

## Three-dimensional photogrammetric mapping of cotton bolls *in situ* based on point cloud segmentation and clustering

Shangpeng Sun<sup>a</sup>, Changying Li<sup>a,\*</sup>, Peng W. Chee<sup>b</sup>, Andrew H. Paterson<sup>b,c</sup>, Yu Jiang<sup>a</sup>, Rui Xu<sup>a</sup>, Jon S. Robertson<sup>c</sup>, Jeevan Adhikari<sup>c</sup>, Tariq Shehzad<sup>c</sup>

<sup>a</sup> School of Electrical and Computer Engineering, College of Engineering, University of Georgia, Athens, GA, United States

<sup>b</sup> Department of Crop and Soil Sciences, College of Agricultural and Environmental Sciences, University of Georgia, Athens, GA, United States

<sup>c</sup> Department of Genetics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States



### ARTICLE INFO

**Keywords:**

Clustering  
Field-based high throughput phenotyping  
LiDAR  
Point cloud  
Segmentation  
Spatial distribution

### ABSTRACT

Three-dimensional high throughput plant phenotyping techniques provide an opportunity to measure plant organ-level traits which can be highly useful to plant breeders. The number and locations of cotton bolls, which are the fruit of cotton plants and an important component of fiber yield, are arguably among the most important phenotypic traits but are complex to quantify manually. Hence, there is a need for effective and efficient cotton boll phenotyping solutions to support breeding research and monitor the crop yield leading to better production management systems. We developed a novel methodology for 3D cotton boll mapping within a plot *in situ*. Point clouds were reconstructed from multi-view images using the structure from motion algorithm. The method used a region-based classification algorithm that successfully accounted for noise due to sunlight. The developed density-based clustering method could estimate boll counts for this situation, in which bolls were in direct contact with other bolls. By applying the method to point clouds from 30 plots of cotton plants, boll counts, boll volume and position data were derived. The average accuracy of boll counting was up to 90% and the  $R^2$  values between fiber yield and boll number, as well as fiber yield and boll volume were 0.87 and 0.66, respectively. The 3D boll spatial distribution could also be analyzed using this method. This method, which was low-cost and provided improved site-specific data on cotton bolls, can also be applied to other plant/fruit mapping analysis after some modification.

### 1. Introduction

Crop demand is expected to increase because of expanding global population (Cordell et al., 2009; Tester and Langridge, 2010; Gerland et al., 2014). New plant breeding approaches that improve rates of genetic gain can help to meet this demand. Cotton (*Gossypium* spp.) is among the most economically important fiber crops and the United States (US) is one of the top counties producing cotton. Around 79% of the world's natural fiber is produced from cotton (Townsend and Sette, 2016). In 2017, an estimated 12 million acres of cotton were grown in the US, resulting in a value of around 7 billion for the US cotton industry (USDA, 2018). Cotton bolls, as a yield-related component, can be used as an indicator for fiber prediction from which physiological and genetic mechanisms of plant growth and development can be derived. Therefore, high throughput phenotyping for cotton bolls plays a vital role in the selection of new cultivars and improvement of management practices. The traditional manual methods of quantification of cotton

bolls is time consuming and labor intensive. Thus, new methods that can improve the throughput capability of cotton boll phenotyping are needed.

Two-dimensional (2D) imaging technologies have been used widely for phenotyping of cotton and other plants during the last several decades (Li et al., 2014). For example, Li et al. (2017) developed a neural network-based data processing pipeline to detect the presence of cotton bolls in color images captured using a digital color camera mounted on a bracket. However, this method could neither count the number of bolls nor estimate the size of the bolls. Image-based methods were used to detect many other plant/crop phenotypic traits such as grape bunch detection (Pérez-Zavala et al., 2018), plant leaf segmentation and tracking (Yin et al., 2018), and mango fruit detection (Stein et al., 2016). All these 2D imaging methods have the same limitations. First, it is difficult to address the occlusion problem, which is very common for plant phenotyping, especially under field conditions, because of the lack of depth information; additionally, it is difficult to

\* Corresponding author at: 712F Boyd Graduate Studies, 200 D. W. Brooks Drive, University of Georgia, Athens, GA 30602, United States.  
E-mail address: [cyl@uga.edu](mailto:cyl@uga.edu) (C. Li).

determine object structure information.

3D remote sensing and computer vision technologies provide opportunities to overcome the limitations of 2D imaging-based methods, and have thus received significant attention for phenotyping by plant scientists during the past decade (Gongal et al., 2015; Vazquez-Arellano et al., 2016; Gibbs et al., 2017). In addition to providing depth information critical for addressing the occlusion problem and obtaining structure information, a wider variety of phenotypic traits can be extracted from 3D data, such as crop biomass estimation (L.A. Wang et al., 2016; Wallace et al., 2017), plant architectural trait analysis (Bao et al., 2019; Conn et al., 2017; Guo et al., 2018), digital elevation model generation (Toda et al., 2017), and dynamic crop growth monitoring (Dong et al., 2016; Duan et al., 2016; Pearse et al., 2016). Specifically, several studies were conducted for cotton plant morphological trait measurements under field conditions (Jiang et al., 2017; Sun et al., 2018b; Xu et al., 2019).

Methodologies for reconstructing 3D plant models fall into two categories: active and passive. Light Detection and Ranging (LiDAR) is one of the most widely used active sensors for field-based phenotyping (Lin, 2015). Sun et al. (2018b) reconstructed 3D canopy surface models for cotton plants by combining a 2D line scan LiDAR with an RTK-GPS which were mounted on a ground vehicle-based mobile platform. Only canopy level morphological traits were measured because the resolution of the point clouds was low so that plant organs such as leaves and branches could not be observed. With the development of sensing technologies, high-resolution 3D LiDAR sensors have been used to obtain dense point clouds and have provided capabilities for the detection of phenotypic traits at the plant organ level, such as sorghum panicles (Malambo et al., 2019), maize main stem and branches (Jin et al., 2018) and plant architecture pattern (Conn et al., 2017; Paulus et al., 2014). However, these high-resolution sensors are typically very expensive. Kinect-v2 is another commonly used active sensor which can capture RGB-D data. It is much less expensive than the high-resolution LiDAR, but the resolution is low, making it challenging for plant phenotyping of organ traits, especially under field conditions (Jiang et al., 2017).

Image-based 3D reconstruction, which belongs to the passive category, provides potential for balancing sensor cost and data resolution. This method captures images of interested objects from different perspectives using digital cameras and then calculates depth information of the targets based on a triangulation principle. Stereo vision is a popular passive method for obtaining 3D data and processes two images captured with slightly different horizontal perspectives (Si et al., 2015). The two images also can be taken in sequence by a single camera at two positions if the scene is static. Particularly, more than two images can be taken by moving the camera or using multi-view cameras, a strategy referred to as structure-from-motion (SfM). Overall, these methods are much less expensive than the active sensor-based methods for 3D reconstruction because they only require standard digital cameras to capture overlapping images of an object. Meanwhile, they can reconstruct dense 3D point clouds because newer cameras can capture high-resolution images with a very high resolution. Dense point clouds were successfully reconstructed using multi-view images under controlled illumination conditions (Nguyen et al., 2015; Duan et al., 2016). Unmanned aerial vehicle (UAV) platforms to which digital camera were attached were used to take images for 3D reconstruction under field conditions (Huang et al., 2016; Xu et al., 2019). Although the resolution of point clouds was too low to detect plant organ traits, particularly traits located at low parts of plants. Fernandez et al. (2017) applied stereo cameras mounted on a ground vehicle-based platform to reconstruct sorghum 3D models under field conditions from which plant architectural traits like stem diameter and plant height were accurately extracted. This study demonstrated that image-based methods could be used to reconstruct plant dense point clouds under field conditions.

Accurate individual cotton boll detection from point clouds is still challenging, especially when plants are planted closely together in the

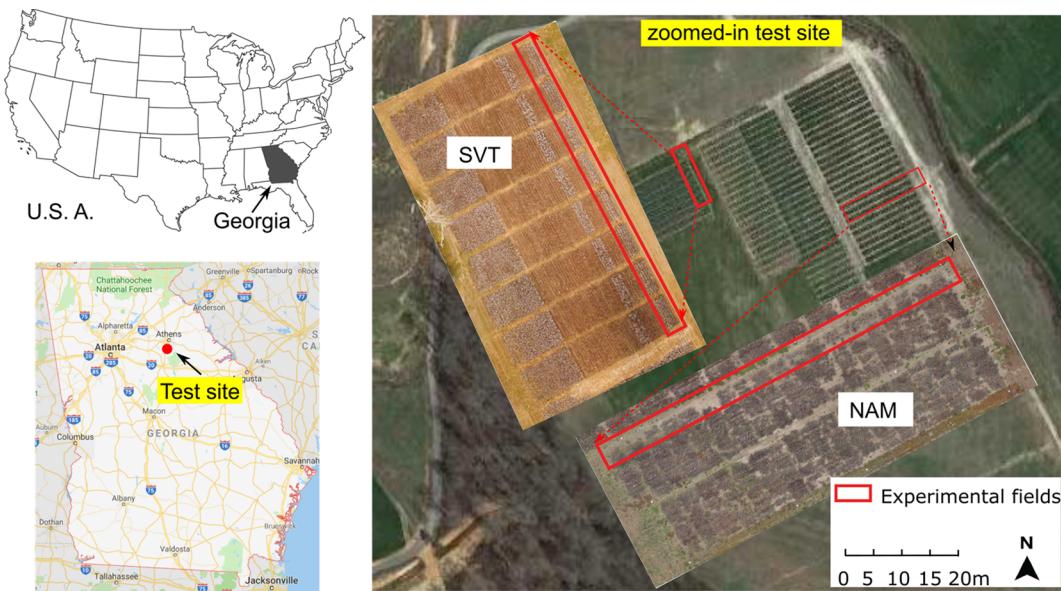
fields. Based on the review of plant phenotyping literature, detection and characterization of single cotton bolls using 3D technologies remains unexplored. Although there have been pioneering studies for other plant organ trait detection, the developed algorithms were designed according to specific phenotypic traits. For example, Velumani et al. (2017) introduced a wheat ear counting method using point clouds. Wheat ears were first segmented from the background using voxel-based segmentation and mean shift segmentation, respectively, based on density and position features, and then an Ostu's threshold method was applied to count wheat ears with an average detection rate of 85%. Malambo et al. (2019) presented a method for individual sorghum panicle detection based on panicle density and color features. While an average accurate rate of ~89% was achieved, the performance of the point cloud filtering operation was not robust because it relied solely upon color features. In addition, as is common under field conditions, objects in direct contact with each other were not accounted for. In this study, cotton boll detection encountered specific challenges. For example, for cotton bolls located around the entire plant and not just at the top of the plant as in wheat ears and sorghum panicles, position features could not be used for segmentation. Furthermore, color can vary because of sunlight effects, which means that non-color related features, such as shape and point spatial distribution, should be explored to reduce sunlight influences. However, these features cannot be extracted from single points, but region-based methods are often used to overcome the limitation. The basic idea is to convert the original point clouds into over-segmented regions. Each region is a group of points with similarly predefined properties. Therefore, color, shape, and other features can be used to describe a region. A color-based region generation method has also been applied for apple detection (Tao and Zhou, 2017). Papon et al. (2013) developed a region generation method called voxel cloud connectivity segmentation that was used by Wolf et al. (2015) to achieve accurate indoor scene point cloud segmentation. Cotton bolls have varied and complex shapes—closed bolls can vary substantially in size, and open bolls can have a variety of shapes based on emerging cotton fibers, resulting in serious overlapping issue compared to other fruits with a more rigid structure, such as apples. The detection of individual bolls from a multi-boll cluster in which several single bolls are physically connected continues to pose challenges.

