Article

Crop Row Segmentation and Detection in Paddy Fields Based on Treble-Classification Otsu and Double-Dimensional Clustering Method

Yue Yu [1,2], Yidan Bao [1,2], Jichun Wang [1,2], Hangjian Chu [1,2], Nan Zhao [1,2], Yong He [1,2] and Yufei Liu [1,2,*]

[1] College of Biosystems Engineering and Food Science, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China; yuyue0416@zju.edu.cn (Y.Y.); ydbao@zju.edu.cn (Y.B.); 3170100671@zju.edu.cn (J.W.); hjchu@zju.edu.cn (H.C.); 3150100510@zju.edu.cn (N.Z.); yhe@zju.edu.cn (Y.H.)
[2] Key Laboratory of Spectroscopy Sensing, Ministry of Agriculture and Rural Affairs, Hangzhou 310058, China
* Correspondence: yufeiliu@zju.edu.cn; Tel.: +86-0571-88982631

Abstract: Visual navigation is developing rapidly and is of great significance to improve agricultural automation. The most important issue involved in visual navigation is extracting a guidance path from agricultural field images. Traditional image segmentation methods may fail to work in paddy field, for the colors of weed, duckweed, and eutrophic water surface are very similar to those of real rice seedings. To deal with these problems, a crop row segmentation and detection algorithm, designed for complex paddy fields, is proposed. Firstly, the original image is transformed to the grayscale image and then the treble-classification Otsu method classifies the pixels in the grayscale image into three clusters according to their gray values. Secondly, the binary image is divided into several horizontal strips, and feature points representing green plants are extracted. Lastly, the proposed double-dimensional adaptive clustering method, which can deal with gaps inside a single crop row and misleading points between real crop rows, is applied to obtain the clusters of real crop rows and the corresponding fitting line. Quantitative validation tests of efficiency and accuracy have proven that the combination of these two methods constitutes a new robust integrated solution, with attitude error and distance error within 0.02° and 10 pixels, respectively. The proposed method achieved better quantitative results than the detection method based on typical Otsu under various conditions.

Keywords: visual navigation; paddy field; image segmentation; crop row detection; Otsu method

1. Introduction

Autonomous navigation for agricultural robots is essential for promoting the automation of modern agriculture, especially in ways to reduce labor intensity and enhance operation efficiency [1,2]. As a branch of autonomous navigation, visual navigation has developed rapidly in recent years, due to the improvements of computer calculation speed and visual sensors [3]. The most important issue of visual navigation primarily concerns extracting a guidance path according to the environment. When the drainage, light, and field management in the process of rice cultivation taken into account, rice is usually planted in rows, especially when transplanted by machines. Therefore, many researchers have tried extracting the guidance path for unmanned agricultural machinery utilizing this feature.

Typically, most of the methods proposed for crop row detection in recent years share the same architecture consisting of four steps, which are image grayscale transformation, image binarization, feature point extraction, and crop row identification. The excess green

Excess green graying method, which was reported to yield good results by Woebbecke et al. [4], is the most widely used grayscale transformation method due to its excellent performance in distinguishing green plants and background under a wide range of illumination conditions. Once the grayscale transformation is completed, the Otsu method [5],

a kind of nonparametric and unsupervised method of automatic threshold selection for image binarization, can be applied. The principle of the Otsu method is to select an optimal threshold by the discriminant criterion, thereby maximizing the separability of the resultant classes in gray levels. As for feature point extraction, the horizontal strip method combined with the vertical projection method serves as the most common solution. Søgaard and Olsen [6] divided a grayscale image into 15 horizontal strips, and then computed the vertical sum of gray values in each strip, with the maximum denoting the center of crop row in each strip. Lastly, crop rows are detected by the center points. On the basis of horizontal strips, Sainz-Costa et al. [7] developed a strategy for identifying crop rows through the analysis of video sequences. Hough transformation [8] is one of the most commonly used machine vision methods for identifying crop rows [9]. Least squares fitting has become another commonly used method to identify crop rows since the separation of weeds and crops has improved. Billingsley and Schoenfisch [10] used least squares fitting on the basis of information from three row segments to detect crop row guidance information. This image processing architecture attached with these classical methods makes it possible to detect crop rows, especially in some simple and specific circumstances.

Various studies on crop rows detection focused on making improvements in some steps of this general architecture. To further distinguish crops and weeds after the binarization using the typical Otsu method, Montalvo et al. [11] designed a method called double thresholding. Considering the crop row arrangement is known in the field, as well as the extrinsic and intrinsic camera system parameters, Guerrero et al. [12] proposed an expert system based on the geometry, and a correction was applied through the well-tested and robust Theil–Sen estimator in order to adjust the detected lines to the real ones. Guoquan Jiang et al. [13] constructed a multi-region of interest method, which integrates the features of multiple rows according to a geometry constraint. In order to enhance the robustness of crop row detection, García et al. [14] divided crop row identification into three steps: extraction of candidate points from reference lines, regression analysis for fitting polynomial equations, and final crop row selection. The method could deal with uncontrolled lighting conditions and unexpected gaps in crop rows. Many scholars have applied crop row detection algorithms to visual navigation systems and conducted field experiments. Guerrero et al. [15] designed a computer vision system involving two modules. The first module aimed to estimate the crop rows as accurately as possible, while the second module used the crop rows to control the tractor guidance and the overlapping. Basso et al. [16] proposed a crop row detection algorithm featuring Hough transform of an embedded guiding system for unmanned aerial vehicle (UAV). Tenhunen et al. [17] proposed a method for recognition of plantlet rows by means of pattern recognition. Li et al. [18] designed a pipeline-friendly crop row detection system using field programmable gate array (FPGA) architecture to reduce the resource utilization and balance the utilization of different onboard resources. Rabab et al. [19] proposed an efficient crop row detection algorithm which functions without the use of templates and most other prior information. The studies mentioned above mainly focused on crop row detection and visual navigation in dry fields.

However, it is difficult to achieve satisfying results of image segmentation in some complex agricultural environments, especially in paddy fields. A paddy field is a kind of open and complex environment, often accompanied by weed and duckweed, especially in areas without proper management. Some typical images in paddy fields are shown in Figure 1. Weed and duckweed are floating on the water, showing a green color similar to the color of the rice seedings. In addition, after fertilizing the paddy field, eutrophication often occurs, which can also cause the water surface to appear with a color similar to the color of the rice seedings. These are the reasons for the significantly increased difficulty in distinguishing crops from background. In this case, the typical Otsu method tends to bring a lot of noise, which totally disturbs the subsequent extraction of the crop row lines. Thus, the traditional architecture of crop row detection does not function well in paddy fields, attributed to the undesirable results of image binarization. Several studies were devoted to identifying crop rows without image segmentation. Aiming at a visual

navigation algorithm for a paddy field weeding robot, Zhang et al. [3] applied the smallest univalue segment assimilating nucleus (SUSAN) corner detection method directly after obtaining the grayscale image without image binarization. This strategy cleverly bypasses the problem of segmentation for paddy field images, but it increases the extent of time-consuming calculation.

(a) (b)
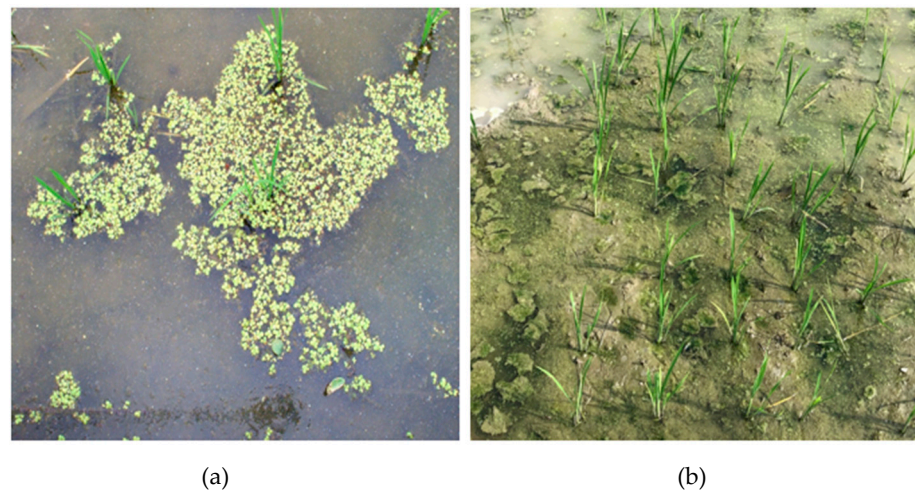
Figure 1. Typical images in paddy field: (**a**) duckweed around the rice seeding; (**b**) eutrophication after fertilizing.

In addition to the general architecture, stereo vision and neural networks have also been tested to detect the crop rows when the heights of the weeds and crop plants above ground are highly visible and when the weeds and crop plants differ in height [20]. Kise and Zhang [21] developed a stereo-vision-based crop-row tracking navigation system for agricultural machinery. Zhai et al. [1] developed a multi-crop-row detection algorithm to locate the three-dimensional (3D) position of crop rows according to their spatial distribution. Fue et al. [22] utilized stereo vision to determine 3D boll location and row detection, and the performance of this method showed promise as a method to assist with the real-time kinematic global navigation satellite system (RTK-GNSS) navigation. Adhikari et al. [23] trained a deep convolutional encoder decoder network to detect crop lines using semantic graphics. Ponnambalam et al. [24] designed a convolution neural network to segment input images based on red, green and blue color system (RGB) into crop and non-crop regions. Although these approaches achieved good results, there is still no superiority of stereo vision and deep learning over traditional architecture in terms of time consumption, and this will cause a significant burden for computation devices. Therefore, there is still a long way to go for the industrialization of crop row detection using stereo vision and neural network.

Although the abovementioned algorithms were proposed for crop row detection, the technical issue of image binarization of paddy fields still remains, which leads to the dilemma that traditional crop row detection methods based on image segmentation may fail to work in a paddy field environment. All these factors demonstrate that the crop row detection method of paddy fields should be originally designed in order to minimize the disturbance caused by the paddy field environment. Guijarro et al. [25] proved that distinguishing objects with different color characteristics by image segmentation is feasible. Thus, in this paper, to reduce the disturbance caused by weed and duckweed, the treble-classification Otsu method and double-dimensional clustering method for paddy fields are proposed, which improve the robustness of separating crop rows from complex paddy fields. The method proposed in this paper is improved on the basis of previous work, such as the typical Otsu method and clustering method. The purpose of this work is to meet the needs of low-price, lightweight computing and real-time performance of the unmanned system in paddy fields. The establishment of a flexible and reliable unmanned

system is of great significance for the realization of large-scale paddy field intelligent unmanned management.

2. Materials and Methods

The method proposed in this paper mainly comprises three modules: image segmentation, feature point extraction, and crop row detection, as described below.

2.1. Image Segmentation

2.1.1. Grayscale Transformation

The original chromatic image contains a large amount of information, the effective part of which is only the location of the green plant. This means that directly processing chromatic images will lead to unnecessary calculations due to information redundancy. To emphasize the living plant tissue, which is the basis of the subsequent steps, and weaken the rest of image [13], existing information needs dimensionality reduction processing. Thus, once the images are captured in the RGB color format, the first step is grayscale transformation.

