

Article

A Crop Canopy Localization Method Based on Ultrasonic Ranging and Iterative Self-Organizing Data Analysis Technique Algorithm

Fang Li ^{1,2}, Xiaohu Bai ^{1,*} and Yongkui Li ^{1,*}

¹ College of Engineering, Shenyang Agricultural University, Shenyang 110866, China; happy-lifang@163.com

² School of Mining Engineering, Liaoning Shihua University, Fushun 113001, China

* Correspondence: baixiaohu@syau.edu.cn (X.B.); SAUYKLEE@163.com (Y.L.)

Received: 26 November 2019; Accepted: 2 February 2020; Published: 3 February 2020



Abstract: To protect crops from diseases and increase yields, chemical agents are applied by boom sprayers. To achieve the optimal effect, the boom and the crop canopy should be kept at an appropriate distance. So, it is crucial to be able to distinguish the crop canopy from other plant leaves. Based on ultrasonic ranging, this paper adopts the fuzzy iterative self-organizing data analysis technique algorithm to identify the canopy location. According to the structural characteristics of the crop canopy, based on fuzzy clustering, the algorithm can dynamically adjust the number and center of clusters so as to get the optimal results. Therefore, the distances from the sensor to the canopy or the ground can be accurately acquired, and the influence of lower leaves on the measurement results can be alleviated. Potted corn plants from the 3-leaf stage to the 6-leaf stage were tested on an experiment bench. The results showed that the calculated distances from the sensor to the canopy using this method had good correlation with the manually measured distances. The maximum error of calculated values appeared at the 3-leaf stage. With the growth of plants, the error of calculated values decreased. The increased sensor moving speeds led to increased error due to the reduced data points. From the 3-leaf stage to the 5-leaf stage, the distances from the sensor to the ground can also be obtained at the same time. The method proposed in this paper provides a practical resolution to localize the canopy for adjusting the height of sprayer boom.

Keywords: canopy localization; fuzzy clustering; self-organizing algorithm; ultrasonic sensor; sprayer boom height control

1. Introduction

Nowadays, boom sprayers have been commonly used for the application of pesticides, herbicides, and fertilizers in agriculture. It is known that an optimal boom height is essential for achieving an even spray distribution [1–3]. In order to achieve a reasonably uniform application of such materials, it is necessary that the boom be maintained at a constant distance from the target area (crop canopy or ground). The performance of the boom height control system is partially dependent on the accuracy of the plant height measurement [4]. To control the boom height, ultrasonic sensors are mounted to the sprayer boom for measuring the distance of the boom either from the canopy or the ground [5–8].

Ultrasonic sensors have been widely used in past decades for agricultural applications because of their low price, simple information processing, and mature technology. Ultrasonic sound is transmitted and then reflected by the measured object. The distance can be calculated based on the sound speed and the travel time between sound transmission and echo return. Ultrasonic sensors for the determination of plant heights have been used for weed detection. The results showed that weed presence was correctly predicted in more than 92% of the cases [9]. Thomas et al. [10] concluded that

the prediction accuracy of forage mass from ultrasonic height measurements is acceptable. Shrestha et al. [11] investigated ultrasonic sensing approach as a kind of technology for vehicle-based corn height estimation. They found a good correlation between the measured and estimated crop plant height using an ultrasonic sensor. Kataoka et al. [12] attached an ultrasonic sensor to the front head of a sprayer that was driven along the crop row to measure the vertical distance between the sensor and plant canopy. They concluded that the ultrasonic sensor performed better for soybean and corn height measurements than a laser beam sensor. Chang et al. [13] developed an ultrasonic plant height measurement using three ultrasonic sensors and incorporated it into a commercial mechanical harvester to measure wild blueberry plant height. Liu et al. [14] designed a precise application system for the sprayer that used an ultrasonic sensor to collect the distance from the boom to the canopy and the ground. The system mainly used the significant difference between the ground height and the crop canopy height to determine whether there is a crop under the nozzle, thereby achieving the purpose of precise spraying. Wei et al. [15] designed an online control system of spray boom height and balance, using ultrasonic sensors to measure the distance from the boom to the ground. In order to eliminate the influence of the leaf occlusion on the ground distance detection, a spray boom height signal filtering algorithm for removing the branches and leaves interference was adopted. Wang et al. [16] designed a boom height automatic adjusting system that used ultrasonic sensors to measure the distance between the boom and the ground to prevent the boom from colliding with the ground.

During the spraying process, the ultrasonic sensor mounted on the boom goes forward along with the sprayer to scan the plants. As a result, the data acquired by the sensor may include the distance from the sensor to different levels of leaves and the ground. An effective method is needed to separate these signals from the sensor. Otherwise, the distance value to the canopy extracted from the data would not be very accurate. However, there have not been any reports about how to avoid the influence of leaves beneath the canopy on the determination of the canopy location under the sensor. In this paper, a method was proposed to distinguish the canopy from the rest of leaves and the ground. In addition, experiments were conducted to evaluate the influence of the plant growth stage and moving speed of the sensor on the precision of the proposed method.

