

Density estimation method of mature wheat based on point cloud segmentation and clustering

Rong Zou^{a,*}, Yu Zhang^a, Jin Chen^a, Jinyan Li^a, Wenjie Dai^a, Senlin Mu^b

^a School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China

^b Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China



ARTICLE INFO

Keywords:

Binocular vision

Wheat

Canopy segmentation

Density detection

ABSTRACT

The sustainable development of agriculture needs to rely on precision agricultural technology. The premise and key to realizing precision agriculture is the research on the characteristics of crops in the field. For wheat plants, the ear is the flower or fruit part at the top of the wheat stem, which is one of the important components of yield, and its density can be said to be one of the most important traits. Its number is arguably one of the most important phenotypic traits. The traditional method of density measurement is manual, which is time-consuming and laborious. Therefore, an efficient and convenient wheat density estimation scheme is needed to provide data support for crop yield monitoring, to achieve a better production management system. Stereo vision is a 3D imaging method that allows rapid measurement of plant structures, and point cloud segmentation is the key to studying the 3D spatial characteristics of plants. In this paper, the three-dimensional point cloud of wheat reconstructed by stereo vision technology is used for segmentation and clustering, and a method for clustering dense wheat rows is proposed. Firstly, a binocular camera was used to record video to reconstruct the wheat point cloud; then, point cloud pre-processing was used to remove noise; then, the octree splitting algorithm and voxel mesh merging algorithm were used to divide the dense wheat, and then clustering algorithm was used to get the point cloud of wheat ears; finally, the relationship model between the number of wheat ears point clouds and the number of wheat plants was established by linear regression analysis with R^2 of 0.97. To verify the effectiveness of the algorithm, the actual field measurements and the predicted values of the algorithm were compared, with R^2 of 0.93. The density estimation method provides a new method for the study of phenotypic information of crop population information and also provides a reference for nondestructive crop measurements.

1. Introduction

With the increasing population, human demand for food and oil crops has increased dramatically, but its output is difficult to increase due to the reduction of available arable land, land desertification, and natural disasters (Chengda et al., 2021). Plant phenotyping is the quantification of physical traits, genetics, and individual development of plants (Mohamed and Dudley, 2019), and plant phenotyping can help meet this need. Wheat is one of the most important food crops in the world and plays an irreplaceable role in ensuring human nutritional health. According to statistics, to meet the needs of the rapidly growing population, the crop yield required is expected to double by 2050 (Yang et al., 2021). Therefore, there is an urgent need to investigate better measurement techniques to automate the collection of large amounts of phenotypic data (Fang et al., 2016) and achieve high throughput

agricultural crop improvement.

For wheat, the phenotypes of interest include crop density, height, ear size, and yield (Thompson et al., 2019). Among them, wheat density is the most important parameter. The detection of crop density can provide data reference for the work of the combined harvester and effectively ensure the stability of the harvester when it is working in the field. Traditional wheat density estimation needs to be done by manual counting, which is also a labor-intensive industry, and the cost is high, and sometimes there are deviations in labor. As the size of the plot and the farm increase, it will be difficult to meet the demand (Dhami et al., 2020). Therefore, new methods are needed to enhance the yield analysis power of wheat density estimation.

Computer vision is currently the main technical means for wheat recognition and counting (Wei et al., 2021). At present, there has been sufficient research on crop plant density detection methods based on

* Corresponding author.

E-mail address: zr@ujs.edu.cn (R. Zou).

two-dimensional images. For example, Reza et al. (2019) used the k-means classification method to classify and segment rice, and calculated the plant density of rice, with a relative error of 6–33 %. Khaki et al. (2022) used a deep learning approach to accurately count wheat ears with mean absolute and root mean square errors of 3.85 and 5.19, respectively. Ayalew et al. (2020) proposed a custom domain adversarial network and trained it using unsupervised domain adaptation techniques for wheat crop density estimation. These two-dimensional imaging methods all have the same limitations. Although the current imaging equipment can obtain higher-resolution images, it is often difficult to fully reflect the true spatial form of crops due to the limitation of the two-dimensional image's dimensions.

Digitization and visualization of 3D plant canopy structure is an important part of calculating plant crop density (Jizhang Wang et al., 2020b). The 3D point cloud adds the stereo coordinate information in the real scene to the 2D image. Based on a review of the literature on plant phenotyping, there are different sensors to generate data for crop density estimation, which can be broadly categorized into active and passive methods (Pankaj et al., 2013). Light detection and ranging (LiDAR) is one of the most widely used active sensors (Lin, 2015). Saeys et al. (2009) proposed a three-dimensional spline fitting method for wheat point cloud using the lidar detection method and generated a point density image, from which the approximate position of the ear can be identified, and the crop density estimation of wheat The coefficient of determination is as high as 0.96. Lidar can provide accurate canopy models, but is expensive and needs to be combined with an RGB camera to obtain accurate color information. Kinect-v2 is another typical active sensor. Bao et al. (2019) used Kinect-V2 to obtain 3D point cloud data of high-density crops under field conditions and divided the point cloud into plant instances with separated stems and leaves. The results show that under the condition of leaf occlusion, plant imaging from only one side has better robustness and accuracy. TOF camera is cheaper than LiDAR, but TOF cameras are suitable for indoor measurements and are difficult to apply under natural light conditions. Multi-view vision technology is an imaging method that uses calibrated cameras to capture images of objects from different angles. Its main applications include Structure From Motion (SFM) and Multi-View Stereo (MVS) techniques. Grenzdörffer (2014) utilizes high-resolution Unmanned Aerial Vehicles (UAV) to acquire images, uses the SFM method to generate point clouds, and then determines the height of wheat through point cloud processing, which helps predict crop yield and biomass. However, SFM technology relies on UAV aerial photography to achieve the 3D reconstruction of large-scale scenes, which increases the cost of experiments (Jizhang Wang et al., 2020b). Stereo vision is a passive imaging technique commonly used in 3D image reconstruction. Salas Fernandez et al. (2017) achieved field phenotyping of sorghum biomass through a camera system mounted on a mobile tractor. Therefore, studies have shown that stereo cameras, as a commonly used method, are not only low-cost and allow for rapid acquisition, but are also capable of operating in a field environment.

