Cereal Grain 3D Point Cloud Analysis Method For Shape Extraction And Filled/Unfilled Grain Identification Based On Structured Light Imaging

Zhijie Qin
Huazhong Agricultural University
Zhongfu Zhang
Huazhong Agricultural University
Xiangdong Hua
Huazhong Agricultural University
Wanneng Yang
National Center of Plant Gene Research(Wuhan), Huazhong Agricultural University
Xiuying Liang
Huazhong Agricultural University
Ruifang Zhai
College of Informatics, Huazhong Agricultural University
Chenglong huang (hcl@mail.hzau.edu.cn)
Huazhong Agricultural University

Research Article

Keywords: phenotypic, method, thickness, accuracy

Posted Date: September 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-882310/v1

Version of Record: A version of this preprint was published at Scientific Reports on February 24th, 2022.
See the published version at https://doi.org/10.1038/s41598-022-07221-4.

Cereal grain 3D point cloud analysis method for shape extraction and filled/unfilled grain identification based on structured light imaging

Zhijie Qin[1], Zhongfu Zhang[1], Xiangdong Hua[1], Wanneng Yang[2], Xiuying Liang[1], Ruifang Zhai[3], Chenglong Huang[1,*]

[1] *College of Engineering, Huazhong Agricultural University, Wuhan 430070, PR China*

[2] *National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research(Wuhan), Huazhong Agricultural University, Wuhan 430070, PR China*

[3] *College of Informatics, Huazhong Agricultural University, Wuhan 430070, PR China*

**Corresponding author (Email: hcl@mail.hzau.edu.cn)*

Authors email:

Zhijie Qin (379936867@qq.com)

Zhongfu Zhang (1742984225@qq.com)

Xiangdong Hua (969326373@qq.com)

Wanneng Yang (ywn@mail.hzau.edu.cn)

Xiuying Liang (nancy@mail.hzau.edu.cn)

Ruifang Zhai (251032630@qq.com)

Chenglong Huang (hcl@mail.hzau.edu.cn)

Abstract:

Cereals are the main food for mankind. The grain shape extraction and filled/unfilled grain recognition are meaningful for crop breeding and genetic analysis. The conventional measuring method is mainly manual, which is inefficient, labor-intensive and subjective. Therefore, a novel method was proposed to extract the phenotypic traits of cereal grains based on point clouds. First, a structured light scanner was used to obtain the grains point cloud data. Then, the single grain segmentation was accomplished by image preprocessing, plane fitting, region growth clustering. The length, width, thickness, surface area and volume was calculated by the specified analysis algorithms for grain point cloud. To demonstrate this method, experimental materials included rice, wheat and corn were tested. Compared with manual measurement results, the average measurement error of grain length, width and thickness was 2.07%, 0.97%, 1.13%, and the average measurement efficiency was about 9.6 seconds per grain. In addition, the grain identification model was conducted with 25 grain phenotypic traits, using 6 machine learning methods. The results showed that the best accuracy for filled/unfilled grain classification was 90.184%.The accuracy for indica and japonica identification was about 99.785%，while for different varieties identification was only 47.252%. Therefore, this method was proved to be an efficient and effective way for crop research.

Introduction

Because of population explosion, global warming, and water shortages, we are facing severe challenges in agricultural production[1-3]. Cereals mainly including rice, wheat, corn, and sorghum have occupied a dominant position in the human's food[4], and cereal production is of great importance to the food security[5,6]. Cereal grain traits including grain shape, grain plumpness have performed direct

45 influence on the final yield, and grain traits measurement are necessary for yield-related research[7].
46 Grain shape is a very important basis of grain classification, and plumpness is the criterion for judging
47 the quality of rice varieties. Therefore the grain trait extraction is essential for cereal research[8].
48 However, the conventional method mainly depends on manual measurement, which is inefficient,
49 labor-intensive and subjective. Therefore, it is urgent to develop a novel method for grain trait
50 extraction with high throughput and high accuracy.

51 The measurement of rice grain size is of great significance in rice breeding and genetic research.
52 With the rapid development of computer technology, machine vision has been applied in grain size
53 measurement[9,10]. Tanabata et al.[11] developed Smart-Grain software for high-throughput measurement
54 of seed shape based on digital images and the open computer vision library (OpenCV). Ma et al.[12]
55 extracted the length and width information of rice grains based on the images taken by smart phones.
56 Le et al.[13] proposed a method to study the morphology of developing wheat grains based on X-ray μCT
57 imaging technique. However, most of the researches focus on the 2D traits[14], and it is not easy to obtain
58 the 3D grain traits such as volume, surface area and thickness. Since the grain size are small, high
59 quality and complete point cloud of which is needed. Point clouds obtained by binocular stereo vision,
60 structure from motion and space carving are relatively sparse[15-18], on the contrary the structured light
61 imaging, an active three-dimensional vision technology, can obtain high-precision point clouds, which
62 is widely used in industrial detection, reverse engineering and cultural relic protection[19], and it provides
63 an effective method for high precision analysis of cereal grain 3D traits.

64 The rice grain plumpness is one of the determinants in yield, which is of great importance to rice
65 breeding. The number of filled grains per panicle is directly related to the crop yield. Therefore,
66 counting of filled and unfilled grains of a panicle is critical to judge the rice quality. Traditionally,

grain counting is performed manually, which is labor-intensive, time-consuming and subjective. Manually, filled grain is distinguished from unfilled grain by water-based or wind-based methods[20,21]. To improve it, some automated methods were developed for identifying and counting the filled grain. Duan et al.[22] proposed a method based on visible light imaging and soft X-ray imaging, which was expensive, and of radiation risk. Kumar et al.[23] built an automated system for discriminating and counting filled and unfilled grains of a rice panicle based on thermal images. Since the system required to monitor the temperature after heating the grains, it was complicated and difficult to achieve high-throughput measurement. Therefore, it is urgent to develop a new method for the recognition of filled/unfilled grains, with high efficiency and low radiation risk.

In this study, cereal grain traits analysis method based on point cloud was proposed. The high-precision point cloud of grains are obtained by structured light scanner, and the specified algorithms were designed for automatic segmentation of the grain point clouds and 3D grain trait extraction. Finally, 25 grain traits were computed, based on which, the model for filled/unfilled grain identification was set up. In conclusion, our research demonstrated a novel method for grain 3D and plumpness information extraction with high throughput and high accuracy, which was definitely helpful to the rice breeding and genetic research.

Results

To verify the accuracy of the algorithm, three experimenters used micrometers to measure the length, width and thickness of 2000 rice (including filled and unfilled grains), 100 wheat and 100 corn grains, and the mean value of the three measurements was taken as ground truth. The accuracy of the error analysis result is evaluated by mean absolute percentage error (MAPE), root mean square error

(RMSE) and determination coefficient (R^2). The relevant formula is as follows:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i}\frac{|x_i - y_i|}{x_i} \times 100\% \tag{10}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_i (x_i - y_i)^2}{n}} \tag{11}$$

$$R^2 = 1 - \frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i - \bar{y})^2} \tag{12}$$

where n is the total number of measurements; x_i is the manual measurement results; y_i is the system measurement results, and $\bar{y}$ is the mean of the system measurements.

Comparison of placement scheme

To verify the measurement accuracy, 100 filled grains of Zhonghua 11 were taken as samples to compare the precision of the horizontal placement scheme with the vertical placement scheme. Figure 9 shows the point cloud comparison in the two schemes. As the results shown in Fig.11, the measurement errors of length, width and thickness of the horizontal placement scheme were 4.55%, 4.05% and 3.82%, while the measurement errors of the vertical placement scheme were 2.15 %, 0.68% and 1.18%. As the Fig.10a-c and Fig.10d-f shown, the grain point clouds obtained by horizontal placement were incomplete due to the restriction of scanning angle, which obviously led to lower measurement accuracy, therefore the vertical placement scheme was proved to be preferable.

Accuracy analysis for length, width, thickness, surface area and volume

Accuracy analysis was performed on all 2200 samples including rice, wheat and corn, and the measuring results were shown in Fig.12. Fig.12a shows that the length measurement results of R^2, RMSE, MAPE was 0.9940, 0.210mm and 2.07% respectively. Figure 11b shows that the width measurement results of R^2, RMSE, MAPE was 0.9960, 0.076mm and 0.97% respectively. And Figure

11c shows that the thickness measurement results of R^2, RMSE, MAPE was 0.9960, 0.048mm and 1.13% respectively. The results showed that the system value was in good consistency with the manual value and the system method was able to extract the grain length, width and thickness of grains with high precision. Meanwhile, as shown in Fig.12d, the measurement errors of wheat and corn were generally smaller than rice, especially in the length, because the wheat and corn were more stable than rice when placed vertically, which led to higher scanning accuracy.

Due to the irregular surface morphology of the grains, the surface area and volume are difficult to measure in a non-destructive way. Therefore, as shown in Fig.13, a standard sphere with a radius of 10mm was adopted to verify the system method validity. The results showed that the surface area and volume measuring error were 2.83% and 1.75% respectively.

Correlation analysis of grain traits

The 25 grain traits extracted in this study could quantitatively describe the geometric shape of grain completely. For example, with the extracted traits in Zhonghua 11, a correlation matrix of Pearson coefficients[38] was calculated to identify inter-relationships. Intergroup correlation analysis was completed based on SPSS, and the results were shown in Fig.14. The results demonstrated that the correlation among the basic traits was strong and all of them were positive except thickness. Thickness as an important trait in grain shape had little correlation with length and width. In particular, the three compactness index were highly independent.

Recognition model of filled and unfilled grains

Filled and unfilled grain identification has great importance to the finally yield evaluation. In this

study, the classification models were studied by 6 different machine learning method with 25 phenotypic traits in Zhonghua 11. All classification models were performed on the Sklearn Tool Kit, and the main parameters were decided by learning curve and grid search method. Then 10-fold cross-validation method was applied to validate each model. The model results for filled and unfilled grains classification were shown in Table.3, the details of which was as follows:

(1) Classification and regression trees (CART): The model was constructed as follows: the information entropy was set as impurity criterion. Meanwhile, the maximum tree depth was 4, and tree branch decision mode was random. The accuracy of model classification was 85.447%.

(2) Random forest (RF): In this model test, the depth of the forest was set to 2, while the gini coefficient was adopted, and the number of base evaluators was set to 24. According to the validation results, the model classification accuracy reached 88.605%. Compared with CART, the model accuracy was significantly improved.

(3) Support vector machines (SVM): Since the distribution of original phenotypic traits is linearly inseparable, an optimal high-dimensional space was constructed by selecting the kernel function and the penalty factor. In this study, Gaussian kernel function was selected, and the penalty factor was set as 6. As a result, the accuracy of model classification was 89.684%.

(4) Naive bayes (NB): in this study, gaussian naive bayes was selected and the classification accuracy rate was 88.079%.

(5) Back propagation (BP) neural networks: The hidden layer was divided into two layers, in which the number of neurons in the first layer is 100 and the second layer is 50. The number of iterations was set to 2000, the initial learning rate was set to 0.0003237, and other parameters were the default values. Eventually, the classification accuracy of the model was 88.105%.