The goal of this study was to develop a data processing pipeline for detecting and counting individual cotton bolls from 3D point clouds reconstructed from multi-view images using SfM under field conditions. The specific objectives of this study were to (1) implement a mobile 3D image acquisition system in the field; (2) develop a region-based semantic point cloud segmentation algorithm for foreground boll voxel recognition; (3) develop a density-based clustering algorithm for individual cotton boll detection and counting; (4) validate the performance of the proposed method as compared to the ground truth.

## 2. Materials and methods

### 2.1. Experimental fields and image collection

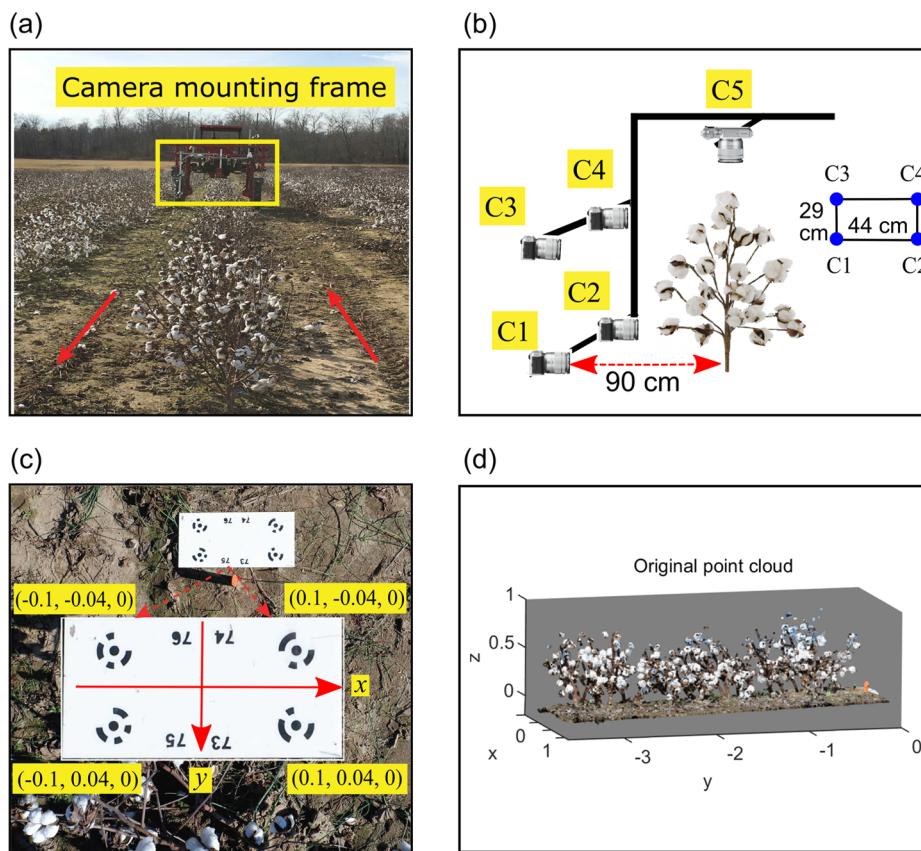
The plant materials used in this study were planted in two experimental fields—a University of Georgia statewide variety testing field (SVT) and a nested association mapping (NAM) field—located at Iron Horse Farm in Greene County, GA, USA (Fig. 1). The NAM field was designed for multiple purposes including phenotyping technique development and proper genotype selection and validation. More than 200 genotypes were in the field, which were randomly chosen members of a population derived from crossing a wild strain of *G. hirsutum* race yucatanenese with either of two elite cultivars, DES 56 or Acala Maxxa. While the SVT field had one genotype and was originally intended to determine whether the genotype was adapted to Georgia growing conditions. Each plot in the fields was a 3-meter row. Fertilization and irrigation practices followed commercial recommendations for the southeastern cotton production region.



**Fig. 1.** Experimental fields. Experiments were conducted in two fields. One was called SVT field and the other was NAM field.

Five Fuji X-A10 cameras (Fujifilm Holdings Corporation, NY, USA) mounted on a tractor platform were used to scan the plants. One camera was set to photograph the plants from the top and the other four were mounted on one side of the plot to scan plants from a side view (Fig. 2a). The mounting height of the top camera was approximately 150 cm, and the distance between the four side view cameras and the plant's main stem was approximately 90 cm (Fig. 2b). A rectangle consisting of camera C1, C2, C3 and C4 had a short edge of 29 cm and a long edge of 44 cm (Fig. 2b). To ensure that high quality images could be obtained for further 3D reconstruction operation, the cameras

needed to be configured according to the guidelines based on field experiments that require the cameras to be set to shutter-priority mode. The specific shutter speed was set according to illumination conditions, but had to be faster than 1/200 s in order to avoid tractor vibration influences. The aperture had to be set accordingly in order to make images have textural features (not too dark or bright). ISO was set to 200 to reduce image noise. The focal length was 16 mm, which was the minimum value for the lens coming with the camera body and allowed the cameras to achieve the largest field of view for the current layout. The sensor size of the camera was  $25.6 \times 15.6$  mm, and the image



**Fig. 2.** Multi-view image acquisition system and plant 3D point cloud reconstruction. (a) Multi-view cameras mounted on a tractor platform. (b) Camera layout information. (c) A scale bar used to build a local 3D Cartesian coordinate system and calibrate the size of reconstructed 3D point cloud; (d) Example of a reconstructed point cloud for a plot with around 4 million points.

resolution was  $4896 \times 3264$ . Therefore, the top camera had a field of view of  $221 \times 146$  cm at the ground plane, resulting in around  $0.2 \text{ mm}^2/\text{pixel}$ , and the four side view cameras had  $133 \times 88$  cm at a distance of 90 cm away, resulting in  $0.07 \text{ mm}^2/\text{pixel}$ . When scanning the plants, the tractor speed was around 0.2–0.3 m/s and a trigger device was developed to enable all five cameras to capture images at the same time with a frequency of 1.5 frames per second, resulting in approximately 10–15 s for one-side scanning (20–30 s for full scanning of one plot) per plot at a length of  $\sim 3$  m. The camera layout and configuration ensured that enough overlap (greater than 60%) between adjacent images, which is optimal for 3D reconstruction. The sensing system continuously took close images of cotton plants from different perspectives. Therefore, the occlusion problem was greatly reduced while detailed information of plant organs was captured simultaneously. The tractor had to go through each row twice in order to scan both sides of the plants, so there were approximately 150–225 images collected for each plot. The first image collection was conducted on Nov 24, 2017 in the SVT field, and a total of 15 plots were scanned. The second collection was on Dec 14, 2017 in NAM field with another 15 plots being scanned. Both days were sunny, but with different illumination levels. Accordingly, the shutter speed was 1/250 s and 1/400 s, respectively, and the aperture was 7.1.

A custom-made scale bar was placed in front of each plot during image collection (Fig. 2c). The rectangular scale bar consisted of four markers (marker 73, 74, 75 and 76 in Fig. 2c). The horizontal edge was 0.2 m, and the vertical edge was 0.08 m. A local 3D Cartesian coordinate system could be built when coordinates of the four markers (73, 74, 75 and 76) were assigned, such as  $(0.1, 0.04, 0)$ ,  $(0.1, -0.04, 0)$ ,  $(-0.1, 0.04, 0)$ , and  $(-0.1, -0.04, 0)$ , respectively. The 3D reconstruction was conducted using the SfM algorithm, which was implemented via Agisoft PhotoScan Professional (version 1.2.5) software (Geoscan. Inc., Moscow, Russia). The upper limit of the feature points and the matching points for each image was set to be 60,000 and 60,000, respectively, for the image alignment operation. For building dense point clouds, a mild depth filter was applied to reduce the likelihood of incorrectly eliminating important features as outliers. The four markers could be automatically detected by the software, and the prebuilt 3D coordinate system could be used as hyper parameters for the software to calibrate the actual size of the reconstructed 3D point clouds.

Hand-harvesting was conducted just after each image collection. The boll numbers for each plot were counted, and the yield was weighed. The summary of the hand harvesting is presented in Table 1.