With the development of agricultural mechanization and information digitization, many scholars have made pioneering research on point cloud segmentation of crops. Most of the existing methods focus on extracting information from individual plants. For example, Jin et al. (2019) proposed a voxel-based convolutional neural network to separate individual corn stalks and leaves, proving that the deep learning approach provides a reference for crop phenotyping. However, due to the large changes in scene density and large scene range, it is difficult to generate and label point cloud data, and a large number of points also lead to huge calculation and memory consumption of deep learning algorithms when adapting to outdoor scenes. Clustering methods have been used to study crop density estimation for clustering tasks with different demands. Zermas et al. (2018) proposed the RAIN (Randomly Intercepted Nodes) clustering algorithm to study the sliding behavior of a randomly placed point cloud on the entire point cloud surface and to separate the leaf and stem according to the different performances of the

leaf and stem, and this method can provide a reference for crop density estimation. Fuli Wang et al. (2020a) proposed a dynamic perspective adaptive k-means clustering algorithm to separate wheat spikes. However, they only used laboratory specimens and did not test on real field data. To calculate the number of wheat spikes, Velumani (2017) used voxel-based and mean shift segmentation methods to segment wheat spikes, with an average detection rate of 87 %. The above method can improve the accuracy of crop density estimation in the corresponding field, however, it is still a challenge to detect wheat crop density from the segmentation of the wheat segmentation population at the field scale.

Given the above problems, the objectives of this paper can be summarized as follows: (1) Study the clustering algorithm for density estimation based on point cloud data separation of ear layers; (2) Analyze the application of 3D methods from models under indoor control conditions to complex natural canopy areas challenges and propose solutions to these challenges; (3) The point cloud density estimation is compared with the results of manual traditional measurements and the performance of the proposed method is verified.

2. Materials and methods

Mature wheat is usually leafy and has ears at the top of the plant, and the ears of wheat are distributed at the top of the plant. In addition, a biological characteristic of wheat is that all grow up along the stalk. Therefore, this study is aimed at the detection of wheat ears in farmland, which is a key technique to achieve crop density estimation. The workflow of the proposed framework for addressing this problem is shown in Fig. 1. Firstly, point cloud data collection and outlier removal are carried out on the field environment through the ZED2 binocular camera. Secondly, an algorithm for the detection and segmentation of dense wheat rows in the field is developed. Thirdly, use the row spacing to judge whether the wheat is dense. If the row spacing is <15 cm, it is a dense point cloud, and the clustering algorithm is used after dividing the rows; if the row spacing is greater than 15 cm, the clustering algorithm is used directly. Finally, a relationship model between the number of wheat plants and the number of point clouds at the ear was constructed, and the performance of the proposed method was verified by comparing the measured values in the field. The following sections describe these techniques in detail.

2.1. Experimental design

We designed and conducted two experiments to investigate the accuracy of the data analysis algorithm to extract wheat density.

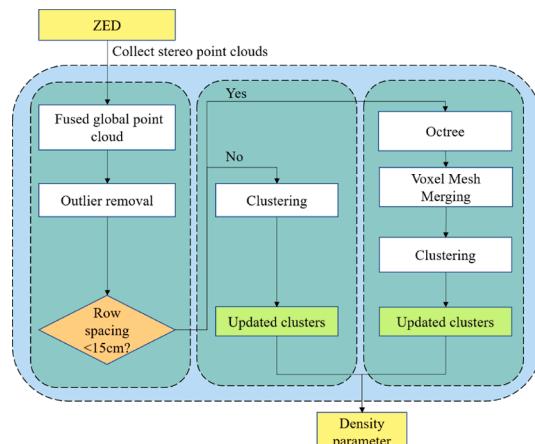


Fig. 1. Flowchart of the wheat ear detection algorithm.

2.1.1. Laboratory wheat samples

To test the performance of the clustering algorithm for the detection of crop canopy ears, experiments were conducted under laboratory conditions. (4)⁵ holes were evenly punched in the foam board, the row spacing was expanded to 20 cm, and 20 wheat crops, each with 10 plants for analysis and research (Fig. 2(c)).

2.1.2. Field wheat experiment

The field trial was conducted in May 2022 on wheat data collection at Shiyezhou Farm, Dantu District, Zhenjiang City, Jiangsu Province (Fig. 2(b)). The ZED2 installed on the harvester is used to scan the plants as it can see several meters ahead without damaging the crop. The camera faces the front of the harvester for density measurement in front of the combine harvester, and the top camera is installed at a height of about 3.4 m. (Fig. 2(a)) Manual data collection was performed after data collection using the camera. The specific parameters of the crops are shown in Table 1.

We set up nine experimental groups in the field using calibration poles and a tape measure and manually cut the excess wheat in the area with densities of 100 m², 150 m², 200 m², 250 m², 300 m², 350 m², 400 m², 450 m², 500 m². To verify the accuracy of the algorithm model in the actual situation in the field, eight control groups were randomly selected and manually measured their density magnitudes with densities of 437 m², 510 m², 479 m², 383 m², 312 m², 466 m², 487 m², and 538 m², respectively. Each sample in the randomized controlled experimental group was randomly selected in the same field and collected by the same sampling method, which can reflect the situation of multiple field plots with inconsistent crop densities. Among them, we selected an area of 400 plants/m² to demonstrate the results of the wheat density analysis

Table 1
Wheat Crop Parameters.

