

Article

A Proposal for Lodging Judgment of Rice Based on Binocular Camera

Yukun Yang ^{1,2}, Chuqi Liang ², Lian Hu ^{2,*} , Xiwen Luo ^{1,2}, Jie He ² , Pei Wang ², Peikui Huang ², Ruitao Gao ² and Jiehao Li ² 

¹ College of Biological and Agricultural Engineering, Jilin University, Changchun 130022, China; ykyang21@mails.jlu.edu.cn (Y.Y.)

² Key Laboratory of Key Technology on Agricultural Machine and Equipment, Ministry of Education, College of Engineering, South China Agricultural University, Guangzhou 510642, China

* Correspondence: lianhu@scau.edu.cn

Abstract: Rice lodging is a crucial problem in rice production. Lodging during growing and harvesting periods can decrease rice yields. Practical lodging judgment for rice can provide effective reference information for yield prediction and harvesting. This article proposes a binocular camera-based lodging judgment method for rice in real-time. As a first step, the binocular camera and Inertial Measurement Unit (IMU) were calibrated. Secondly, Census and Grayscale Level cost features are constructed for stereo matching of left and right images. The Cross-Matching Cost Aggregation method is improved to compute the aggregation space in the LAB color space. Then, the Winner-Takes-All algorithm is applied to determine the optimal disparity for each pixel. A disparity map is constructed, and Multi-Step Disparity Refinement is applied to the disparity map to generate the final one. Finally, coordinate transformation obtains 3D world coordinates corresponding to pixels. IMU calculates the real-time pose of the binocular camera. A pose transformation is applied to the 3D world coordinates of the rice to obtain its 3D world coordinates in the horizontal state of the camera (pitch and roll angles are equal to 0). Based on the distance between the rice and the camera level, thresholding was used to determine whether the region to be detected belonged to lodging rice. The disparity map effect of the proposed matching algorithm was tested on the Middlebury Benchmark v3 dataset. The results show that the proposed algorithm is superior to the widely used Semi-Global Block Matching (SGBM) stereo-matching algorithm. Field images of rice were analyzed for lodging judgments. After the threshold judgment, the lodging region results were accurate and could be used to judge rice lodging. By combining the algorithms with binocular cameras, the research results can provide practical technical support for yield estimation and intelligent control of rice harvesters.

Keywords: stereo matching; binocular camera; image processing; lodging judgment; rice



Citation: Yang, Y.; Liang, C.; Hu, L.; Luo, X.; He, J.; Wang, P.; Huang, P.; Gao, R.; Li, J. A Proposal for Lodging Judgment of Rice Based on Binocular Camera. *Agronomy* **2023**, *13*, 2852.

<https://doi.org/10.3390/agronomy13112852>

Academic Editor: Enrico Corrado Borgogno Mondino

Received: 30 October 2023

Revised: 13 November 2023

Accepted: 15 November 2023

Published: 20 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rice is one of the world's major food crops, and 1/3 of the world's population consumes rice as a staple food [1]. The increasing world population places higher demands on rice yield and quality [2], and rice production should increase by 50% by 2025 to accommodate the increase in world population [3]. In China, rice production is highly affected by unfavorable meteorological conditions, such as typhoons, heavy rains, hailstorms, and improper crop management, which can result in rice losses immediately [4]. The earlier crop lodging occurs, the more serious the impact on crop growth. When rice undergoes severe lodging, the yield reduction rate may be higher than 20% [5]. Therefore, rice lodging judgment has essential research value for predicting rice yield, disaster analysis, and mechanical harvesting.

Researchers and producers judge rice lodging as a primary concern. The existing methods for assessing lodging are usually statistical or traditional ground surveys [6–8]. It is less automated and more time-consuming to use these methods. Additionally, assessment

results are subjective since different inspectors conducting evaluations on the same scope may produce different results. There is a lower level of accuracy and a more significant influence of human factors.

Besides traditional methods of lodging determination, crop spectral analysis of satellite images and optical remote sensing based on radar systems have also been used [9,10]. Using Gaofen-2 satellite images with high spatial resolution, Tang et al. proposed a Pyramid Transposition Convolutional Network (PTCNet) model for large-scale wheat lodging extraction and detection [11]. Chauhan et al. demonstrated the potential of Sentinel-1 and Sentinel-2 data for real-time detection of wheat lodging [12]. Sun et al. developed a remote sensing method for monitoring rice lodging rank based on Sentinel-2A change vector analysis before and after lodging. A model based on the magnitude of change in vegetation index can effectively monitor rice lodging rank using remote sensing [13]. Zhao et al. used polarization features of polarized Synthetic Aperture Radar (SAR) data to characterize lodging wheat and oilseed rape. Their detection results depend on the crop's canopy structure [14]. Chauhan et al. estimated crop inclination through SAR remote sensing data to indicate lodging severity at the lodging stage [15].

Generally, lodging judgment can be realized through spectral analysis of lodging and non-lodging regions using satellite images, but lodging assessment is still limited by spectral band characteristics [9,10]. It is difficult to determine the lodging region when there is little spectral difference between lodging and non-lodging crops [16]. Additionally, the spectral characteristics of the rice region can be affected by changes in soil conditions or other crop stresses [17]. The SAR provides data for all weather conditions and is sensitive to changes in crop structure. Optical remote sensing based on radar systems suits large and relatively homogeneous areas. The subsequent analysis of images in small collocated regions will be complex and inaccurate when the spatial resolution of the image and scanner is relatively small [18,19].

Using unmanned aerial vehicles (UAV) in lodging judgments has also been extensively studied [20]. Varela et al. analyzed sorghum lodging by using 3D-CNNs with an accuracy of 0.88, suggesting that spatiotemporal convolutional neural network (CNN) architectures based on UAV time series images are promising for lodging judgments [21]. Su et al. proposed an end-to-end, pixel-to-pixel semantic segmentation method to identify rice lodging, which can process input multi-band images, and the accuracy of the proposed lodging detection model on rice lodging images is 97.30% [22]. Based on UAV and CNN, Zhao et al. developed a method for wheat lodging detection, and the average accuracy of the model was 89.23% [23]. A study by Zhang et al. evaluated and compared the performance of traditional machine learning and deep learning methods in detecting lodgings and found that the Random Forest algorithm performed the best in detecting lodgings [24]. However, given the additional cost of UAVs, farmers may not be willing to bear the extra expense. In addition, UAVs, satellites, and radars cannot provide real-time and practical lodging information for the ground machine, and UAV lodging information is not synchronous or real-time. It is tough to synchronize and cooperate with information from field management and harvesting machinery.