(6) Extreme gradient boosting (XGBoost): The classifier was constructed based on tree model. After the logistic regression loss function was selected, the number of weak classifiers was set as 20, while the maximum tree depth was set as 5, and the learning rate was set as 0.3. As a result, the classification accuracy of the model was 90.184%, which was the best in all the models.

In order to explore the contribution of phenotypic traits, the XGBoost classifier was analyzed in detail and the results were shown in Table 4. From the results, the thickness weight had reached 0.34, which was proved to be dominant in filled and unfilled grain classification. Furthermore, the traits including volumetric-width ratio, volume, length-thickness ratio and surface area-length radio were all related to length, the weight of which were greater than 4%. Moreover, 4 varieties of rice grains were selected to verify the traits significance in the filled and unfilled grain classification. As shown in Fig.15, the results indicated that the thickness had higher difference than width and length. The result also proved that the length had higher difference than width, especially in indica.

Classification of different rice varieties and classification of indica and japonica

The classification models for 10 different varieties of rice grains were built by the above machine learning methods. However, the best performance was 47.252% by the SVM model, and the worst performance was only 37.027% by the CART model. What's more, the classification models for indica and japonica were also constructed, and the results showed that all the models performed great, and the worst performance was 99.785% by the CART model.

Efficiency evaluation

To obtain the complete point clouds, 25 cereal grains would be scanned 8 times, and it took about

14 seconds for each time, while the sample turntable rotated 45 degree. Therefore, it took about 2 minutes for the point clouds acquisition. Meanwhile it took about 2 minutes for point clouds segmentation and phenotypic traits computation. Thus 25 grains measurement totally cost about 4 minutes, and the average efficiency was 9.6 seconds per grain. However the manual measurement efficiency was about 120 seconds per grain, which was one-twelfth of the system efficiency.

DISCUSSION

Cereal grain traits have important impact on the final yield, which are also necessary for crop breeding and genetic analysis. Phenotypic traits such as length, width, thickness, volume and surface area are of great significance. In this study, a novel method for grain trait extraction by 3D structured light imaging was invented with high-throughput and high-accuracy. In addition, the grain identification model was conducted with 25 grain phenotypic traits, using 6 machine learning methods. The results indicated that the thickness was dominant in filled and unfilled grain classification. The result also proved that the length had higher difference than width, especially in indica.

At present, distinguishing filled grain from unfilled grain mainly relies on water-based or wind-based methods which are inaccurate and destructive. There are few researches on the filled/unfilled grain distinction. Therefore, there is an urgent need for a method that can accurately identify filled and unfilled grains. Liu et al.[39] designed a method based on image analysis to measure grain plumpness by the grain shadow in four directions. In addition, some methods were proposed based on X-ray and thermal imaging[22,23], but all these methods were identified in 2D imaging and could not provide more phenotypic information. Hua et al. [24] extracted the point cloud of rice grains based on a laser scanner to calculate phenotypic information. However, it was not suitable for requirements of

high throughput. The method of this study can obtain the phenotypic information of grains with high precision and high efficiency, which provides a method for crop breeding research.

In the research of the placement method, it was confirmed that the vertical placement was more accurate than the horizontal placement. Also, it is worth noting that during the scanning process, the stability of the vertical placement played great effects on the measuring result. From the results, the measurement errors of wheat and corn were generally smaller than rice, especially in the length, because the wheat and corn were more stable than rice when placed vertically.

Conclusion

Based on the 3D structured light imaging, a novel method for cereal grain shape extraction and filled/unfilled grain identification was proposed. The results showed that the system measurement had high consistency with the manual measurement and the system method was able to extract the grain length, width and thickness of grains with high precision. Filled and unfilled grain identification were studied by 6 different machine learning method with 25 phenotypic traits. In conclusion, our research demonstrated a novel method for grain 3D and plumpness information extraction with high throughput and high accuracy, which was definitely helpful to the rice breeding and genetic research. Based on the experiment results, the following conclusions are drawn.

(1) Considering grain placement methods, the vertical placement scheme performed better results than the horizontal placement scheme. The measurement errors of length, width and thickness in the horizontal placement were 4.55%, 4.05% and 3.82%, while the measurement errors in the vertical placement were only 2.15%, 0.68% and 1.18%.

(2) 25 phenotypic traits of cereal grains could obtained automatically in batch, including 11 basic traits,

14 derived traits. And the average efficiency for single grain measurement was about 9.6 seconds, including 3D structure light imaging and point clouds analysis.

(3) 2200 samples including rice, corn and wheat were tested to evaluate this method, and the results showed that the average relative errors of length, width and thickness were 2.07%, 0.97% and 1.13%.

(4) With the extracted traits, a correlation matrix of Pearson coefficients was calculated to identify inter-relationships. The results demonstrated that thickness as an important trait in grain shape had little correlation with length and width. In particular, the three compactness index were highly independent.

(5) 6 machine learning methods were used to classify the phenotypic traits of the filled/unfilled grains of 10 kinds of grains. The results showed that XGBoost was the best in all the models, with the classification accuracy of the model was 90.184%, while the thickness was proved to be dominant in filled and unfilled grain classification. And for the classification among 10 different varieties of rice grains, the best performance was 47.252% by the SVM model, and the worst performance was only 37.027% by the CART model. What's more, all the models performed great to classify indica and japonica, and the worst performance was 99.785% by the CART model.

Material and Methods

Material

In this study, the test materials included rice, wheat, and corn three types of cereals, which were purchased from the market and rice was the main part. According to the classification of rice type, rice could be divided into two categories: indica and japonica. In this experiment, 5 varieties of indica and japonica were selected, each of which contained 100 filled grains and 100 unfilled grains, and a total of 2000 rice grains were used as experimental materials. Since the distinction between filled and unfilled

grains was still based on manual criteria in practice. Thus, in the experiment, three experimenters judged the same grain. When the judgment was consistent, the grain type was determined to be filled. Moreover, for wheat and corn, 100 grains were selected as samples. Varieties of experimental materials include: Japonica--Zhonghua 11, Wuyunjing 3, Nanjing 2728, Zhenghan 10, Nipponbare, Indica--C Liangyou Huazhan, Zhuliangyou 211, Liangyou 336, Fengliangyou No. 4, Guangliangyouxiang 66, Wheat--Jimai 22, Corn--Zhengdan 958. The following Figure 1 showed the experimental materials.

System design

3D Structured light scanner

The 3D structured light scanner (Reeyee Pro, China) was adopted in the study, which was based on white light LED raster scanning technology. Combing the advantages of structured light and binocular stereo vision, the scanner can achieve a single-sided accuracy of 0.05mm within 2s, which is suitable for high-precision scanning of small-sized work pieces, plastic products, and medical equipment. The main equipment is composed of a projector, two cameras and an internal modulated light source. Based on the principle of triangulation and sinusoidal grating image, it can obtain the dense point cloud data of objects. The detailed parameters of Reeyee Pro scanner are listed in Table 1. The structure of the scanner is shown in Fig.2b.

Cereal grain scanning system

As shown in Fig.2a, the whole system consists of 6 parts: structured light scanner, robot, scanner fixture, object platform, industrial computer and control unit. AUBO i5 robot was adopted, which was a 6 degrees of freedom (DOF) collaborative robot with a positioning accuracy of ±0.02mm and a

maximum load of 5kg. The working range of the robotic arm was a sphere with a radius of 886.5 mm, which ensured sufficient scanning space. In order to fix the scanner on the robot, a fixture was designed and 3D printed with ABS material, and the entire weight of the scanner and the fixture was less than 2.5kg. The object platform was designed to fix the robot and place samples. The industrial computer was connected with the control unit and the scanner, to achieve the cooperative operation of robot movement and the scanner imaging.

Cereal grain point cloud acquisition

The cereal grain point cloud acquisition is shown as Fig.3, which could be divided into 4 steps: the scanner calibration, the selection of the placement schemes, the scanning path determination, and the batch scanning.

(1) Calibration of the scanner. The structured light scanner needed to be calibrated and corrected before working. When calibrating the camera, the calibration board need to be set in four positions including the directions of 0°, 90°, 180°, and 270°. Then the distance between the scanner and the calibration board should be adjusted from 350mm to 450mm, while collecting images.

(2) Selection of the placement schemes. At present, there are mainly two kinds of three-dimensional scanning schemes for grains. One way is to spread the grains flatly on a platform, and another is to fix the grains through the seed holder[24]. The former has high efficiency, but the accuracy is low because the scanning grain is not complete. The latter obtains the complete point cloud of grain with high accuracy, but the disadvantage is that it can only scan a single grain, which is too time-consuming. To improve it, the grains were directly fixed vertically on the stage, and multiple grains could be scanned completely in the study.

(3) Determination of the scanning path point. As shown in Table 1, the minimum space point distance of Reeyee Pro is 0.16mm. To achieve as high spatial resolution as possible, the robot was studied to obtain proper scanning path point. In this study, the average minimum point distance of the grain point cloud was capable of reaching 0.1731mm.

(4) Batch scanning. Due to the limitation of scanning area and rotation effect, the grain placement range was set to $100 \times 100\text{mm}$ in the center. In addition, the distance between adjacent grains was set as 20mm to avoid grain shading. Meanwhile, the grain placement strategy was $\mathbf{6 \times 4 + 1}$ (4 rows for every 6 grains in a row, and the last one is placed separately), which is helpful for matching the manual and automatic values. What is more, the scanning strategy of rotating 8 times and scanning 45 degree a time was adopted.

System development environment

The configuration of the industrial computer is I5 3470 and GTX1050TI. The development environment is Windows 7 Pro, Visual Studio 2015, cross-platform open source Point Cloud library (PCL) based on C++, QT 5.9.8, Python 3.7.6 and Visualization Toolkit (VTK). In addition, there is a software, Reeyee-Pro_V2.6.1.0, which can display the 3D data collection of the point cloud in real time. And the robot is controlled by the Robot Operating System (ROS) system in the Ubuntu environment.

Cereal grain point cloud processing pipeline

The overall processing pipeline of cereal grain point cloud is shown in Fig.4. It mainly includes 4 steps: point cloud preprocessing, point cloud segmentation, phenotypic traits calculation, and filled/unfilled grain recognition.

306 **Preprocessing of point clouds**

307 The preprocessing procedure of grain point cloud was shown in Fig.5, mainly including 3 steps:
308 coordinate transformation, down sampling, and filtering.

309 (1) The coordinate transformation were conducted as equation 1, 2. Firstly, move the original
310 coordinates (T0) to the centroid point of point cloud. Then, based on principal component analysis
311 (PCA)[25], the covariance matrix (MT) was computed to generate the new coordinate. The transformed
312 result was shown as Fig.5b.

$$T_A = M_T \times (T_0 - A) \quad (1)$$

$$M_T = \{e_1, e_2, e_3\}^T \quad (2)$$

313 Where $\boldsymbol{e_1, e_2, e_3}$ are the three unit eigenvectors of the covariance matrix $\boldsymbol{M_T}$; $\boldsymbol{T_0}$ is the original point
314 cloud coordinates; $\boldsymbol{T_A}$ is the new coordinates after coordinate transformation; A is the translation
315 matrix from the original coordinates (T0) to the centroid point of point cloud.