Parameter	Value
crop variety	Yuanliang No.1
Average plant height (cm)	80
Average line spacing (cm)	14
Average ear length (cm)	12
Height of stubble left (cm)	25–30
Average yield (kg)	1100–1200

algorithm (Fig. 2(d)). This plot was chosen because during the experimental analysis, the wheat density in this plot was large enough, and the possibility of intersecting with each other was very high, so it had high representativeness.

Table 2
Main parameters of the ZED2 stereo camera for this research.

Parameter	Value
RGB resolution and frame rate	2560 × 720, 60 fps
Baseline	120 mm
Focal length	2.12 mm
Measurable range	0.5–20 m
Field of view (vertical × horizontal)	110° × 70°
Accuracy	<1% up to 3 m <5% up to 15 m



(a)



(b)



(c)



(d)

Fig. 2. 3D point cloud acquisition system. (a) the ZED2 camera installed on the harvester platform; (b) the harvester installed with the ZED2 camera is harvesting wheat in the field; (c) laboratory wheat samples; (d) experimental samples of wheat in the field.

2.2. Data acquisition

In this study, we used a Stereolabs ZED2 stereo camera for data acquisition. The main parameters of the camera are shown in Table 2. By adjusting the ZED2 camera angle to collect wheat plant videos (Fig. 3.), while capturing detailed information of plant organs, the occlusion problem is greatly reduced. The signal is transferred to a laptop computer for data acquisition, SVO data was generated from the ZED2 stereo camera video and compressed using the H265 format, and the Open3D library is used for 3d data processing. According to the stereo matching principle, given a pair of stereo images, the pixels in the left image are matched with the corresponding pixels in the right image, and then the depth value of each pixel is calculated according to the camera parameters (Wang et al., 2019), and the point cloud P_K is obtained. We generate the corresponding point cloud, RGB images, Dense fused point cloud global map, Pose history and Camera calibration parameters, and other files from the SVO file to form the three-dimensional data information of wheat crops. The obtained laboratory 3D point cloud images and field 3D point cloud images are shown in Fig. 4.

2.3. Data processing

When acquiring point cloud data, some noise points will inevitably appear in the point cloud data due to the influence brought by equipment accuracy, operator experience, and environmental factors. In addition, due to the influence of external interference such as line of sight occlusion, obstacles, and other factors, there are often some discrete points far away from the main point cloud in the point cloud data, namely outliers (Sun et al., 2021). This leads to inaccurate segmentation of point cloud data. The 3D point cloud preprocessing includes outlier removal and denoising. In this paper, the Statistical Outlier Removal (SOR) algorithm is used to remove outlier point clouds that are far from the main part of the wheat point cloud. Let $P_K = \{P_1, P_2, \dots, P_n\}$ be the set of point clouds, and $P(x_i, y_i, z_i)$ be any point in the point cloud. Statistical analysis is performed on the neighborhood of each point, and some outlier points that do not meet the requirements are filtered out based on the distance distribution characteristics of the point to all neighboring points. The algorithm performs two iterations over the entire input via the K-nearest searching method: In the first iteration, it computes the average distance from each point to the k-nearest neighbors, yielding a result that fits a Gaussian distribution. Next, the mean μ and standard deviation σ of all these distances are calculated according to Equation (1) to determine the distance threshold d (Wu et al., 2018).

$$d = \mu + k^* \sigma \quad (1)$$

where k is a multiple of the standard deviation. In the next iteration,

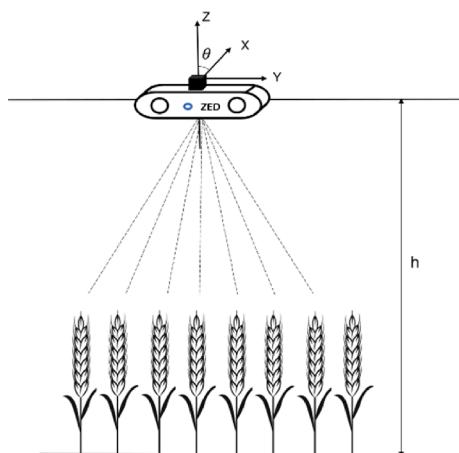


Fig. 3. A schematic diagram of the data acquisition process of the ZED2 camera.

the points will be classified as inliers or outliers if their average neighborhood distance is below or above this threshold, respectively. All measurements with intensities below a given threshold or measurement distances above a given threshold were excluded from the data. The laboratory 3D point cloud image and the field 3D point cloud image after processing to remove outliers are shown in Fig. 5.

2.4. Point cloud segmentation

2.4.1. Wheat row segmentation

Firstly, the octree-based splitting algorithm is used to divide the collected point cloud grid into many small cubes, and then the small cubes are merged according to the feature consistency of the row where the wheat is located by using the voxel grid merging algorithm, and finally, each row of wheat wrapped in the space bounding box is obtained.

2.4.1.1. Octree-based splitting algorithm. An octree is a tree-like data structure used to describe three-dimensional space. Each node of an octree represents a volume element of a square, and each node has eight child nodes. The volume elements represented by the eight child nodes are added together to equal the volume of the parent node. For data collection with a dense set of objects, the octree enables fast data management, visual cropping, ray tracing, and other 3D spatial operations (Mosa et al., 2012).

The criterion for constructing an octree for each voxel to determine whether to continue splitting is whether the points located in the grid can satisfy the feature consistency, and the number of points must be greater than a certain number. If it is satisfied, the process stops, if not, the grid needs to be split. The steps of the octree-based splitting algorithm are as follows:

a. Generate an initial grid containing the entire volume data, the grid starting position is $(0, 0, 0)$; the end position is the maximum value of the volume data $(x_{max}, y_{max}, z_{max})$.

b. Determine whether the initial grid can meet the feature consistency and whether the number of points is greater than N_{min} . If the feature consistency can be satisfied, the voxel grid is not divided; otherwise, the grid is evenly divided into 8 small cubes and divided into corresponding small cubes according to the coordinates of the points.

c. For each obtained small cube, follow steps a and b to determine whether to continue splitting until the points in each small block can satisfy the feature consistency and the number of points is less than the set threshold.