With the advancement of machine vision technology, more and more ground-working machines are equipped with vision sensors. The visual sensor can capture images of the environment surrounding the working machine and detect the state and information of the operating object. Haridasan et al. developed a rice disease detection network using ReLU and softmax functions and achieved the highest verification accuracy of 0.9415 [25]. He et al. proposed a MobileV2-UNet deep learning model, which enables automatic detection and navigation of rice fields [26]. Ma et al. developed an algorithm based on linear clustering and supervised learning to detect rice root rows. In the experiments, the accuracy rates of root-row detection during day 6, day 20, and day 35 after transplantation were 96.79%, 90.82%, and 84.15%, respectively [27]. Machine vision has many applications in rice production, but studies have yet to be conducted on lodging judgments. Wen et al. developed a binocular vision-based method to detect wheat lodging positions, contours,

and areas in the harvesting region of a combined harvester [28]. Wen's study shows that binocular vision can be used in rice lodging research. It provides real-time and reliable information about rice lodging for ground machines such as rice harvesters. The rice harvester controls the cutting table to complete the harvest of lodging rice, thus reducing the yield reduction caused by lodging.

Satellites, radars, and UAVs cannot transmit lodging information in real-time to ground vehicles. Machine vision provides high accuracy and real-time synchronization and is more cost-effective than the three common methods of lodging determination. Machine vision technologies include deep learning [29,30] and image processing [31]. Lodging can be detected through a monocular camera combined with deep learning [32]. Data collection costs and deployment difficulties must be considered. Binocular cameras combined with deep learning are mainly used in large outdoor objects [33]. Obtaining standard disparity maps of small moving objects for training neural networks is difficult. According to Wen et al. study [28], binocular cameras with image processing can be used to detect lodgings. In this study, Wen et al. [28] used the SGBM stereo-matching algorithm and assumed that the camera's pose remained constant. In the actual process, the camera's pose will change due to bumps on the road. The SGBM algorithm's processing accuracy is also easily affected by external conditions such as repetitive textures and light changes [34].

Based on this research background, this article proposes an algorithm for rice lodging judgment based on binocular vision. The work has three main objectives: (1) The stereo-matching binocular vision algorithm is studied and improved. (2) Compare our improved matching algorithm with the commonly used matching algorithm regarding disparity maps and metrics. (3) A feasibility test was conducted with field images using the algorithm for rice lodging judgment.

2. Materials and Methods

2.1. Algorithm Flow

The rice lodging judgment algorithm proposed in this article consists of four steps: camera and IMU calibration, matching cost calculation, cost aggregation, disparity calculation, and lodging judgment. A flow chart of the proposed rice lodging judgment is shown in Figure 1.

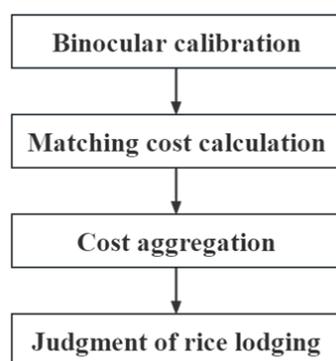


Figure 1. Algorithm flow.

The detailed steps are as follows: (1) Binocular calibration. Calibrate the binocular camera to obtain internal and external parameters. (2) Matching cost calculation. Based on Census matching costs, a Grayscale Level matching cost is proposed. For subsequent cost aggregation, Census and Grayscale Level matching costs are merged. (3) Cost aggregation. Optimize the Cross-Matching Cost Aggregation method, determine the aggregation range based on the LAB color space of pixels, and calculate all disparity aggregation costs for pixels within the disparity range one by one. (4) Judgment of rice lodging. With the Winner-Takes-All algorithm, the optimal disparity is obtained at the lowest aggregation cost within the pixel's disparity range. The disparity map optimization eliminates the

wrong matching points and fills in part of the weak texture area. The IMU determines the camera's three-axis pose angle. A coordinate transformation determines the distance from the rice to the camera level. A threshold judgment of the distance between the rice and the camera level determines whether the acquisition region belongs to the lodging region. Figure 2 shows the relative positions of the camera and the rice. The distance from the rice to the camera level is y .

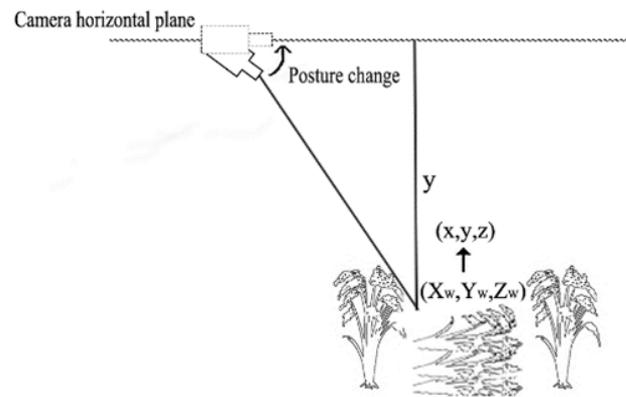


Figure 2. Schematic diagram of coordinate transformation.

2.2. Binocular Calibration

When the binocular camera is calibrated, the internal parameters of the camera, distortion parameters, and spatial position parameters of the two cameras can be obtained, which is of great importance for the subsequent correction of distortion and coordinate conversion [35]. In this study, a ZED 2 binocular camera is used. The collected left and right tessellated images are calibrated using MATLAB's Stereo Camera Calibrator Binocular Calibration Toolbox. To ensure calibration accuracy, a glass substrate checkerboard grid is selected. Image pairs with significant calibration errors are removed. The calibration results are output to obtain the internal and external parameter matrices and the right camera's translation and rotation matrices. The images captured by the left and right cameras are preprocessed using Gaussian filters. OpenCV library performs line and distortion correction on the left and right images. The images processed in the subsequent steps are polar-corrected. The polar corrected images ensure that the sliding window search is conducted on a polar line, improving the algorithm's efficiency and accuracy.