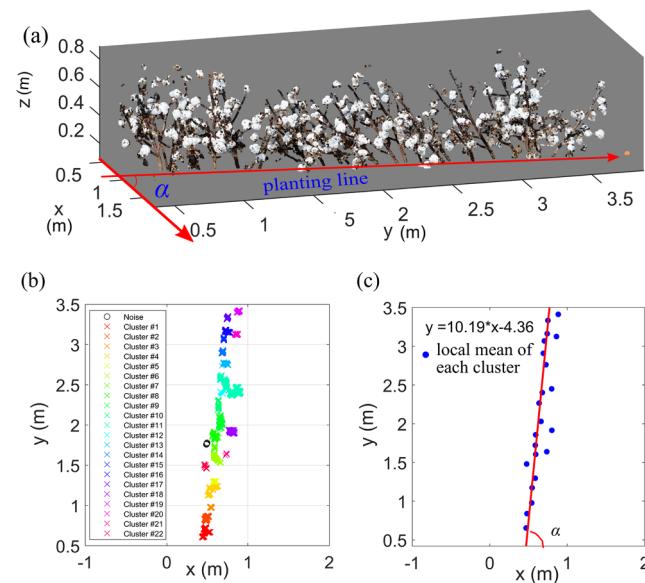
## 2.2. Cotton boll detection and characterization

The 3D Cotton boll detection and characterization algorithm developed in this study worked on plot-level point clouds and involved three main steps. The first phase consisted of point cloud pre-processing. Next, point cloud segmentation was performed to exclude branch points from point clouds. This operation relied on characteristic differences such as color and shape features observed between branches and bolls. Finally, individual bolls were detected based on the point density and boll size features, deriving boll number as well as volume, and position traits.

**Table 1**

Ground truth data of boll number and fiber yield. A total of 15 plots for each dataset.

Field	Boll number			Yield (g)		
	Min	Max	Mean	Min	Max	Mean
SVT	85	349	218	330	1474	845.5
NAM	37	377	206	113	1633	747.3



**Fig. 3.** Rotation of a 3D point cloud. (a) An angle  $\alpha$  between the x axis and the planting line. (b) Clustering results of the 3 cm thick slice points. (c) A line (in red) was fit using RANSAC to calculate the angle  $\alpha$  for rotation operation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 2.2.1. Point cloud preprocessing

Point cloud preprocessing included ground plane removal and point cloud rotation. The ground plane was removed by setting a threshold at  $z_h = 10$  cm along the z axis direction because it was not useful for the subsequent data processing. Although every precaution was taken to place a scale bar, the y axis of the local coordinate created by the scale bar (Fig. 2c) may not be aligned adequately with the planting line (root base) (Fig. 3a). Therefore, the planting line needed to be estimated in order to conduct point cloud rotation. To do this, a 3 cm thick slice was cut starting from the bottom of the point cloud and a 2D density-based spatial clustering of applications with noise (DBSCAN) (Wierzchoń and Kłopotek, 2018) was then applied to cluster the sliced points with only x and y coordinates (Fig. 3b); the two parameters—*minPts* which specifies how many neighbors a point should have to be included into a cluster, and *Eps* which specifies how close points should be to each other for a cluster, were set to be 5 and 2 cm for the 2D DBSCAN operation, respectively. The local mean x, y coordinates were used to represent the location of each cluster (Fig. 3c), and a line was fit by random sample consensus (RANSAC). The fit line was considered to be the planting line. Finally, the point cloud was rotated  $(90 - \alpha)$  around the z axis by Eq. (1) to ensure that the planting line was at  $x = 0$ .

$$R_z(90 - \alpha) = \begin{pmatrix} \cos(90 - \alpha) & \sin(90 - \alpha) & 0 & 0 \\ -\sin(90 - \alpha) & \cos(90 - \alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

### 2.2.2. Point cloud segmentation

**2.2.2.1. Over-segmentation of point clouds.** Regions, which can result from over-segmentation methods, provide better spatial support from which to compute richer feature representations as compared to single points (Rusu, 2010; Tao and Zhou, 2017). The purpose of over-segmentation methods is to merge the points that are close enough in terms of predefined attributes. The output is a set of clusters, each one is a set of points. In this study, a voxel cloud connectivity segmentation (VCCS) algorithm (Papon et al., 2013), which is a region growing method incrementally expanding supervoxels from a set of seed voxels, is applied to produce over-segmented regions in order to perform boll

points segmentation from the branch points. The general process of VCCS is as follows: the original point cloud is voxelized with a resolution of  $R_{\text{voxel}}$ . Each voxel can be represented in a 39-dimensional space (Eq. (2)).

$$F = \{x, y, z, L, a, b, \text{FPFH}_{1,2,\dots,33}\} \quad (2)$$

where  $F$  is the feature vector for each voxel,  $x, y$  and  $z$  are values of a local Cartesian coordinate system,  $L$ ,  $a$  and  $b$  are intensities in the CIELab color space, which expresses color as three values:  $L$  for lightness from black to white,  $a$  from green to red, and  $b$  from blue to yellow.  $\text{FPFH}_{1,\dots,33}$  are the 33 elements of a fast point feature histogram (FPFH), which is feature vector describing the voxel geometrical characteristics (Rusu et al., 2009). Then the voxel cloud space is divided into a voxelized grid with a resolution of  $R_{\text{seed}}$ .  $R_{\text{seed}}$  is significantly higher than  $R_{\text{voxel}}$ .  $R_{\text{seed}} = 0.08$  m and  $R_{\text{voxel}} = 0.002$  m according to the size of cotton bolls in this study. For each grid of voxels, a seed voxel is initialized by selecting the voxel nearest to the center of the grid. Seed voxels are then moved to the connected voxel which has the smallest gradient within  $R_{\text{seed}}$  of the seed center. Supervoxels are grown by assigning voxels to seed voxels iteratively according to a distance,  $D$ , calculated in  $F$  using Eq. (3).

$$D = \sqrt{\omega_c D_c^2 + \frac{\omega_s D_s^2}{3R_{\text{seed}}^2} + \omega_n D_n^2} \quad (3)$$

where  $D_c$  is the Euclidean distance in CIELab space,  $D_s$  is the spatial distance, which is normalized by the seeding resolution, and where  $D_n$  is an angle distance between the surface normal vectors in the FPFH space. Parameters  $\omega_c$ ,  $\omega_s$  and  $\omega_n$  are color, spatial, and normal weights, respectively. The three weights were set equally ( $\omega_c = \omega_s = \omega_n = 1$ ) in this study. For a cluster center, the nearest voxel is selected, and starting from which adjacent voxels selected through flowing outwards, the distances between the voxels and the center are calculated. If the distance is the smallest this voxel has been seen, its label is set. Once all voxels are assigned, the center of each supervoxel is updated by taking the mean of all its constituents. This is repeated until either the cluster centers stabilize or for a fixed number of iterations. VCCS takes advantage of 3D geometry so that the over-segmentations conform to object boundaries better than existing methods while maintaining processing efficiency. A typical single boll is spherical, with a diameter ranging from 5 to 10 cm (i.e., volume from 133 to 523 cm<sup>3</sup>). So,  $R_{\text{seed}}$  and  $R_{\text{voxel}}$  were selected as 8 cm and 2 mm, respectively, and the color, spatial, and normal weights were set to be equal importance, i.e.,  $\omega_c = 1$ ,  $\omega_s = 1$  and  $\omega_n = 1$ .

Color-based region growing segmentation (CRGS) is another commonly used region growth method and can produce over-segmented regions from point clouds. CRGS generates regions based on RGB color differences and Euclidean distances between the seed point and its neighboring points (Tao and Zhou, 2017). The general process of CRGS is as follows: a point with the minimum curvature value is picked from the original point cloud and then is added to the set called seeds. The color differences in RGB space are calculated between the seed point and its neighboring points. If the difference is less than a predefined threshold,  $t_c$ , the neighboring point is added to the current region. After every neighboring point is tested, the spatial differences between seed point and the neighboring points are calculated. If the difference is less than a threshold  $t_s$ , then the point is added to the seeds and the current seed point is removed. This process is repeated until the seeds set is empty, resulting in a region. The procedure iterates until there are no labeled points in the original points. The parameters  $t_c$  were selected as a small value,  $t_c = 10$ , in order to get small scale over-segmented regions, and  $t_s$  was set to be 8 cm, which was the same to the parameter  $R_{\text{seed}}$  in VCCS. We compared the two methods for cotton boll counting.

**2.2.2.2. Feature extraction.** A 97-dimensional feature vector including color features, spatial features, and shape features was extracted from

each over-segmentation to describe the point cloud data denoting  $R = [p_1, p_2, \dots, p_n]$  as an over-segmented region, where each point  $p_i$  contains Cartesian coordinates  $(x_i, y_i, z_i)$  and point intensity in an RGB color space  $(R_i, G_i, B_i)$ . We observed significantly noticeable differences between foreground boll points and the points in the background in terms of color. Therefore, color information is a good descriptor for the region features. In this study, a 64-dimensional vector was extracted from RGB and HSV color spaces, which was a histogram distributions of sixteen bins in channels R, G, B and S (Li et al., 2016; Sun et al., 2018a). Although the color of the points could be described in an RGB color space, the points in the HSV color space showed relatively effective light-interference resistant capabilities (Tao and Zhou, 2017).

Geometric features, including spatial features (Zhou and Tuzel, 2017) and shape features (Munoz, 2013; Verdoja et al., 2017), were required because some branches had almost the same color as the bolls because of the effects of sunlight during field image collection. For spatial features, the local mean was computed as the centroid of a region, denoted as  $(v_x, v_y, v_z)$ ; then, the distance between each point,  $p_i$ , and the centroid was computed and denoted as  $S = (x_i - v_x, y_i - v_y, z_i - v_z)$ ,  $i = 1, 2, \dots, n$ . For each axis, a histogram distribution with 10 bins was computed, resulting in a 30-dimension feature vector for x, y, and z coordinates. Shape features were extracted from the covariance matrix of a region. Let  $K_{XX}$  be the covariance matrix of a region X with  $N_p$  points (Eqs. (4) and (5)), and  $\lambda_0, \lambda_1, \lambda_2$  ( $\lambda_0 > \lambda_1 > \lambda_2$ ) be the eigenvalues of  $K_{XX}$ .

$$K_{XX} = \frac{1}{N_p} \sum_{q \in N_p} (q - \mu_p)(q - \mu_p)^T \quad (4)$$

$$\mu_p = \frac{1}{N_p} \sum_{q \in N_p} q \quad (5)$$

The three-dimensional eigen value vector  $\{\lambda_0, \lambda_1, \lambda_2\}$  can be used as the region shape feature descriptor. If  $\lambda_0 \gg \lambda_1, \lambda_2$ , the region has a linear structure, such as branches of cotton plants; if  $\lambda_0 \approx \lambda_1 \approx \lambda_2$ , the region has a spherical structure, such as the cotton bolls in this study. Because a boll could be split into several regions, and each region may not obtain rigid spherical shape. For these boll regions, the eigen values have the property of  $\lambda_0 \approx \lambda_1 \gg \lambda_2$ .