2.4.1.2. Voxel-based mesh merging algorithm. For the small cubes divided by the octree, we use the grid merging algorithm to obtain a single row of wheat. In the process of generating the bounding box, the grid of each area is allocated according to the line spacing, the lowest point of the Z-coordinate of the cube is extracted as the bottom end of the voxel grid, and the highest point of the Z-coordinate of the cube is extracted as the top of the voxel grid. The steps of the voxel grid-based (Su et al., 2016) merging algorithm are as follows:

a. Find the subregion where the cube is located by the coordinate position of the row spacing division; then create a queue to be merged and put the subregion just found into this queue; then put this subregion into the set of merged regions.

b. Take out a subregion node from the queue to be merged; if the queue to be merged is empty, it means that the merge is complete, and jump to step d.

c. Consistency detection is performed on the adjacent cubes of the nodes taken out in step b in turn. If the detection passes, the adjacent cubes are placed in the queue to be merged and the merged area set respectively.

d. Complete the merge operation.

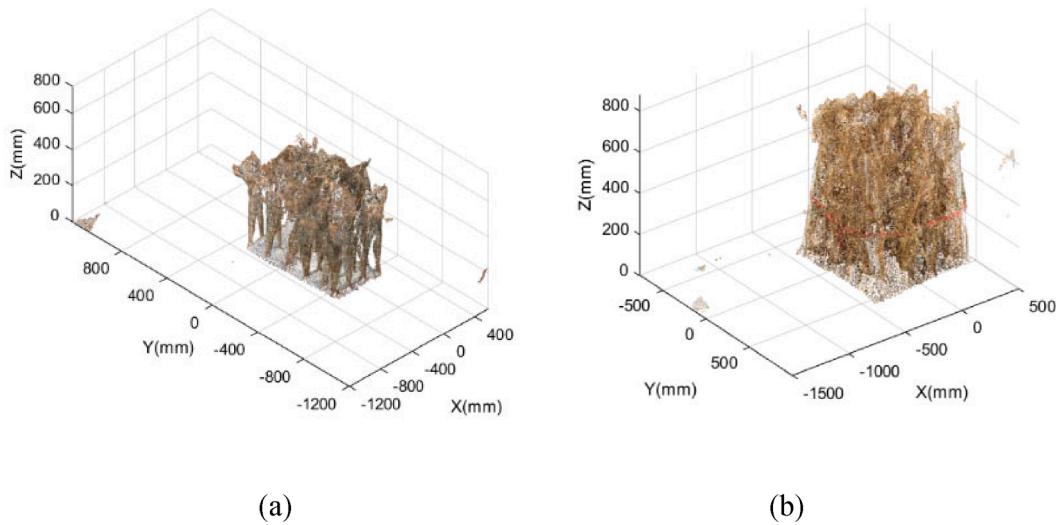


Fig. 4. (a) point cloud of laboratory wheat samples; (b) point cloud of field wheat samples.

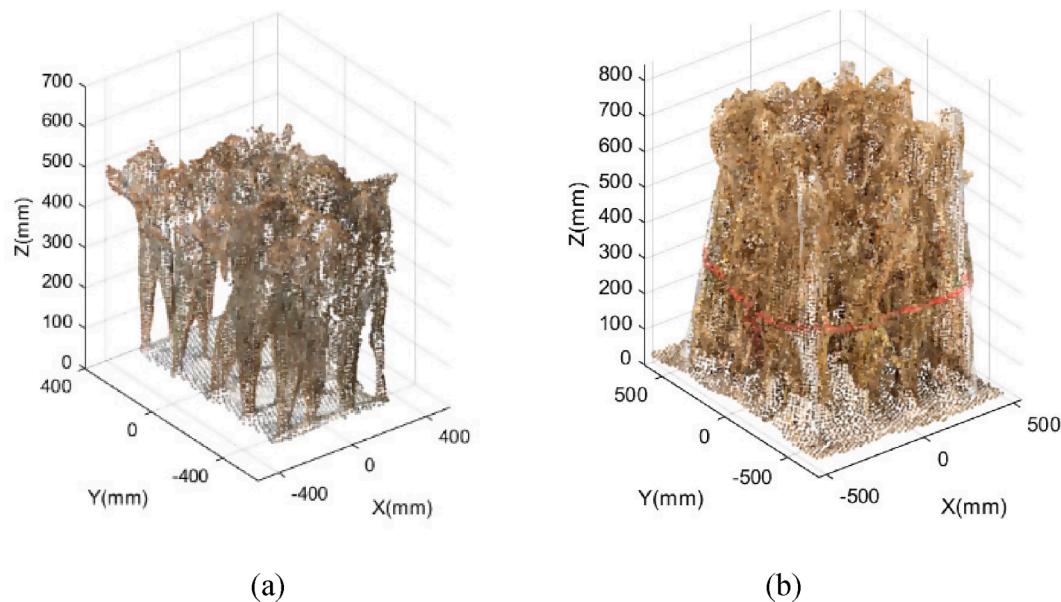


Fig. 5. (a) point cloud of laboratory wheat samples after denoising; (b) point cloud of field wheat samples after denoising.