2.3. Traditional Census Transform

A Census transform [36] compares the gray values of all pixels within a window with the gray value of the center pixel p . When the gray value of a pixel point is greater than the gray value of the center pixel p , the position of the pixel point is recorded as 0, and vice versa. The formula is shown below:

$$\sigma(I(p), I(q)) = \begin{cases} 0 & I(p) \leq I(q) \\ 1 & I(p) > I(q) \end{cases} \quad (1)$$

where $\sigma(I(p), I(q))$ is the comparison function, and the result of the completed comparison is a bit string, which is used as the Census feature of the window centered on center pixel p . The following formula calculates the Census feature:

$$\text{Census}(p) = \bigotimes_{q \in N} \sigma(I(p), I(q)) \quad (2)$$

where \bigotimes represents the connection by bit, and N represents the total number of pixels in the window centered on pixel p . Census features are computed for the window corresponding to the left and right images captured by the binocular camera. A dissimilarity operation is

performed on the computed left and right bit strings. Census feature matching costs can be calculated using the Hamming distance similarity metric as follows:

$$Cost_{census}(P_L, P_R) = \text{Hamming}(Census(P_L), Census(P_R)) \tag{3}$$

where $Census(P_L)$ and $Census(P_R)$ denote the Census bit-string features computed in the left and right windows, respectively.

Consider a 3×3 window as an example (shown in Figure 3). Census feature calculation results for the left and right windows are shown in Figure 3. 00001011 for the left window and 10010100 for the right window represent Census features in bit-string format. The bit-string Hamming distance of the left and right windows is calculated as the Census feature matching cost value under the center pixel. The Hamming distance calculation for this left and right window is 6.

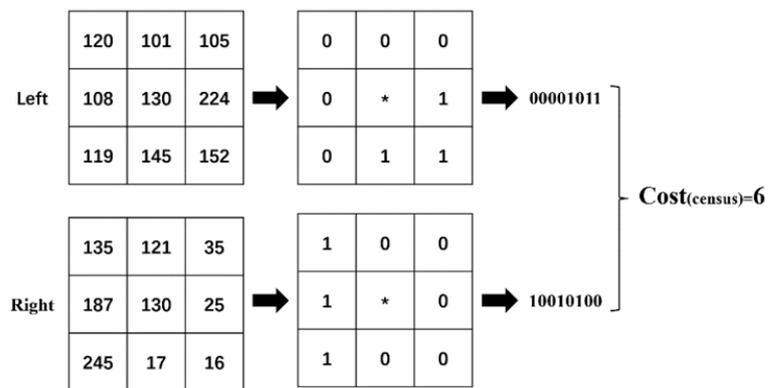


Figure 3. Schematic diagram of Census cost calculation (* represents the center pixel).

2.4. Grayscale Level

The Census transform is robust to luminance changes and image noise. However, Census feature costs can be easily miscalculated in similar localized windows. Figure 4 shows the field rice image with four windows, including soil, leaf, tassel, and road. In the four windows, the gray value of the center pixel is similar to the gray value of the surrounding pixels. The Census transform does not produce significant differences in Census matching feature costs, resulting in matching errors if the objects within the four windows are considered identical. We introduce a Grayscale Level feature and the corresponding cost to address this issue to compensate for the lack of Census matching costs.



Figure 4. Image of rice in the field.

The Grayscale Level divides the grayscale values of pixel points in the left and right images into five ranges. The cost function is constructed by determining whether the

Grayscale Level values of the same pixel points within the left and right image windows are the same. For a single pixel point, the Grayscale Level feature is calculated as follows:

$$\text{Grayscale Level}(q) = \begin{cases} 1 & \text{when Pixel } (q) \in [0, 51) \\ 2 & \text{when Pixel } (q) \in [51, 102) \\ 3 & \text{when Pixel } (q) \in [102, 153) \\ 4 & \text{when Pixel } (q) \in [153, 204) \\ 5 & \text{when Pixel } (q) \in [204, 255] \end{cases} \quad (4)$$

where pixel (q) denotes the gray value at any pixel point q , and $\text{Grayscale Level}(q)$ denotes the Grayscale Level feature value at pixel point q . A cost value is calculated for each unequal Grayscale Level value within the left and right image windows, and the formula is as follows:

$$\text{Cost}_{\text{Grayscale Level}}(P_L, P_R) = \sum (\text{Grayscale Level}(P_L) \neq \text{Grayscale Level}(P_R)) \quad (5)$$

where $\text{Cost}_{\text{Grayscale Level}}(P_L, P_R)$ denotes the cost of the Grayscale Level, $\text{Grayscale Level}(P_L)$ and $\text{Grayscale Level}(P_R)$ denote the Grayscale level feature value at the same location within the left and right image windows, respectively.

Consider the Grayscale Level cost calculation for a 3×3 window. In Figure 5, the left and right windows of 3×3 size are filled with random grayscale values. The Grayscale Level values for the pixel points in the left and right windows are calculated based on Equation (4). The Grayscale Level values at the same pixel points in the left and right windows are compared using Equation (5). The number of unequal Grayscale Level values is used as a cost value for the center pixel of the window. The Grayscale Level feature cost value was calculated for the left and right windows as 6.

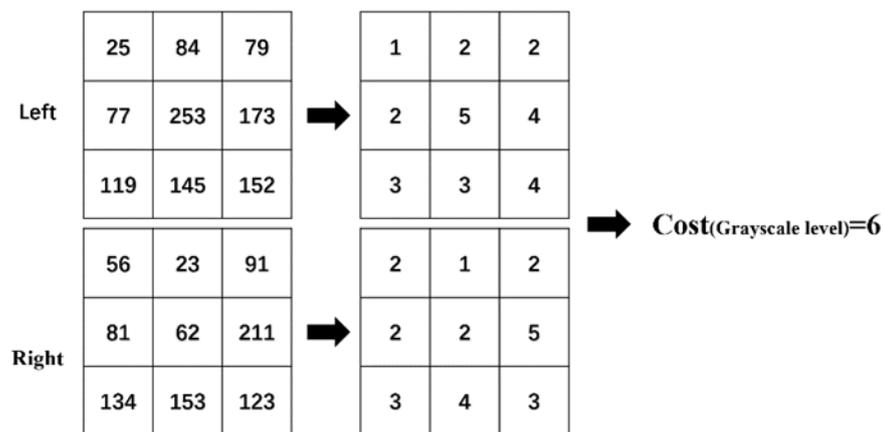


Figure 5. Schematic diagram of Grayscale Level cost calculation.

2.5. Cost Aggregation

A single-pixel window has a limited neighborhood range, and its matching cost can easily be disturbed by noise. It is necessary to aggregate neighboring pixels' matching costs to achieve more accurate matching results. In this article, we optimize the Cross-Matching Cost Aggregation method proposed by Ke Zhang et al. [37]. The cost aggregation space is based on a simple but valid assumption that neighboring pixels with similar colors should belong to the same region. However, RGB colors in the outdoors are highly affected by lighting. Therefore, we implement cost aggregation in LAB color space.