Color is one of the most common indices used to discriminate plants from background clutter in computer vision [26]. A pixel where the predominant spectral component is the green is considered vegetation [12]. Through this strategy, color index-based approaches are resorted to achieve grayscale transformation. Generally, common green indices methods include normalized difference index (NDI) [27], excess green index (ExG) [4], color index of vegetation extraction (CIVE) [28], and vegetative index (VEG) [29].

The original images are obtained from the paddy field. Because of the water situation, it is necessary to take into account the reflections that frequently appear on the water surface. When only the intensity of the light source changes, the components of the light reflected on the surface of the same material are the same. Hence, the following formula is defined:

$$I_1 = CI_2, \tag{1}$$

where I_1 is the reflection intensity of the water surface under strong light, I_2 is the reflection intensity of the water surface without strong light reflection, and C is a constant greater than 1.

For the same water surface, the reflective components under strong light reflection and the reflective components without strong light reflection are the same, where only the intensity differs. Thus, for each RGB color channel, the following formula is defined:

$$\frac{\Delta R}{R} = \frac{\Delta G}{G} = \frac{\Delta B}{B} = \frac{I_1 - I_2}{I_2} = C - 1, \tag{2}$$

where ΔR, ΔG, and ΔB are the changes in the RGB channels due to changes in light intensity, respectively. R, G, and B are the values of RGB channels without strong light reflection, respectively.

This means that if the values of color channels can be expressed in the form of rates, the green index will not be bothered by the intensity of light. Under the same light source conditions, the reflection intensity of the water surface in the paddy field is much greater than that in the upland field. Although the above methods are all equipped with robustness to various lighting conditions, *CIVE* and *VEG* will fluctuate within a certain range as the reflected light intensity changes because the channel values of RGB cannot be expressed in the form of rates [13,30]. Therefore, the two methods of *CIVE* and *VEG* were eliminated from the candidate list. Additionally, the result of *NDI* is a near-binary image [26] with little capability to cope with the separation of weed and duckweed from rice seedings. Therefore, after comparing the above green indices, the *ExG* index was selected to process the images of paddy fields.

2.1.2. Thresholding with Treble-Classification Otsu Method

After the grayscale transformation is completed, large amounts of invalid information still remain in the image, while only the green plant tissues need to be considered. In addition to multi-channel color information, it is also necessary to refine the grayscale information. Therefore, image binarization, which means to reduce a multi-value digital signal into a two-value binary signal [31], is the second step of image segmentation. The Otsu method [5] is one of the best thresholding techniques for image binarization. The basic idea of the Otsu method is to dichotomize the pixels into two classes (background and objects) using a selected optimal threshold. This binarization method has been proven to be adaptively effective in different studies related to image segmentation between crop and background [32,33]. However, the dichotomy between objects and background is too rough to distinguish the real crops and green distractors, which will be both identified as objects in a binary image, especially in complex conditions such as paddy fields.

For the paddy field environment, the existence of green distractors can be explained as weed, duckweed, and cyanobacteria [3]. Furthermore, the water in the paddy field undergoes eutrophication after fertilization, resulting in a green paddy field environment. According to the above issues, the typical Otsu method needs to be improved in order to deal with the separation of real rice seedings and green distractors, rather than simply classify them as objects. To further classify the objects which include both rice seedings and green distractors, the typical Otsu method based on dichotomy was improved to be based on the trichotomy.

The pixels of a given greyscale image can be represented in L gray levels $[0, 1, \cdots, L-1]$. The number of pixels at level i is denoted by n_i, and the total number of pixels is denoted by $N = n_1 + n_2 + \cdots + n_L$. To simplify the discussion, the gray-level histogram is normalized and regarded as a probability distribution [5].

$$p_i = \frac{n_i}{N}, \ p_i \geq 0, \ \sum_{i=1}^{L} p_i = 1. \tag{3}$$

Now, suppose that two thresholds k_1 and k_2 are selected, which divides the pixels into three classes C_0, C_1, and C_2 (background, green distractor, and crop). C_0 denotes pixels with levels $[0, \cdots, k_1]$, C_1 denotes pixels with levels $[k_1, \cdots, k_2]$, and C_2 denotes pixels with levels $[k_2, \cdots, L-1]$. Then, the probabilities of class occurrence and the class mean levels, respectively, are given by

$$P_0(k_1, k_2) = \sum_{i=0}^{k_1} p_i. \tag{4}$$

$$P_1(k_1, k_2) = \sum_{i=k_1}^{k_2} p_i. \tag{5}$$

$$P_2(k_1, k_2) = \sum_{i=k_2}^{L-1} p_i. \tag{6}$$

According to the Bayes theorem, the mean levels of the pixels assigned to classes are given by

$$\begin{aligned}
m_0(k_1, k_2) &= \sum_{i=0}^{k_1} i P(i/C_0) \\
&= \sum_{i=0}^{k_1} i P(C_0/i) P(i) / P(C_0) \\
&= \sum_{i=0}^{k_1} i p_i / P_0(k_1, k_2) \\
&= \mu_0(k_1, k_2) / P_0(k_1, k_2),
\end{aligned} \tag{7}$$

$$
\begin{aligned}
m_1(k_1, k_2) &= \sum_{i=k_1}^{k_2} iP(i/C_1) \\
&= \sum_{i=k_1}^{k_2} iP(C_1/i)P(i)/P(C_1) \\
&= \sum_{i=k_1}^{k_2} ip_i/P_1(k_1, k_2) \\
&= \mu_1(k_1, k_2)/P_1(k_1, k_2),
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
m_2(k_1, k_2) &= \sum_{i=k_2}^{L-1} iP(i/C_2) \\
&= \sum_{i=k_2}^{L-1} iP(C_2/i)P(i)/P(C_2) \\
&= \sum_{i=k_2}^{L-1} ip_i/P_2(k_1, k_2) \\
&= \mu_2(k_1, k_2)/P_2(k_1, k_2),
\end{aligned}
\tag{9}
$$

where the total mean level of the grayscale image is

$$
\mu_0(k_1, k_2) = \sum_{i=0}^{k_1} ip_i,
\tag{10}
$$

$$
\mu_1(k_1, k_2) = \sum_{i=k_1}^{k_2} ip_i,
\tag{11}
$$

$$
\mu_2(k_1, k_2) = \sum_{i=k_2}^{L-1} ip_i.
\tag{12}
$$

The following relationships can be easily verified for any combination of k_1 and k_2:

$$
P_0(k_1, k_2)m_0(k_1, k_2) + P_1(k_1, k_2)m_1(k_1, k_2) + P_2(k_1, k_2)m_2(k_1, k_2) = m_G,
\tag{13}
$$

$$
P_0(k_1, k_2) + P_1(k_1, k_2) + P_2(k_1, k_2) = 1,
\tag{14}
$$

where m_G is the total mean level of the greyscale image, defined as

$$
m_G = \sum_{i=0}^{L-1} iP_i.
\tag{15}
$$

Referring to the evaluation of "goodness" of the threshold at a selected level in the Otsu method, the discriminant criterion is introduced.

$$
\eta = \sigma_B{}^2/\sigma_G{}^2,
\tag{16}
$$

where $\sigma_G{}^2$ is the total variance, defined as

$$
\sigma_G{}^2 = \sum_{i=0}^{L-1} (i - m_G)^2 P_i,
\tag{17}
$$

and $\sigma_B{}^2$ is the between-class variance, defined as

$$
\sigma_B{}^2 = P_0(k_1, k_2)(m_0 - m_G)^2 + P_2(k_1, k_2)(m_1 - m_G)^2 + P_2(k_1, k_2)(m_2 - m_G)^2.
\tag{18}
$$

Considering that the total variance $\sigma_G{}^2$ is a constant once the image is defined, the only way to maximize η is to maximize $\sigma_B{}^2$. Equation (18) can be converted into the following form using Equations (14) and (15):

$$
\begin{aligned}
\sigma_B(k_1,\ k_2)^2 &= \mu_0(k_1, k_2)^2/P_0(k_1, k_2) + \mu_1(k_1, k_2)^2/P_1(k_1, k_2) \\
&\quad + \mu_2(k_1, k_2)^2/P_2(k_1, k_2) - m_G{}^2.
\end{aligned}
\tag{19}
$$

Thus, the issue of maximizing discriminant criterion η is reduced to an optimization problem to search for a combination of k_1 and k_2 that maximizes between-class variance $\sigma_B{}^2$. The optimal threshold combination of $k_1{}^*$ and $k_2{}^*$ is

$$\sigma_B(\,k_1{}^*,\,k_2{}^*)^2 = \max \sigma_B(\,k_1,\,k_2)^2,\ 0 \leq k_1 \leq k_2 \leq L - 1. \tag{20}$$

The processing steps of the treble-classification Otsu method are as follows:

1. Calculate the normalized histogram of the input greyscale image, and record the minimum gray value and the maximum gray value as k_{min} and k_{max}.
2. Traverse k_1 from k_{min} to k_{max}, and calculate $P_0(k_1, k_2)$ and $\mu_0(k_1, k_2)$.
3. Traverse k_2 from k_{min} to k_{max}, and calculate $P_2(k_1, k_2)$ and $\mu_2(k_1, k_2)$.
4. Traverse k_1 from k_{min} to $k_{max} - 1$, then traverse k_2 from k_1 to k_{max}, and calculate $P_0(k_1, k_2)$, $\mu_0(k_1, k_2)$, and $\sigma_B(k_1,\,k_2)^2$.
5. Record $k_1{}^*$, $k_2{}^*$, which maximize $\sigma_B(k_1, k_2)^2$. If the combination of $k_1{}^*$, $k_2{}^*$ is not unique, calculate the mean value of $k_1{}^*$, $k_2{}^*$.

2.1.3. Filtering Operations

Generally, the initial binary image obtained using the thresholding method does not clearly represent the original information of the crop row. Some noise pixels are distributed among the crop rows in the form of islands, leading to interference with crop row detection. Although subsequent algorithms are not sensitive to the noise pixels of small island shapes, it is a wise choice to remove as much noise as possible that may cause interference. According to the theory in this research, all of the white pixels should represent the position of the crops, rather than the islands of noise. Therefore, to remove the small, discrete, and insignificant white patches, an extra filtering process is applied after image binarization. In this paper, isolated connected domains are traversed, and those with an area less than 30 pixels should be eliminated.

Figure 2a displays a typical image of a paddy field with duckweed. Figure 2b displays the result of grayscale transformation by applying the *ExG* index in Figure 1a. After image binarization through a treble-classification Otsu method, green crops are identified as white pixels and green distractors are significantly removed, as shown in Figure 2c. Lastly, the filtering operation is performed, and the result is as shown in Figure 2e. In contrast, the result of image binarization through the typical Otsu method is shown in Figure 2d, and the result of the filtering operation after the typical Otsu method is shown in Figure 2f.

### 2.2. Feature Point Extraction

In the process of image processing, every step is applied to refine the information attached to the image. The essence of refinement is to ensure that the remaining information is effective and the useless information is eliminated. Until now, the binary image obtained after morphological operation, in which the white connected domain represents green crops, has already roughly displayed the position of the crop rows. To further determine the location of the crop rows through quantitative assessment, the white connected domain should be identified as a serial of feature points with exact coordinates, which is called feature point extraction.