**2.2.2.3. SVM-based classifier.** A support vector machine (SVM), which effectively generalizes pattern recognition and provides robust calculation capabilities (Vapnik, 2013), was used to train a classification model to discriminate between boll points in the front and the background points. A total of 651 over-segmentations with two categories—bolls and branches—were labeled manually based on a custom-developed tool using MATLAB 2018b (The Math Works Inc., Natick, MA, USA). There were 328 boll samples and 323 branch samples, respectively. The dataset was divided into training and validation subsets with a ratio of 0.7:0.3. During model training, a radial basis function (RBF) kernel was used and punishment parameter  $C$  and kernel parameter  $\gamma$  were tuned based on a grid-search and 5-fold cross validation methods. After SVM classification, a color filter was used to remove black exocarps of cotton bolls.

### 2.2.3. Point cloud clustering

After obtaining boll points from the point cloud segmentation operation, individual cotton bolls were detected by a clustering-based algorithm relying on point density and boll size features. For two single bolls, there was typically a gap between them if they were not in direct contact with each other, resulting in different point density distributions.

The 3D DBSCAN clustering algorithm was applied first to group boll points into clusters. The parameter  $Eps$  of DBSCAN was set to be 2 cm because most bolls were larger; the other parameter,  $MinPts$ , could be selected from a wide range of 5–15, which did not affect the result

significantly.  $\text{MinPts} = 10$  was used in this study. The output clusters were divided into three categories according to cluster volume: noisy clusters, single boll clusters, and multi-boll clusters. A cluster was considered to be noisy if its volume was less than  $1 \text{ cm}^3$ , so it was removed by a size filter. For single boll clusters, each cluster was considered as a single boll, while multi-boll clusters included multiple bolls

---

```

Input: pt    # boll points
Output: N    # the total boll number
1  Initialize N=0
2  C = DBSCAN(pt)
3  C ← denoise(C)
4  C = {c1, c2, ..., cn}  # C: clusters output from DBSCAN with denoise operation
5  V = {v1, v2, ..., vn}  # V: clusters volume calculated using convex hull
6  vm = Ω(V)    # Ω(•) is the function to calculate average volume of regular bolls
7  for i = 1 to size(V) do
8    q = ⌊vi/vm⌋
9    if q ≤ 1 then
10   | N ← N + 1
11   else
12   | N ← N + q
13   end
14 end

```

---

inside, which had a larger volume than a single boll. Once boll numbers in a multi-boll cluster were estimated, the whole boll counts could be obtained.

The boll numbers in a multi-boll cluster could be estimated through the quotient between the contact cluster volume and the average single boll volume,  $v_m$  (Miller et al., 2017). We assumed that most clusters consisted of single bolls with a typical volume (hereafter, regular bolls), though some were multi-boll clusters with a larger volume and some were single bolls with a smaller volume.

Let  $V$  be a set of volumes for clusters.

$$V = \{v_1, v_2, \dots, v_n\} \quad (6)$$

A square matrix  $M$  could be obtained by eq (7).

$$M_{ij} = \lfloor v_i/v_j \rfloor, \quad i, j = 1, 2, \dots, n \quad (7)$$

where  $\lfloor \cdot \rfloor$  was a round function; each row of  $M$  was the rounded volume ratio between  $v_i$  and  $v_j$ . The number of unit values (ratio = 1) in each row of  $M$  was counted and stored in a vector  $K$ .

$$K = \{k_1, k_2, \dots, k_n\} \quad (8)$$

The mode (the value in a set of data appearing most often) of the elements in  $K$  was calculated and is denoted by  $e$ . The index of elements whose value equals  $e$  in  $K$  was calculated using Eq. (9), creating a vector  $d = \{d_1, d_2, \dots, d_n\}$ . If  $d_i = 1$ , the  $i$ th cluster is a regular boll.

$$d_i = \begin{cases} 1 & k_i = e \\ 0 & k_i \neq e \end{cases} \quad (9)$$

The average volume  $v_m$  of the regular bolls was calculated by Eq. (10).

$$v_m = \frac{\sum_{i=1}^n v_i \times d_i}{\sum_{i=1}^n d_i} \quad (10)$$

Then, the regular boll index and the corresponding average volume  $v_m$  were updated using Eqs. (11) and (10), respectively.

$$d_i = \begin{cases} 1 & |v_i/v_m| = 1 \\ 0 & |v_i/v_m| \neq 1 \end{cases} \quad (11)$$

The updating steps were repeated until  $v_m$  did not change. For a multi-boll cluster, the volume was divided by  $v_m$  and the quotient was rounded to produce the boll counts. Then, a 3D k-means clustering method was applied to obtain the boundaries of individual bolls.

#### Algorithm 1 (Boll counting).

In addition to boll numbers, other cotton boll data, such as boll volume and boll position, could be calculated when the individual bolls were detected. The volume of a cotton boll was estimated using the convex hull method, and the local mean of the x, y, and z coordinates was used to represent the position of the boll.

### 2.3. Validation and statistical analysis

In order to validate the performance of the proposed boll counting algorithm, the mean absolute percentage error (MAPE) was computed between the algorithm measuring boll numbers  $B_i$  and the manual measurements  $m_i$  by Eq. (12).  $P$  was the number of plots.

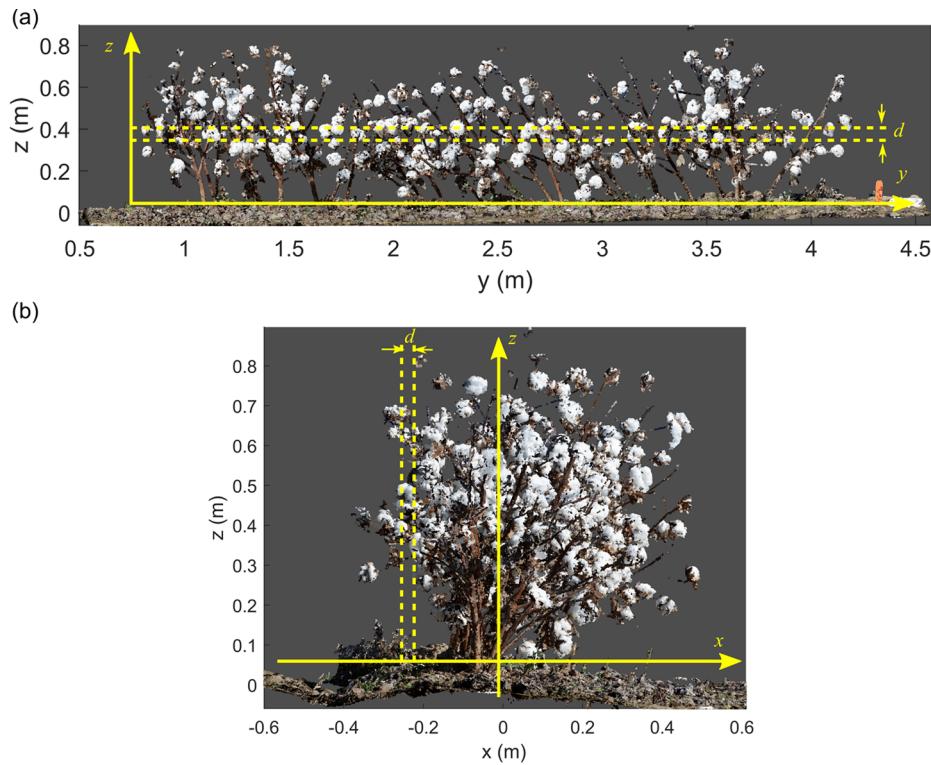
$$\text{MAPE} = \frac{\sum_{i=1}^P \frac{|B_i - m_i|}{m_i}}{P} \times 100\% \quad (12)$$

Additionally, in order to test the correlation between the algorithm measurements and the ground truth, the coefficient of determination ( $R^2$ ) and the root mean squared errors (RMSE) were calculated.

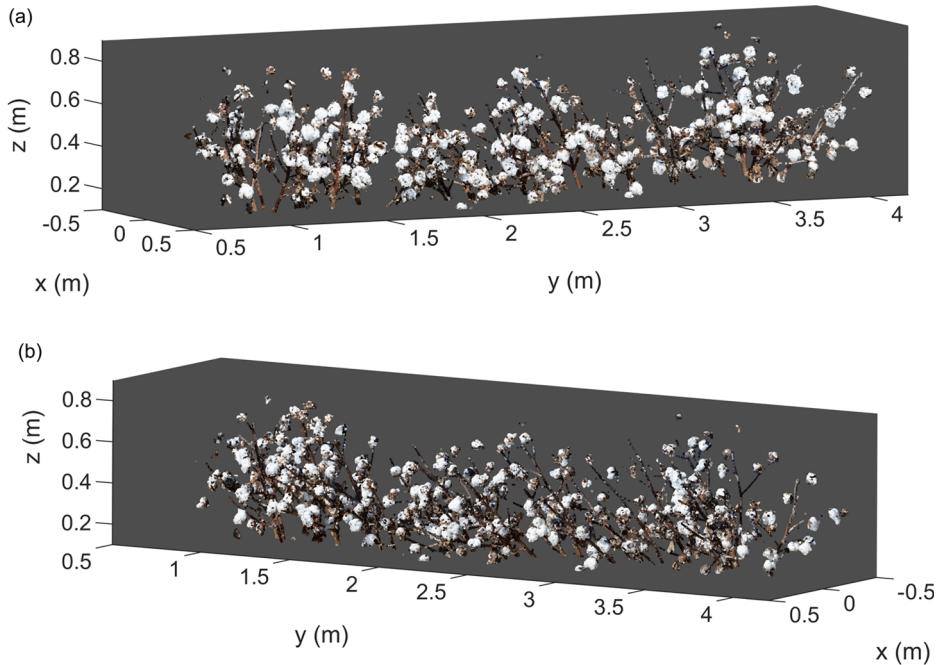
A linear regression analysis was conducted to explore the relationship between fiber yield and boll number as well as boll volume. Boll position additionally can enable boll spatial density distribution analysis (Fig. 4). For height (vertical direction) spatial analysis (Fig. 4a), we divided the space between the ground and the maximum height of the bolls into several slices at consistent intervals, whereby the number of bolls in each slice could be calculated and the distribution histogram could be obtained. For the x-y plane (horizontal direction) analysis (Fig. 4b), the space was divided into slices along the x axis at consistent intervals, and the bolls located at each slice were calculated.