For a given pixel, the cost aggregation space is calculated as follows:

$$\begin{aligned}
 D_c(p_l, p) &< \tau_1 \text{ and } D_c(p_l, p_l + (1,0)) < \tau_1 \\
 D_s(p_l, p) &< L_1 \\
 D_c(p_l, p) &< \tau_2, \text{ if } L_2 < D_s(p_l, p) < L_1
 \end{aligned}
 \tag{6}$$

where p is the center of the cost space, p_l is the extension point of the cost space, $D_s(p_l, p)$ is the straight line distance between pixels p_l and p , τ_1 , and τ_2 are the LAB color thresholds ($\tau_1 > \tau_2$), L_1 and L_2 are the distance thresholds of the adaptive window ($L_1 > L_2$), $p_l + (1,0)$ is the neighboring pixel of p_l , and $D_c(p_l, p)$ is the maximal difference between pixels p_l and p in the LAB color space, which is calculated as shown in the formula below:

$$D_c(p_l, p) = \max_{i=L,A,B} |I_i(p_l) - I_i(p)|
 \tag{7}$$

where I_i denotes the intensity value of this pixel under component i . Equation (6) restricts the color difference between p_l and p at a distance less than L_2 by τ_1 and the color difference between p_l and the neighboring pixel, effectively ensuring the aggregation space's boundary stability. When the distance exceeds L_2 , a more stringent color threshold τ_2 is given to ensure the effects of similar colors and lighting.

Figure 6 illustrates the steps involved in cost aggregation in this article. In the first step, the practical length information is calculated for the four horizontal and vertical directions at pixel center A and saved. Once the length information for the four directions has been determined, the second step calculates the horizontal direction lengths of all pixels based on the vertical direction lengths. This calculation is completed and saved to get the adaptive cost aggregation region for pixel center A. The third step involves adding all pixels in the vertical direction of pixel A, along with their corresponding horizontal arms. The horizontal cost aggregation is completed. The fourth step consists of summing the costs of pixel center A after aggregation in the horizontal direction along the vertical path. The final aggregated cost of this pixel A is obtained after the operation is completed.

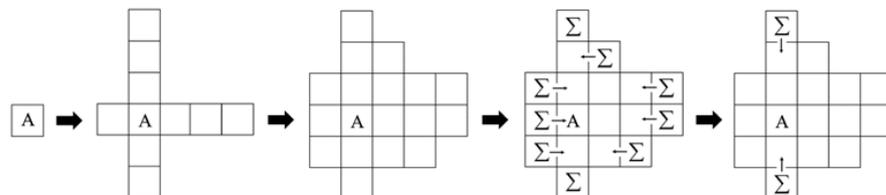


Figure 6. Schematic diagram of the cost aggregation step.

Cost aggregation will be performed on all pixels within a given disparity range. After the cost aggregation is completed, the Winner-Takes-All algorithm is applied to find the disparity with the lowest aggregation cost within the disparity range as the final disparity for each pixel. For the disparity optimization stage, we use the Multi-Step Disparity Refinement proposed by Xing Mei et al. [38] to eliminate false matches and fill some regions with weak textures.

After the disparity calculation is completed, a disparity map corresponding to the left camera image can be obtained. Furthermore, the disparity information of each pixel can be used to determine the scale information corresponding to its space. Comparatively to the coordinate transformation of a monocular camera, the disparity allows us to determine the distance between the object and the camera. The world coordinates of the object in the binocular camera can be obtained. Based on the disparity map, the following formula can be used to calculate the transformation of an image pixel into the world coordinate system:

$$[X, Y, Z, W]^T = Q^*[x, y, d, 1]^T
 \tag{8}$$

where X , Y , and Z are the spatial 3D coordinates corresponding to the image pixels, Q is the reprojection matrix obtained through stereo correction, x and y are the pixel position coordinates on the disparity map, and d represents the disparity value associated with the pixel coordinates. The reprojection matrix formula is shown below:

$$Q = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f \\ 0 & 0 & \frac{-1}{T_x} & \frac{c_x - c'_x}{T_x} \end{bmatrix} \quad (9)$$

where c_x and c_y denote the coordinates of the left camera's principal point in the image, f denotes the focal length, T_x denotes the translation between the projection centers of the two cameras (negative), and c'_x is the coordinates of the right camera's principal point. The reprojection matrix parameters in Equation (9) are all obtained from binocular calibration. When Equation (9) is substituted for Equation (8), the concrete expression for (X, Y, Z) is obtained.

$$Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} x - c_x \\ y - c_y \\ f \\ \frac{-d + c_x - c'_x}{T_x} \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \quad (10)$$

As shown in Equation (11), the corresponding 3D coordinates of pixels under the left camera coordinate system are calculated. This article establishes the world coordinate system for the left camera. At this time, the 3D coordinates of the object in space are the 3D world coordinates (X_W, Y_W, Z_W) under the left camera coordinate system.

$$(X_W, Y_W, Z_W) = (X/W, Y/W, Z/W) \quad (11)$$

2.6. Judgment of Rice Lodging

The difference between lodging and non-lodging is posture. By comparing the height values of rice in the region, we can determine if the rice region is lodging or not. However, it is not easy to directly measure rice height using a binocular camera due to the density and occlusion of the rice region. Furthermore, the camera's field of view is insufficient to capture rice's entirety. At the same time, the ability to acquire ground height is also limited. Based on this, an IMU and binocular camera method for rice lodging judgment was proposed in this article. By converting the rice canopy height relative to the ground, the camera plane was used as the "ground" to measure the rice height indirectly, and the lodging of the rice was assessed.

By calculating the disparity map and the reprojection matrix, the 3D coordinates under the world coordinate system correspond to the pixels in the image coordinate system. The 3D world coordinate corresponds to the camera's tilted pose. We calculate the distance between the canopy and the camera plane by acquiring the image while obtaining the camera pose information. Camera coordinates in the tilted state are converted into camera coordinates in the horizontal attitude (no pitch or roll angle). At this time, the y-coordinate value indicates the distance between the target and the camera's horizontal plane. Figure 2 illustrates the schematic.

An IMU module is pre-installed on the selected ZED 2 camera. This IMU module has no relative rotation or displacement to the binocular camera. Binocular cameras and IMU modules are considered rigid bodies. The IMU's three-axis orientation was calibrated and rotated through matrix variation to align with the left camera. The IMU's three-axis pose angle is used as the camera's three-axis pose angle. Camera pose measurements by IMU are retrieved using the API and expressed as quaternions. Scipy converts the quaternions into a rotation matrix. The rotation matrix corresponds to the camera's three-axis pose. The inverse of this rotation matrix is obtained and left-multiplied to the target's 3D world

coordinates. The result is the 3D world coordinates of the target in the camera's horizontal pose. The calculation formula is shown below:

$$[x, y, z]^T = R@[X_w, Y_w, Z_w]^T \quad (12)$$

where x, y, z are the 3D world coordinates of the camera in the horizontal pose, R is the rotation matrix corresponding to the current camera pose measured by the IMU, $@$ denotes matrix multiplication, and (X_w, Y_w, Z_w) denote the 3D world coordinates of the camera in the current pose.