2. Materials and Methods

2.1. Experimental Apparatus

The experimental apparatus consists of a motion control system and a data acquisition system.

As shown in Figure 1, a horizontal guide rail was mounted on the top of the frame to allow a trolley to move along it. Crop plants were placed below the rail at an interval of 0.3 m. An ultrasonic sensor pointing straight down to the plants was fixed to the trolley. The distance from the ultrasonic sensor to the ground was 0.81 m. A fixed pulley was placed on one end of the rail. An electrical motor was used to pull the trolley through a hauling rope. One end of the hauling rope was fixed to the rope-winding wheel, which was installed on the motor shaft, while the other end was mounted to the trolley via the fixed pulley. As the trolley moved, the ultrasonic sensor on the trolley can scan the crop plants under it to measure the distance from the sensor to the plants. The motor is a speed adjustable one (80YT25DV22, Jingyan Automation Component Co. Ltd., Xiamen, China) that can provide a rotation speed of 900–1400 r/min with the support of a SF25E controller. During the experiment, the speed of the sensor was in the range of 0.5–6 km/h, which is comparable to the actual operation speed of the sprayer in the field.

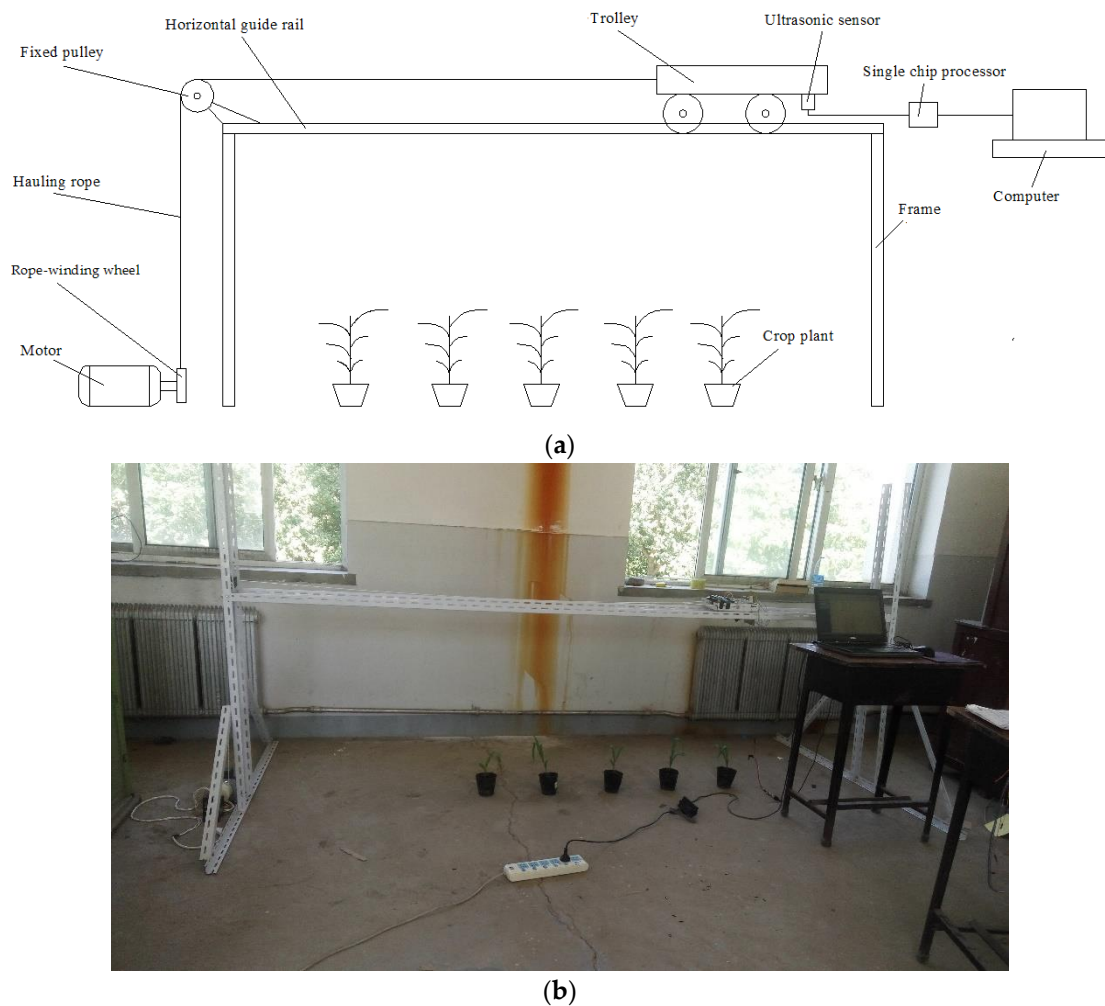


Figure 1. Experimental apparatus: (a) Schematic diagram; (b) Actual photo.