316 (2) Point cloud down sampling and filtering was shown as Fig.5c. Based on voxel grids, all points in
317 the voxel were replaced by the gravity center to reduce the point cloud, which can effectively improve
318 the processing efficiency[26]. Then statistical filtering algorithm was applied to remove point data[27], in
319 which the point distance is abnormal.

320

321 **Segmentation of point cloud**

322 The segmentation of point cloud was conducted as Fig.6. After the preprocessing, the random
323 sample consensus algorithm (RANSAC) was adopted to fit the sample stage plane[28] and separate the
324 grain point clouds from the background. Then, based on curvature and normal angle, the single grain
325 point cloud was identified by region growing algorithm[29].

Phenotypic traits calculation

After single grain was obtained, phenotypic traits were extracted, including length, width, thickness, volume, surface area, projected area and perimeter in the main direction. Figure 7 shows the processing steps for grain trait extraction.

(1) Grain length, width and thickness extraction

As shown in Fig.7a-c, the extraction of grain length, width and thickness was mainly achieved by constructing a bounding box. Firstly, the coordinate system of the segmented single grain point clouds were transformed to convert axis-aligned bounding box (AABB)[30] into orientation bounding box (OBB)[31]. Secondly, the maximum and minimum values of the transformed single grain point cloud in the new coordinate system were calculated as $\boldsymbol{x_{max}}$、$\boldsymbol{x_{min}}$、$\boldsymbol{y_{max}}$、$\boldsymbol{y_{min}}$、$\boldsymbol{z_{max}}$、$\boldsymbol{z_{min}}$ respectively. Finally, the grain length, width and thickness were computed as following equations.

$$l = x_{max} - x_{min} \tag{3}$$

$$w = y_{max} - y_{min} \tag{4}$$

$$h = z_{max} - z_{min} \tag{5}$$

Where $\boldsymbol{l}$, $\boldsymbol{w}$ and $\boldsymbol{h}$ are the length, width and thickness of a grain, respectively.

(2) Grain surface area extraction

Firstly, the triangular mesh model of the point clouds was established by greedy projection triangulation algorithm[32].Secondly, the holes were filled by reconstructing the mesh boundary edges, which were generated by the grain segmentation. As shown in Fig.7e, the length of the side of the triangle was calculated by the coordinates of the three vertices of the triangle. Then, based on Helen's

formula[33], the areas of all the triangular surfaces were calculated and the sum of them was used to approximate the surface area of the grain. The calculation formula is as equation 6-7.

$$S_0 = \sum_{i=1}^{k} s_i \tag{6}$$

$$s_i = \sqrt{p_i(p_i - a_i)(p_i - b_i)(p_i - c_i)} \tag{7}$$

where $\boldsymbol{S_0}$ is surface area of a grain, k is total number of triangles, $\boldsymbol{s_i}$ is area of the i-th triangle, $\boldsymbol{p_i}$ is half the perimeter of the triangle, $\boldsymbol{a_i}$, $\boldsymbol{b_i}$ $\boldsymbol{and}$ $\boldsymbol{c_i}$ represent the length of each side of the triangle.

(3) Grain Volume extraction

The grain volume was extracted as Fig.8. Firstly, the convex pentahedrons were constructed by the triangular mesh and central plane projection, and then grain volume V was the sum of their volumes. Fig.8a is the central plane of the triangular mesh projection. And as shown in the Fig.8b, $\boldsymbol{A_1}$, $\boldsymbol{B_1}$ and $\boldsymbol{C_1}$ are the three vertices of a triangular mesh. It is assumed that the volume of the straight triangular prism $\boldsymbol{A_0B_0C_0ABC}$ is equal to the volume of this convex pentahedron, then the height of the straight triangular prism could be approximated as the height of the gravity center of $\triangle \boldsymbol{A_1B_1C_1}$.

$$V_{A_1B_1C_1\mathrm{ABC}} = V_{A_0B_0C_0\mathrm{ABC}} = S_{\triangle ABC} \times \mathrm{h} \approx S_{\triangle ABC} \times h_0 \tag{8}$$

where $\boldsymbol{h}$ is height of the straight prism, $\boldsymbol{h_0}$ is height of the center of gravity of $\triangle \boldsymbol{A_1B_1C_1}$

(4) Projected area and perimeter of grain in the main direction extraction

In this study, three main directions of grain point cloud were projected, and the projected area and perimeter of cross section, longitudinal section, and horizontal section were obtained as the shape description of grain (Fig.7d). Firstly, the point cloud of a single grain after coordinate transformation was projected on the plane of x=0, y=0, z=0 respectively. Then, based on the greedy projection

triangulation algorithm[32], the areas of the projected triangular mesh and the perimeter of the mesh edges were calculated.

Filled/unfilled grain analysis

A total of 25 phenotypic traits were extracted in the study, including 11 basic traits and 14 derived traits, as shown in Table 2. Compactness index, as a comprehensive grain shape description factor[34], is calculated by the following formula:

$$c = \frac{C^2}{4\pi A} \tag{9}$$

Where c is the compactness index, $\boldsymbol{C}$ is perimeter of cross-section, $\boldsymbol{A}$ is area of cross-section

With the rice grain phenotypic dataset, the models of recognition between filled and unfilled grains, distinction between indica and japonica, and classification of rice grain sub-varieties were established by six different machine learning algorithms including decision tree, random forest, support vector machine, Naive Bayes, XGBoost, and BP neural network[35-37].

System software design

In order to facilitate grain 3D point cloud analysis, A specific user software was designed based on QT Designer, PCL, QVTKWidget and XGBoost as shown in Fig.9, in which the above algorithms including grain point cloud processing, grain traits calculation and analysis were integrated. The segmentation window displayed the original point cloud and the grain segmentation result as shown in Fig.9a. Meanwhile, in order to predict the grain category and the plumpness, the python script was adopted to load the filled/unfilled grain classification model, and the result window displayed the single grain point cloud, 11 basic traits, categories and plumpness as shown in Fig.9b. Moreover, the software

parameters of plane segmentation threshold and cluster point cloud range were able to be easily modified by users to optimize the grain segmentation result. Finally the results including grain point cloud and traits would be saved, and the software operation was shown as Supplementary Video S1.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Approval for plant experiments

We confirmed that all experiments were performed in accordance with relevant named guidelines and regulations.

References

1 He, T. & Li, C. Harness the power of genomic selection and the potential of germplasm in crop breeding for global food security in the era with rapid climate change. The Crop Journal 8, 688-700, doi:10.1016/j.cj.2020.04.005 (2020).

2 Tester, M. & Langridge, P. Breeding technologies to increase crop production in a changing world. Science 327, 818-822, doi:10.1126/science.1183700 (2010).

3 Wang, Z. et al. Identification of QTLs with main, epistatic and QTL x environment interaction effects for salt tolerance in rice seedlings under different salinity conditions. Theor Appl Genet 125, 807-815, doi:10.1007/s00122-012-1873-z (2012).

4 Jiang, L. et al. Increased grain production of cultivated land by closing the existing cropping intensity gap in Southern China. Food Security, doi:10.1007/s12571-021-01154-y (2021).

5 Fan, M. et al. Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. J Exp Bot 63, 13-24, doi:10.1093/jxb/err248 (2012).

6 Zhou, W.-b., Wang, H.-y., Hu, X. & Duan, F.-y. Spatial variation of technical efficiency of cereal production in China at the farm level. Journal of Integrative Agriculture 20, 470-481, doi:10.1016/s2095-3119(20)63579-1 (2021).

7 Upadhyaya, H. D., Reddy, K. N., Singh, S. & Gowda, C. L. L. Phenotypic diversity in Cajanus species and identification of promising sources for agronomic traits and seed protein content. Genetic Resources and Crop Evolution 60, 639-659, doi:10.1007/s10722-012-9864-0 (2012).

8 Yang, W. et al. Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. Mol Plant 13, 187-214, doi:10.1016/j.molp.2020.01.008 (2020).

9 Tellaeche, A., BurgosArtizzu, X. P., Pajares, G., Ribeiro, A. & Fernández-Quintanilla, C. A new vision-based approach to differential spraying in precision agriculture. Computers and Electronics in Agriculture 60, 144-155, doi:10.1016/j.compag.2007.07.008 (2008).

10 Yang, W., Duan, L., Chen, G., Xiong, L. & Liu, Q. Plant phenomics and high-throughput phenotyping: accelerating rice functional genomics using multidisciplinary technologies. Curr Opin Plant Biol 16, 180-187, doi:10.1016/j.pbi.2013.03.005 (2013).

11 Tanabata, T., Shibaya, T., Hori, K., Ebana, K. & Yano, M. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. Plant Physiol 160, 1871-1880, doi:10.1104/pp.112.205120 (2012).

12 Zhihong., M., Yuhan., M. & Chengliang., G. L. L. Smartphone-Based Visual Measurement and Portable Instrumentation for Crop Seed Phenotyping. Ifac Papersonline 49(16), 259-264, doi:10.1016/j.ifacol.2016.10.048 (2016).

13 Le, T. D. Q., Alvarado, C., Girousse, C., Legland, D. & Chateigner-Boutin, A. L. Use of X-ray micro computed tomography imaging to analyze the morphology of wheat grain through its development. Plant Methods

15, 84, doi:10.1186/s13007-019-0468-y (2019).

14 Ducournau, S. et al. High throughput phenotyping dataset related to seed and seedling traits of sugar beet genotypes. Data Brief 29, 105201, doi:10.1016/j.dib.2020.105201 (2020).

15 An, P., Fang, K., Jiang, Q., Zhang, H. & Zhang, Y. Measurement of Rock Joint Surfaces by Using Smartphone Structure from Motion (SfM) Photogrammetry. Sensors (Basel) 21, doi:10.3390/s21030922 (2021).

16 Gutierrez, A., Jimenez, M. J., Monaghan, D. & O'Connor, N. E. Topological evaluation of volume reconstructions by voxel carving. Computer Vision and Image Understanding 121, 27-35, doi:10.1016/j.cviu.2013.11.005 (2014).

17 Jay, S., Rabatel, G., Hadoux, X., Moura, D. & Gorretta, N. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. Computers and Electronics in Agriculture 110, 70-77, doi:10.1016/j.compag.2014.09.021 (2015).

18 Kim, W.-S. et al. Stereo-vision-based crop height estimation for agricultural robots. Computers and Electronics in Agriculture 181, doi:10.1016/j.compag.2020.105937 (2021).

19 Paulus, S. Measuring crops in 3D: using geometry for plant phenotyping. Plant Methods 15, 103, doi:10.1186/s13007-019-0490-0 (2019).

20 Liu, T. et al. A shadow-based method to calculate the percentage of filled rice grains. Biosystems Engineering 150, 79-88, doi:10.1016/j.biosystemseng.2016.07.011 (2016).