An image of a rice field collected for lodging judgment is easily interfered with by noise, such as rice movement or shadows. To minimize noise interference, the rice region to be detected was divided into equally spaced regions of interest (ROI). Calculate the average distance between all targets in the ROI and the camera level. A distance threshold TH ($0 < TH$) is given. Define the ROI as an un-lodging region when $\bar{y} \in (0, TH)$, and define the ROI as a lodging region when $\bar{y} > TH$. Figure 7 illustrates the judgment schematic for one of the rice field images collected. The middle and lower parts of the image were divided into ROIs with equal spacing. The distance between the rice in each ROI and the camera level can be determined. Given the distance threshold TH, the lodging regions are ROI 4, and ROI 3 can be obtained by threshold judgment.

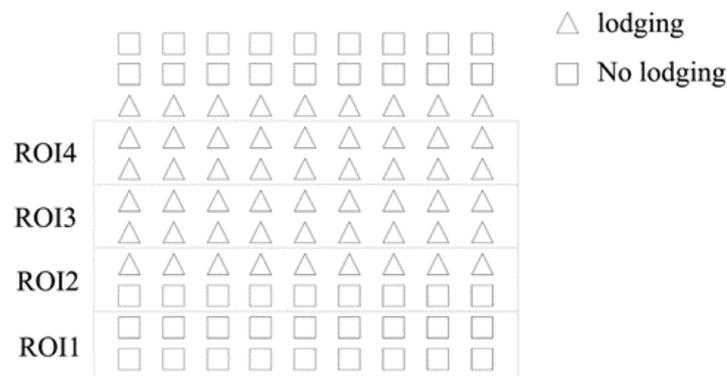


Figure 7. Schematic diagram of lodging judgment.

3. Results and Analysis

3.1. Disparity Map Acquisition Effect

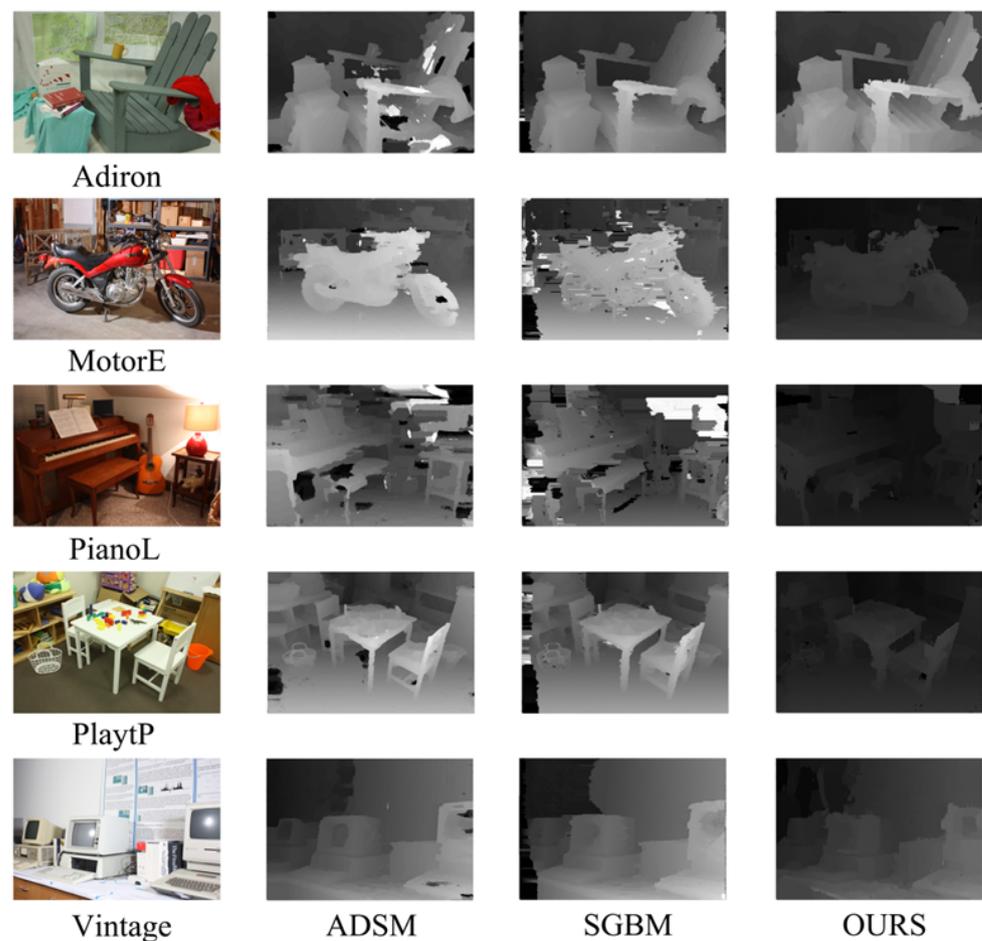
Five standard test images from Middlebury Benchmark v3 were selected for testing to verify the improved algorithm's effectiveness. These images included Adiron, MotorE, PianoL, PlaytP, and Vintage. This laptop runs on Intel Core i7-12700H, GPU NVIDIA GeForce RTX 3050, RAM 16 G, Windows 11 (64-bit), program development software PyCharm 2023, and the algorithm is written in Python.

We compare the proposed algorithm with the ADSM algorithm [39] and the SGBM algorithm integrated with the OpenCV library to verify its accuracy and effectiveness. Among them, ADSM is a more advanced stereo-matching algorithm improved upon Census. SGBM is a more mature stereo-matching algorithm. The root-mean-square disparity error in pixels (RMS) and average absolute disparity error in pixels (Avgerr) are used as effect measures. The results of the three algorithms are shown in Table 1. Table 1 shows that the algorithm proposed in this article are better than the SGBM algorithm.

Table 1. Algorithm test results.