Our choice of the ultrasonic sensor KS109 (manufactured by Guide Electromechanical Technology Co. Ltd., Shenzhen, China) comes with noise reduction functionality. The transducer ultrasonic wave frequency is 40 kHz with the beam angle of 10° . The sensing range is 0.08–11 m with the allowing temperature of 0–60 °C. A temperature sensor (DS18B20) is embedded in the ultrasonic sensor to compensate for the effect of temperature fluctuation on the electronics. A single chip processor (STC12C5A60S2, Hongjing Technology Co. Ltd., Shenzhen, China) was used to control the transmission and receiving of the ultrasonic wave and calculate the travel time. Then, the travel time was sent to a computer to conclude the corresponding distance, which was stored for further analysis. The canopy height can be acquired by subtracting the distance between the sensor and the canopy from the distance between the sensor and the ground. For every plant, its canopy height was also measured with a measuring tape manually for comparisons. At the 3-leaf, 4-leaf, 5-leaf, and 6-leaf growth stages respectively, five corn plants were randomly chosen and tested.

2.2. Data Processing Method

Since the data acquired by the ultrasonic sensor include the distance to the crop canopy, the distance to the underlying leaves, the distance to the soil, and so on, the data in different classes are far from each other. Fuzzy clustering analysis can establish a fuzzy similarity to classify data based on the characteristics and similarity [17,18]. Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA) can dynamically modify the clustering center and get local optimal results according to the sample characteristics. Furthermore, it can dynamically adjust the number of clusters based on the

local data to obtain the best clustering results [19–21]. Therefore, this paper proposed an algorithm of fuzzy ISODATA that used ISODATA to identify parameters based on fuzzy clustering.

2.2.1. Basic Principle

First, the data samples are divided into several clusters. Then, the fuzzy matrix is modified repeatedly according to the objective function until the matrix has converged.

The set of samples to be clustered is $X = \{X_1, X_2, \dots, X_n\}$. If each sample has d characteristic indicators, i.e., $X_i = (x_{i1}, x_{i2}, \dots, x_{id})$, then the characteristic indicator matrix is

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix}. \quad (1)$$

Define c cluster center vectors as

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_c \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1d} \\ v_{21} & v_{22} & \cdots & v_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ v_{c1} & v_{c2} & \cdots & v_{cd} \end{bmatrix}. \quad (2)$$

Divide the sample X into c clusters ($2 \leq c \leq n$), then the fuzzy clustering matrix is

$$U_{c \times n} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_c \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{c1} & u_{c2} & \cdots & u_{cn} \end{bmatrix} \quad (3)$$

where u_{ij} is the membership degree of the j th sample to the i th cluster, which satisfies the following conditions: (1) $u_{ij} \in [0, 1], \forall i, j$; (2) $\sum_{i=1}^c u_{ij} = 1, \forall j$; (3) $0 < \sum_{j=1}^n u_{ij} < n, \forall i$.

In order to obtain the optimal clustering, the following objective function should be minimized.

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^q \|X_j - V_i\|^2 \quad (4)$$

where $\|X_j - V_i\|^2$ is the distance between the j th sample and the i th cluster center, and q is the weighted index. Bezdek has proved that $q = 2$ is the optimal by experiments [22].

Construct a Lagrange function:

$$L(U, V, \lambda) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^q \|X_j - V_i\|^2 + \sum_{i=1}^c \lambda_i \left(\sum_{j=1}^n u_{ij} - 1 \right) \quad (5)$$

where λ is Lagrange multiplier.

Taking the partial derivative of Equation (5) with respect to u_{ij} and V_i , and then setting it to zero, we obtain the expression of the membership degree and the cluster center when Equation (5) takes a minimum value.

The expression of membership degree is:

$$u_{ij} = \frac{1}{\sum_{p=1}^c \left[\frac{\|X_j - V_p\|^2}{\|X_j - V_i\|^2} \right]^{\frac{1}{q-1}}} \quad (6)$$

where X_j is the j th sample, V_i is the i th cluster center, and V_p is the p th cluster center.

The expression of the cluster center is

$$V_i = \frac{\sum_{j=1}^n (u_{ij})^q X_j}{\sum_{j=1}^n (u_{ij})^q}. \quad (7)$$

2.2.2. Split and Merge

The two main factors affecting the spray quality are the distances from the boom to the crop canopy and to the ground. In our analysis, these two distances were extracted from the two clusters whose centers are the closest and the farthest from the boom, respectively. In order to achieve more accurate distance measures, the members of the two clusters were further adjusted with merge and split operation based on ISODATA.

(1) Split

Calculate the average distance from each sample in a cluster to the cluster center \bar{D}_i using the following formula:

$$\bar{D}_i = \frac{1}{n_i} \sum_{X \in \omega_i} \|X - V_i\|, i = 1, 2, \dots, c \quad (8)$$

where n_i is the number of samples in the i th cluster, X is the sample in the i th cluster, V_i is the i th cluster center, and ω_i represents the i th cluster.