2.4.2. Wheat canopy segmentation

After each row was obtained by segmentation, a clustering-based algorithm is used to detect the wheat canopy. Clustering algorithms are important tools for separating wheat crop ears and stalks in point clouds. As mentioned in the introduction, the wheat ears are located at the top of the plant and are interlaced complexly. The position of the ears and stems can be easily observed from the side view, and the number of wheat canopy point clouds is often larger than that of wheat stalks. Therefore, ear grains and stems can be effectively separated according to different point density distributions. We introduce the above idea into the Quick-shift algorithm, project the point cloud data to the X-Z plane, and perform the Quick-shift in the 2D side view. Quick-shift is improved from Mean shift. Quick-shift improves the computational complexity by not using gradients to find patterns in the probability density, but simply by moving each point to the closest point which increases the probability density (Vedaldi and Soatto, 2008). The formula is as follows:

$$y_i(1) = \underset{j|P_j > P_i}{\operatorname{argmin}} D_{ij}, P_i = \frac{1}{N} \sum_{j=1}^N \phi(D_{ij}) \quad (2)$$

where $y_i(1)$ represents the next position of the point in the feature space, $D_{ij} = d^2(x_i, x_j)$ represents the distance between two points, $\phi(\bullet)$ is the kernel function, usually, the Gaussian kernel function is selected, and N is the number of points in the feature space.

The Quick-shift method inherits the advantages of Mean shift and does not need to specify cluster centers, but at the same time improves its slow speed, and the time complexity is $O(nlogn)$. In this algorithm, we implement clustering by setting two parameters: K is the kernel size parameter of the density function and the neighborhood distance threshold D in three-dimensional space, and finally output all point labels, and use labels to mark all 3D points.

2.5. Validation and statistical analysis

The relationship between the number of wheat ear point clouds and

the number of wheat plants after segmentation was explored by linear regression analysis. To test the correlation between the predicted values of the algorithm and the true values measured manually, the coefficient of determination (R^2) and the root mean square error (RMSE) were calculated (Chai and Draxler, 2014). Where n denotes the number of comparison objects, y_i is the true value measured manually and \hat{y}_i denotes the predicted value.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

3. Results

3.1. Laboratory wheat ear segmentation

Fig. 6(a) shows the results of clustering wheat ear detection from 2D side views using the Quick-shift algorithm. The different parts of wheat are distinguished by color. The purple area in the picture is the ear part, and the yellow area is the stalk part. Overall, in the side view of the laboratory sample, when the row spacing was increased so that the wheat stalk could be identified, all ears were successfully detected, confirming the effectiveness of the Quick-shift algorithm in detecting wheat ears from the side view.

3.2. Field wheat ear segmentation

3.2.1. Quick-shift algorithm segmentation

Directly use the Quick-shift algorithm to apply to field wheat results as shown in **Fig. 6(b)**. It can be seen that objects of the same type are directly identified in the clustering process. Because the wheat crops in the field are closely connected, the wheat at the mature stage has more serious occlusion and crossover between the wheat ears, making it impossible to distinguish the specific categories, and a large number of misclassifications occur. After parameter optimization and adjustment, modify the value of the neighborhood distance threshold D in the algorithm parameters. For K , we experimented with a few values and showed a value ($K = 200$) with reasonable performance on all point clouds. due to the complexity of wheat point cloud data, accurate segmentation results cannot be obtained. Therefore, the Quick-shift algorithm cannot be directly applied in a fairly dense field crop environment.

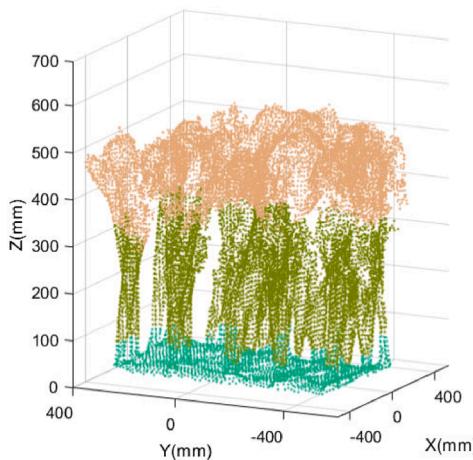


Fig. 6. (a) Segmentation map of laboratory wheat samples using Quick-shift clustering algorithm; (b) Mis-segmentation map of wheat samples in the field using the Quick-shift clustering algorithm.

3.2.2. Field wheat segmentation

To solve the above problem to achieve field application, we first perform voxel-based row segmentation of the region (**Fig. 7**). **Fig. 7 (a)** shows the result of using octree-based segmentation for the 1 m*1m area plants in the field. The region is partitioned by setting Octree Level and Min points per component to generate the small cubic grid in yellow in **Fig. 7 (b)** shows the results of the analysis of plant rows in the 1 m*1m area of the field. The average row spacing of wheat in the test field is 14 cm, and for the 1 m*1m wheat area, seven enclosing boxes were used. The wheat rows were displayed in envelope boxes to ensure that the stalk part of the plant in each row fell completely in the corresponding voxel grid. Therefore, the yellow boxes in the display look different in height and front, and behind. **Fig. 7 (c)** shows a single row of wheat extracted after voxel grid segmentation, and **Fig. 7 (d)** shows the point cloud segmentation effect obtained by using the Quick-shift algorithm in **Fig. 7 (c)**.

Fig. 8 shows the classification effect of the wheat ear layer in the regional interior block after the merging of row segmentation and cluster segmentation. The yellow area in the figure is the point cloud of wheat ears, and the green area is the stalk and ground. The relationship between different wheat densities and the number of wheat ear point clouds in the experimental and validation groups was obtained by the algorithm. From **Table 3**, it can be concluded that the number of ear layer point clouds has a high correlation with the number of wheat plants collected manually.

The linear regression analysis (**Fig. 9**) for the experimental group can be used to obtain the functional relationship between the number of wheat plants and the number of wheat ear points, as shown in equation (5).