Algorithm	Adiron		MotorE		PianoL		PlaytP		Vintage	
	RMS	Avgerr	RMS	Avgerr	RMS	Avgerr	RMS	Avgerr	RMS	Avgerr
ADSM	38.1	14.3	26.6	8	41.9	20.4	18.7	5.84	34	11.1
SGBM	23.3	7.07	52.5	21.3	55.2	29	31.5	9.97	42.3	16.1
OURS	20.6	6.94	23.6	8.7	33.7	15.7	19.3	8.77	25.1	14.3

Regarding index parameters, ADSM has similar results to the algorithm presented in this article. However, the index parameters of some ADSM images are substantially lower than our algorithm. Table 1 indicates that the algorithms of this article are more robust on different images. Figure 8 illustrates the disparity map generation results of the three algorithms. Overall, the disparity map results are more satisfactory. It is worth noting that on the MotorE and PianoL images, both SGBM and ADSM exhibit a significant degree of disparity mismatch. In addition, they show large void regions. This article's algorithm can better match weak and repeated texture regions. The disparity map indicates that Grayscale Level cost can improve image matching effectiveness.

**Figure 8.** Results of different algorithms for disparity map generation.

3.2. Lodging Judgment Test

The algorithm presented in this article was used for the rice lodging judgment test. A tripod is used to fix the binocular camera in the rice field. Binocular cameras are connected via USB to a laptop computer that serves as the host computer. Figure 9 illustrates

the field. The right and left images were 1920×1080 (width \times height). A red box indicates the lodging rice region in Figure 10a. There were five equal-spaced regions to be detected. The five regions were kept the same width. The height ranges of the five regions are $[300, 400]$, $[400, 500]$, $[500, 600]$, $[600, 700]$, and $[700, 800]$, respectively. Calculate the 3D world coordinates of all pixels in the five regions in the current camera pose. By transforming the rotation matrix coordinates of the IMU, the 3D world coordinates of the camera in horizontal posture can be obtained. In Figure 10b,d,f, the X-axis represents the width of the image, the Y-axis represents the height of the image, and the Z-axis indicates the distance between each pixel target in the five regions and the horizontal plane of the camera (which corresponds to the distance y in the schematic of Figure 2).



Figure 9. Image acquisition.

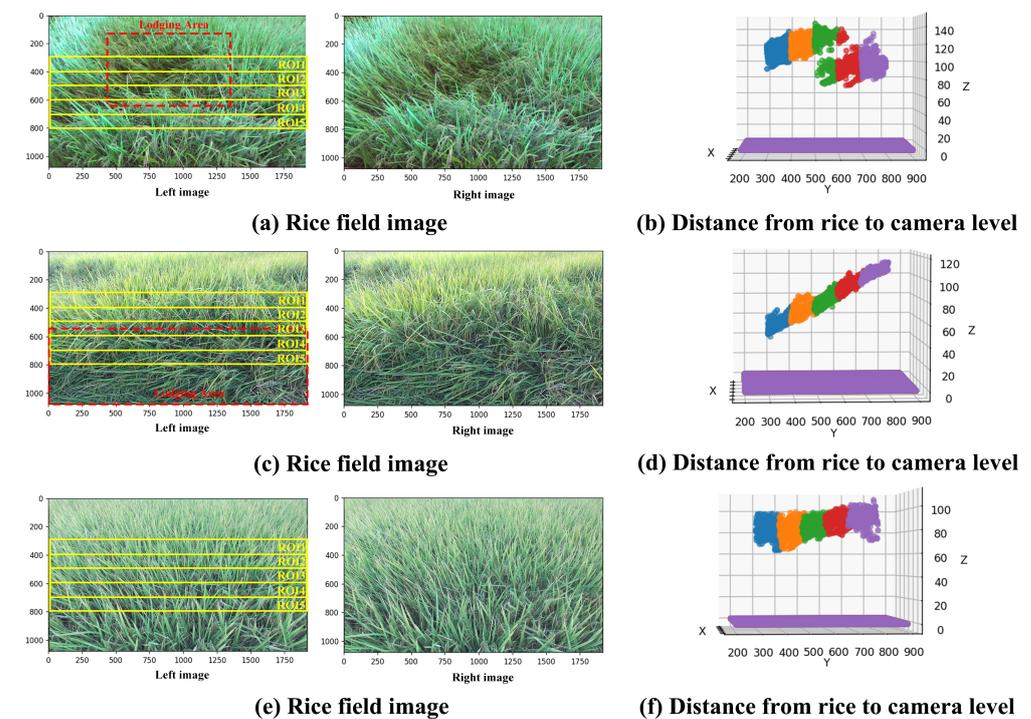


Figure 10. Rice field image and rice to camera horizontal height map.

The demarcation line between the lodging and un-lodging regions on the Z-axis can be seen in Figure 10b. A larger Z-value indicates that the object is farther from the camera level. Each of the five regions calculates the average distance from all pixel points to the camera level. Given the distance threshold TH, the ROI is defined as an un-lodging region when $\bar{y} \in (0, TH)$. Define the ROI as a lodging region when $\bar{y} > TH$. The lodging regions in Figure 10 are [300, 400], [400, 500], and [500, 600] for a 100 distance threshold. The non-lodging regions are [600, 700] and [700, 800]. Figure 10a shows that the heights around 600 correspond to the dividing line between lodging and non-lodging regions on the image.

Figure 10c shows that another scene of lodging images has been collected. A red box indicates the spatial extent of the lodging region. The difference between Figure 10a and this scene is that the degree of rice lodging gradually increases, which is not the case for all rice lodging. Figure 10d shows the distance from the rice to the camera level. Figure 10d shows that as the rice's lodging degree increases, the distance to the camera level also increases. In this case, selecting an appropriate threshold value also determines the lodging location. Using the same distance threshold of 100, we can determine that [700, 800] is part of the lodging region.

It should be noted that based on the results of the red box labeling of the lodging range in Figure 10c, the two regions of [500, 600] and [600, 700] also have mild lodging but are considered non-lodging regions by the threshold value. Consequently, threshold selection should be analyzed and selected based on the specific task. It is possible to use a larger distance threshold when the only goal is to determine whether lodging is complete. A smaller distance threshold can be chosen when judging moderate or light lodging.

Field images of rice without lodging were selected for the lodging judgment test to verify the algorithm's robustness. Figure 10e shows a field image of rice without lodging, while Figure 10f shows the distance from the rice to the camera level. In Figure 10f, there is no obvious line of demarcation between the rice and camera level, and the distance from the rice to the camera level is stable. Several fluctuations occur in the distance values in the [700, 800] image height range, caused by the spacing between rice plants evident in the [700, 800] height range. The fluctuations of a small part do not affect the overall average. Non-lodging regions may also be determined using the 100 distance threshold.