Calculate the average distance \bar{D} from every sample in sample set X to the cluster center in which the sample is located using the following formula:

$$\bar{D} = \frac{1}{n} \sum_{i=1}^c n_i \bar{D}_i \quad (9)$$

where n is the number of samples in sample set X .

Calculate the standard deviation σ_i of the i th cluster ω_i using the following formula:

$$\sigma_i = \sqrt{\frac{1}{n_i} \sum_{X \in \omega_i} \|X - V_i\|^2}. \quad (10)$$

If $\sigma_i > \theta_S$, $\bar{D}_i > \bar{D}$, and $n_i > 2(\theta_n + 1)$ where θ_S is a preset standard deviation value, θ_n is the minimum number of samples in a cluster that can be set according to the sprayer speed; then, ω_i can be split into two clusters of which the cluster center are the followings:

$$\begin{aligned} V_i^{(2)} &= V_i - 0.5\sigma_i \\ V_i^{(1)} &= V_i + 0.5\sigma_i. \end{aligned} \quad (11)$$

(2) Merge

Calculate the distance D_{ij} between two different cluster centers. If $D_{ij} < \theta_c$ where θ_c is the minimum distance between two cluster centers, then the two clusters are merged into a new cluster. Its center is as follows:

$$V_k = \frac{1}{n_i + n_j} (n_i V_i + n_j V_j) \quad (12)$$

where V_k is the center of the new cluster, n_i and n_j are the numbers of samples in the i th and j th cluster, and V_i and V_j are the centers of the i th and j th cluster.

2.2.3. Algorithm Procedures

(1) Set the cluster number c ($2 \leq c \leq n$), the initial fuzzy clustering matrix $U^{(0)}$, the number of iterations k ($k = 0$), and the threshold value ε ($\varepsilon > 0$).

(2) Calculate the cluster center of the k th iteration using Formula (7):

$$V^{(k)} = [V_1^{(k)}, V_2^{(k)}, \dots, V_c^{(k)}]^T. \quad (13)$$

(3) Modify the fuzzy clustering matrix $U^{(k)}$ of the k th iteration applying Formula (6).

(4) Compare $U^{(k)}$ with $U^{(k+1)}$. If $\max\{\|u_{ij}^{(k+1)} - u_{ij}^{(k)}\|\} \leq \varepsilon$, then stop the iteration, and the current clusters are optimal. Otherwise, set $k = k + 1$ and go back to Step (2).

(5) Split or merge the clusters according to Formulas (11) and (12).

(6) Make the clusters distinct. For $\forall X_j \in X$, if $\|X_j - V_k\| = \min_{1 \leq i \leq c} \|X_j - V_i\|$, then the sample X_j is classified into the k th cluster.

2.2.4. Clustering Effect Test

Evaluate the effect of the fuzzy ISODATA method using the following formula:

$$v_{PE}(U) = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{ij} \ln(u_{ij}) \quad (14)$$

where $v_{PE}(U)$ is the average fuzzy entropy of the fuzzy clustering matrix.

The smaller the value of $v_{PE}(U)$, the less the uncertainty in clustering and the better the clustering effect.