Considering that the noise attached to the obtained binary image is not significant, the horizontal strip method [6], which determines the feature points by investigating the number of white pixels on each horizontal strip, is applied. The size of the binary image is assumed as $H \times W$, where H denotes the height of the image while W denotes the width of the image, and the binary image is divided into N strips. In order to appropriately reduce subsequent calculations, the size of each horizontal strip can be expressed as $h \times W$, where h denotes the height of the strip and $h = H/N$. According to Zhang et al. [34], 30 horizontal strips can provide a good result for a wide variety of conditions. In this paper, to match 30 horizontal strips, h was adopted as 20. For each point of the binary image, $Gv(i, j)(i = 1, \ldots, W$ and $j = 1, \ldots, H)$ denotes the gray value of point (i, j). For the points

on the medial horizontal line in each horizontal strip, the number of white pixels at each column i is denoted as $S_k(i)$ [34], as shown in Equation (21).

$$S_k(i) = \sum_{j=kh}^{k(h-1)} Gv(i,j), \ k = 1, 2, \ldots N, \tag{21}$$

where k denotes the index of horizontal strips.
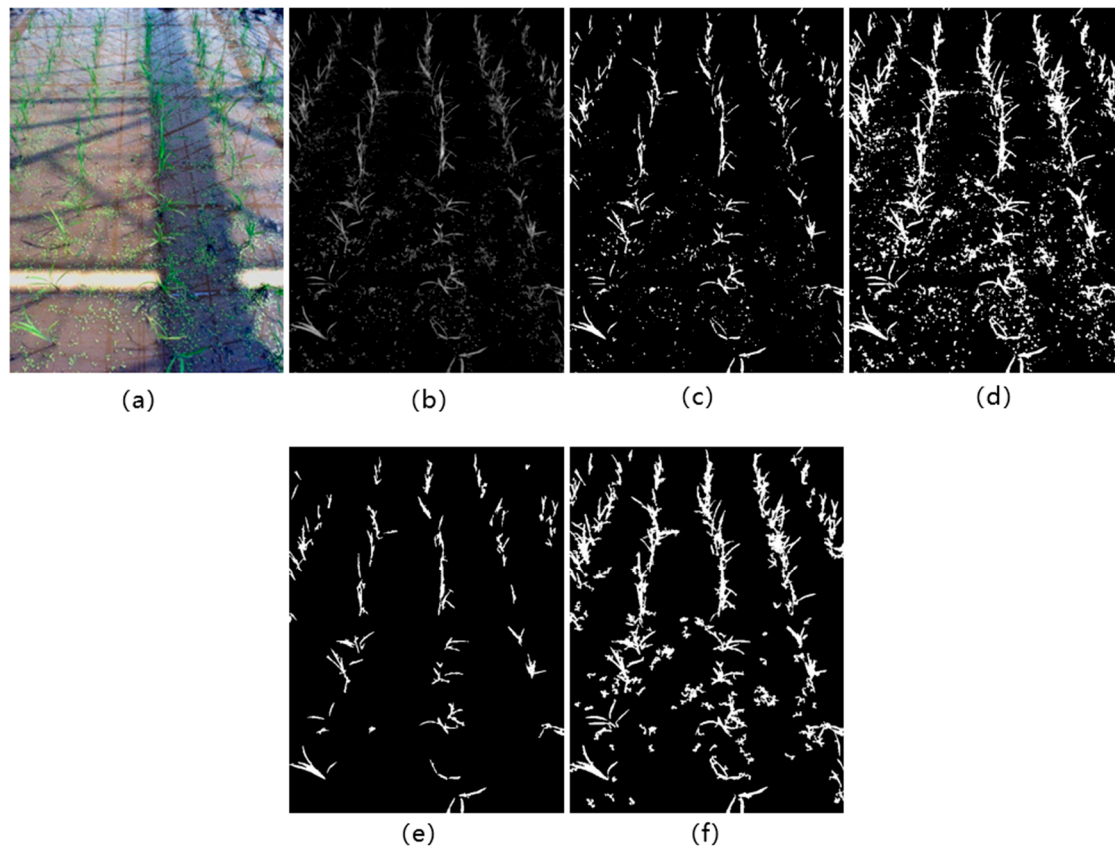
Figure 2. The typical image of a paddy field and the processing results after applying the treble-classification Otsu method and the typical Otsu method: (**a**) the original paddy field image with duckweed; (**b**) grayscale transformation applying the *ExG* index; (**c**) binary image obtained after the treble-classification Otsu method; (**d**) binary image obtained after the typical Otsu method; (**e**) filtering result after the treble-classification Otsu method; (**f**) filtering result after the typical Otsu method.

Theoretically, if $S_k(i) > 0$, it means that the set of pixels in the i-th column on the k-th horizontal strip shows implicit crop information, but the feature point cannot be determined accordingly. The small patches of white pixels actually representing noise may be mistaken for feature points, since the area of the horizontal strip where only noise exists would also provide a positive $S_k(i)$. Thus, some restrictions should be imposed on the judgement of feature points. To prove that it is reliable to recognize a certain area as a feature point, the information of green crops attached to the area should be relatively bigger; thus, the $S_k(i)$ of the area should be higher than a given threshold T(k), as shown in Equation (22).

$$T(k) = \mu * h * e^{k/100}, \tag{22}$$

where μ is the thresholding coefficient with value $\mu = 0.3$, and e is the natural logarithm.

Due to the clairvoyant principle of three-dimensional space, the densities of crop rows above and below the image are slightly different. To make the threshold T(k) adaptive to every horizontal strip, T(k) is constructed as a monotone increasing function, for the crop row in the upper part of the image is narrower while the lower part is wider. Through the

threshold T(k), each column of pixels in the k-th horizontal strip is traversed, but the result is a series of intervals, not points. The next step is to find the starting and ending points of these intervals, and the midpoints between the starting and ending points are the feature points. A judge function $J_k(i)$ is defined to search the boundary points of these intervals, as shown in Equation (23).

$$J_k(i) = \begin{cases} 1, & S_k(i-1) < T(k), \ S_k(i) \geq T(k) \\ -1, & S_k(i-1) \geq T(k), \ S_k(i) < T(k) \end{cases}. \tag{23}$$

If $J_k(i) = 1$, the abscissa of the starting point of a certain interval in the k-th horizontal strip is i. If $J_k(i) = -1$, the abscissa of the ending point of a certain interval in the k-th horizontal strip is i. Each start point and the next adjacent end point form an interval, and the midpoint of them can be identified as a feature point. In Figure 3, all steps mentioned to extract feature points are illustrated.
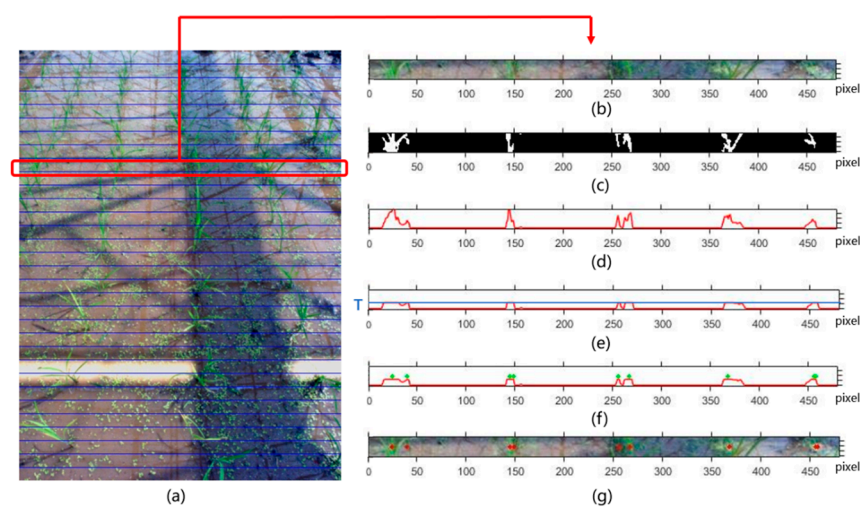
Figure 3. Feature point extraction: (**a**) original image of paddy field; (**b**) strip from the original image; (**c**) strip from the binary image; (**d**) curve of pixel value; (**e**) curve of pixel value after threshold; (**f**) feature point extraction; (**g**) feature point locations in the original image.

2.3. Crop Row Detection

The feature points extracted from the binary image are scattered; hence, the next step is their classification based on coordinate information. In order to sort out these scattered feature points and dig out information of crop row position as accurately as possible, the proposed double-dimensional adaptive clustering algorithm is applied.

According to the principle and results of feature point extraction, it can be found that, in each horizontal strip, a single crop row may be identified with more than one feature point. That is, the feature points belonging to the same crop row have both horizontal and vertical extensions. Therefore, it is difficult to take the information of all these feature points into account at the same time if only the horizontal or vertical traversal is applied for clustering analysis. Additionally, there may be gaps in the distribution of feature points from the same crop row or pseudo feature points caused by green information distractors between two adjacent crop rows. In this case, the adoption of traditional clustering analysis will easily lead to over-clustering which means the feature points on the same crop row are divided into multiple clusters, or under-clustering which means the feature points that do not belong to any crop row are classified into a cluster of crop row. In the proposed double-dimensional adaptive clustering algorithm, firstly, a horizontal clustering analysis is performed, through which feature points in each horizontal strip are clustered according to their abscissa, and then a vertical clustering is performed to assign the horizontal clustering results to each corresponding crop row. This clustering method is proposed according to the relative positions between the feature points and the approximate direction of the

crop rows in the image; therefore, prior knowledge about the number of crop rows is not required.

In horizontal clustering, a horizontal strip formed in the feature point extraction step is used as a unit to traverse. Initially, each feature point in a horizontal strip represents a single cluster. Now that the number of feature points in the k-th horizontal strip is assumed to be n_k $(k = 1, 2, \ldots n)$, and $P_{k,m}$ $(m = 1, 2, \ldots n_k)$ denotes the m-th feature point of the k-th horizontal strip, then the distance between adjacent feature points in the same horizontal strip can be expressed as $d_{k,m}$ $(m = 1, 2, \ldots n_k - 1)$, and the average distance between all these adjacent feature points is expressed as $d_{k,avg}$, as shown in Equation (24).

$$d_{k,avg} = \sum_{m=1}^{n_k-1} d_m / (n_k - 1). \tag{24}$$

The number of clusters in the k-th horizontal strip is assumed as n_{c_k} $(k = 1, 2, \ldots n)$. $C_{k,l}$ $(l = 1, 2, \ldots, n_{c_k})$ denotes the l-th cluster of the k-th horizontal strip. $n_{k,l}$ denotes the number of feature points in $C_{k,l}$, and $P_{k,l,m}$ $(m = 1, 2, \ldots n_{k,l})$ denotes the m-th feature point in $C_{k,l}$. The distance between adjacent feature points in $C_{k,l}$ can be expressed as $d_{k,l,m}$ $(m = 1, 2, \ldots n_{k,l} - 1)$ and the average distance between all these adjacent feature points is expressed as $d_{k,l,avg}$, as shown in Equation (25).

$$d_{k,l,avg} = \sum_{m=1}^{n_{k,l}-1} d_m / (n_{k,l} - 1). \tag{25}$$

All the feature points are scanned from left to right in each horizontal strip, and the horizontal strips are scanned from top to bottom. In each horizontal strip, all feature points are traversed and, according to the relative position, feature points that meet the clustering conditions are merged. The above process is repeated until there is no incomplete clustering. The procedures of the horizontal clustering method are shown in Figure 4. The specific steps of the horizontal clustering method are as follows:

1. Initialize $k = 0$.
2. For the k-th horizontal strip, determine the value of n_k. If $n_k = 0$, skip to step 9. If $n_k \geq 0$, calculate $d_{k,avg}$.
3. For the first feature point $P_{k,0}$ in each horizontal strip, initialize $C_{k,0}$, make $n_{k,0} = 0$, and push $P_{k,0}$ into it. Make $n_{k,0} = n_{k,0} + 1$. Make $m = 0$, $l = 0$.
4. Define $P_{k,m}$ as the current feature point. Define $C_{k,l}$ as the current cluster. Calculate the distance $d_{k,m}$ between $P_{k,m}$ and the next adjacent feature point $P_{k,m+1}$. If $d_{k,m} < \alpha * d_{k,avg}$, push $P_{k,m+1}$ into $C_{k,l}$ and make $n_{k,l} = n_{k,l} + 1$. If $d_{k,m} \geq \alpha * d_{k,avg}$, make $l = l + 1$, initialize $C_{k,l}$, push $P_{k,m+1}$ into $C_{k,lm}$ and make $n_{k,l} = n_{k,l} + 1$. Practical experience has shown that 0.8 is suitable for α.
5. Make $m = m + 1$. If $m < n_k$, return to step 4. If $m = n_k$, the first round of the k-th horizontal clustering is completed. Record l at this time as $n_{c_{k,l}}$. Make $l = 0$.
6. If $l > n_{c_k}$, skip to step 9. Define $C_{k,l}$ as the current cluster, and then calculate $d_{k,l,avg}$. If $d_{k,l,avg} * \left(n_{c_{k,l}} - 1 \right) < H/10$, make $l = l + 1$ and repeat step 6. If $d_{k,l,avg} * \left(n_{c_{k,l}} - 1 \right) \geq H/10$, make $m = 0$ and define $P_{k,l,m}$ as the current feature point. Initialize $a = 0$.
7. Calculate the distance $d_{k,l,m}$ between $P_{k,l,m}$ and the next adjacent feature point $P_{k,l,m+1}$. If $d_{k,l,m} < \alpha * d_{k,l,avg}$, push $P_{k,l,m+1}$ into $C_{k,l}$ and make $n_{c_k} = n_{c_k} + 1$. If $d_{k,l,m} \geq \alpha * d_{k,l,avg}$, make $a = a + 1$, initialize $C_{k,n_{c_{k,l}}+a}$, make $n_{k,n_{c_{k,l}}+a} = 0$, push $P_{k,m+1}$ into $C_{k,n_{c_{k,l}}+a}$, and make $n_{k,n_{c_{k,l}}+a} = n_{k,n_{c_{k,l}}+a} + 1$. Practical experience has shown that 0.8 is suitable for α.
8. Make $m = m + 1$. If $m < n_{k,l}$, return to step 7. If $m = n_{k,l}$, make $n_{c_k} = n_{c_k} + a$. Return to step 6.
9. Make $k = k + 1$. If $k \leq n$, return to step 2. If $k > n$, the horizontal clustering method ends.

k-th horizontal strip
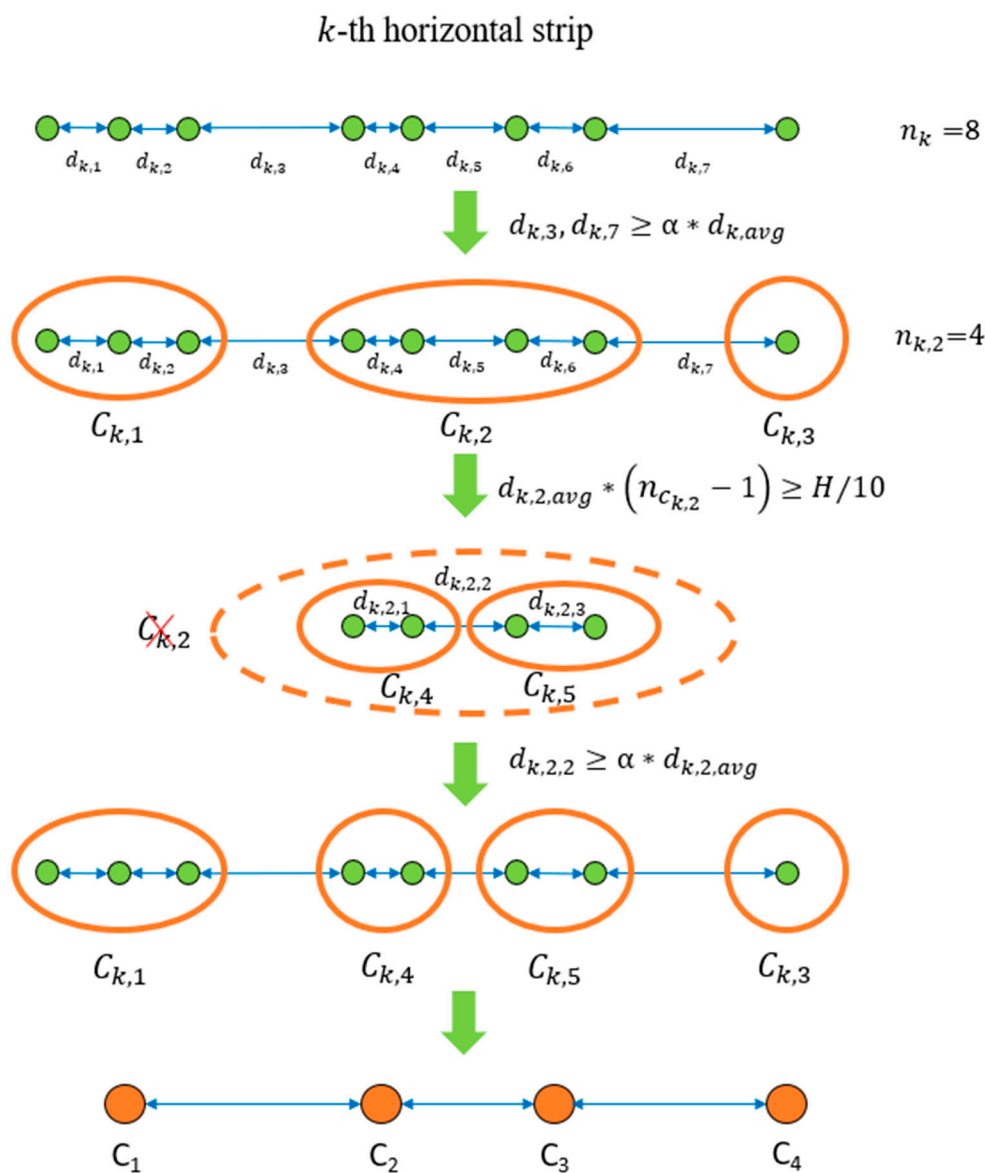
Figure 4. Schematic diagram of the processing steps in horizontal clustering.

In this manner, the feature points on each horizontal strip are clustered on the basis of their relative position of the abscissa, and the mean value $d_{k,l,avg}$ of the abscissa of feature points in $C_{k,l}$ is regarded as the abscissa of a new feature point, as shown in Figure 5. After the horizontal clustering, feature points belonging to the same crop row are merged into new feature points horizontally. For each crop row, there is nearly only one new feature point remaining in each horizontal strip, and the current distribution of new feature points makes the crop rows appear clearer.

Vertical clustering is applied to new feature points, i.e., the results of horizontal clustering. Since the pitch angle of the camera is $60°$, the closer to the top of the image, the closer the distance between crop rows, and the closer to the bottom of the image, the farther the distance between crop rows. For each cluster, the selection of an initial feature point is pivotal. To achieve better results at the initial stage, the vertical clustering is performed from bottom to top.
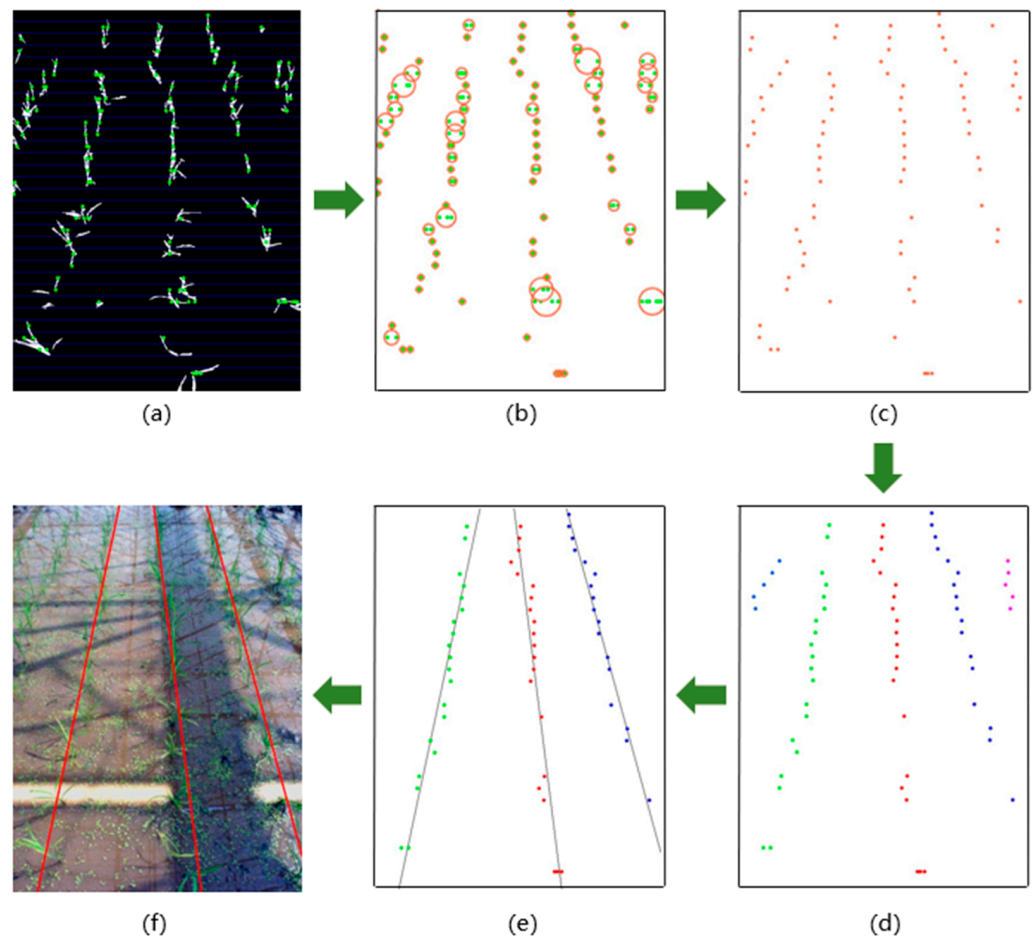
Figure 5. Crop row detection steps: (**a**) original feature points in binary image; (**b**) horizontal clustering of original feature points; (**c**) result of horizontal clustering; (**d**) result of vertical clustering; (**e**) crop row detection based on valid clusters; (**f**) final result presented in original image.