21 Yang, J. et al. Grain filling pattern and cytokinin content in the grains and roots of rice plants. Plant Growth Regulation 30(3), 261-270, doi:10.1023/A:1006356125418 (2000).

22 Tirol-Padre, A. et al. Grain yield performance of rice genotypes at suboptimal levels of soil N as affected by N uptake and utilization efficiency. Field Crops Research 46, 127-143, doi:10.1016/0378-4290(95)00095-x (1996).

23 Duan, L. et al. Fast discrimination and counting of filled/unfilled rice spikelets based on bi-modal imaging.

Computers and Electronics in Agriculture 75, 196-203, doi:10.1016/j.compag.2010.11.004 (2011).

24 Kumar, A. et al. Discrimination of filled and unfilled grains of rice panicles using thermal and RGB images. Journal of Cereal Science 95, doi:10.1016/j.jcs.2020.103037 (2020).

25 li, H. et al. Calculation method of surface shape feature of rice seed based on point cloud. Computers and Electronics in Agriculture 142, 416-423, doi:10.1016/j.compag.2017.09.009 (2017).

26 Wold, S., Esbensen, K. & Geladi, P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems 2, 37-52, doi:10.1016/0169-7439(87)80084-9 (1987).

27 Chiang, P.-Y. & Kuo, C. C. J. Voxel-based shape decomposition for feature-preserving 3D thumbnail creation. Journal of Visual Communication and Image Representation 23, 1-11, doi:10.1016/j.jvcir.2011.07.008 (2012).

28 Rusu, R. B., Marton, Z. C., Blodow, N., Dolha, M. & Beetz, M. Towards 3D Point cloud based object maps for household environments. Robotics and Autonomous Systems 56, 927-941, doi:10.1016/j.robot.2008.08.005 (2008).

29 Schnabel, R., Wahl, R. & Klein, R. in John Wiley & Sons, Ltd Vol. 26 214-226 (2007).

30 Date, H., Kanai, S. & Kawashima, K. As-built modeling of piping system from terrestrial laser-scanned point clouds using normal-based region growing. Journal of Computational Design and Engineering 1, 13-26, doi:10.7315/jcde.2014.002 (2014).

31 Bergen, G. v. d. Efficient Collision Detection of Complex Deformable Models using AABB Trees. Journal of Graphics Tools 2, 1-13, doi:10.1080/10867651.1997.10487480 (1997).

32 Dimitrov, D., Knauer, C., Kriegel, K. & Rote, G. Bounds on the quality of the PCA bounding boxes. Computational Geometry 42, 772-789, doi:10.1016/j.comgeo.2008.02.007 (2009).

33 Marton, Z. C., Rusu, R. B. & Beetz, M. in IEEE International Conference on Robotics & Automation. IEEE

(2009).

34 Connelly, R. Comments on generalized Heron polynomials and Robbins' conjectures. Discrete Mathematics 309, 4192-4196, doi:10.1016/j.disc.2008.10.031 (2009).

35 Shouche, S. P., Rastogi, R., Bhagwat, S. G. & Sainis, J. K. Shape analysis of grains of Indian wheat varieties. Computers and Electronics in Agriculture, 55-76, doi:10.1016/S0168-1699(01)00174-0 (2001).

36 Kurt, I., Ture, M. & Kurum, A. T. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. Expert Systems with Applications 34, 366-374, doi:10.1016/j.eswa.2006.09.004 (2008).

37 Narayan, Y. Comparative analysis of SVM and Naive Bayes classifier for the SEMG signal classification. Materials Today: Proceedings 37, 3241-3245, doi:10.1016/j.matpr.2020.09.093 (2021).

38 Wang, X.-W. & Liu, Y.-Y. Comparative study of classifiers for human microbiome data. Medicine in Microecology 4, doi:10.1016/j.medmic.2020.100013 (2020).

39 Edelmann, D., Móri, T. F. & Székely, G. J. On relationships between the Pearson and the distance correlation coefficients. Statistics & Probability Letters 169, doi:10.1016/j.spl.2020.108960 (2021).

Acknowledgements

We would like to thank all the colleagues in the Crop Phenotyping Center, Huazhong Agricultural University for their helping the experiments. This work was supported by grants from National Key Research and Development Plan of China (2016YFD0100101-18), the National Natural Science Foundation of China (31600287, 31770397), and the Fundamental Research Funds for the Central Universities (2662018JC004).

Author's contributions

Z.Q. designed the research, performed the experiments, analyzed the data and wrote the manuscript. Z.Z., X.H. also performed experiments. W.Y, X. L, R. Z Revised the manuscript. C.H. supervised the project.

Competing interests

The authors declare that they have no competing interests.

Figure Legends

Figure 1 Display of experimental materials, including wheat grains, corn grains and 10 different varieties of rice grains.

Figure 2 Schematic diagram of cereal grain scanning system. (a) The overall structure, (b) the structured light scanner

Figure 3 Flow chart of obtaining point cloud of cereal grains using structured light scanning system.

Figure 4 Cereal grain point cloud processing pipeline.

Figure 5 The process and result of preprocessing. (a) Original point cloud position, (b) Transformed point cloud position, (c) Single grain point cloud

Figure 6 Cereal grain segmentation pipeline

Figure 7 Overview of the proposed method for extraction of grain traits.

Figure 8 Grain volume calculation method in this study. (a) The central plane of triangular mesh projection, (b) the projected area integration method.

Figure 9 The user software for grain 3D point cloud analysis. (a) Grain 3D point cloud processing, (b) grain traits extraction.

Figure 10 Comparison of two placement schemes, (a) the horizontal placement scheme (b) the scanned point clouds in horizontal placement (c) the single grain point cloud in horizontal placement (d) the vertical placement scheme, (e) the scanned point clouds in vertical placement, (f) the single grain point cloud in vertical placement.

Figure 11 The measuring results in different placement scheme. (a) Length (vertical) (b) Width (vertical) (c) Thickness (vertical) (d) Length (horizontal) (e) Width (horizontal) (f) Thickness (horizontal)

Figure 12 The sample accuracy analysis. (a) Length (b) Width (c) Thickness (d) japonica, indica, wheat and corn grains mean relative error.

Figure 13 The sample sphere. (a) Real object (b) Point cloud (c) Triangular mesh

Figure 14 The result of grain traits correlation analysis.

Figure 15 The Comparison of main traits between filled grain and unfilled grain. (a) Zhonghua 11 (b) Wuyunjing 3 (c) C Liangyou Huazhan (d) Zhuliangyou 211

Table 1. Reeyee Pro scanner detailed parameters.

Parameter	Value
Light source	White LED
Point distance	0.16mm
Spatial resolution	0.05mm
Scanning area	210×150mm
Working distance	290-480mm
Maximum scan size	$200 \times 200 \times 200$mm

Table 2. 25 phenotypic traits.

No.	**Symbol**	**Trait**	**No.**	**Symbol**	**Trait**
1	l	Length	14	w/h	Width-thickness ratio
2	w	Width	15	V_{obb}	Box volume
3	h	Thickness	16	S/V	Specific surface area
4	V	Volume	17	S/l	Surface area-length

					ratio
5	S	Surface area	18	S/w	Surface area-width ratio
6	C_{yz}	Perimeter of cross section	19	S/h	Surface area-thickness ratio
7	S_{yz}	Area of cross section	20	V/l	Volume-length ratio
8	C_{xz}	Perimeter of longitudinal section	21	V/w	Volume-width ratio
9	S_{xz}	Area of longitudinal section	22	V/h	Volume-thickness ratio
10	C_{xy}	Perimeter of horizontal section	23	c_{yz}	Compactness index of cross section
11	S_{xy}	Area of horizontal section	24	c_{xz}	Compactness index of longitudinal section
12	l/w	Length-width ratio	25	c_{xy}	Compactness index of horizontal section
13	l/h	Length-thickness ratio			

527

528 Table 3. Comparison of test results of each classification method based on 25 phenotypic traits.

Classification method	Accuracy	F1 score	AUC value
CART	85.447%	0.85706	0.91667
RF	88.605%	0.89145	0.95511
SVM	89.684%	0.90371	0.94822

NB	88.079%	0.89363	0.93611
BP	88.105%	0.88811	0.94478
XGBoost	90.184%	0.90615	0.96933

529

530 Table 4. Weight rank of characteristic traits (> 4%).

Rank	Trait	Importance weight
1	Thickness	0.342219
2	Length	0.067255
3	Perimeter of horizontal section	0.062472
4	Volume-width ratio	0.056376
5	Compactness index of horizontal section	0.053502
6	Volume	0.049749
7	Length-thickness ratio	0.042486
8	Surface area-length ratio	0.042199

531

Figures

Figure 1

Display of experimental materials, including wheat grains, corn grains and 10 different varieties of rice grains.

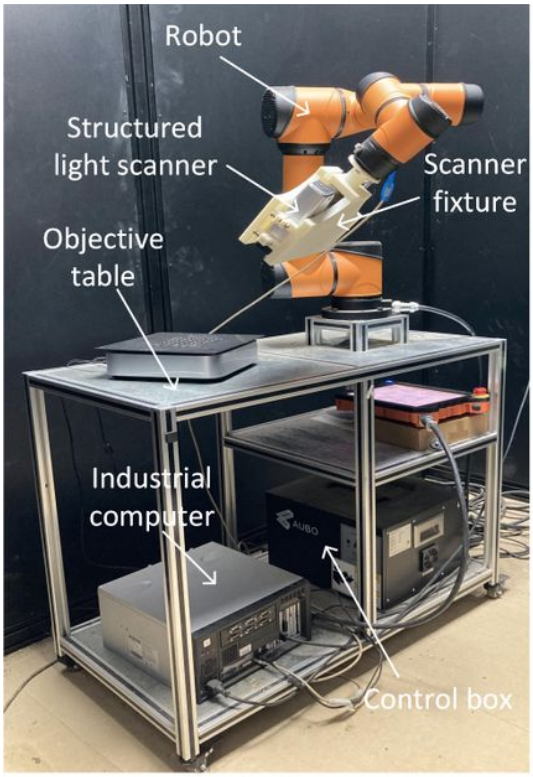

Figure 2

Schematic diagram of cereal grain scanning system. (a) The overall structure, (b) the structured light scanner