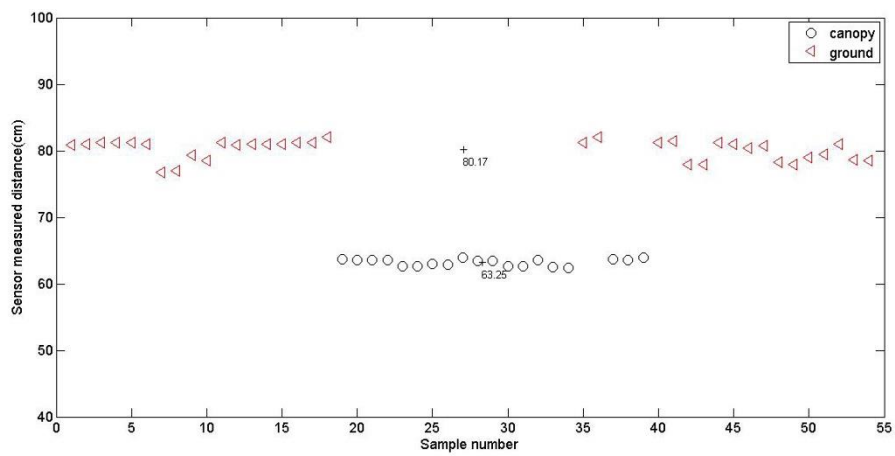
3. Results and Discussion

3.1. Algorithm Verification

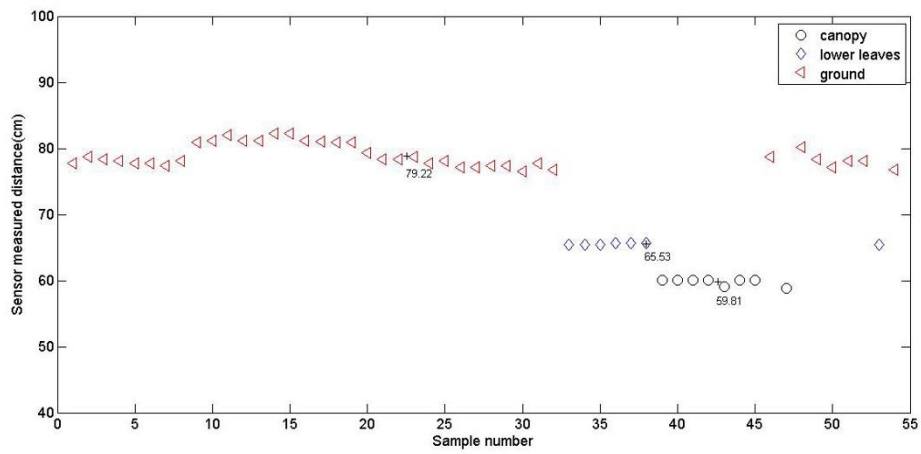
At the beginning, the data measured by an ultrasonic sensor was divided into 3 clusters that represented the crop canopy, the lower leaves, and the ground, respectively. Therefore, the initial number of clusters c is 3. To ensure the reliability of the clustering results, the threshold value ε is generally between 10^{-4} and 10^{-6} . So, in this paper, it was set to 0.0001 in order to reduce the number of iterations. As for the minimum number of samples θ_n , the standard deviation θ_s and the minimum distance between two cluster centers θ_c , we performed tests by setting different values and verified the accuracy of the clustering results with some known sample data sets. Then, we selected the values that provided the best clustering results as the preset values. Finally, they were set to 20, 5, and 2, separately. The corn plants from the 3-leaf stage to the 6-leaf stage were scanned by an ultrasonic sensor and the data were clustered with the proposed algorithm. The results are shown in Table 1 and in Figure 2.

Table 1. Cluster center and test index.

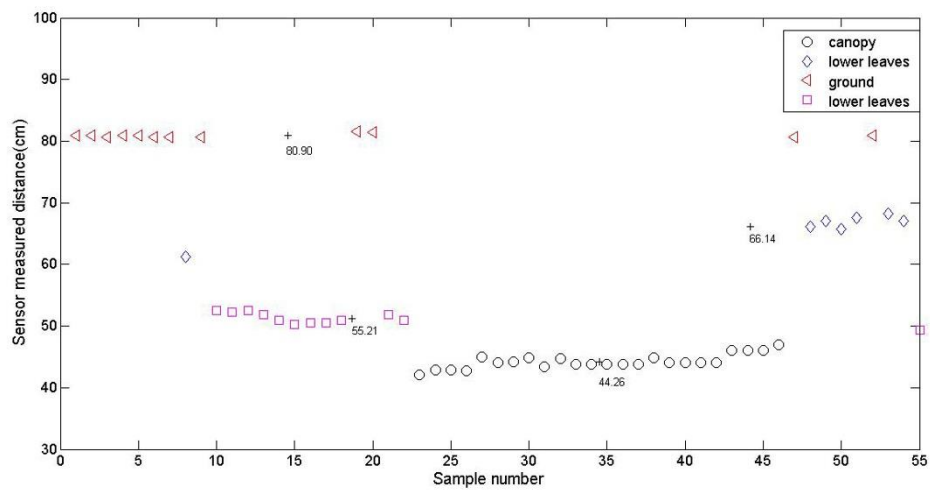
Growth Stage	Number of Clusters	Cluster Center Value (cm)			
3-leaf stage	2	80.17	63.25		
4-leaf stage	3	79.22	65.53	59.81	
5-leaf stage	4	80.90	66.14	51.21	44.26
6-leaf stage	4	66.46	46.87	43.19	39.71



(a)

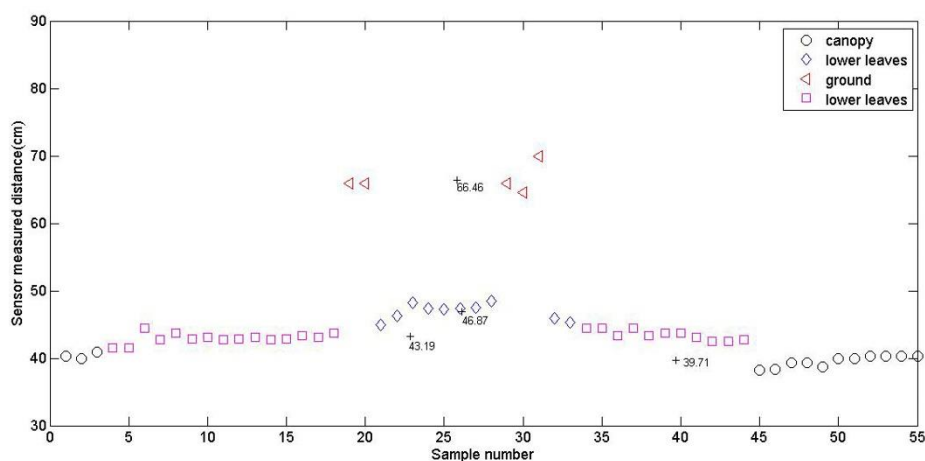


(b)



(c)

Figure 2. Cont.