Overall, the distance threshold determines whether lodging is present in an image. The experiment's ROI position can be adjusted according to specific needs. The threshold selection needs to consider different lodging situations. Furthermore, rice height and camera installation height must also be considered. Since rice heights and camera mounting heights differ, the same distance threshold cannot be applied to all field scenes. When performing a specific task, the distance value from the un-lodging rice region to the camera level can be read first. This distance value is used as a reference for determining the distance threshold and adaptive adjustment to complete the subsequent lodging judgment.

4. Conclusions and Discussion

An algorithm for rice lodging judgment based on a binocular camera is presented in this article. The algorithm improves the two steps of matching cost and cost aggregation. Five standard test images from Middlebury Benchmark v3 are selected for testing. The improved algorithm performs better than the commonly used SGBM algorithm. The overall index is better than the ADSM algorithm. A threshold judgment can be used to accurately determine the rice lodging region from images collected in the field.

The algorithm has a certain degree of robustness but needs to adaptively find the threshold value in practice. Different rice varieties and camera mounting heights can affect the threshold value. Second, the deployment of this algorithm on agricultural machinery is not discussed in this article. Compared to the current common lodging judgment scheme, binocular cameras can provide ground agriculture equipment with low-cost and real-time lodging information. Nevertheless, the actual field operation environment is

complex. Vehicle body shaking and weather may affect lodging judgments. In the next step, the algorithm will be deployed on agricultural machinery and tested in the field.

Author Contributions: Funding acquisition, L.H.; Investigation, C.L.; Methodology, Y.Y., C.L. and P.H.; Project administration, L.H. and X.L.; Software, Y.Y.; Validation, Y.Y.; Writing—original draft, Y.Y.; Writing—review & editing, L.H., X.L., J.H., P.W., P.H., R.G. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Laboratory of Science and Technology Innovation 2030—“New Generation Artificial Intelligence” Major Project (2021ZD01109).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

References

- Iqbal, J.; Qamar, Z.U.Q.; Yousaf, U.; Asgher, A.; Dilshad, R.; Qamar, F.; Sajida, B.; Rehman, S.; Haroon, M. Sustainable Rice Production Under Biotic and Abiotic Stress Challenges. In *Sustainable Agriculture in the Era of the OMICs Revolution*; Springer: Cham, Switzerland, 2023; pp. 241–268. [\[CrossRef\]](#)
- Jin, L.; Lu, Y.; Xiao, P.; Sun, M.; Corke, H.; Bao, J.S. Genetic diversity and population structure of a diverse set of rice germplasm for association mapping. *Theor. Appl. Genet.* **2010**, *121*, 475–487. [\[CrossRef\]](#)
- Khush, G. Green revolution: The way forward. *Nat. Rev. Genet.* **2001**, *2*, 815–822. [\[CrossRef\]](#)
- Shah, L.; Yahya, M.; Shah, S.M.A.; Nadeem, M.; Ali, A.; Ali, A.; Wang, J.; Riaz, M.W.; Rehman, S.; Wu, W.X.; et al. Improving Lodging Resistance: Using Wheat and Rice as Classical Examples. *Int. J. Mol. Sci.* **2019**, *20*, 4211. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dai, X.M.; Chen, S.S.; Jia, K.; Jiang, H.; Sun, Y.S.; Li, D.; Zheng, Q.; Huang, J.X. A Decision-Tree Approach to Identifying Paddy Rice Lodging with Multiple Pieces of Polarization Information Derived from Sentinel-1. *Remote Sens.* **2023**, *15*, 240. [\[CrossRef\]](#)
- Chauhan, S.; Darvishzadeh, R.; Boschetti, M.; Pepe, M.; Nelson, A. Remote sensing-based crop lodging assessment: Current status and perspectives. *Isprs J. Photogramm. Remote Sens.* **2019**, *151*, 124–140. [\[CrossRef\]](#)
- Tian, M.L.; Ban, S.T.; Yuan, T.; Ji, Y.B.; Ma, C.; Li, L.Y. Assessing rice lodging using UAV visible and multispectral image. *Int. J. Remote Sens.* **2021**, *42*, 8840–8857. [\[CrossRef\]](#)
- Zhao, X.; Yuan, Y.T.; Song, M.D.; Ding, Y.; Lin, F.F.; Liang, D.; Zhang, D.Y. Use of Unmanned Aerial Vehicle Imagery and Deep Learning UNet to Extract Rice Lodging. *Sensors* **2019**, *19*, 3859. [\[CrossRef\]](#)
- Li, X.; Wang, K.; Ma, Z.; Wang, H. Early detection of wheat disease based on thermal infrared imaging. *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 183–189. [\[CrossRef\]](#)
- Li, Z.; Chen, Z.; Wang, L.; Liu, J.; Zhou, Q. Area extraction of maize lodging based on remote sensing by small unmanned aerial vehicle. *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agric. Eng.* **2014**, *30*, 207–213. [\[CrossRef\]](#)
- Tang, Z.Q.; Sun, Y.Q.; Wan, G.T.; Zhang, K.F.; Shi, H.T.; Zhao, Y.D.; Chen, S.; Zhang, X.W. Winter Wheat Lodging Area Extraction Using Deep Learning with GaoFen-2 Satellite Imagery. *Remote Sens.* **2022**, *14*, 4887. [\[CrossRef\]](#)
- Chauhan, S.; Darvishzadeh, R.; Lu, Y.; Boschetti, M.; Nelson, A. Understanding wheat lodging using multi-temporal Sentinel-1 and Sentinel-2 data. *Remote Sens. Environ.* **2020**, *243*, 111804. [\[CrossRef\]](#)
- Sun, Q.; Gu, X.H.; Chen, L.P.; Xu, X.B.; Pan, Y.C.; Hu, X.Q.; Xu, B. Monitoring rice lodging grade via Sentinel-2A images based on change vector analysis. *Int. J. Remote Sens.* **2022**, *43*, 1549–1576. [\[CrossRef\]](#)
- Zhao, L.L.; Yang, J.; Li, P.X.; Shi, L.; Zhang, L.P. Characterizing Lodging Damage in Wheat and Canola Using Radarsat-2 Polarimetric SAR Data. *Remote Sens. Lett.* **2017**, *8*, 667–675. [\[CrossRef\]](#)
- Chauhan, S.; Darvishzadeh, R.; Boschetti, M.; Nelson, A. Estimation of crop angle of inclination for lodged wheat using multi-sensor SAR data. *Remote Sens. Environ.* **2020**, *236*, 111488. [\[CrossRef\]](#)
- Schaepman, M.E.; Ustin, S.; Plaza, A.; Painter, T.; Verrelst, J.; Liang, S. Earth system science related imaging spectroscopy—An assessment. *Remote Sens. Environ.* **2009**, *113* (Suppl. S1), S123–S137. [\[CrossRef\]](#)
- Miphokasap, P.; Kiyoshi, H.; Vaiphasa, C.; Souris, M.; Nagai, M. Estimating Canopy Nitrogen Concentration in Sugarcane Using Field Imaging Spectroscopy. *Remote Sens.* **2012**, *4*, 1651–1670. [\[CrossRef\]](#)
- Liu, T.; Li, R.; Zhong, X.; Jiang, M.; Jin, X.; Zhou, P.; Liu, S.; Sun, C.; Guo, W. Estimates of rice lodging using indices derived from UAV visible and thermal infrared images. *Agric. For. Meteorol.* **2018**, *252*, 144–154. [\[CrossRef\]](#)
- Somers, B.; Asner, G.; Tits, L.; Coppin, P. Endmember variability in Spectral Mixture Analysis: A review. *Remote Sens. Environ.* **2011**, *115*, 1603–1616. [\[CrossRef\]](#)
- Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging. *Remote Sens.* **2014**, *6*, 10395–10412. [\[CrossRef\]](#)
- Varela, S.; Pederson, T.L.; Leakey, A.D.B. Implementing Spatio-Temporal 3D-Convolution Neural Networks and UAV Time Series Imagery to Better Predict Lodging Damage in Sorghum. *Remote Sens.* **2022**, *14*, 733. [\[CrossRef\]](#)