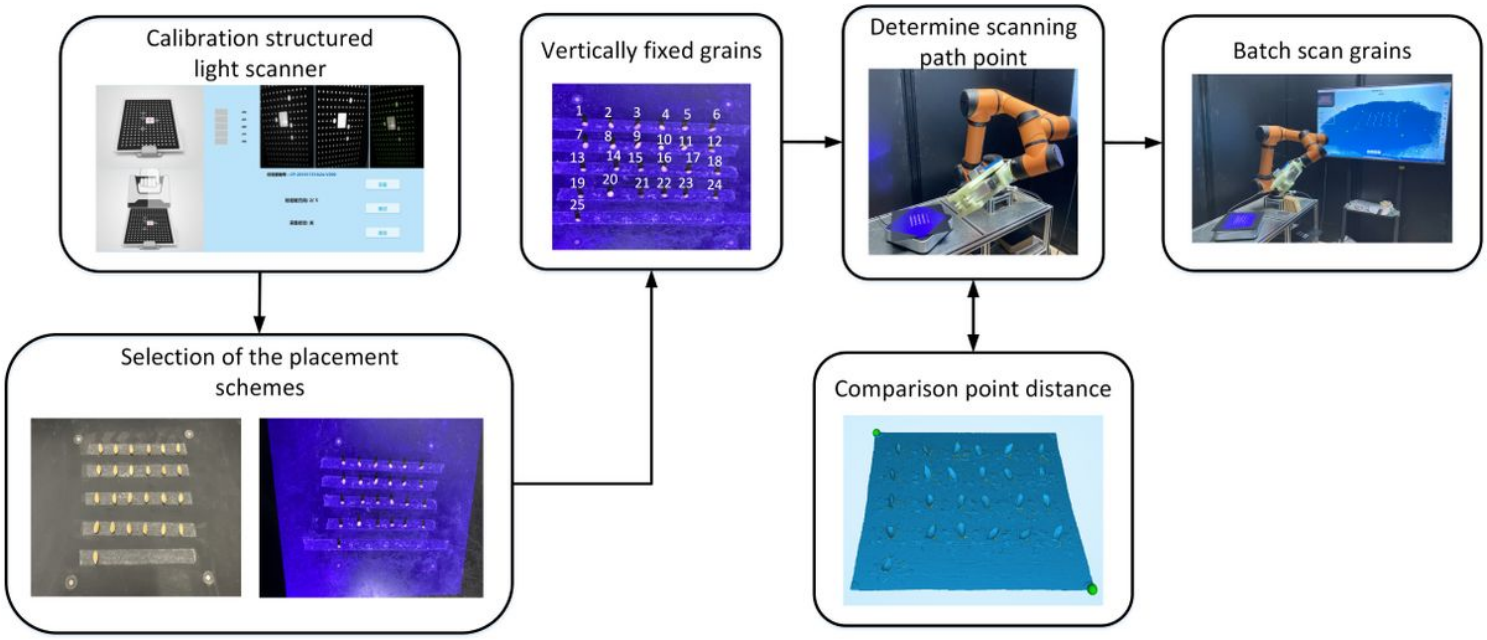

Figure 3

Flow chart of obtaining point cloud of cereal grains using structured light scanning system.

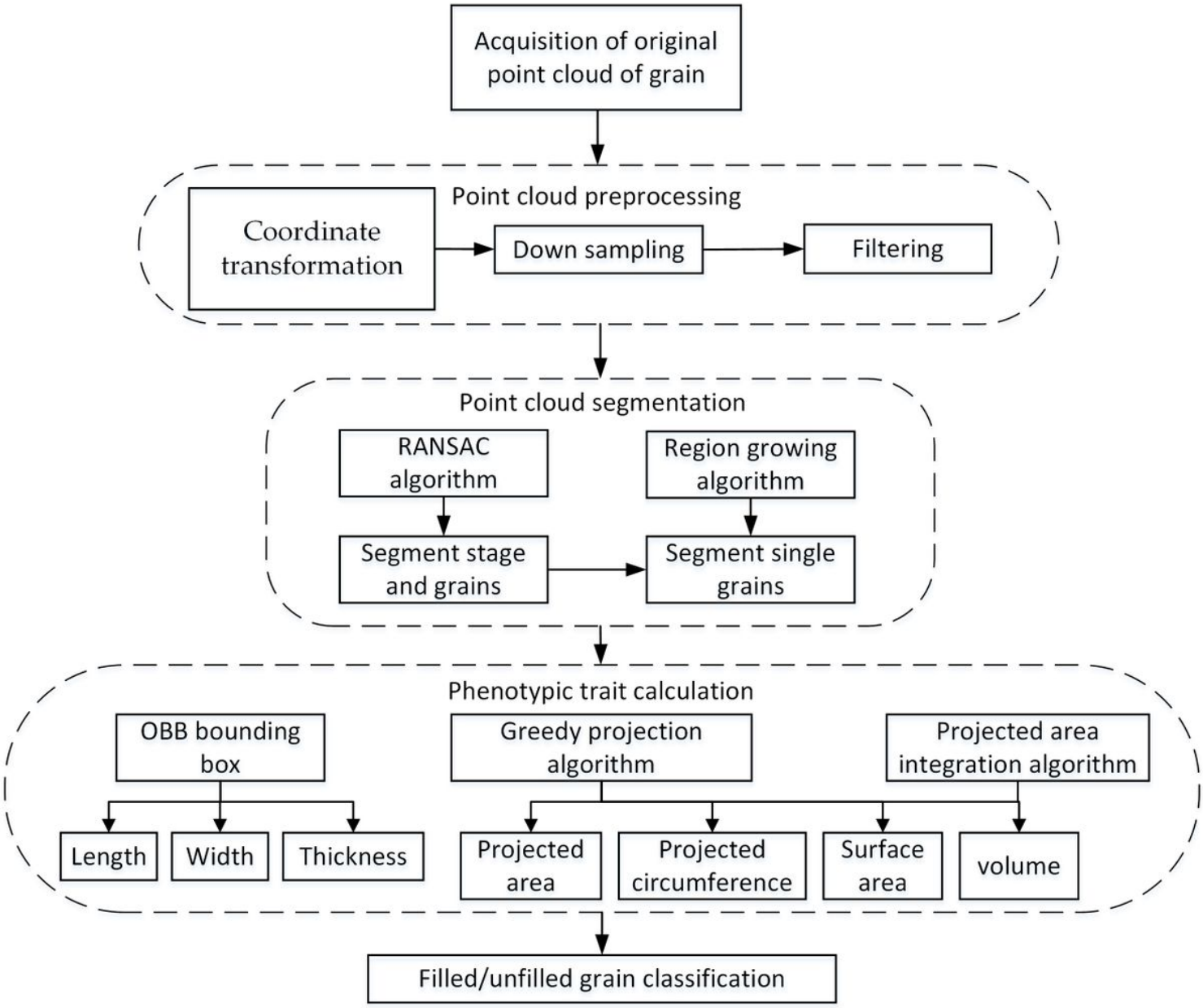

Figure 4

Cereal grain point cloud processing pipeline.

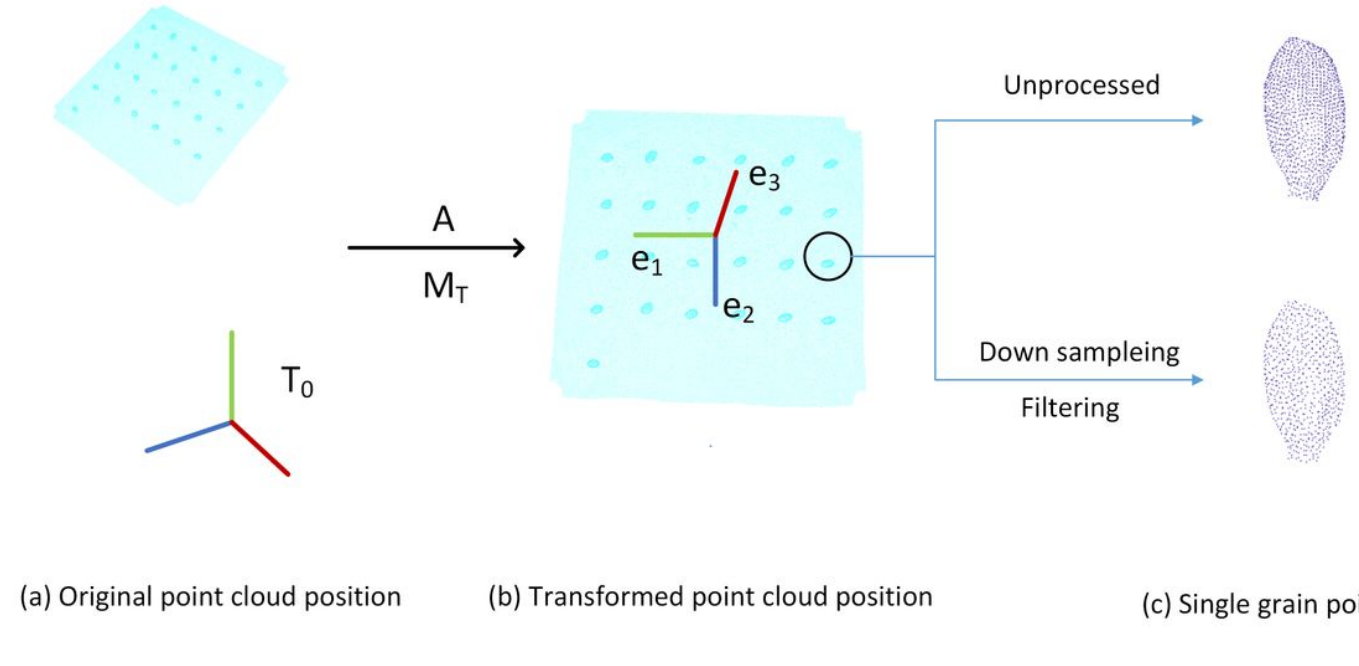

Figure 5

The process and result of preprocessing. (a) Original point cloud position, (b) Transformed point cloud position, (c) Single grain point cloud