(d)

Figure 2. Clustering results for corn plants at different growth stages: (a) 3-leaf stage; (b) 4-leaf stage; (c) 5-leaf stage; and (d) 6-leaf stage.

For corn plants of the 3-leaf stage, the cluster center value 63.25 and 80.17 represented the distances from the sensor to the canopy and the ground, respectively. For corn plants at the 4-leaf stage, the cluster center values 59.81, 65.53, and 79.22 were the distances from the sensor to the canopy, to the lower leaves, and to the ground. For corn plants at the 5-leaf stage, there were 4 clusters. Except for the distances from the sensor to the canopy and the ground, the distance from the sensor to the lower leaves was classified into 2 clusters. Their centers were 51.21 and 66.14. For corn plants at the 6-leaf stage, there were also 4 clusters. Nevertheless, the maximum value of cluster center 66.46 can not represent the distance from the sensor to the ground. Due to the growth of the plant, the canopy volume increases and prevents the ultrasonic wave from reaching the ground; as a result, the distance from the sensor to the ground can not be measured. The distance from the sensor to the lower leaves was classified into 3 clusters. Their centers were 43.19, 46.87, and 66.46. All test indexes were very small and proved that the results were reliable. So, we can use the fuzzy ISODATA to deal with the distance data to acquire the locations of the crop canopy and the ground, which are very important for controlling the boom height. However, if the crop was fully covered and no ground was observed, the cluster representing the ground would not be acquired. However, that would not affect the cluster representing the canopy.

The calculation results of fuzzy ISODATA were compared with ones of other methods, such as k-means clustering, mean, and median. The results are listed in Table 2.

Table 2. Calculated and manually measured values of distance from the sensor to the canopy with different algorithms. ISODATA: Iterative Self-Organizing Data Analysis Technique Algorithm.

Growth Stage	Manually Measured Value (cm)	Fuzzy ISODATA (cm)	K-Means Clustering (cm)	Mean (cm)	Median (cm)
3-leaf stage	60	63.25	63.65	74.22	78.6
4-leaf stage	57	59.81	65.52	74.33	77.70
5-leaf stage	46	44.26	43.89	55.05	50.75
6-leaf stage	38	39.71	42.00	45.18	43.1

Obviously, the values of fuzzy ISODATA are the closest to the manually measured values. Compared with the k-means clustering algorithm, fuzzy ISODATA has better clustering results. The reason is that the number of clusters is difficult to estimate in k-means clustering. Moreover, the choice of initial cluster center has a greater impact on the clustering results. However, fuzzy ISODATA can dynamically adjust the number and center of clusters to make the clustering results closer to the objective results. As for the methods of computing the mean and median, they all have

big errors because they can not exclude the points of lower leaves and ground when calculating the mean and median. However, for fuzzy ISODATA, there are relatively more parameters to be set, and the parameter values are not easy to determine. To get good clustering results, good initial values are required.

3.2. Influence of Plant Growth Stage on Calculation Accuracy

When the sensor speed was 1 km/h, for corn plants from the 3-leaf stage to the 6-leaf stage (Figure 3), the calculated distances from the sensor to the canopy were compared with the manually measured value. The results are shown in Table 3.

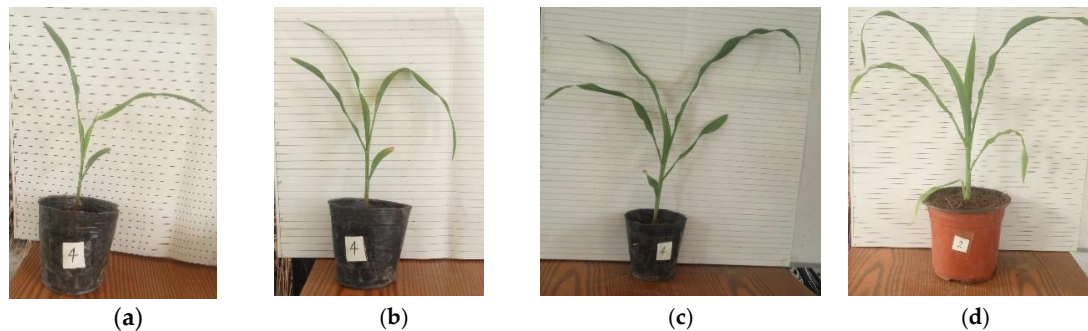


Figure 3. Corn plants at different growth stages: (a) 3-leaf stage; (b) 4-leaf stage; (c) 5-leaf stage; and (d) 6-leaf stage.

Table 3. Calculated and manually measured values of distance from the sensor to the canopy at different growth stages.