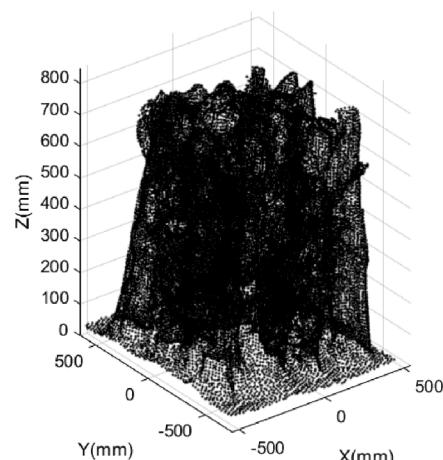
$$y = 64.59*x - 2802.01 \quad (5)$$

Where y is the number of point clouds in the wheat ear layer and x is the number of wheat plants, the coefficient of determination was $R^2 = 0.97$, and the correlation was highly significant. It can be concluded that the number of point clouds in the wheat ear layer has a high correlation with the results of manually collected wheat plants.

4. Discussion

4.1. Comparison of experimental results with manual measurements

According to the obtained relationship model between wheat density and ear point cloud number, the density of each block in the verification group is obtained. The wheat plant number results from the method proposed in this paper are accurate and highly correlated with manual



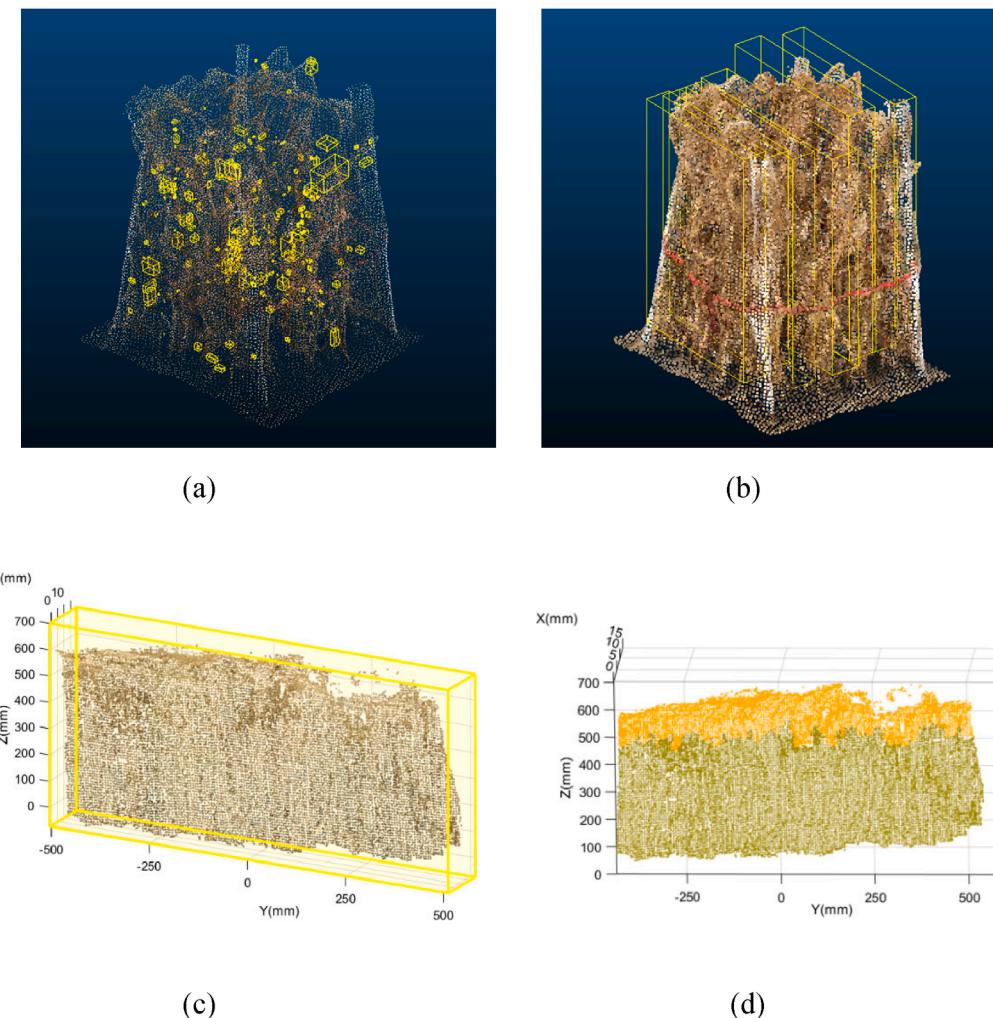


Fig. 7. Detailed view of row clustering segmentation of wheat samples in the field. (a) octree-based splitting segmentation view; (b) row view based on voxel grid merging; (c) single row wheat extraction; (d) single row wheat segmentation view using Quick-shift clustering algorithm.

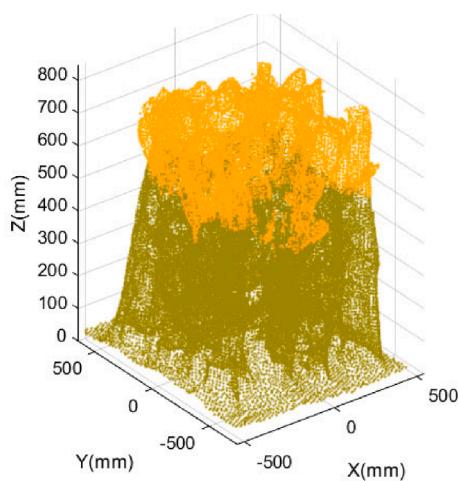


Fig. 8. Field wheat samples Split-row clustering segmentation point cloud data.

measurements, with a coefficient of determination of $R^2 = 0.93$ (Fig. 10). Compared with manual density measurements, the method in this paper can greatly reduce the workload and provide more information about canopy development.

Table 3
Relationship between wheat density and number of ear point clouds.