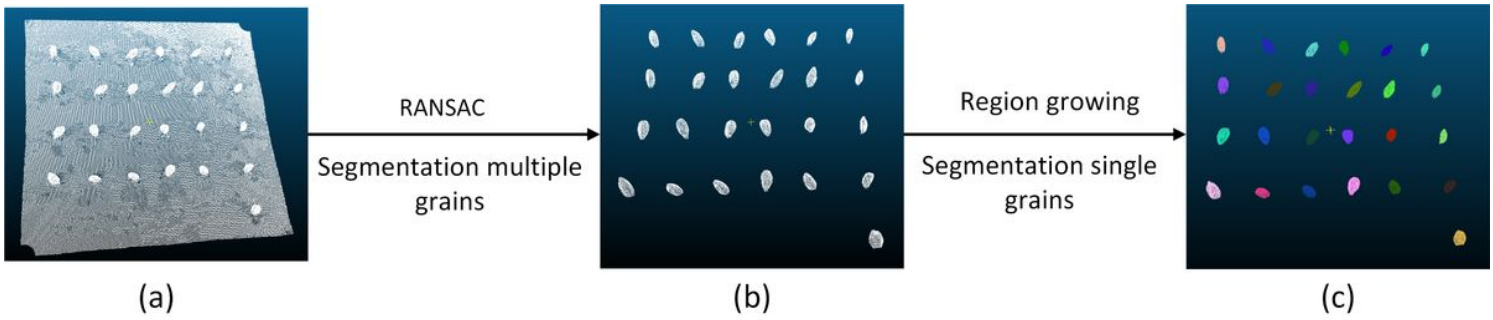

Figure 6

Cereal grain segmentation pipeline

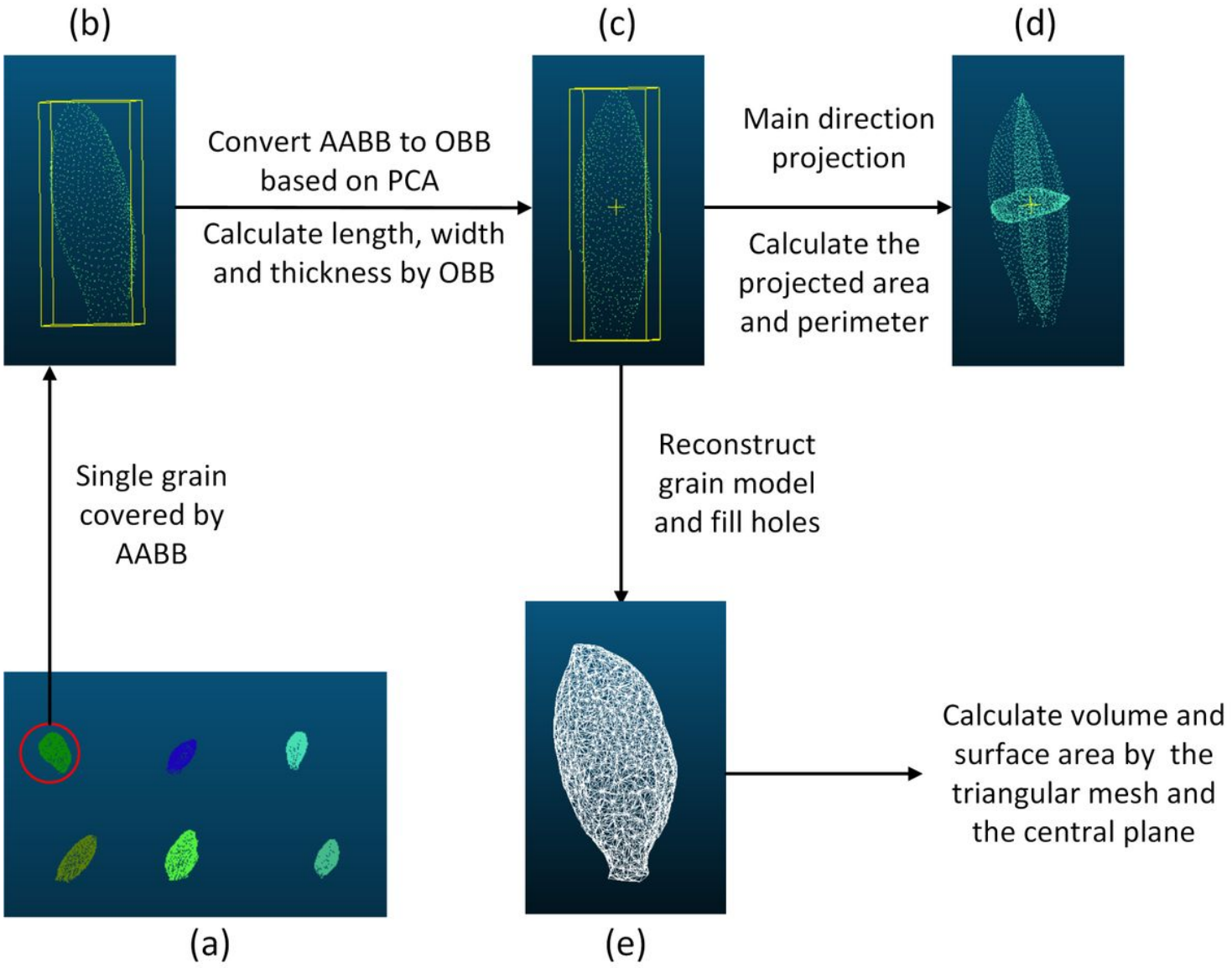

Figure 7

Overview of the proposed method for extraction of grain traits.

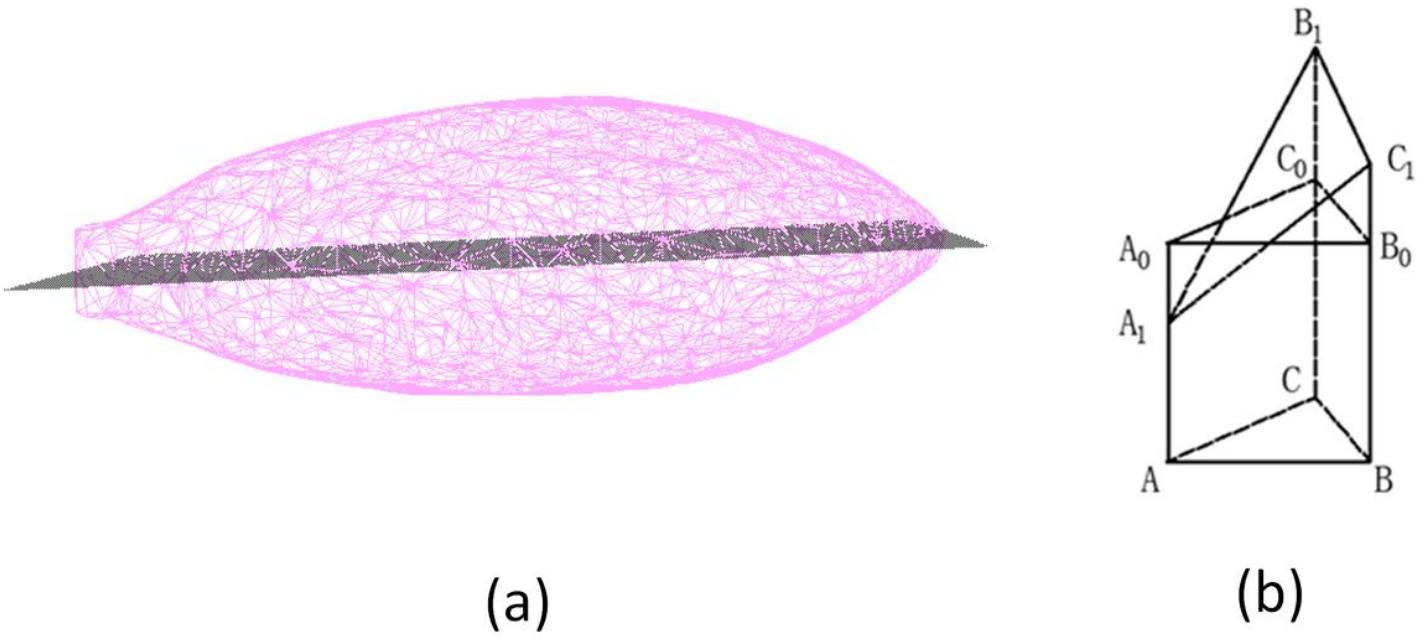

Figure 8

Grain volume calculation method in this study. (a) The central plane of triangular mesh projection, (b) the projected area integration method.

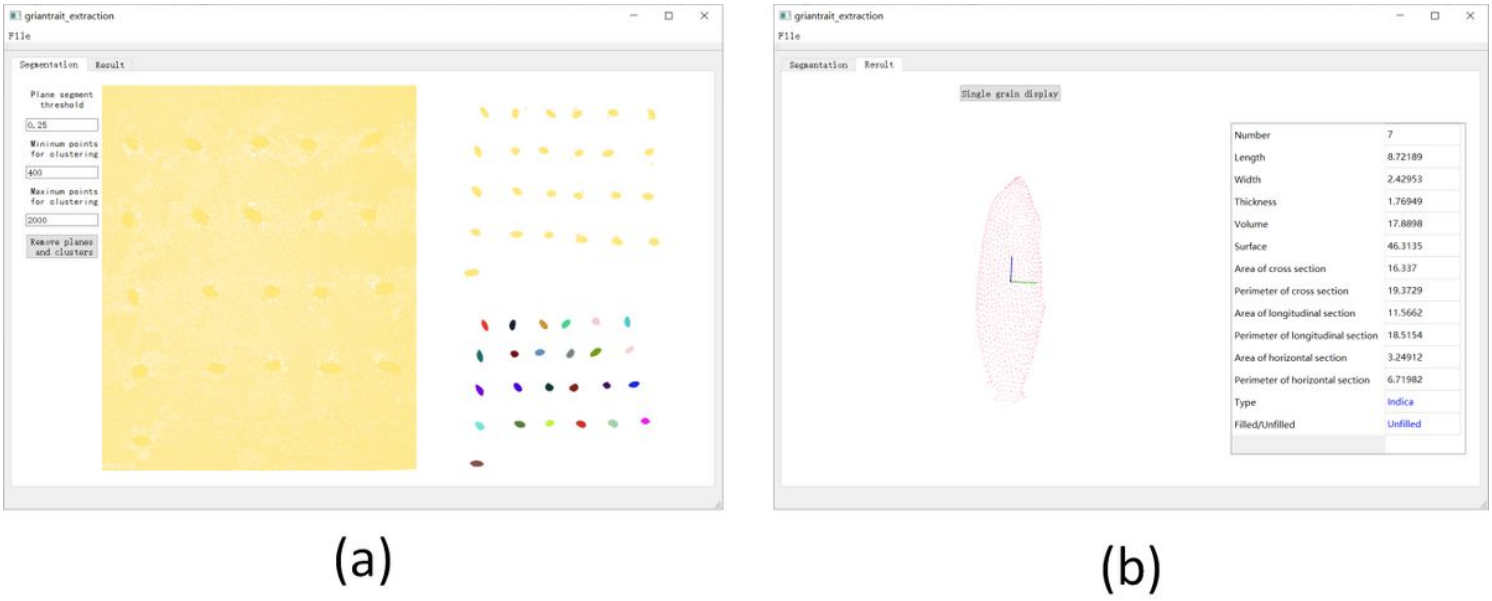

Figure 9

The user software for grain 3D point cloud analysis. (a) Grain 3D point cloud processing, (b) grain traits extraction.