Make n_l denote the number of clusters, and make C_l ($l = 1, 2, \ldots, n_l$) denote the l-th cluster of crop row. In vertical clustering, once a new feature point is pushed into C_l, the fitting line parameters of the feature points in C_l need to be calculated using the least square method. Make n_{p_k} denote the number of feature points of the k-th horizontal strip, and make $P_{k,m}$ ($m = 1, 2, \ldots, n_{p_k}$) denote the m-th feature point of the k-th horizontal strip. Make $d_{p_{k,m,l}}$ denote distance between $P_{k,m}$ and the last point in C_l, and make $d_{l_{k,m,l}}$ denote distance between $P_{k,m}$ and the fitting line of C_l. The thresholds of $d_{p_{k,m,l}}$ and $d_{l_{k,m,l}}$ are represented as T_p and T_l. If the ordinate distance between the current feature point and the last point in C_l is greater than h, this situation is defined as a gap. The basic judgment of vertical clustering is divided into two cases according to whether or not a gap is encountered. If a gap is encountered, $d_{l_{k,m,l}}$ is the judging criterion. If there is no gap, $d_{p_{k,m,l}}$ is the judging criterion. If $d_{l_{k,m,l}} < T_l$ or $d_{p_{k,m,l}} < T_p$, the current feature point is pushed into the current cluster C_l. For a feature point, after traversing all the existing clusters and finding that none meets the judgment criterion, initialize a new cluster and push this feature point into it. Lastly, filter out those clusters with fewer than six feature points. The process of the vertical clustering method is shown in Algorithm 1, and the flow chart is shown in Figure 6.

Algorithm 1. The process of the vertical clustering method.

Input: n which denotes the number of horizontal strips.

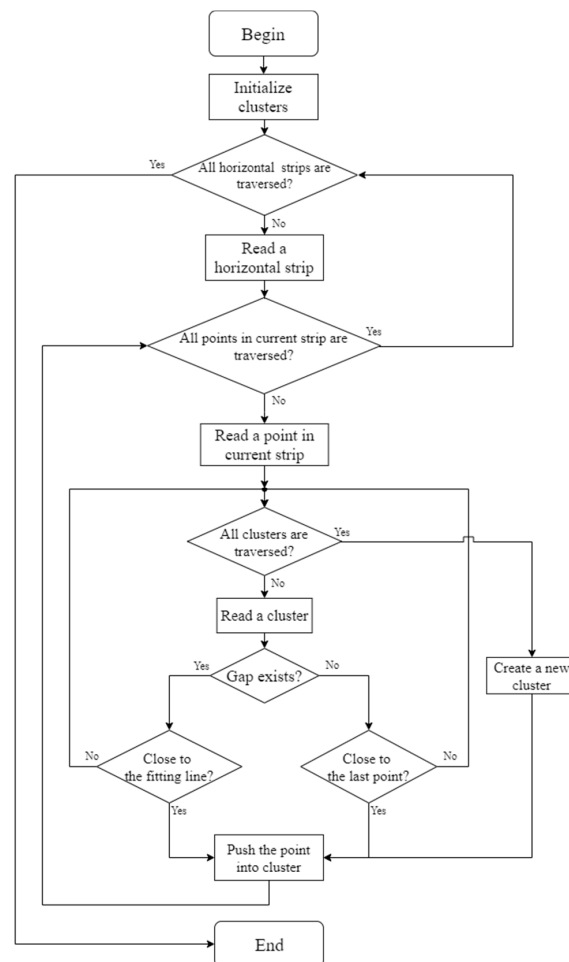n_{p_k} $(k = 1, 2, \ldots, n)$ which denotes the number of feature points of the k-th horizontal strip.

$P_{k,m}$ $(k = 1, 2, \ldots, n; m = 1, 2, \ldots, n_{p_k})$ which denotes the collection of feature points.

Outputs: Horizontal clusters C_l $(l = 1, 2, \ldots, n_l)$.

1: initialize $l = 1$, $n_l = 0$, $C_l = \varnothing$
2: for $k = n$: 1
3:    for $m = 1$: n_{p_k}
4:    for $l = 1$: n_l
5:     if gap
6:      if $d_{l_{k,m,l}} < T_l$
7:        push $P_{k,m}$ into C_l
8:        continue
9:     else
10:      if $d_{p_{k,m,l}} < T_p$
11:        push $P_{k,m}$ into C_l
12:        continue
13:    $n_l = n_l + 1$,
14:     push $P_{k,m}$ into C_{n_l}

Figure 6. Vertical clustering flow chart.

3. Experimental Results

3.1. Image Acquisition

In the experiments of image acquisition, an industrial camera (DFK-23U445, IMAG-ING SOURCE) was selected to capture images. The proposed algorithm was developed us-

ing Microsoft Visual C++ and the free computer vision library OpenCV 4.1.0. Furthermore, the digital images were stored as 24 bit color images with resolutions of 1280 × 960 pixels and saved in RGB (red, green, and blue) color space in the JPEG format. The camera settings were as follows: pitch and roll angles of 60° and 0° with the camera placed at a height of 1.5 m from the water surface of the paddy field. To verify the effectiveness of the proposed method under different conditions, several representative kinds of experimental images were captured at the China National Rice Research Institute in August 2020, the experimental field in the west area of Zhejiang University in June 2020, and Zhejiang Province (Xiaoshan) Modern Agriculture Innovation Park in October 2020. A total of 100 experimental images, including strong interference with eutrophication, moderate interference with disturbed weed or gaps, and weak interference, were selected to test the accuracy, efficiency, and reliability of the proposed method.

The ultimate purpose of crop row detection is to provide a guiding basis for automatic navigation; thus, the real-time requirement of common automatic navigation systems must be considered. In order to improve the real-time performance of image processing, the amount of calculation required should be reduced. Hence, downsampling processing was performed on the image, and then the image size was shrunk to 640 × 480 pixels.

### 3.2. Validation of Image Segmentation

In this paper, image segmentation based on treble-classification Otsu method is the most important step. A paddy field is a kind of open and complex environment, in which the water surface always presents a color close to that of rice seedings because of the existence of weed, duckweed, and eutrophication. The proposed treble-classification Otsu method should dig out the crop information be as little disturbed by the complex paddy field environment as possible. Thus, the performance of the proposed treble-classification Otsu method is crucial during the whole image processing. To validate the capability of the treble-classification Otsu method, accuracy validation tests and efficiency validation tests were conducted.

3.2.1. Accuracy Validation Tests of the Treble-Classification Otsu Method

In order to verify the accuracy performance of the proposed treble-classification Otsu method under various interference environments, several representative images of paddy fields were randomly selected for a validation test. Figure 7a displays the eutrophication in a paddy field, which mostly occurs after fertilization and in the early growth stage of the rice seeding. Early rice seedlings have smaller leaves and lighter colors, which are more likely to be confused with the color of the eutrophic water surface. Firstly, grayscale transformation was performed on Figure 7a, and the result is shown in Figure 7b. Subsequently, the treble-classification Otsu method and typical Otsu method were applied to Figure 7b, and the results are shown in Figure 7c,d, respectively. From the images shown in Figure 7, it can be observed that the proposed treble-classification Otsu method eliminated most of the interference information caused by eutrophication and left a small amount of noise, while the typical Otsu method could hardly distinguish between eutrophic water surface and real crops.

Figure 8a displays the disturbed weed located in crop rows. Some weeds that are not rice seedings appeared in the crop rows, which could interfere with the identification of the real rice crop row direction. Figure 8b displays the result of grayscale transformation which was performed on Figure 8a, while Figure 8c,d display the results of the treble-classification Otsu method and typical Otsu method, respectively. From the images shown in Figure 8, it can be observed that the leaves and overall shape of the disturbed weed were almost totally preserved and adhered together to form a connected domain. This means that it is difficult to remove these interference regions through filtering or morphological processing. Compared to the typical Otsu method, the treble-classification Otsu method adopted a higher threshold. The white area representing the crop in the binary image was significantly

more refined. The pixels of the weed were almost filtered, merely leaving some isolated noise, which reduced the interference of weeds on crop row identification.

(**a**)   (**b**)   (**c**)   (**d**)
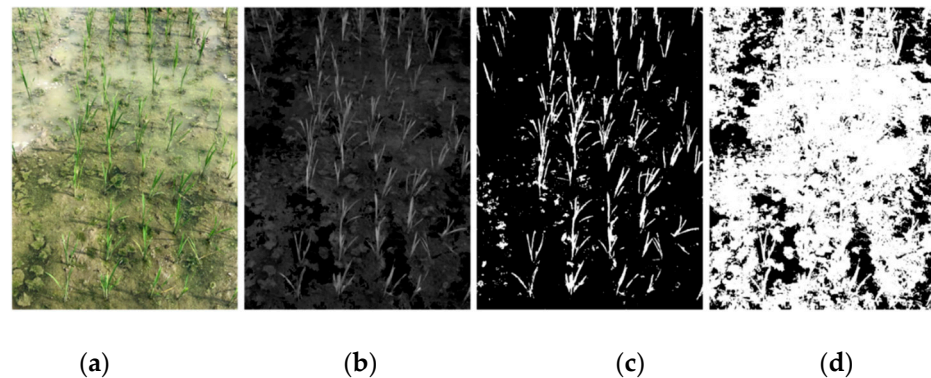
Figure 7. Validation tests of treble-classification Otsu method under eutrophication condition: (**a**) eutrophication after fertilizing; (**b**) grayscale transformation; (**c**) binary image obtained using treble-classification Otsu method; (**d**) binary image obtained using typical Otsu method.

(**a**)   (**b**)   (**c**)   (**d**)
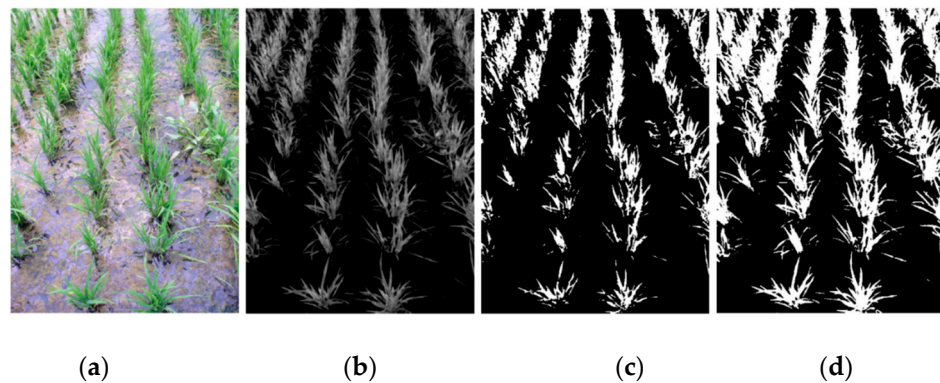
Figure 8. Validation tests of treble-classification Otsu method under conditions with disturbed weed: (**a**) disturbed weed located in crop rows; (**b**) grayscale transformation; (**c**) binary image obtained using treble-classification Otsu method; (**d**) binary image obtained using typical Otsu method.

Figure 9a displays an image of a paddy field with little duckweed and eutrophication and without weeds. Under an environment with weak interference, the processing results of grayscale transformation and binarization are shown in Figure 9b–d, respectively. It is evident that the two methods all performed well in extracting the crops from the background. The treble-classification Otsu method eliminated green distractors, and the white area representing the crop in the binary image was slightly more refined, which is consistent with the conditions of weak interference. In summary, under conditions with weak interference, the treble-classification Otsu method could obtain more refined crop information than typical Otsu method, although the results of both met all the requirements of image segmentation.

As the result of Figure 2, it can be observed that the treble-classification Otsu method removed most of the noise caused by duckweed, while the typical Otsu method retained almost all noise presented in the grayscale image. The aim of image binarization is to lay the foundation for the subsequent steps to achieve the final detection of crop rows. Therefore, after comparing the results of image binarization, the filtering results of binary image need to be further compared and discussed. After the filtering operation, crop information is well preserved in the binary image obtained using the treble-classification Otsu method, and the noise is almost completely eliminated. However, a large amount of

noise still remains in the binary image obtained using the typical Otsu method even after the filtering operation.

