtained by preprocessing the corn crop row image;

2. The maximum rhombic skeleton extraction principle is used to extract the crop row skeleton from the binary image;
3. Taking the image centerline as the symmetry axis, the extracted skeleton image is divided into left and right regions;

4. Based on the axis of symmetry, scan the skeleton line from bottom to top, calculate the distance between each row of skeleton line and the center line in the left and right areas, retain the skeleton points closest to the center line, and corrode other skeleton points at the same time, so as to obtain the corn crop row skeleton line. The above extraction step image is shown in Fig. 5 (a-c).

![Image](image_url)

(a) Maximum Diamond Skeleton Extraction (b) Region Segmentation (c) Central Crop Row Skeleton Line

**Figure 5**: Extraction of central crop row skeleton line.

5 **DEVIATION ANALYSIS**

5.1 **CENTRAL CROP LINE STRAIGHT LINE FITTING**

The purpose of extracting the crop row skeleton is a prerequisite for determining the machine vision navigation path. Therefore, in this study, further linear fitting was performed on the extracted central crop row skeleton, and error analysis was carried out by comparing it with the actual navigation route to verify the accuracy of precise pesticide spraying. Scholars both at home and abroad have conducted extensive research on crop row fitting methods, mainly including least squares method [19],[27],[18], Hough transform method [12],[13],[23], vertical projection method [26],[20], etc., each with its own characteristics as shown in Table 1.

<table>
<thead>
<tr>
<th>Fitting method</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least square</td>
<td>It can find the best function match of the data, but it is only suitable for linear fitting function.</td>
</tr>
<tr>
<td>Hough transform</td>
<td>Simple calculation, strong anti-jamming ability, not sensitive to noise.</td>
</tr>
<tr>
<td>Vertical projection</td>
<td>Time complexity and space complexity are very high, time-consuming and affected by weeds and other factors.</td>
</tr>
</tbody>
</table>

**Table 1**: Characteristics of crop row fitting methods.
The basic principle of the least squares method is to find the best function match of the data by calculating the sum of the squares of the minimized errors. Its algorithm for identifying crop rows is simple but not effective under high weed pressure, and usually requires prior knowledge to improve the recognition accuracy. The crop row extraction algorithm based on vertical projection segments the farmland image into several rectangular images, and then performs pixel accumulation in the column direction. The crops are displayed as peaks on the vertical projection image, and the gaps between crop rows are displayed as valleys. The positioning points of the crop rows can be obtained by calculating the coordinates of the peaks. This method is heavily influenced by weeds and other factors. On the other hand, the crop row recognition based on Hough transform has strong anti-interference ability and is not sensitive to the missing parts, noise, and other coexisting non-linear structures in the image. It can tolerate gaps in the feature boundary description. Therefore, in this study, the Hough transform algorithm was used to perform linear fitting on the extracted central crop row skeleton. The algorithm steps are as follows:

Assume that the parameter equation of the line is:

$$\rho = acos\theta + bsin\theta$$

where \((a, b)\) is the image space coordinate; \((\rho, \theta)\) is the coordinate in parameter space; \(\rho\) is the distance from the origin to the line segment; \(\theta\) is the angle between the line and the x-axis. In each step of the Hough transform, randomly select two points \((a_i, b_i)\) and \((a_{i+1}, b_{i+1})\) and substitute them into the parameter equation to obtain a point \((\rho_i, \theta_i)\) in the parameter space. Then, set the corresponding unit of the accumulator in the parameter space to zero. If \((\rho_i, \theta_i)\) exists in the parameter space, the corresponding accumulator unit will be incremented by one; otherwise, it will be inserted into the parameter space. When a certain accumulator unit in the parameter space reaches a certain threshold, the corresponding \((\rho_i, \theta_i)\) is considered as the detected line parameter. The linear fitting results are shown in Figure 6.