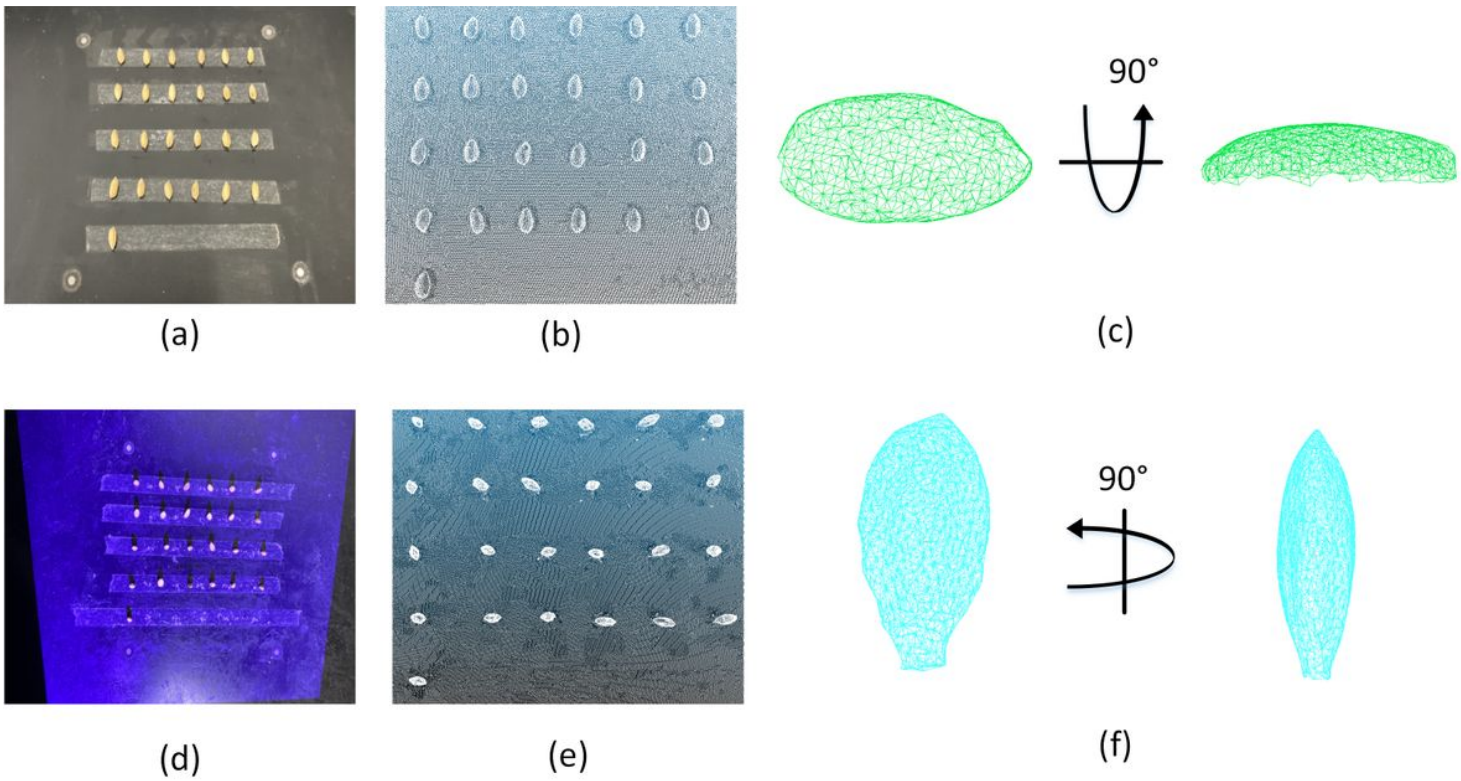

Figure 10

Comparison of two placement schemes, (a) the horizontal placement scheme (b) the scanned point clouds in horizontal placement (c) the single grain point cloud in horizontal placement (d) the vertical placement scheme, (e) the scanned point clouds in vertical placement, (f) the single grain point cloud in vertical placement.

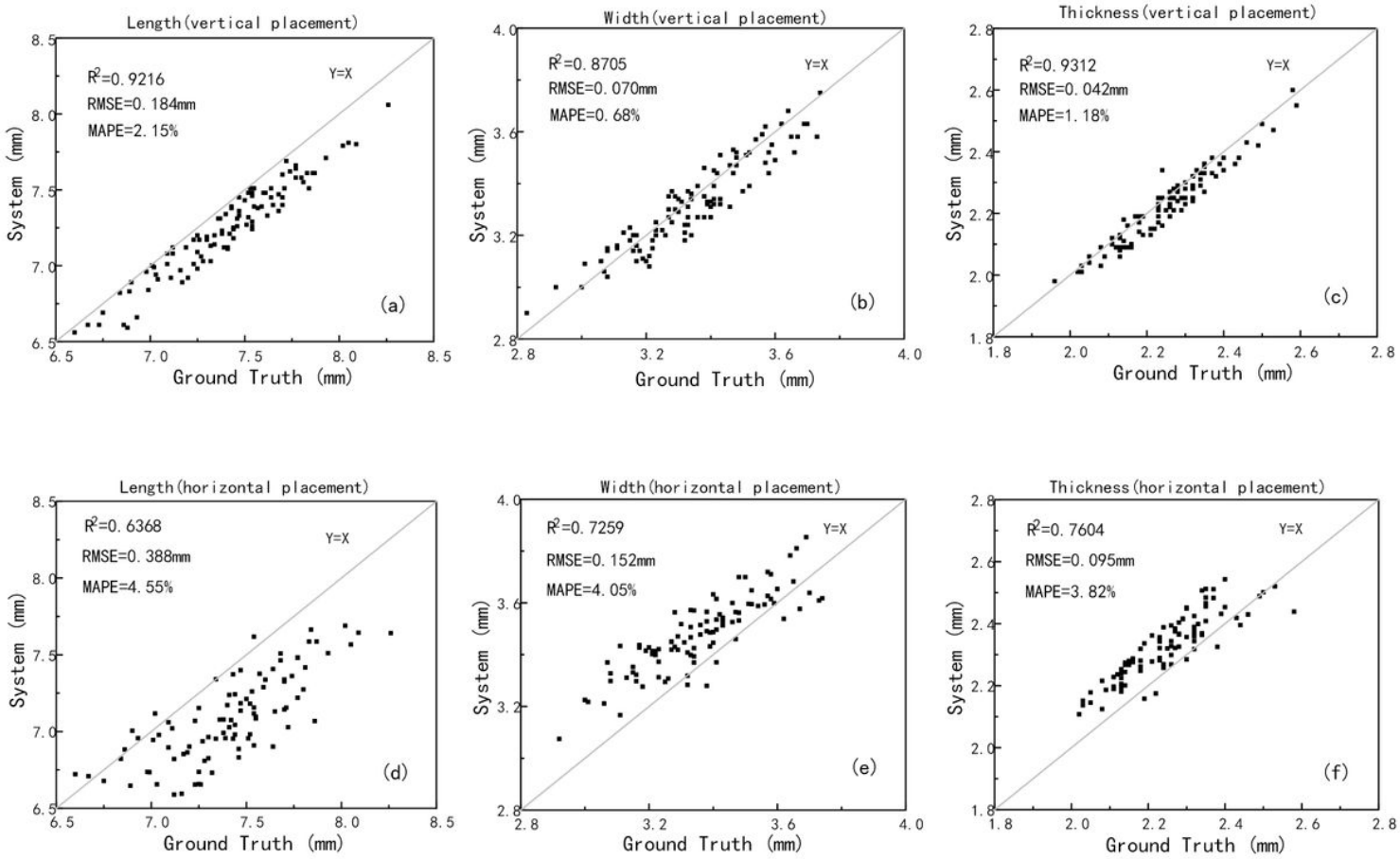

Figure 11

The measuring results in different placement scheme. (a) Length (vertical) (b) Width (vertical) (c) Thickness (vertical) (d) Length (horizontal) (e) Width (horizontal) (f) Thickness (horizontal)

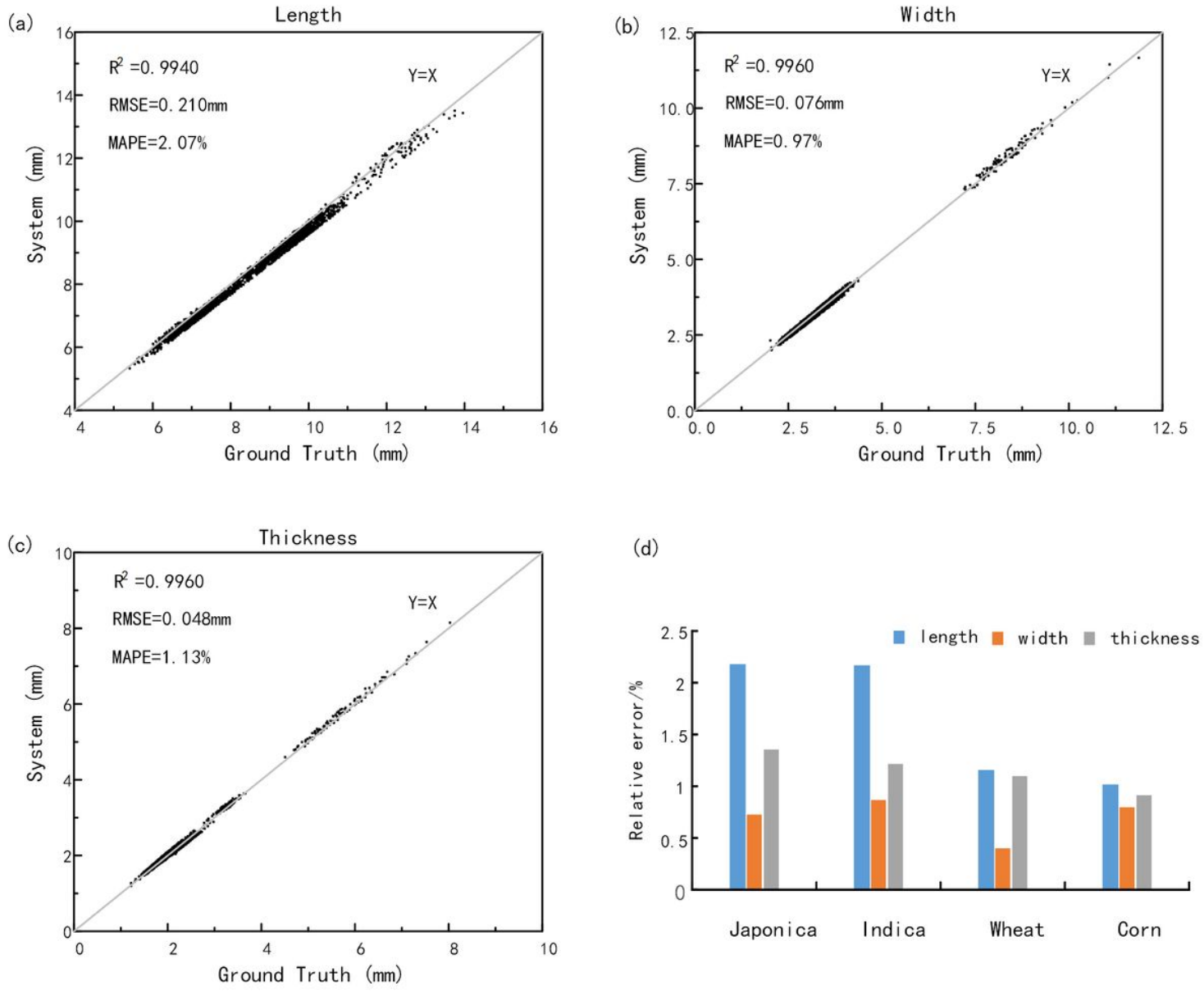

Figure 12

The sample accuracy analysis. (a) Length (b) Width (c) Thickness (d) japonica, indica, wheat and corn grains mean relative error.