(a)   **(b)**   **(c)**   **(d)**
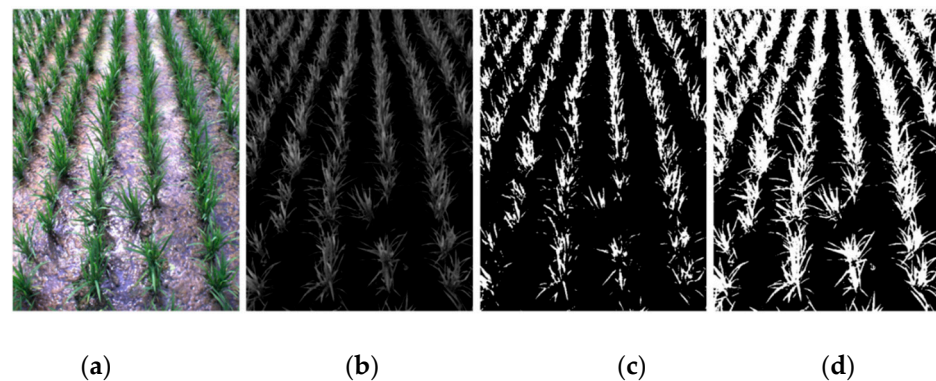
Figure 9. Validation tests of treble-classification Otsu method under weak interference: (**a**) paddy field image with weak interference (**b**) grayscale transformation; (**c**) binary image obtained using treble-classification Otsu method; (**d**) binary image obtained using typical Otsu method.

3.2.2. Efficiency Validation Tests of the Treble-Classification Otsu Method

To verify the efficiency of the treble-classification Otsu method, 100 original images under various conditions were used for testing. Firstly, all the original images were transformed to grayscale images. Subsequently, these grayscale images were processed using the treble-classification Otsu method and typical Otsu method. Strictly speaking, the complete binarization algorithm can be divided into two steps: calculating the threshold and binarizing the image. Once the threshold is obtained, the subsequent image binarization steps of treble-classification Otsu method and typical Otsu method are the same. Therefore, only the time consumed in calculating the threshold should be recorded in this validation test. After testing the 100 original images, the results of the efficiency validation test were as shown in Table 1. As can be seen, the average value of time consumed in calculating the threshold through treble-classification Otsu was 2.89 ms, while that of typical Otsu was 2.07 ms. When the size of the image was 1280×960, the average deviation of time consumed through the two methods was 0.82 ms, which would have little effect on the efficiency requirements of common autonomous navigation systems. During the experiments of image processing, original images were downsampled to the size of 640×480. When the size of the image was 640×480, the average deviation of time consumed through the two methods was further reduced to 0.41 ms.

**Table 1.** Time consumed using treble-classification Otsu method and typical Otsu method (ms).

Binarization Method	Average Value of Time Consumed 1280×960 **(ms)**	Average Value of Time Consumed 640×480 **(ms)**
Treble-classification Otsu method	2.89	0.93
Typical Otsu method	2.07	0.52

3.3. Results of Crop Row Identification

During the visual navigation, image segmentation lays the foundation for the subsequent crop row detection. To further verify the function and significance of the treble-classification Otsu method, validation experiments of crop row identification were also carried out. Figure 5 shows the result of detected crop rows tested in Figure 2a. As shown in Figure 5a, the binary image obtained using treble-classification Otsu method was divided into a series of horizontal strips to extract feature points, and the feature points extracted clearly and accurately represented the location of the green plants. The original feature points were clustered twice, and the process of horizontal clustering is shown in Figure 5b.

Figure 5c shows the result of horizontal clustering, from which it can be seen that feature points which were relatively close horizontally were clustered and the position of the cluster center could adequately represent the crop center. A slight defect in this process is that if the feature points in a horizontal strip are in the same cluster, they are forced into several clusters, such as the situation displayed at the bottom of Figure 5b. However, the interval of feature points that should have been in the same cluster is relatively small; thus, even if they are forced into several clusters, the vertical clustering can still filter out a suitable one and push it into the correct cluster. The result of vertical clustering is shown in Figure 5d. Owing to the appropriate selection criteria during vertical clustering, some feature points located between the crop rows representing duckweed or weed were eliminated. To ensure the accuracy of the fitting line and ensure that clusters represent real crop rows, the clusters with fewer than six feature points were eliminated. Thus, the final clusters with their fitting line are shown in Figure 5e, and the result of crop row detection in the original image is shown in Figure 5f. Clustering results and detection results of five illustrative original images under various conditions are shown in Figure 10. Due to errors in transplanting, some gaps may occur inside a single crop row or a rice seeding may be located between two crop rows. As shown in Figure 10, the proposed double-dimensional adaptive clustering method is not disturbed by gaps or misleading crops.
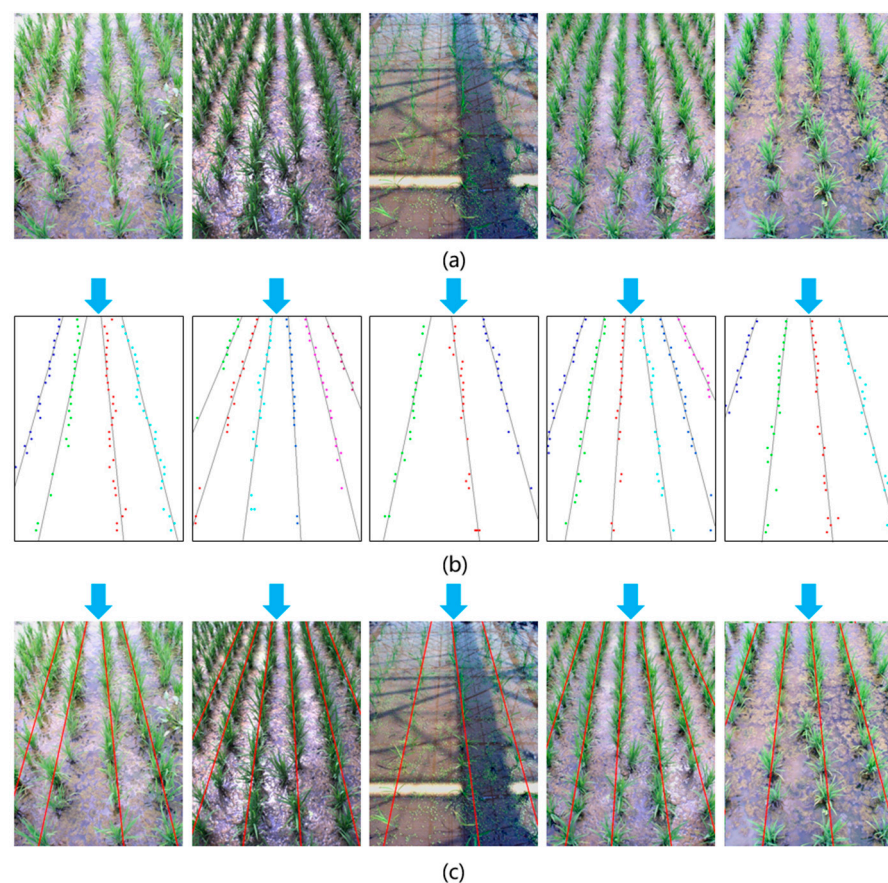
Figure 10. The clustering results and detection results: (**a**) original input images; (**b**) crop row detection; (**c**) final results.

4. Discussion

From the results of accuracy validation tests of the treble-classification Otsu method, it is clear that crop information is better distinguished and preserved in the binary image using treble-classification Otsu method. Under paddy field environments with strong interference, the grayscale image actually comprises three kinds of objects: background (non-green parts such as clear water), green distractors (duckweed, light-green weeds, and

the water surface during eutrophication), and real rice seedings. Due to their different degrees of green color, these three kinds of objects obtain different degrees of gray values. Thus, the treble-classification Otsu method divides the pixels in the greyscale image into three clusters according to their gray value, in order to distinguish the green distractors and real crops. The typical Otsu method only divides the pixels into two clusters: foreground and background; therefore, the green distractors will mix into the real crops, together in the foreground. The experimental results and analyses mentioned above verified that, under various interference environments, the treble-classification Otsu method has superior performance to the typical Otsu method.

During the efficiency validation tests of the treble-classification Otsu method, the average value of time consumed in calculating the threshold through the treble-classification Otsu method was slightly larger. By analyzing the theories of the two methods, it can be found that treble-classification Otsu requires more nested loops compared to typical Otsu; hence, the amount of calculation is larger, which will inevitably lead to lower efficiency. From the results in Table 1, it can be concluded that, when the size of images matches the industrial requirements for visual navigation [34], although the efficiency of the treble-classification Otsu method is inevitably lower, the deviation of time consumed through the two methods is small enough and can definitely meet the efficiency requirement of visual navigation.

Under the perspective of qualitative analysis, the crop row detection method achieved good visual results, as shown in Figure 10. However, the quantitative measurement of the detection accuracy is not quite straightforward because it is difficult to get true position and direction for the center lines of crop rows due to natural variations in the crop growth stage [13]. To establish a quantitative evaluation standard, a simple evaluation method is proposed.

A schematic diagram of the mechanism is given in Figure 11. In an image, assume that line l_1 is a straight line which has been detected and line l_2 is a known correct line of the same crop row. The straight line l_1 intersects the upper and lower boundaries of the image at two points T_1 and B_1, while l_2 intersects the upper and lower boundaries of the image at T_2 and B_2. In order to rigorously evaluate the similarity between these two line segments, the evaluation of both angle and distance should be considered. Make θ denote the angle between l_1 and l_2. Make d_1 denote the distance between T_1 and l_2. Make d_2 denote the distance between B_1 and l_2. The linear equations of l_1 and l_2 are assumed as follows:

$$l_1 : y = k_1 x + b_1, \tag{26}$$

$$l_2 : y = k_2 x + b_2, \tag{27}$$

where k_1 and k_2 are the slopes of l_1 and l_2, respectively, and b_1 and b_2 are the y-intercepts of l_1 and l_2, respectively. Then, the calculation formula of θ can be expressed as

$$\theta = \arctan\left(\frac{|k_2 - k_1|}{|1 + k_{1k_2}|}\right). \tag{28}$$

θ is used to evaluate the similarity of postures of l_1 and l_2. A smaller value of θ denotes more similar postures of l_1 and l_2. The calculation formulas of d_1 and d_2 can be expressed as

$$d_1 = \left|\frac{k_2 x_{T_1} - y_{T_1} + b_2}{\sqrt{k_2{}^2 + 1}}\right|, \tag{29}$$

$$d_2 = \left|\frac{k_2 x_{B_1} - y_{B_1} + b_2}{\sqrt{k_2{}^2 + 1}}\right|, \tag{30}$$

where x_{T_1} and y_{T_1} are the horizontal and vertical coordinates of respectively, and x_{B_1} and y_{B_1} are the horizontal and vertical coordinates of B_1, respectively. In order to combine the

results of d_1 and d_2, make $\bar{d}$ denote the average of d_1 and d_2. A smaller $\bar{d}$ denotes that l_1 and l_2 are closer in the distance scale. The calculation formula of $\bar{d}$ can be expressed as

$$\bar{d} = \frac{d_1 + d_2}{2}. \tag{31}$$
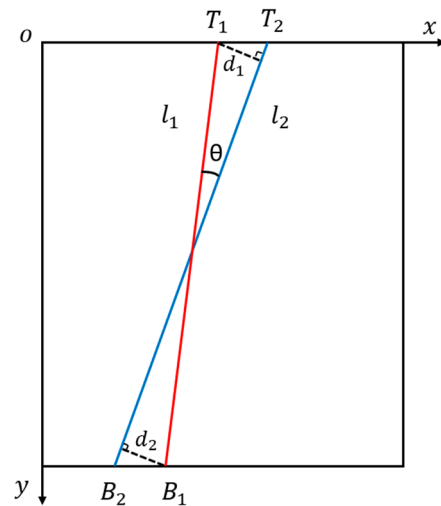
Figure 11. Schematic diagram of accuracy evaluation method.