## 3. Results

A plot with 312 cotton bolls (Genotype: P011.608.GA2012F2) was selected to demonstrate the 3D boll mapping results (Fig. 5). This plot was selected because the bolls were dense enough that they had a high possibility of being connected. After pose adjustment, the ground plane was located on the x-y plane with  $z = 0$ , and the planting line was located at the line  $x = 0$ . The ground plane was removed for the following processing.



**Fig. 4.** Schematic of 3D cotton boll spatial density distribution analysis. (a) z-axis (vertical direction) distribution. (b) x-y plane (horizontal direction) distribution.

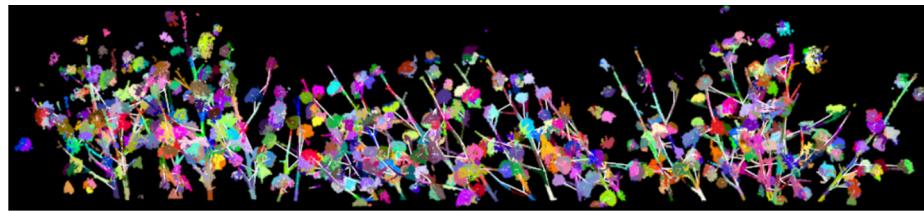


**Fig. 5.** Point cloud of a plot after preprocessing (pose adjustment and ground plane removal) from two different perspectives. (a) Point cloud with the viewing angle of azimuth and elevation [120, 10] degree. (b) Point cloud with the viewing angle of azimuth and elevation [60, 10] degree.

### 3.1. Boll point segmentation

Over-segmentations of the point cloud were generated by the VCCS algorithm (Fig. 6), where each over-segmentation (supervoxel) contained several individual points and was indicated using a color different from surrounding supervoxels. A boll might be divided into several supervoxels, as were branches. Similar results could be generated using CRGS algorithm.

The optimal parameters  $C$  and  $\gamma$  for VCCS-generated region dataset were 5.28 and 4.60, respectively. The trained SVM classifier achieved classification accuracies of 95.6% and 91.6% in training and validation datasets, respectively. Regarding the CRGS-generated region dataset, the optimal parameters  $C$  and  $\gamma$  were 1 and 2.30, and the training and validation accuracies were 94.7% and 91.4%, respectively. The results indicated that both classifiers achieved good performance for boll region and branch region classification. However, the VCCS algorithm



**Fig. 6.** Over-segmented regions using VCCS algorithm for a plot.

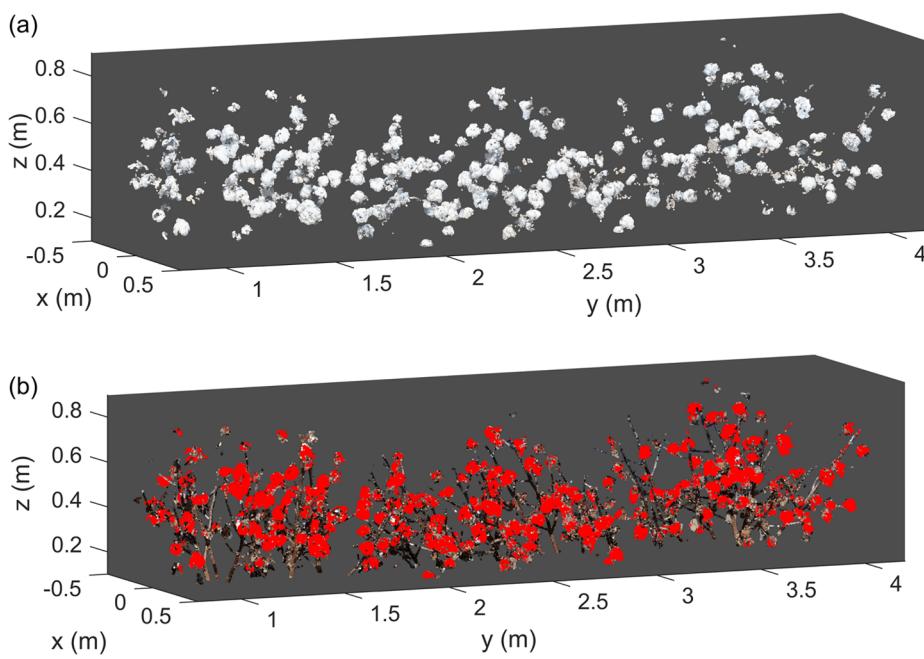
was much faster (~5 times) than the CRGS algorithm. This advantage was mainly attributed to two factors: the VCCS applied more efficient data structures kd-tree which is a binary search tree and very efficient for range and nearest neighbor search operation (Papon et al., 2013); and an operation merged small regions together for the CRGS method, which was time consuming. Fig. 7 demonstrated an example of obtained cotton boll points after classification operation using a VCCS-generated region dataset. Overall, it was observed that almost all of the boll points were segmented successfully, most of the branches with white color were correctly classified because of the contribution of the region spatial and shape features. However, we found that some branch regions and boll regions were misclassified. This did not cause significantly negative effects on boll detection because typically only a small part of a boll was misclassified. Detailed results would be presented in Section 3.2.

### 3.2. 3D boll mapping

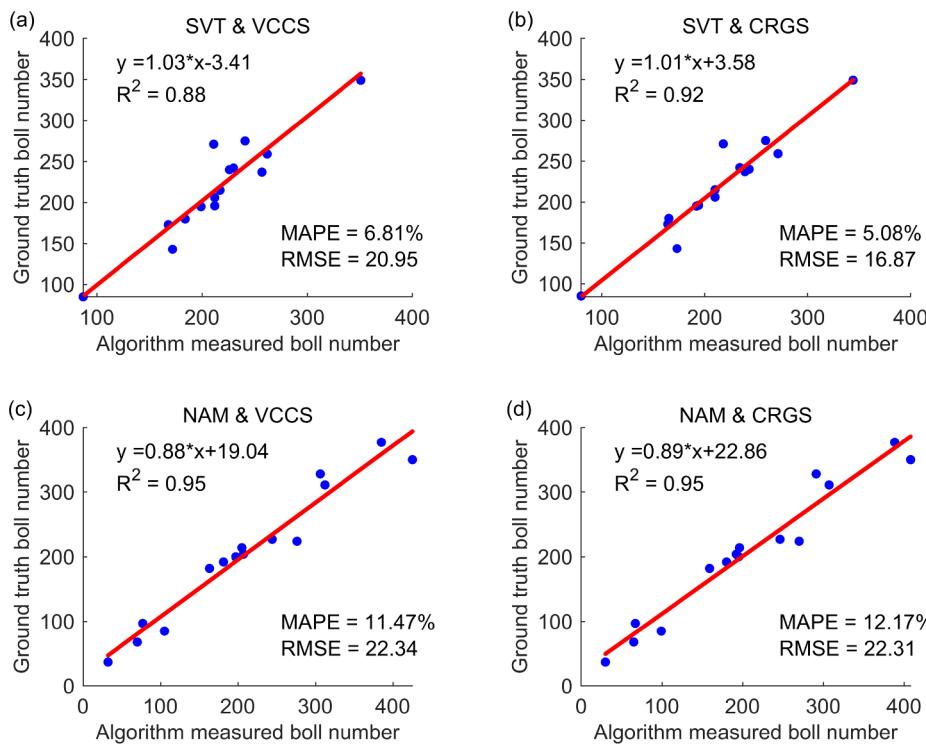
The boll counting results from the proposed method were precise and highly correlated with manual harvesting results (Fig. 8). An average accuracy of ~90% for boll counts for all 30 plots was achieved, with the best accuracy of ~95% (MAPE = 5.08%) and RMSE of 16.87. It was observed that for both boll segmentation methods, the MAPE and RMSE for data collected on November 22, 2017 (Fig. 8a and b) were better than those for data collected on Dec 14, 2017 (Fig. 8c and d), suggesting that sampling date affected boll counting performance. The most likely reason was that at the early stage of the harvesting season, cotton bolls retained relatively regular shapes (tight spheres) and have few occlusion issues. However, after prolonged exposure to wind and rain at a later stage in the harvesting season, the cotton fiber would not stay tightly attached to the burr and became strung out, resulting in

cotton bolls that were droopy and have complex shapes and a higher probability occlusion. However, the  $R^2$  values for the NAM field were better than the ones for the SVT field since the boll number for NAM field was distributed over a large range while it was restricted to a relatively small range for SVT. Each plot in the NAM field was a different genotype with diverse morphology and yield, while the plots in SVT field were commercial cultivars with more homogenous morphology and high yield. For point clouds obtained from each experimental field, similar boll counting performance was obtained using both VCCS- and CRGS-generated region datasets.