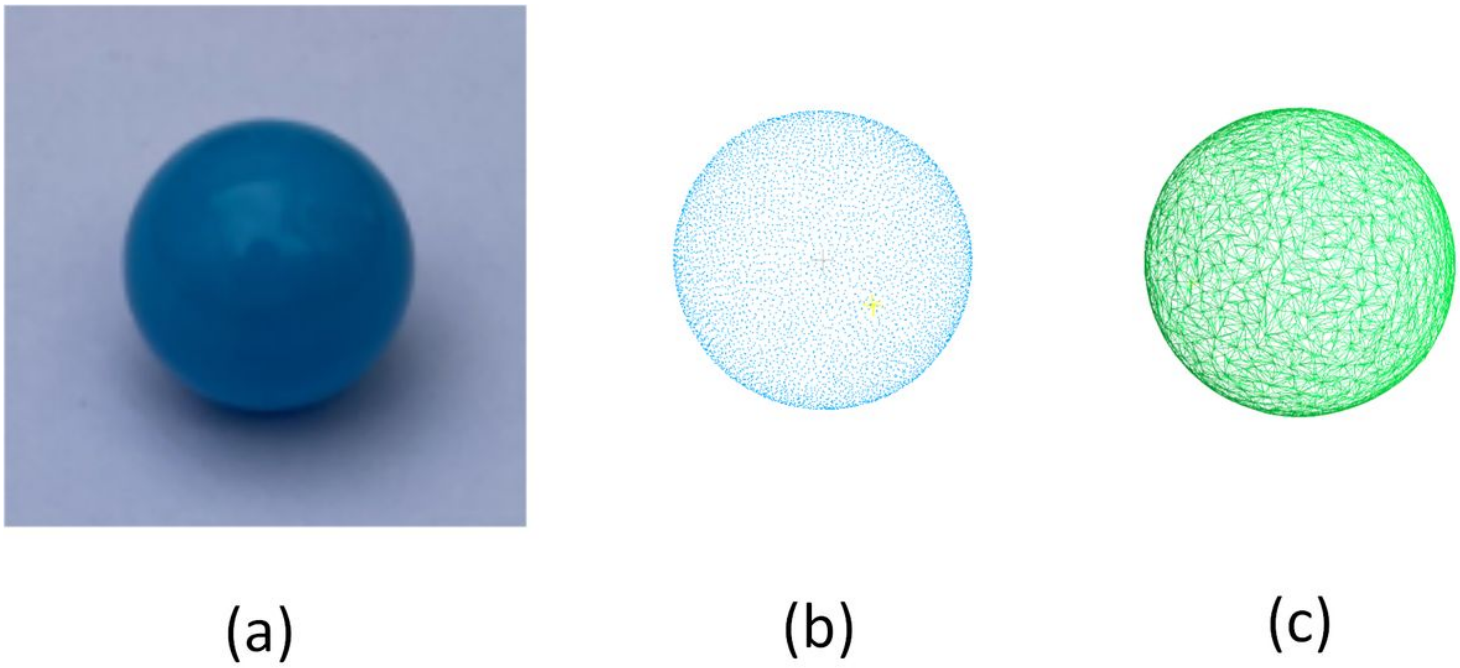

Figure 13

The sample sphere. (a) Real object (b) Point cloud (c) Triangular mesh

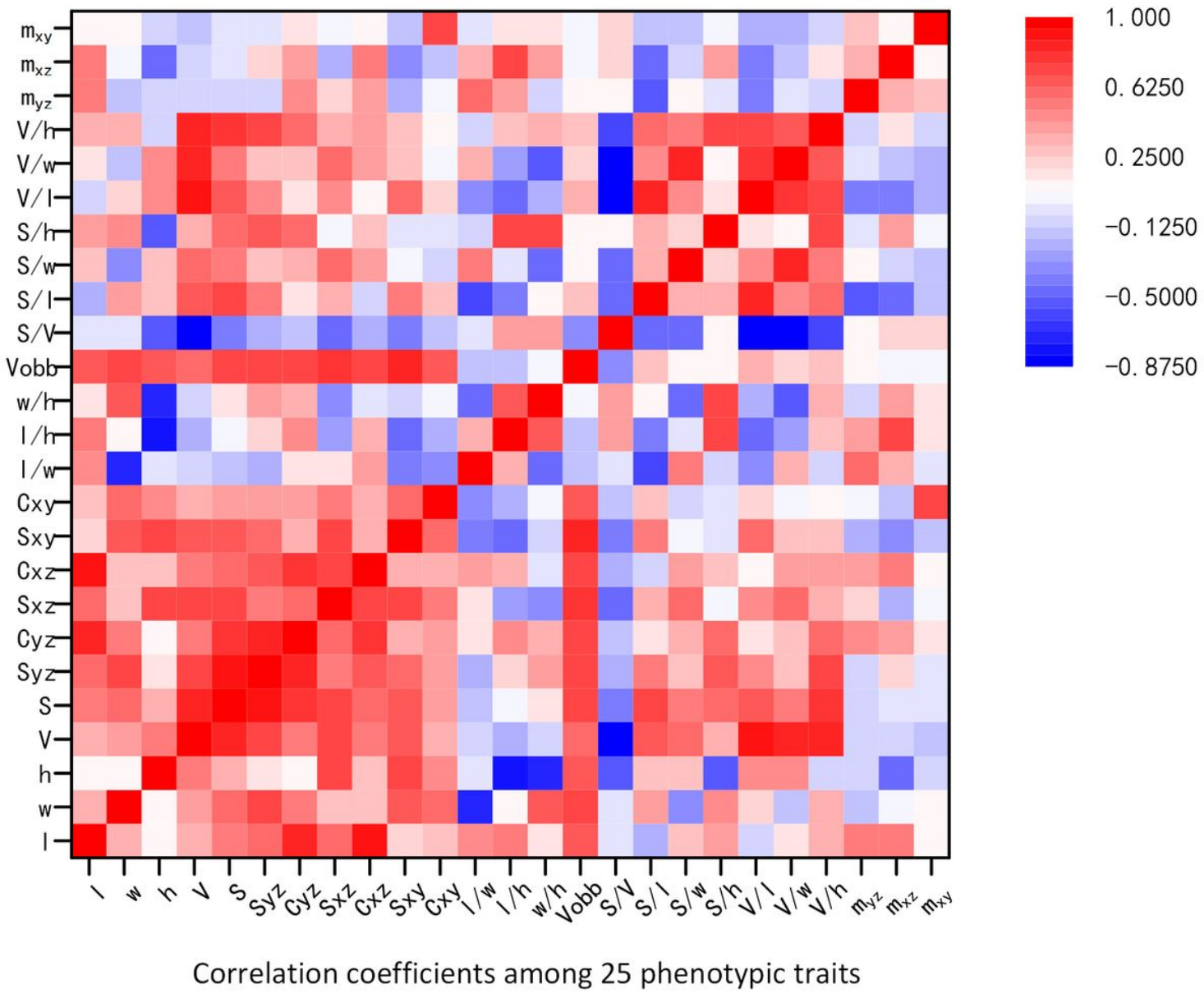

Figure 14

The result of grain traits correlation analysis.

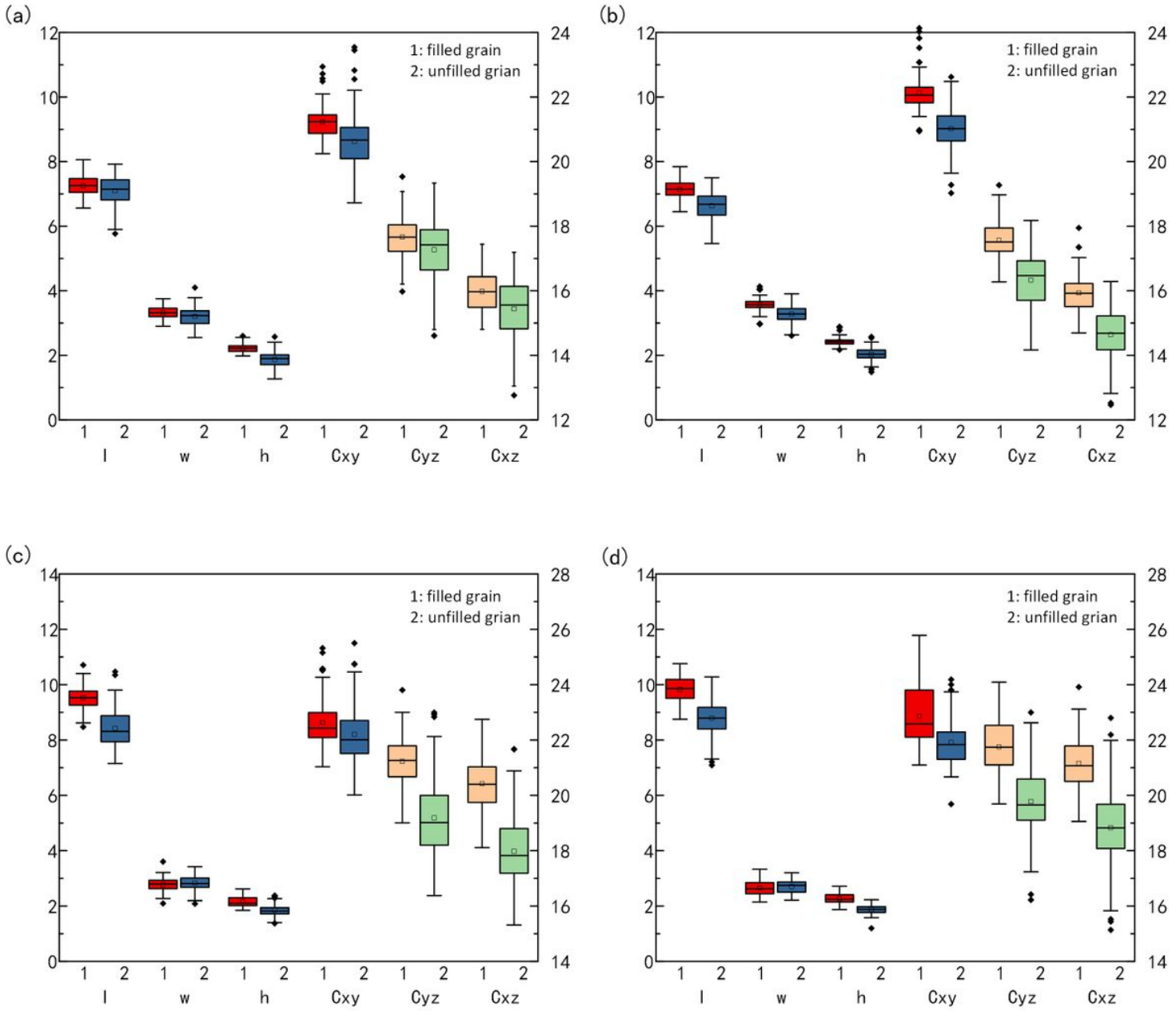

Figure 15

The Comparison of main traits between filled grain and unfilled grain. (a) Zhonghua 11 (b) Wuyunjing 3 (c) C Liangyou Huazhan (d) Zhuliangyou 211

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryAppendixS1.docx
- SupplementaryAppendixS2.docx
- SupplementaryInformation.pdf

- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryVideoS1.wmv