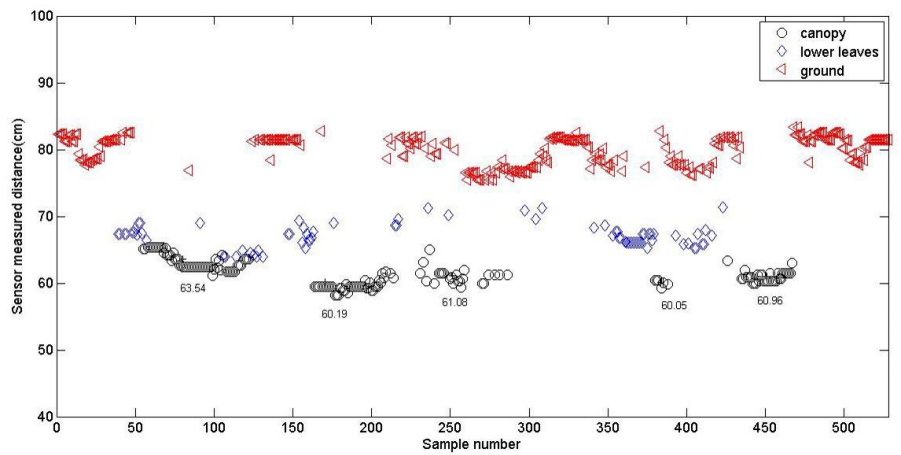
Growth Stage	Calculated Value (cm)	Manually Measured Value (cm)	Absolute Error (cm)
3-leaf stage	63.25	60	3.25
4-leaf stage	59.81	57	2.81
5-leaf stage	44.26	46	1.74
6-leaf stage	39.71	38	1.71

As shown in Figure 3, at the 3-leaf stage, the plant leaves were short and narrow, so the leaf area was small. Furthermore, the inclination angle of leaves was also small. The ultrasonic echoes bounced from the erect leaves were not detected by the sensor, because these leaves reflected the ultrasonic signal away from the sensor. Consequently, this caused the larger estimation errors of the distance to the canopy at the 3-leaf stage than the ones at other stages. With the growth of the plants, the leaf area increased and the canopy volume became larger. Leaves were gradually extended and tended to be more perpendicular to the sound beam from the sensor. The ultrasonic wave reflected from the canopy increased, and more distance data were able to be used for clustering. Therefore, the calculated error gradually decreased from the 3-leaf stage to the 6-leaf stage. Although the error reached 3.25 cm at the 3-leaf stage, it was not more than 5 cm and met the requirement of boom height control.

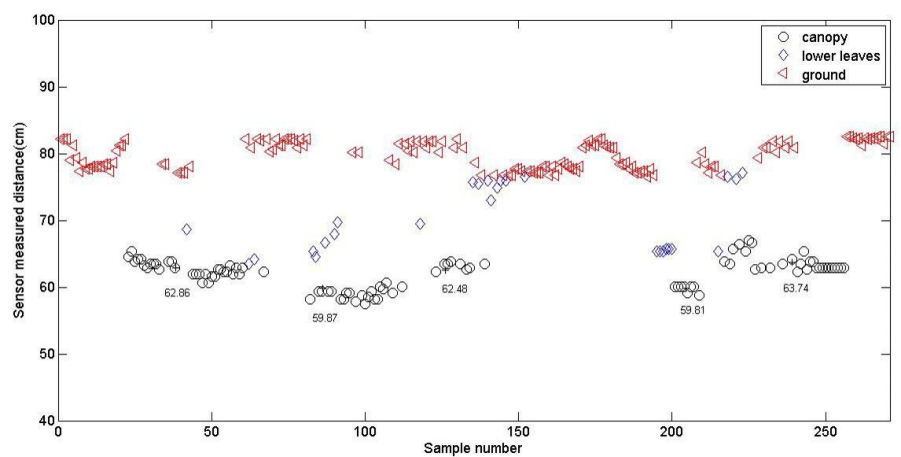
3.3. Influence of Sensor Moving Speed on Calculation Accuracy

At the 3-leaf stage, the measurement error was larger, owing to the larger inclination angle of leaves. That would make the conclusions less reliable. From the 4-leaf to 6-leaf stage, the leaves became flatter and flatter, and the measurement error was smaller. Moreover, the experiments results showed that the conclusions for the 4-leaf stage are the same with that for the 5-leaf stage and 6-leaf stage. So, we only present the results for the 4-leaf stage.