The error of cotton boll counting (over estimation and under estimation) was mainly attributed to the splitting operation of multi-boll clusters. As the reproductive organ, cotton bolls on a cotton plant reach physiological maturity and open at a wide range of times because flowering time usually begins about 60–70 days after planting and new flowers continue to develop throughout the 4–6 weeks of the reproductive season depending on the interaction between cultivar and environmental factors such as temperature and moisture conditions (Sharma et al., 2015). The boll size and shape varied significantly among plots grown even in the same field. If all bolls were opened while maintaining regular shapes for a plot, the multi-boll cluster splitting method would achieve good counting performance (Fig. 9a). However, for a plot in which some bolls were located in higher canopy, they were not opened fully and had small volumes, while other bolls typically located at lower canopy were opened fully and were droopy and had a much larger volume (Fig. 9b), the calculated average individual boll volume would be similar to the incompletely opened bolls. Therefore, the cluster splitting operation would generate a number much greater than the ground truth number for a multi-boll cluster. The highest over estimation plot in our dataset was demonstrated in Fig. 9b and the zoomed-in part showed that a multi-boll cluster was over divided. If all



**Fig. 7.** A detailed view of boll segmentation for a plot using a classifier trained by a support vector machine. (a) Segmented boll points without branch points. (b) Segmented boll points with branch points. Boll points were indicated using red color and other points using their original color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** Scatter plots of references versus estimated boll numbers. (a) SVT field data with VCCS algorithm. (b) SVT field data with CRGS algorithm. (c) NAM field data with VCCS algorithm. (d) NAM field data with CRGS algorithm.

cotton bolls were fully opened, they had large size, resulting in a higher possibility to create multi-boll clusters (Fig. 9c). Overall, the calculated average individual boll volume would be large, resulting in that less number (under estimation) would be obtained from the splitting operation, because some two-boll clusters had similar volume to the calculated average boll volume as demonstrated in the zoomed-in part in Fig. 9c. Another possible situation was that the identified individual bolls during the cluster splitting operation would be clusters consisting of two or more bolls, thereby leading to under estimation would happen. Fig. 9c demonstrated the plot with the highest under estimation.

### 3.3. Correlation with yield

Overall, boll number and boll volume presented a noticeable correlation with fiber yield for the SVT and NAM fields ( $R^2 > 0.87$  for boll number,  $R^2 > 0.66$  for boll volume) (Table 2, Figs. S1 and S2). Boll number gave better correlation than boll volume with fiber yield for both fields, resulting in better MAPE and RMSE. This was not surprising, as mentioned boll size can vary substantially and is affected by the length of time exposed to the weather. Cotton bolls developed from flowers that opened early in the growing season, often located in the lower canopy, tend to mature and open at an earlier time. These bolls have a higher likelihood to become droopy, which can result in increased size variation at maturity. Similar to the explanation for boll number counting performance analysis, the timing of data collection related to the maturity levels of cotton bolls affected yield correlations. Given that the data collection in the SVT field was performed in a more favorable time window (i.e., plants were less weathered) than in the NAM field, a better performance was obtained in the SVT field than in the NAM field for boll numbers and volume measurements.

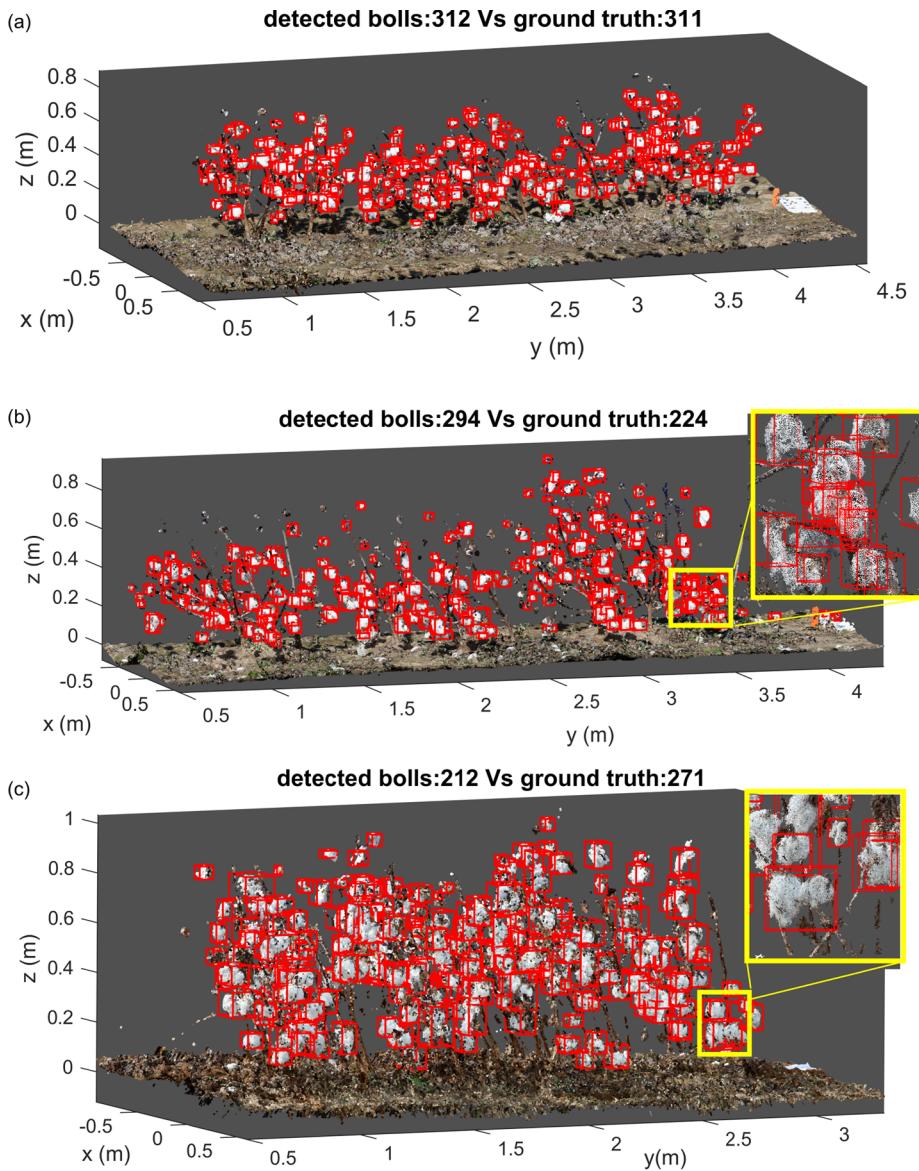
### 3.4. Spatial density distribution analysis

Boll distribution, a phenotypic trait highly useful for yield distribution analysis, could also be analyzed using such techniques. For a

typical cotton plant, the first position bolls between nodes 7 and 20 generally contributed the majority of fiber yield (Ritchie et al., 2007). As an example, the bolls detected from a plot in Fig. 9a represented using local mean coordinates in Fig. 10a were distributed within a bounding box with a size of 0.83 m, 3.35 m, 0.78 m along x, y, z direction, respectively. The bolls located at the longest distance from the planting line ( $x = 0$ ) were 0.37 m and 0.48 m for the left and right sides, respectively (Fig. 10b). The bolls were not distributed evenly on the two sides of the plots: around 70% of the bolls were located to the right side of the plot (Fig. 10c). This was mainly caused by hurricane Irma, which made landfall on September 11th, 2017 and resulted in all the plants in the plots becoming tilted. 91.4% of bolls were located within a distance of 0.3 m from the planting line. In the vertical direction, the highest and lowest bolls were located at heights of 0.78 m and 0.11 m, respectively, and the average height of all bolls' location was 0.42 m (Fig. 10d). Around 85% of the bolls were located within a height range of [0.2, 0.6] m (Fig. 10e).

## 4. Discussion

We demonstrated that high quality, color 3D point clouds can be reconstructed from multi-view images collected using a ground-based mobile platform under field conditions. This provides the potential to extract plant organ-level phenotypes from the point cloud, which has advantages over UAV-based 3D model platforms (Colomina and Molina, 2014; Pajares, 2015; Sankaran et al., 2015). A multi-view scanning strategy greatly reduces the that significantly influences plant phenotyping, especially under field conditions (Vazquez-Arellano et al., 2016). The tractor platform was very heavy, however, which might have caused soil disturbance and compaction. In addition, a person was needed to drive it for data collection. These limitations could be solved by replacing the platform with a motorized, narrow-wheeled and lightweight robot which can further promote easier transportation (Thompson et al., 2018). Another potential solution to further improve image collection efficiency is to use a UAV platform that attaches a tilted digital camera to take low altitude images from different perspectives.



**Fig. 9.** Representative results of 3D boll counting, each boll was indicated using a red bounding box. (a) The plot achieved the best boll counting result. Cotton bolls in this plot were fully opened while maintaining regular shapes. (b) The plot had the highest over estimation of boll numbers. Boll size had a wide variation in this plot, bolls at the high plant parts were not fully opened, resulting in small volume, while bolls at the bottom plant parts were fully opened, resulting in much larger volume. (c) The plot had the highest under estimation of boll numbers. In this plot all bolls were fully opened, resulting in a higher possibility to generate more multi-clusters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

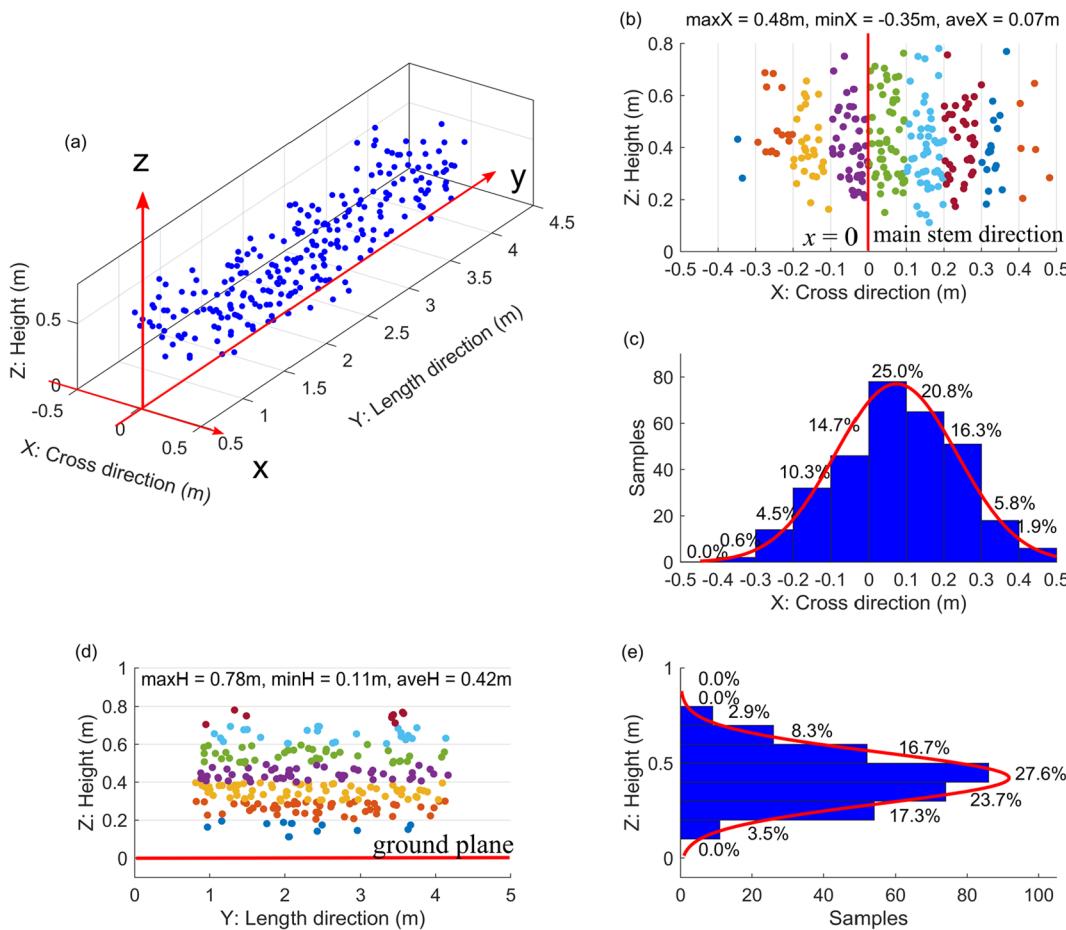
Illumination conditions play an important role in the quality of raw images, so noontime is preferred for field plant scanning with this platform. At midday, there is peak sunlight, which allows configuration of the camera with a high shutter speed that is helpful to reduce blurry images. Another advantage is that the shadow effect can be reduced since the sunlight is almost perpendicular to the ground plane. Wind is another factor influencing image quality, and breezy conditions can be detrimental to 3D point cloud reconstruction. One potential solution is to add an enclosure with artificial lights (Andrade-Sánchez et al., 2013; Barker et al., 2016). The enclosure itself could block winds, and the artificial light could improve consistency of light condition and reduce