5.2 DEVIATION CALCULATION

In the precision spraying system, the crop row image is first acquired through an industrial camera, and then transmitted in real-time to the crop row recognition system in the PC. The deviation calculation is then performed by the crop row recognition system, and the deviation information is sent to the precision spraying controller. The controller then issues the corresponding control commands to the spray nozzles to achieve automatic alignment. Assuming that the distance between the camera and the center of the crop row (camera deviation) is \(x\), when the center of the crop row is to the right of the camera, \(x>0\); otherwise, \(x<0\). If the distance between two nozzles, i.e., the row width, is \(z\), and the distance from the nozzle to the camera is \(y\), then the distance \(s\) moved by the nozzle is:

$$s = \frac{z}{2} + |x| \cdot y$$
Based on the deviation information sent by the row recognition system and the deviation calculation of the precision spraying system controller, real-time correction information can be fed back to control the offset of the precision spraying mechanical nozzle, thereby achieving automatic alignment of the precision spraying system nozzle. During the experiment, assuming y=z/2, the distance moved by the nozzle is equal to the deviation between the camera and the crop row. When x>0, the nozzle moves to the right, when x<0, the nozzle moves to the left, and when x=0, the nozzle does not move. The selected image was processed using the algorithm proposed in this paper, and the results of the image processing are shown in Figure 7, with a camera deviation information of 4.1mm and a processing time of 229ms per frame.

![Figure 7: Camera deviation data and processing time.](image)

**6 COMPARISON AND ANALYSIS OF ERROR OF OTHER ALGORITHMS**

In order to verify that the proposed algorithm can meet the needs of navigation path selection in actual agricultural precision spraying processes, this paper selected two commonly used skeleton extraction algorithms, namely, morphological thinning and topological thinning, to perform skeleton extraction on the same maize crop row image, and then performed linear fitting, followed by analysis of the deviation of the navigation path for the two algorithms and the proposed algorithm. The extraction results of the two algorithms are shown in Figures 8-9, and the results of the error analysis are shown in Table 2-3. Through comparative analysis, it can be seen that the morphological thinning algorithm extracts the central crop row skeleton with more redundant branches, resulting in the largest deviation of the navigation path; although the topological thinning algorithm can effectively extract the central path, it takes a long time to do so. The proposed algorithm not only maintains good connectivity of the crop rows, but also consumes the least amount of time and has the smallest error. The experiment has verified that it can meet the navigation needs of agricultural machinery.

![Figure 8: Morphological thinning skeleton extraction algorithm.](image)
7 CONCLUSION

In this paper, a central crop row skeleton extraction algorithm based on maximum rhombus is proposed for the study of maize crop behavior, the conclusions are as follows:

1. the improved over-green grayscale (1.65 g-r-b) algorithm not only greatly reduces the noise interference, but also clearly segmented the background and target image;
2. improved the median filter algorithm;
3. combining the $3 \times 1$ linear structure and the $3 \times 3$ square structure elements to carry on the morphological close operation, not only has maintained the break point plant crop row skeleton connectivity, but also has saved the running time, in addition, the redundant branches smaller than this structure element are eliminated, thus effectively ensuring the linear characteristics of crop rows.
4. the algorithm can keep the uniformity and connectivity of the skeleton through the comparison with the commonly used skeleton extraction algorithm and the analysis of field experiments, the algorithm has strong anti-interference ability to the noise of crop rows, and the redundant branches of skeleton extracted by the algorithm are few, the running time is short, which can meet the needs of modern precision agricultural navigation, but its memory consumption is large, the algorithm needs further improvement.

Beibei Wu, https://orcid.org/0009-0001-4779-611X

ACKNOWLEDGEMENTS
This work was supported by Henan Provincial Science and Technology Research Project: Research on the key technology of cancer cell pathological image recognition based on dual-path isomeric neural network (No.222102210074); Key Teacher Training Program of Sanquan College of Xinxiang Medical College, (No.SQ2021GGJS08).

REFERENCE


[29] Zhang, W.; Li, X.; Wan, H.: etc. . Segmentation Method of Corn Plant Image Stem and Leaf Based on Skeleton Extraction and Binary Tree Analysis, Henan Agricultural Sciences, 49(9), 2020, 166-172.