Due to the complexity of a paddy field, traditional methods do not work well for crop row detection. Thus, the known correct crop lines can be drawn by experts to establish accuracy criterion. The quantitative evaluation method can be used to compare the results of the proposed crop row detection method and the results of an expert. As a comparison, experiments of crop row detection using the method based on typical Otsu were also conducted. The method based on typical Otsu used an image processing flow similar to the proposed method. The only difference between the two methods was the binarization process, whereby the proposed method used treble-classification Otsu and the traditional method used typical Otsu.

The comparison results and accuracy of the proposed method and detection method based on typical Otsu in eutrophication condition are presented in Figure 12 and Table 2. The average value of θ and the average value of $\bar{d}$ of each crop row detected under eutrophication conditions are presented. Obviously, the proposed method was better than the traditional method in terms of the quantitative accuracy index. From the images of crop row detection based on valid clusters, it can be found that the proposed method finished the clustering process by fewer valid points than the method based on typical Otsu. The method based on typical Otsu retained more valid feature points, but these valid points were not enough to represent the position information of the crop rows. Thus, the final results of traditional method were relatively poor. However, since the proposed method has higher screening requirements, when the image quality is not high enough and the number of feature points that can be screened out is small, the detection accuracy is likely to be greatly reduced. In contrast, the method based on typical Otsu can retain more feature points; therefore, the detection accuracy is relatively stable.

The comparison results and accuracy of the proposed method and detection method based on typical Otsu with disturbed weed are presented in Figure 13 and Table 3. The average value of θ and the average value of $\bar{d}$ of each crop row detected with disturbed weed are presented, and the proposed method is shown to be better than traditional method in terms of the quantitative accuracy index. From the images of crop row detection based on valid clusters, it can be found that feature points of the traditional method were more susceptible to interference by disturbed weed. The area where the disturbed

weed was located was mixed with the crop row area, which affected the accuracy of crop row detection.
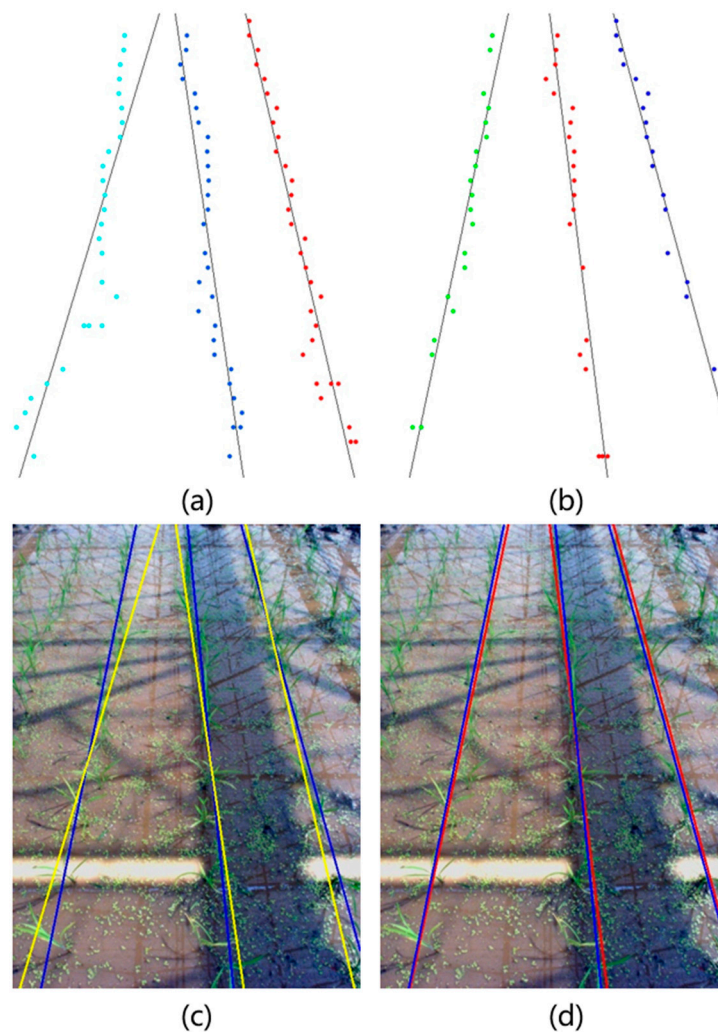
Figure 12. Comparison results of proposed method and detection method based on typical Otsu under eutrophication condition: (**a**) crop row detection based on valid clusters using proposed method; (**b**) crop row detection based on valid clusters using detection method based on typical Otsu; (**c**) detection result using proposed method; (**d**) detection result using detection method based on typical Otsu.
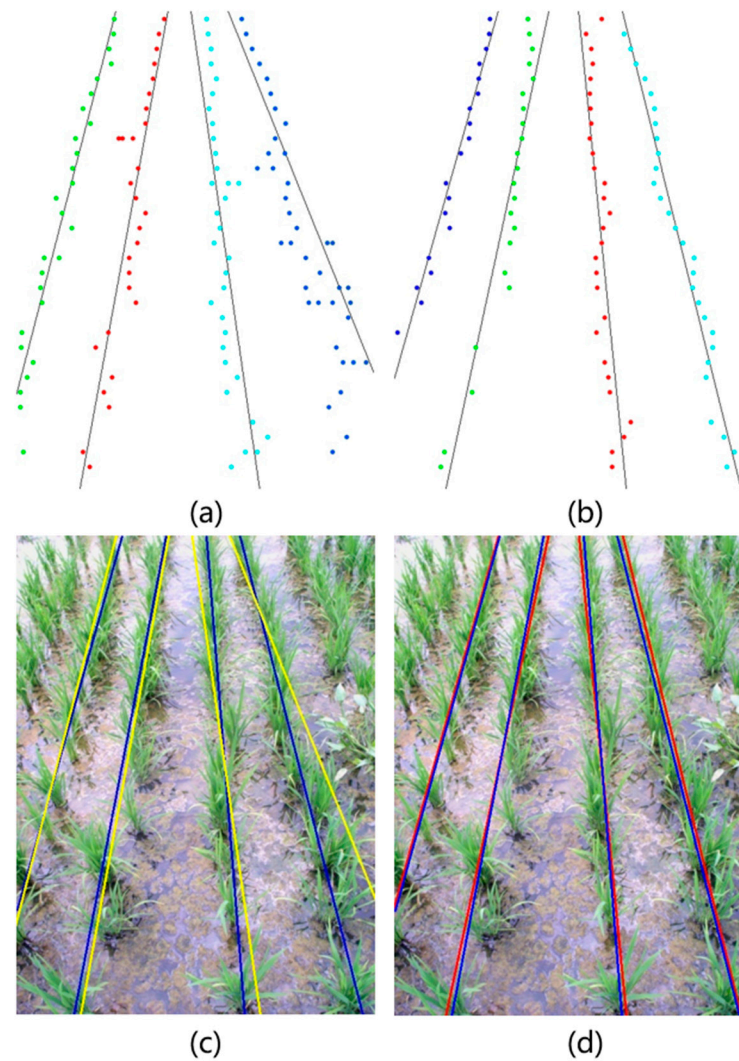
Table 2. Accuracy of the proposed crop row detection method and method based on typical Otsu under eutrophication conditions.

Detection Accuracy	Proposed Method	Detection Method Based on Typical Otsu
θ (°)	0.0125	0.0608
$\bar{d}$ (pixel)	5.12	19.99

The comparison results of five illustrative images under various conditions are shown in Figure 14. In this research, to get more convincing results, 60 images of three different conditions were tested, and the results are shown in Table 4.

In Table 4, the average value of θ and the average value of $\bar{d}$ of each crop row detected under three different conditions are presented. Through quantitative analysis, it can be seen that the detection accuracy under weak interference was the highest among all three conditions, and this is consistent with our expectation. For a total of 60 images, the average

values of θ and $\bar{d}$ were within $0.02°$ and 10 pixels, respectively. The results of the detection method based on the typical Otsu method are also shown in Table 4. For the traditional method, although good results could be achieved under weak interference, the accuracy increasingly declined when interference increased. The proposed method performed better than traditional method especially under strong interference.

Figure 13. Comparison results of proposed method and detection method based on typical Otsu with disturbed weed: (**a**) crop row detection based on valid clusters using proposed method; (**b**) crop row detection based on valid clusters using detection method based on typical Otsu; (**c**) detection result using proposed method; (**d**) detection result using detection method based on typical Otsu.

Table 3. Accuracy of the proposed crop row detection method and method based on typical Otsu with disturbed weed.

Detection Accuracy	Proposed Method	Detection Method Based on Typical Otsu
θ (°)	0.0115	0.0561
$\bar{d}$ (pixel)	4.21	19.48
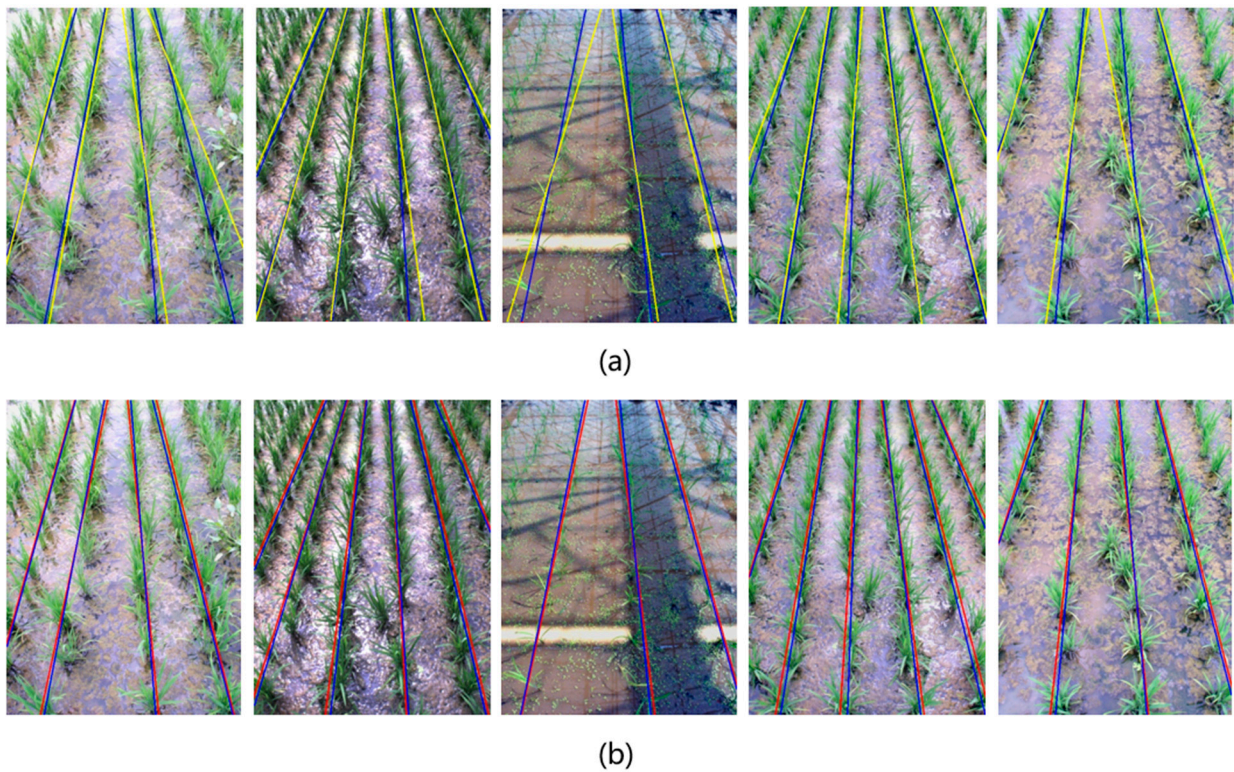
Figure 14. Comparison of the detected crop rows and the drawn rows: (**a**) detection result using detection method based on typical Otsu; (**b**) detection result using proposed method. The blue lines were drawn by experts, the yellow lines were obtained using detection method based on typical Otsu, and the red lines were obtained using the proposed approach.