Number	Density	Number of ear point clouds
Test group	1	100
	2	150
	3	200
	4	250
	5	300
	6	350
	7	400
	8	450
	9	500
		29,345
Verification group	1	437
	2	510
	3	479
	4	383
	5	312
	6	466
	7	487
	8	538
		30,989

4.2. Comparison with other methods

Experimental results show that the detection accuracy of this method is better than the wheat ear counting method of 2D images (Pietragalla et al., 2012), because the 3D point cloud provides richer dimensional

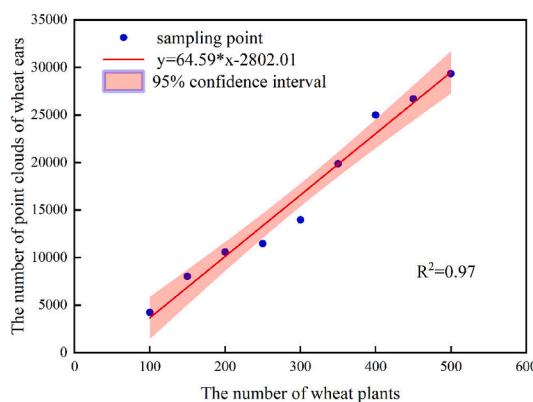


Fig. 9. Analysis of wheat density estimation results.

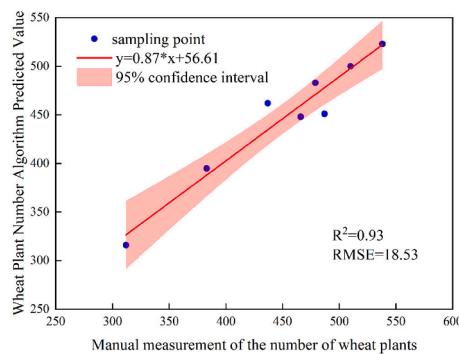


Fig. 10. Comparison of algorithm-calculated and field-measured values.

information, which greatly reduces the problem of false detection and missed detection. Compared with methods based on deep learning of 2D images (Gong et al., 2020), the method based on 3D models in this study can not only improve the performance of density estimation, but also effectively and efficiently obtain more phenotypic traits, such as spatial density distribution of wheat ears, the height of wheat, etc., are of great significance for future plant breeding and combine harvesting operations.

Currently, the real competitor to the stereo vision for multivalent 3D measurements in natural conditions is lidar. Stereo vision may be more popular than lidar, and binocular camera devices are inexpensive and can provide color and shape information at the same time without the need to fuse two sensors of different natures. Studies have shown that crop density can be estimated by analyzing crop data using stereo vision technology. This is consistent with the conclusion of Dandriofosse et al. (2020) using binocular camera sensors to detect plant traits such as wheat canopy height and leaf area, that is, stereo vision technology is an inexpensive and flexible method to study wheat canopy structure under natural conditions. The literature shows that Blanquart et al. (2020) use lidar to detect the estimation of wheat and barley crop height and density, and the R^2 is 0.77. Compared with this paper, the detection method of this paper can achieve higher detection accuracy. The method proposed in this paper can better estimate the density of crops such as wheat in the field, avoid time-consuming and tedious manual measurement, and provide a reference for crop density estimation.

4.3. Future work

The current 3D point cloud reconstruction and point cloud analysis are performed offline, and it is not possible to realize the online experiment of real-time detection in complex agricultural environments (Luo et al., 2022). In the experiment, some point clouds may be missing in the

point cloud collection, resulting in experimental errors. For example, some wheat ears were not detected in the field, mainly because the experimental weather was sunny and the strong light directly hit the wheat ears and the camera, which affected the accuracy of the ZED2 camera in the field. The point cloud of the stalk part will also be missed because the stalk part is sparser than the ear part. In addition, wind speed will also have a greater impact on point cloud scanning (Shangpeng Sun et al., 2020). The impact of different weather conditions on data collection needs to be considered in the future, and an attempt can be made to input the algorithm into the cloud, where real-time density detection of the crop can be performed.

5. Conclusion

In order to improve the accuracy of detection and counting of wheat ears in the field, this paper proposes a point cloud segmentation algorithm, which can be used to deal with the complex environment of the dense growth of wheat ears by batch processing of dense point clouds. The measured number of ear point clouds and the number of wheat plants were highly correlated. The results show that the algorithm has certain robustness. This method not only achieves accurate and non-destructive acquisition of mature wheat density, but also provides data support for field crop information management, combine harvester feeding amount estimation, and genotype and phenotypic traits research. In this paper, only the density parameters of crops are obtained, and more crop attribute parameters are not studied. Future work will focus on extracting more phenotypic traits from 3D point clouds that involve processing other sensor data. Although this system has only been tested on wheat plants, it is expected to apply to other crops with fruit located at the top of the plant, such as rice, maize, and sorghum.

CRediT authorship contribution statement

Rong Zou: Conceptualization, Methodology, Formal analysis. **Yu Zhang:** Software, Validation, Writing – original draft. **Jin Chen:** Resources, Supervision. **Jinyan Li:** Investigation, Visualization. **Wenjie Dai:** Data curation, Writing – review & editing. **Senlin Mu:** .

Declaration of Competing 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 paper.

Data availability

The data that has been used is confidential.

Acknowledgment

This work was supported by Jiangsu Province and the Education Ministry Co-sponsored Synergistic Innovation Center of Modern Agricultural Equipment [grant number XTCX2020].

References

- Ayalew, T.W., Ubbens, J.R., Stavness, I., 2020. Unsupervised domain adaptation for plant organ counting[C]. European conference on computer vision 330–346.
- Bao, Y., Tang, L., Srinivasan, S., et al., 2019. Field-based architectural traits characterisation of maize plant using time-of-flight 3D imaging[J]. Biosystems Engineering 178, 86–101.
- Blanquart, J.-E., Sirignano, E., Lenaerts, B., et al., 2020. Online crop height and density estimation in grain fields using LIDAR[J]. Biosystems Engineering 198, 1–14.
- Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)—Arguments against avoiding RMSE in the literature[J]. Geoscientific model development 7 (3), 1247–1250.
- Chengda, L., Jing, H., Liangyi, X., et al., 2021. Cylinder space segmentation method for field crop population using 3D point cloud[J]. Transactions of the Chinese Society of Agricultural Engineering 37 (07), 175–182.