22. Su, Z.B.; Wang, Y.; Xu, Q.; Gao, R.; Kong, Q.M. LodgeNet: Improved rice lodging recognition using semantic segmentation of UAV high-resolution remote sensing images. *Comput. Electron. Agric.* **2022**, *196*, 106873. [[CrossRef](#)]
23. Zhao, B.Q.; Li, J.T.; Baenziger, P.S.; Belamkar, V.; Ge, Y.F.; Zhang, J.; Shi, Y.Y. Automatic Wheat Lodging Detection and Mapping in Aerial Imagery to Support High-Throughput Phenotyping and In-Season Crop Management. *Agronomy* **2020**, *10*, 1762. [[CrossRef](#)]
24. Zhang, Z.; Flores, P.; Igathinathane, C.; Naik, D.L.; Kiran, R.; Ransom, J.K. Wheat Lodging Detection from UAS Imagery Using Machine Learning Algorithms. *Remote Sens.* **2020**, *12*, 1838. [[CrossRef](#)]
25. Haridasan, A.; Thomas, J.; Raj, E.D. Deep learning system for paddy plant disease detection and classification. *Environ. Monit. Assess.* **2023**, *195*, 120. [[CrossRef](#)] [[PubMed](#)]
26. He, Y.; Zhang, X.Y.; Zhang, Z.Q.; Fang, H. Automated detection of boundary line in paddy field using MobileV2-UNet and RANSAC. *Comput. Electron. Agric.* **2022**, *194*, 106667. [[CrossRef](#)]
27. Ma, Z.H.; Tao, Z.Y.; Du, X.Q.; Yu, Y.X.; Wu, C.Y. Automatic detection of crop root rows in paddy fields based on straight-line clustering algorithm and supervised learning method. *Biosyst. Eng.* **2021**, *211*, 63–76. [[CrossRef](#)]
28. Wen, J.Q.; Yin, Y.X.; Zhang, Y.W.; Pan, Z.L.; Fan, Y.D. Detection of Wheat Lodging by Binocular Cameras during Harvesting Operation. *Agriculture* **2023**, *13*, 120. [[CrossRef](#)]
29. Li, J.H.; Dai, Y.P.; Su, X.H.; Wu, W.B. Efficient Dual-Branch Bottleneck Networks of Semantic Segmentation Based on CCD Camera. *Remote Sens.* **2022**, *14*, 3925. [[CrossRef](#)]
30. Li, J.; Li, J.H.; Zhao, X.; Su, X.H.; Wu, W.B. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4. *Comput. Electron. Agric.* **2023**, *211*, 107955. [[CrossRef](#)]
31. Yang, Y.K.; Nie, J.; Kan, Z.; Yang, S.; Zhao, H.X.; Li, J.B. Cotton stubble detection based on wavelet decomposition and texture features. *Plant Methods* **2021**, *17*, 113. [[CrossRef](#)]
32. Sun, J.W.; Zhou, J.; He, Y.Q.; Jia, H.B.; Liang, Z. RL-DeepLabv3+: A lightweight rice lodging semantic segmentation model for unmanned rice harvester. *Comput. Electron. Agric.* **2023**, *209*, 107823. [[CrossRef](#)]
33. Laga, H.; Jospin, L.V.; Boussaid, F.; Bennamoun, M. A Survey on Deep Learning Techniques for Stereo-Based Depth Estimation. *Ieee Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 1738–1764. [[CrossRef](#)] [[PubMed](#)]
34. Deng, C.G.; Liu, D.Y.; Zhang, H.D.; Li, J.R.; Shi, B.J. Semi-Global Stereo Matching Algorithm Based on Multi-Scale Information Fusion. *Appl. Sci.* **2023**, *13*, 1027. [[CrossRef](#)]
35. Ren, J.; Guan, F.; Wang, T.; Qian, B.; Luo, C.; Cai, G.; Kan, C.; Li, X. High Precision Calibration Algorithm for Binocular Stereo Vision Camera using Deep Reinforcement Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 6596868. [[CrossRef](#)]
36. Hou, Y.; Liu, C.; An, B.; Liu, Y. Stereo matching algorithm based on improved Census transform and texture filtering. *Optik* **2021**, *249*, 168186. [[CrossRef](#)]
37. Zhang, K.; Lu, J.; Lafruit, G. Cross-Based Local Stereo Matching Using Orthogonal Integral Images. *Circuits Syst. Video Technol. IEEE Trans.* **2009**, *19*, 1073–1079. [[CrossRef](#)]
38. Mei, X.; Sun, X.; Zhou, M.; Jiao, S.; Wang, H.; Zhang, X. On building an accurate stereo matching system on graphics hardware. In Proceedings of the 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, Spain, 6–13 November 2011; pp. 467–474. [[CrossRef](#)]
39. Ma, N.; Men, Y.B.; Men, C.G.; Li, X. Accurate Dense Stereo Matching Based on Image Segmentation Using an Adaptive Multi-Cost Approach. *Symmetry* **2016**, *8*, 159. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.