Table 4. Accuracy of the proposed crop row detection method and method based on typical Otsu under three different conditions.

Detection Accuracy		Weak Interference	Moderate Interference	Strong Interference
The proposed method	θ (°)	0.0117	0.0125	0.0169
	$\bar{d}$ (pixel)	4.53	5.12	7.09
Traditional method	θ (°)	0.0196	0.0248	0.0552
	$\bar{d}$ (pixel)	8.49	13.27	17.13

Compared with previous studies [3,34], the proposed method does not require prior knowledge about the number of crop rows and does not occupy a lot of computing resources. As a result of the proposed method, the screening criteria for feature points are indirectly improved. Therefore, the field of view of the image to which this method is applicable should not be too narrow; otherwise, it would be more susceptible to interference from local extreme values than traditional methods. In short, it can be inferred from the quantitative results that, after applying the treble-classification Otsu method, the proposed feature point extraction method and the clustering algorithm could achieve better performance than the method based on typical Otsu under various interferences.

5. Conclusions

This work presented the proposal of a new integrated solution of crop row detection to deal with complex paddy field conditions. In this paper, an improved treble-classification Otsu method which can distinguish the green distractors and real crops was applied in image segmentation, and a designed double-dimensional clustering method was proposed to arrange feature points belonging to each crop row. The combination of these two methods constituted the new integrated solution. The performance of the proposed method was tested using a set of illustrative images. The efficiency validation tests showed that, when the image size was 640×480, the proposed treble-classification Otsu method only con-

sumed 0.41 ms more time than the typical Otsu method. The proposed treble-classification Otsu method was verified to meet the efficiency requirements of common autonomous navigation systems. The accuracy validation tests of the proposed method showed that the average values of θ and $\bar{d}$ were within $0.02°$ and 10 pixels, respectively, which verified that the proposed method performed better than traditional method under various conditions. In the future, this integrated solution will be developed within an embedded system to extract the guidance line for visual navigation of unmanned agricultural vehicles working in complex paddy fields. With this crop row detection solution, the robustness of visual navigation systems in paddy fields could be enhanced.

Author Contributions: Conceptualization, Y.L. and Y.H.; methodology, Y.L. and Y.Y.; software, Y.Y.; investigation, H.C. and N.Z.; writing—original draft preparation, Y.Y.; writing—review and editing, Y.L.; visualization, J.W.; supervision, Y.B. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhai, Z.; Zhu, Z.; Du, Y.; Song, Z.; Mao, E. Multi-crop-row detection algorithm based on binocular vision. *Biosyst. Eng.* **2016**, *150*, 89–103. [CrossRef]
2. Liu, Y.; Noguchi, N.; Liang, L. Development of a positioning system using UAV-based computer vision for an airboat navigation in paddy field. *Comput. Electron. Agric.* **2019**, *162*, 126–133. [CrossRef]
3. Zhang, Q.; Shaojie Chen, M.E.; Li, B. A visual navigation algorithm for paddy field weeding robot based on image understanding. *Comput. Electron. Agric.* **2017**, *143*, 66–78. [CrossRef]
4. Woebbecke, D.M.; Meyer, G.E.; Bargen, K.V.; Mortensen, D.A. Color Indices for Weed Identification Under Various Soil, Residue, and Lighting Conditions. *Trans. Asae* **1995**, *38*, 259–269. [CrossRef]
5. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. ManCybern.* **1979**, *9*, 62–66. [CrossRef]
6. Søgaard, H.T.; Olsen, H.J. Determination of crop rows by image analysis without segmentation. *Comput. Electron. Agric.* **2003**, *38*, 141–158. [CrossRef]
7. Sainz-Costa, N.; Ribeiro, A.; Burgos-Artizzu, X.P.; Guijarro, M.; Pajares, G. Mapping Wide Row Crops with Video Sequences Acquired from a Tractor Moving at Treatment Speed. *Sensors* **2011**, *11*, 7095–7109. [CrossRef]
8. Hough, P.V.C. A Method and Means for Recognizing Complex Patterns. U.S. Patent 3,069,654, 18 December 1962.
9. Wera, W.; Veronika, F.F.; Christian, D.; Wolfram, B. Crop Row Detection on Tiny Plants With the Pattern Hough Transform. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3394–3401.
10. Billingsley, J.; Schoenfisch, M. Vision-guidance of agricultural vehicles. *Auton. Robot.* **1995**, *2*, 65–76. [CrossRef]
11. Montalvo, M.; Pajares, G.; Guerrero, J.M.; Romeo, J.; Guijarro, M.; Ribeiro, A.; Ruz, J.J.; Cruz, J.M. Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Syst. Appl. Int. J.* **2012**, *39*, 11889–11897. [CrossRef]
12. Guerrero, J.M.; Guijarro, M.; Montalvo, M.; Romeo, J.; Emmi, L.; Ribeiro, A.; Pajare, G. Automatic expert system based on images for accuracy crop row detection in maize field. *Expert Syst. Appl.* **2013**, *40*, 656–664. [CrossRef]
13. Jiang, G.; Wang, Z.; Liu, H. Automatic detection of crop rows based on multi-ROIs. *Expert Syst. Appl.* **2015**, *42*, 2429–2441. [CrossRef]
14. García-Santillán, I.; Guerrero, J.M.; Montalvo, M.; Pajares, G. Curved and straight crop row detection by accumulation of green pixels from images in maize fields. *Precis. Agric.* **2018**, *19*, 18–41.
15. Guerrero, J.M.; Ruz, J.J.; Pajares, G. Crop rows and weeds detection in maize fields applying a computer vision system based on geometry. *Comput. Electron. Agric.* **2017**, *142*, 461–472. [CrossRef]
16. Basso, M.; Edison, P.D.F. A UAV Guidance System Using Crop Row Detection and Line Follower Algorithms. *J. Intell. Robot. Syst.* **2019**, *97*, 605–621. [CrossRef]

17. Tenhunen, H.; Pahikkala, T.; Nevalainen, O.; Teuhola, J.; Mattila, H.; Tyystjärvi, E. Automatic detection of cereal rows by means of pattern recognition techniques. *Comput. Electron. Agric.* **2019**, *162*, 677–688. [CrossRef]
18. Li, S.; Zhang, Z.; Du, F.; He, Y. A New Automatic Real-time Crop Row Recognition based on SoC-FPGA. *IEEE Access* **2020**, *8*, 37440–37452. [CrossRef]
19. Rabab, S.; Badenhorst, P.; Chen, Y.P.P.; Daetwyler, H.D. A template-free machine vision-based crop row detection algorithm. *Precis. Agric.* **2021**, *22*, 124–153. [CrossRef]
20. Romeo, J.; Pajares, G.; Montalvo, M.; Guerrero, J.M.; Guijarro, M.; Ribeiro, A. Crop Row Detection in Maize Fields Inspired on the Human Visual Perception. *Sci. World J.* **2012**. [CrossRef] [PubMed]
21. Kise, M.; Zhang, Q. Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance. *Biosyst. Eng.* **2008**, *101*, 191–198. [CrossRef]
22. Fue, K.; Porter, W.; Barnes, E.; Li, C.; Rains, G. Evaluation of a Stereo Vision System for Cotton Row Detection and Boll Location Estimation in Direct Sunlight. *Agronomy* **2020**, *10*, 1137. [CrossRef]
23. Adhikari, S.P.; Kim, G.; Kim, H. Deep Neural Network-Based System for Autonomous Navigation in Paddy Field. *IEEE Access* **2020**, *8*, 71272–71278. [CrossRef]
24. Ponnambalam, V.R.; Bakken, M.; Moore, R.J.D.; Gjevestad, J.G.O.; From, P.J. Autonomous Crop Row Guidance Using Adaptive Multi-ROI in Strawberry Fields. *Sensors* **2020**, *20*, 5249. [CrossRef] [PubMed]
25. Guijarro, M.; Pajares, G.; Riomoros, I.; Herrera, P.J.; Burgos-Artizzu, X.P.; Ribeiro, A. Automatic segmentation of relevant textures in agricultural images. *Comput. Electron. Agric.* **2011**, *75*, 75–83. [CrossRef]
26. Hamuda, E.; Glavin, M.; Jones, E. A survey of image processing techniques for plant extraction and segmentation in the field. *Comput. Electron. Agric.* **2016**, *125*, 184–199. [CrossRef]
27. Woebbecke, D.; Meyer, G.; Von Bargen, K.; Mortensen, D. Plant Species Identification, Size, and Enumeration Using Machine Vision Techniques on Near-Binary Images. *Proc. SPIE* **1993**, *1836*.
28. Kataoka, T.; Kaneko, T.; Okamoto, H.; Hata, S. Crop growth estimation system using machine vision. In Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Kobe, Japan, 20–24 July 2003.
29. Hague, T.; Tillett, N.D.; Wheeler, H. Automated Crop and Weed Monitoring in Widely Spaced Cereals. *Precis. Agric.* **2006**, *7*, 21–32. [CrossRef]
30. Zhang, D.; Mansaray, L.R.; Jin, H.; Sun, H.; Kuang, Z.; Huang, J. A universal estimation model of fractional vegetation cover for different crops based on time series digital photographs. *Comput. Electron. Agric.* **2018**, *151*, 93–103. [CrossRef]
31. Matsui, N.; Mita, K. Method of Reducing Digital Images. U.S. Patent 4,931,881, 05 June 1990.
32. Guoquan, J.; Cuijun, Z. A vision system based crop rows for agricultural mobile robot. In Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China, 22–24 October 2010; pp. V11-142–V111-145.
33. Meyer, G.E.; Neto, J.O.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **2008**, *63*, 282–293. [CrossRef]
34. Zhang, X.Y.; Li, X.N.; Zhang, B.H.; Zhou, J.; Tian, G.Z.; Xiong, Y.J.; Gu, B.X. Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method. *Comput. Electron. Agric.* **2018**, *154*, 165–175. [CrossRef]