shadows. Although general guidelines were proposed to take high quality images for 3D reconstruction, the camera configuration methods may vary for other plant applications because of different trait features like scale and surface. Human experience is very important for data collection, and no strict formulae can ensure high quality images. This is another limitation for the automated sensing system.

Current 3D point cloud reconstruction and data analyses were performed off-line. Around three to five hours were needed for a plot, depending on how many images used, based on a desktop equipped with 16 GB RAM memory and a 4 GB memory GPU card with the computational capability of 5.0. Data acquisition and processing could

**Table 2**  
Comparison of fiber yield correlation analysis with boll number and boll volume.

Attributes	Methods	SVT			NAM		
		R <sup>2</sup>	MAPE (%)	RMSE (g)	R <sup>2</sup>	MAPE (%)	RMSE (g)
Boll number	VCCS	0.91	7.99	77.09	0.90	19.29	143.84
	CRGS	0.87	8.15	93.57	0.90	17.75	141.01
Boll volume	VCCS	0.66	17.10	152.05	0.84	37.86	179.94
	CRGS	0.66	17.59	151.70	0.85	34.36	173.25



**Fig. 10.** Cotton boll 3D spatial density distribution analysis. (a) Localization of all detected bolls, each represented using a blue point. (b) Boll distribution on x-z plane. (c) Histogram of boll localization on x-z plane. (d) Boll distribution on y-z plane. (e) Histogram of boll localization on y-z plane. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

become online if 3D point clouds were acquired by using high resolution 3D LiDAR, although the cost of instrumentation would increase significantly (Kirchgessner et al., 2017; Q.H. Guo et al., 2018).

The boll counting accuracy of the method outperformed the studies of fruit counting using 2D image based methods (Malik et al., 2016; C.L. Wang et al., 2016; Bargoti and Underwood, 2017), because more information such as actual size and depth was included for 3D point clouds, and the occlusion effect was much reduced due to multi-view images. In addition, rich features such as spatial and shape features, which are not affected by varied illumination conditions, can be extracted from over-segmented regions. This greatly improved detection performance (Sun et al., 2018b; Malambo et al., 2019). Although these hand-designed features performed well in this study, performance could be further improved if more advanced or complex features from 3D point clouds can be extracted and more representative labeled samples are added into the training and testing datasets. In addition, the performance of the classification model could be improved by using more specific machine learning methods for training. For example, the algorithm high-order Markov random fields was reported to achieve better performance than standard SVM for the application of point cloud classification (Muñoz, 2013), because it can efficiently model the interaction of a group of points. VSSC is preferred for over-segmentation region generation considering the results of segmentation and algorithm computing efficiency in the future. 3D deep learning is another potential method to be explored in the future. The method can extract automatically point cloud features from low-level information, such as edges and corners, to high-level semantic information by using a hierarchical multi-stage structure that enables end-to-end ‘boll mapping’

(Qi et al., 2017; Zhou and Tuzel, 2017), which can overcome the limitation of hand-designed feature extraction operation. Compared to 2D image-based methods, in addition to improving boll counting performance, 3D model-based methods permit more phenotypic traits such as boll volume and boll spatial density distribution to be derived effectively and efficiently. Using such a method, it could be feasible to delineate cotton boll locations on each branch if a cotton plant branch pattern is obtained (Jin et al., 2018). Identification of each individual cotton boll and its location would be highly useful to plant breeding and cotton-picking robot in the future.

This study provides a new method to explore and understand plant growth using 3D mapping *in situ*, which is a grand challenge in plant biology. On one hand, the extracted traits such as the boll distribution on a cotton plant would be helpful for manipulating plant architecture, which are increasingly being emphasized in crop yield improvement (Mathan et al., 2016). If gin turnout data were available, cotton yield could be more accurately estimated by associating with number of bolls and boll size obtained using the proposed method (Sharma et al., 2015). Therefore, it would not be necessary for breeders to harvest cotton in order to determine which genotypes will produce the highest yield. On the other hand, 3D mapping *in situ* provides opportunities to extract more features such as plant nodes and branch architecture from point clouds (Conn et al., 2017).

## 5. Conclusions

A multi-view, camera-based system was developed to acquire high quality colored 3D point clouds of cotton plants *in situ*. High accuracy

boll segmentation and counting was achieved based on 3D point cloud segmentation and clustering algorithms. The developed segmentation algorithm was robust for varied illumination conditions, and the individual cotton boll detection algorithm accurately estimated boll numbers from multi-boll clusters. As a result, reliable 3D boll mapping and counting was obtained. Overall, the proposed method not only automated cotton boll counting, but also extracted new phenotypic traits such as boll size and 3D boll spatial distribution. These traits can help speed up new cultivar selection and be used for yield prediction and robotic harvest. Future work will focus on improving the throughput of data acquisition and processing, as well as exploring 3D deep learning for boll detection under field conditions.

## Acknowledgements

This work was supported by the National Robotics Initiative grant (NIFA grant No. 2017-67021-25928), partially supported by NSF Growing Convergence Research (1934481), Cotton Incorporated (17-510GA), and Presidential Interdisciplinary Seed Grant at the University of Georgia, United States.

The authors gratefully thank Gary pierce, Lisa Rainville, Susan Auckland, Mengyun Zhang and Liujun Li for their assistance in cotton boll hand harvesting, and Kyle Hunady for his assistance in sample labeling.

## Author contributions

S.S., C.L. and P.W.C. conceived the idea and designed the experiments; A.H.P. and P.W.C. contributed to the preparation of materials and equipment; S.S., Y.J., R.X., J.S.R., J.A. and T.S. conducted the experiments; S.S. analyzed the data; S.S., C.L., A.H.P., and P.W.C. interpreted results and wrote the paper. All authors read and approved the final manuscript.

## Declaration of Competing Interest

The authors declared that there is no conflict of interest.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.isprsjprs.2019.12.011>.