- Dandrifosse, S., Bouvry, A., Leemans, V., et al., 2020. Imaging wheat canopy through stereo vision: Overcoming the challenges of the laboratory to field transition for morphological features extraction[J]. *Frontiers in Plant Science* 11, 96.
- Dhami, H., Yu, K., Xu, T., et al., 2020. Crop height and plot estimation for phenotyping from unmanned aerial vehicles using 3D LiDAR[C]. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2020, 2643–2649.
- Fang, W., Feng, H., Yang, W., et al., 2016. High-throughput volumetric reconstruction for 3D wheat plant architecture studies[J]. *Journal of Innovative Optical Health Sciences* 9 (05), 1650037.
- Gong, B., Ergu, D., Cai, Y., et al., 2020. Real-time detection for wheat head applying deep neural network[J]. *Sensors* 21 (1), 191.
- Grenzdörffer, G., 2014. Crop height determination with UAS point clouds[J]. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 40 (1), 135.
- Jin, S., Su, Y., Gao, S., et al., 2019. Separating the structural components of maize for field phenotyping using terrestrial LiDAR data and deep convolutional neural networks[J]. *IEEE Transactions on Geoscience and Remote Sensing* 58 (4), 2644–2658.
- Khaki, S., Safaei, N., Pham, H., et al., 2022. Wheatnet: A lightweight convolutional neural network for high-throughput image-based wheat head detection and counting [J]. *Neurocomputing* 489, 78–89.
- Lin, Y., 2015. LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? [J]. *Computers and electronics in Agriculture* 119, 61–73.
- Luo, Y., Wei, L., Xu, L., et al., 2022. Stereo-vision-based multi-crop harvesting edge detection for precise automatic steering of combine harvester[J]. *Biosystems Engineering* 215, 115–128.
- Mohamed, I., Dudley, R., 2019. Comparison of 3D Imaging Technologies for Wheat Phenotyping[C]. IOP Conference Series: Earth and Environmental Science, 012002.
- Mosa, A.S.M., Schön, B., Bertolotto, M., et al., 2012. Evaluating the benefits of octree-based indexing for LiDAR data[J]. *Photogrammetric Engineering & Remote Sensing* 78 (9), 927–934.
- Pankaj, D.S., Nidamanuri, R.R., Prasad, P.B., 2013. 3-D imaging techniques and review of products[C]. Proceedings of international Conference on innovations in Computer Science and Engineering.
- Pietragalla, J., Mullan, D., Reynolds, M., 2012. Physiological breeding II: a field guide to wheat phenotyping[M]. Cimmyt.
- Reza, M.N., Na, I.S., Baek, S.W., et al., 2019. Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images[J]. *Biosystems engineering* 177, 109–121.
- Saeys, W., Lenaerts, B., Craessaerts, G., et al., 2009. Estimation of the crop density of small grains using LiDAR sensors[J]. *Biosystems Engineering* 102 (1), 22–30.
- Salas Fernandez, M.G., Bao, Y., Tang, L., et al., 2017. A high-throughput, field-based phenotyping technology for tall biomass crops[J]. *Plant physiology* 174 (4), 2008–2022.
- Su, Y.-T., Bethel, J., Hu, S., 2016. Octree-based segmentation for terrestrial LiDAR point cloud data in industrial applications[J]. *ISPRS Journal of Photogrammetry and Remote Sensing* 113, 59–74.
- Sun, S., Li, C., Chee, P.W., et al., 2020. Three-dimensional photogrammetric mapping of cotton bolls in situ based on point cloud segmentation and clustering[J]. *ISPRS Journal of Photogrammetry and Remote Sensing* 160, 195–207.
- Sun, Y., Luo, Y., Chai, X., et al., 2021. Double-threshold segmentation of panicle and clustering adaptive density estimation for mature rice plants based on 3D point cloud [J]. *Electronics* 10 (7), 872.
- Thompson, A., Livina, V., Harris, P., et al., 2019. Model-based algorithms for phenotyping from 3D imaging of dense wheat crops[C]. IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor) 2019, 95–99.
- Vedaldi, A., Soatto, S., 2008. Quick shift and kernel methods for mode seeking[C]. European conference on computer vision 705–718.
- Velumani K. Wheat ear detection in plots by segmenting mobile laser scanner data[D]. University of Twente, 2017.
- Wang, Y., Chao, W.-L., Garg, D., et al., 2019. Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving[C]. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8445–8453.
- Wang F, Mohan V, Thompson A, et al. Dimension fitting of wheat spikes in dense 3D point clouds based on the adaptive k-means algorithm with dynamic perspectives [C]. 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), 2020a: 144–148.
- Wang, J., Zhang, Y., Gu, R., 2020b. Research status and prospects on plant canopy structure measurement using visual sensors based on three-dimensional reconstruction[J]. *Agriculture* 10 (10), 462.
- Wei, L., Luo, Y., Xu, L., et al., 2021. Deep Convolutional Neural Network for Rice Density Prescription Map at Ripening Stage Using Unmanned Aerial Vehicle-Based Remotely Sensed Images[J]. *Remote Sensing* 14 (1), 46.
- Wu, J., Xue, X., Zhang, S., et al., 2018. Plant 3D reconstruction based on LiDAR and multi-view sequence images[J]. *International Journal of Precision Agricultural Aviation* 1 (1).
- Yang, Y., Zhang, J., Wu, K., et al., 2021. 3d point cloud on semantic information for wheat reconstruction[J]. *Agriculture* 11 (5), 450.
- Zermas, D., Morellas, V., Mulla, D., et al., 2018. Extracting phenotypic characteristics of corn crops through the use of reconstructed 3D models[C]. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2018, 8247–8254.