## References

- Andrade-Sanchez, P., Gore, M.A., Heun, J.T., Thorp, K.R., Carmo-Silva, A.E., French, A.N., Salvucci, M.E., White, J.W., 2013. Development and evaluation of a field-based high-throughput phenotyping platform. *Funct. Plant Biol.* 41 (1), 68–79.
- Bao, Y., Tang, L., Srinivasan, S., Schnable, P.S.J.B.E., 2019. Field-based architectural traits characterisation of maize plant using time-of-flight 3D imaging. *Biosyst. Eng.* 178, 86–101.
- Bargoti, S., Underwood, J.P., 2017. Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Rob.* 34 (6), 1039–1060.
- Barker III, J., Zhang, N., Sharon, J., Steeves, R., Wang, X., Wei, Y., Poland, J., 2016. Development of a field-based high-throughput mobile phenotyping platform. *Comput. Electron. Agric.* 122, 74–85.
- Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: a review. *ISPRS J. Photogramm. Remote Sens.* 92, 79–97.
- Conn, A., Pedmale, U.V., Chory, J., Stevens, C.F., Navlakha, S., 2017. A statistical description of plant shoot architecture. *Curr. Biol.* 27 (14), 2078–2088 e1–e3.
- Cordell, D., Drangert, J.-O., White, S., 2009. The story of phosphorus: global food security and food for thought. *Global Environ. Change* 19 (2), 292–305.
- Dong, J., Burnham, J.G., Boots, B., Rains, G.C., Dellaert, F., 2016. 4D Crop Monitoring: Spatio-Temporal Reconstruction for Agriculture. arXiv preprint arXiv:1610.02482.
- Duan, T., Chapman, S.C., Holland, E., Rebetzke, G.J., Guo, Y., Zheng, B., 2016. Dynamic quantification of canopy structure to characterize early plant vigour in wheat genotypes. *J. Exp. Bot.* 67 (15), 4523–4534.
- Fernandez, M.G.S., Bao, Y., Tang, L., Schnable, P.S., 2017. A high-throughput, field-based phenotyping technology for tall biomass crops. *Plant Physiol.* 174 (4), 2008–2022.
- Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J., Lalic, N., 2014. World population stabilization unlikely this century. *Science* 346 (6206), 234–237.
- Gibbs, J.A., Pound, M., French, A.P., Wells, D.M., Murchie, E., Pridmore, T., 2017. Approaches to three-dimensional reconstruction of plant shoot topology and geometry. *Funct. Plant Biol.* 44 (1), 62–75.
- Gongal, A., Amatya, S., Karkee, M., Zhang, Q., Lewis, K., 2015. Sensors and systems for fruit detection and localization: a review. *Comput. Electron. Agric.* 116, 8–19.
- Guo, J., Xu, S., Yan, D., Cheng, Z., Jaeger, M., Zhang, X., 2018a. Realistic procedural plant modeling from multiple view images. *IEEE Trans. Visual Comput. Graph* 1–1.
- Guo, Q.H., Wu, F.F., Pang, S.X., Zhao, X.Q., Chen, L.H., Liu, J., Xue, B.L., Xu, G.C., Li, L., Jing, H.C., et al., 2018b. Crop 3D-a LiDAR based platform for 3D high-throughput crop phenotyping. *Sci. China-Life Sci.* 61 (3), 328–339.
- Huang, Y., Brand, H.J., Sui, R., Thomson, S.J., Furukawa, T., Ebelhar, M.W., 2016. Cotton yield estimation using very high-resolution digital images acquired with a low-cost small unmanned aerial vehicle. *Trans. ASABE* 59 (6), 1563–1574.
- Jiang, Y., Li, C., Paterson, A., Sun, S., Xu, R., Robertson, J., 2017. Quantitative analysis of cotton canopy size in field conditions using a consumer-grade RGB-D camera. *Front. Plant Sci.* 8, 2233.
- Jin, S., Su, Y., Wu, F., Pang, S., Gao, S., Hu, T., Liu, J., Guo, Q., 2018. Stem-leaf segmentation and phenotypic trait extraction of individual maize using terrestrial LiDAR data. *IEEE Trans. Geosci. Remote Sens.* 57 (3), 1336–1346.
- Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A., Walter, A., 2017. The ETH field phenotyping platform FIP: a cable-suspended multi-sensor system. *Funct. Plant Biol.* 44 (1), 154.
- Li, L., Zhang, Q., Huang, D.F., 2014. A review of imaging techniques for plant phenotyping. *Sensors* 14 (11), 20078–20111.
- Li, Y., Cao, Z., Lu, H., Xiao, Y., Zhu, Y., Cremers, A.B., 2016. In-field cotton detection via region-based semantic image segmentation. *Comput. Electron. Agric.* 127, 475–486.
- Li, Y., Cao, Z., Xiao, Y., Cremers, A.B., 2017. DeepCotton: in-field cotton segmentation using deep fully convolutional network. *J. Electron. Imaging* 26 (5), 053028.
- Lin, Y., 2015. LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? *Comput. Electron. Agric.* 119, 61–73.
- Malambo, L., Popescu, S., Horne, D., Pugh, N., Rooney, W., 2019. Automated detection and measurement of individual sorghum panicles using density-based clustering of terrestrial lidar data. *ISPRS J. Photogramm. Remote Sens.* 149, 1–13.
- Malik, Z., Ziauddin, S., Shahid, A.R., Safi, A., 2016. Detection and counting of on-tree citrus fruit for crop yield estimation. *Int. J. Adv. Comput. Sci. Appl.* 7 (5), 519–523.
- Mathan, J., Bhattacharya, J., Ranjan, A., 2016. Enhancing crop yield by optimizing plant developmental features. *Development* 143 (18), 3283–3294.
- Miller, N.D., Haase, N.J., Lee, J., Kaepller, S.M., de Leon, N., Spalding, E.P., 2017. A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. *Plant J.* 89 (1), 169–178.
- Munoz, D., 2013. Inference Machines Parsing Scenes via Iterated Predictions. PhD thesis. Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.
- Nguyen, T.T., Slaughter, D.C., Max, N., Maloof, J.N., Sinha, N., 2015. Structured light-based 3D reconstruction system for plants. *Sensors* 15 (8), 18587–18612.
- Pajares, G., 2015. Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogramm. Eng. Remote Sens.* 81 (4), 281–329.
- Papon, J., Abramov, A., Schoeler, M., Wörgötter, F., 2013. Voxel cloud connectivity segmentation-supervoxels for point clouds. In: 2013 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, Portland, Oregon, USA, pp. 2027–2034.
- Paulus, S., Schumann, H., Kuhlmann, H., Léon, J., 2014. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosyst. Eng.* 121, 1–11.
- Pearse, G.D., Watt, M.S., Morgenroth, J., 2016. Comparison of optical LAI measurements under diffuse and clear skies after correcting for scattered radiation. *Agric. For. Meteorol.* 221, 61–70.
- Pérez-Zavalá, R., Torres-Torriti, M., Cheein, F.A., Troni, G., 2018. A pattern recognition strategy for visual grape bunch detection in vineyards. *Comput. Electron. Agric.* 151, 136–149.
- Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in Neural Information Processing Systems. Long Beach, California, USA, pp. 5105–5114.
- Ritchie, G.L., Bednarz, C.W., Jost, P.H., Brown, S.M., 2007. Cotton growth and development. Bulletin 1252. Cooperative Extension Service and the University of Georgia College of Agricultural and Environmental Sciences, Athens, GA.
- Rusu, R.B., 2010. Semantic 3D object maps for everyday manipulation in human living environments. *KI-Künstliche Intelligenz* 24 (4), 345–348.
- Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast point feature histograms (FPFH) for 3D registration. In: IEEE International Conference on Robotics and Automation. IEEE, Kobe, Japan, pp. 3212–3217.
- Sankaran, S., Khot, L.R., Espinoza, C.Z., Jarolmasjed, S., Sathuvalli, V.R., Vandemark, G.J., Miklas, P.N., Carter, A.H., Pumphrey, M.O., Knowles, N.R., et al., 2015. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. *Eur. J. Agron.* 70, 112–123.
- Sharma, B., Mills, C.I., Snowden, C., Ritchie, G.L., 2015. Contribution of boll mass and boll number to irrigated cotton yield. *Agron. J.* 107 (5), 1845–1853.
- Si, Y., Liu, G., Feng, J., 2015. Location of apples in trees using stereoscopic vision. *Comput. Electron. Agric.* 112, 68–74.
- Stein, M., Bargoti, S., Underwood, J., 2016. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. *Sensors* 16 (11), 1915.
- Sun, S., Li, C., Paterson, A., Jiang, Y., Robertson, J., 2018a. 3D computer vision and machine learning based technique for high throughput cotton boll mapping under field conditions. In: 2018 ASABE Annual International Meeting. American Society of Agricultural and Biological Engineers, p. 1–10.
- Sun, S., Li, C., Paterson, A., Jiang, Y., Xu, R., Robertson, J., Snider, J., 2018b. In-field high throughput phenotyping and cotton plant growth analysis using LiDAR. *Front. Plant Sci.* 9, 1–17.
- Tao, Y., Zhou, J., 2017. Automatic apple recognition based on the fusion of color and 3D

- feature for robotic fruit picking. *Comput. Electron. Agric.* 142, 388–396.
- Tester, M., Langridge, P., 2010. Breeding technologies to increase crop production in a changing world. *Science* 327 (5967), 818–822.
- Thompson, A., Conrad, A., Conley, M., Shrock, H., Taft, B., Miksch, C., Mills, T., Dyer, J., 2018. Professor: a motorized field-based phenotyping cart. *HardwareX* 4, 1–10.
- Toda, L.L., Yokingco, J.C.E., Paringit, E.C., Lasco, R.D.J.A.g., 2017. A LiDAR-based flood modelling approach for mapping rice cultivation areas in Apalit, Pampanga. *Appl. Geogr.* 80, 34–47.
- Townsend, T., Sette, J., 2016. Natural fibres and the world economy. In: *Natural Fibres: Advances in Science and Technology towards Industrial Applications*. Springer, pp. 381–390.
- USDA, 2018. Crop Production Historical Track Records. National Agricultural Statistics Service. [https://www.nass.usda.gov/Publications/Todays\\_Reports/reports/cropr18.pdf](https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropr18.pdf).
- Vapnik, V., 2013. *The Nature of Statistical Learning Theory*. Springer Science & Business Media, New York, USA.
- Vazquez-Arellano, M., Grieppentrog, H.W., Reiser, D., Paraforos, D.S., 2016. 3-D imaging systems for agricultural applications-a review. *Sensors* 16 (5), 24.
- Velumani, K., Elberink, S.O., Yang, M.Y., Baret, F., 2017. Wheat ear detection in plots by segmenting mobile laser scanner data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inform. Sci. IV-2/W4*, 149–156.
- Verdoja, F., Thomas, D., Sugimoto, A., 2017. Fast 3D point cloud segmentation using supervoxels with geometry and color for 3D scene understanding. In: IEEE International Conference on Multimedia and Expo (ICME). IEEE, Hong Kong, pp. 1285–1290.
- Wallace, L., Hillman, S., Reinke, K., Hally, B., 2017. Non-destructive estimation of above-ground surface and near-surface biomass using 3D terrestrial remote sensing techniques. *Meth. Ecol. Evol.* 8, 1607–1616.
- Wang, C.L., Zou, X.J., Tang, Y.C., Luo, L.F., Feng, W.X., 2016a. Localisation of litchi in an unstructured environment using binocular stereo vision. *Biosyst. Eng.* 145, 39–51.
- Wang, L.A., Zhou, X.D., Zhu, X.K., Dong, Z.D., Guo, W.S., 2016b. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J.* 4 (3), 212–219.
- Wierzchon, S.T., Kłopotek, M., 2018. *Modern Algorithms of Cluster Analysis*. Springer, Switzerland.
- Wolf, D., Prankl, J., Vincze, M., 2015. Fast semantic segmentation of 3D point clouds using a dense CRF with learned parameters. In: 2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 4867–4873.
- Xu, R., Li, C.Y., Paterson, A.H., 2019. Multispectral imaging and unmanned aerial systems for cotton plant phenotyping. *PLoS ONE* 14 (2).
- Yin, X., Liu, X., Chen, J., Kramer, D.M., 2018. Joint multi-leaf segmentation, alignment, and tracking for fluorescence plant videos. *IEEE Trans. Patt. Anal. Mach. Intell.* 40 (6), 1411–1423.
- Zhou, Y., Tuzel, O., 2017. VoxelNet: end-to-end learning for point cloud based 3D object detection. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Salt Lake City, UT, USA, pp. 1–10.