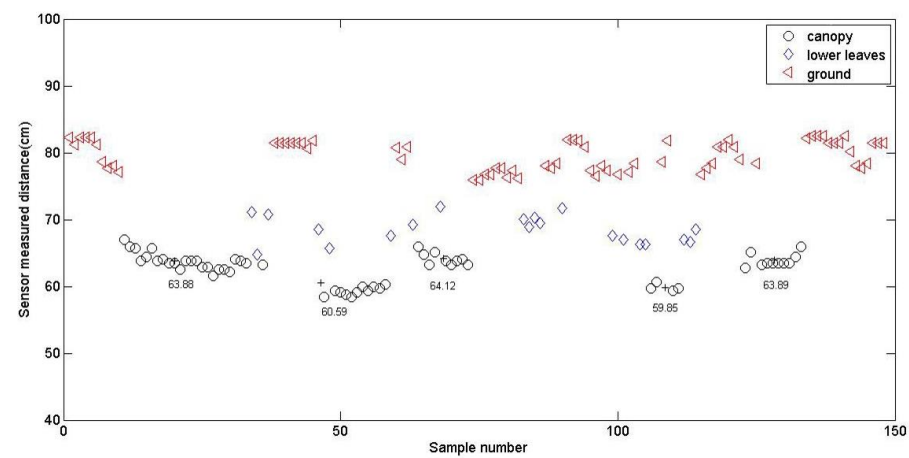
For five randomly chosen corn plants at the 4-leaf stage, the distances from the sensor to the canopy were calculated at different sensor moving speeds from 0.5 km/h to 6 km/h. The clustering results are demonstrated in Figure 4, and the detailed values are shown in Table 4.



(a)

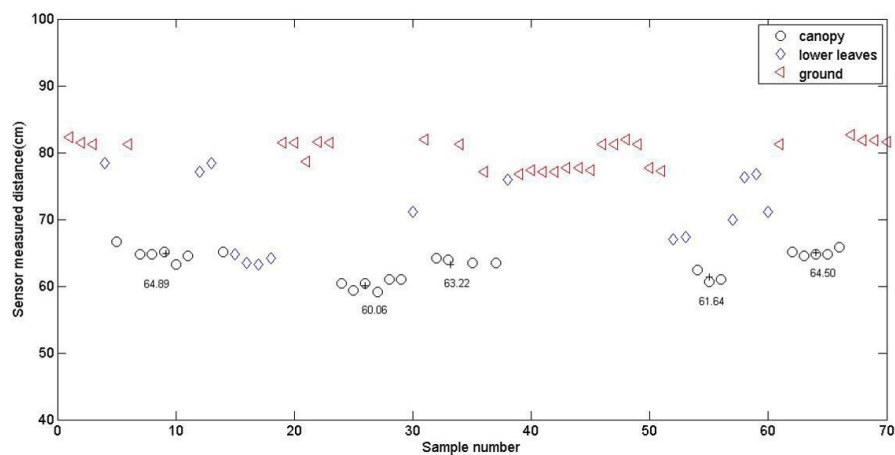


(b)

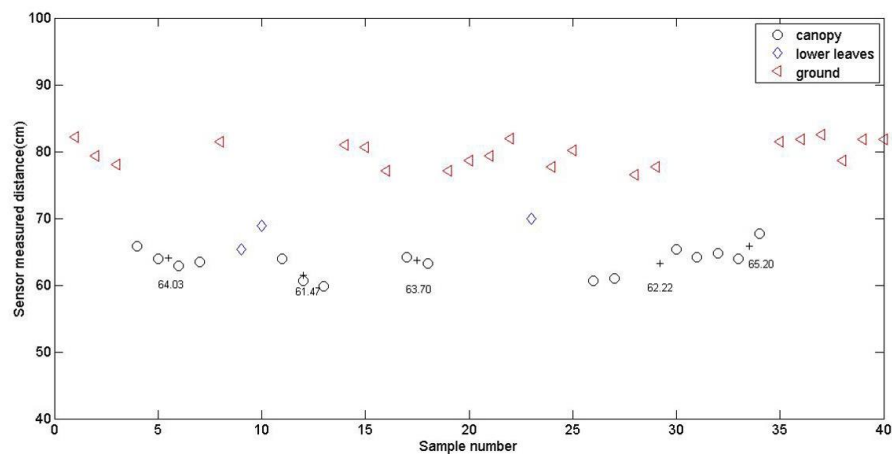


(c)

Figure 4. Cont.



(d)



(e)

Figure 4. Clustering results for corn plants at the 4-leaf stage at different speeds: (a) $v = 0.5$ km/h; (b) $v = 1$ km/h; (c) $v = 2$ km/h; (d) $v = 4$ km/h; (e) $v = 6$ km/h.

Table 4. Calculated and manually measured values of distance from the sensor to the canopy at different sensor moving speeds.

Plant Number	Manually Measured Value (cm)	Calculated Value at Different Speed (cm)				
		0.5 km/h	1 km/h	2 km/h	4 km/h	6 km/h
Plant no. 1	61	63.54	62.86	63.88	64.89	64.03
Plant no. 2	59	60.19	59.87	60.59	60.06	61.47
Plant no. 3	60	61.08	62.48	64.12	63.22	63.7
Plant no. 4	57	60.05	59.81	59.85	61.64	62.22
Plant no. 5	61	60.96	63.74	63.89	64.5	65.2
Mean values	59.6	61.164	61.752	62.466	62.862	63.324
Absolute error		1.564	2.152	2.866	3.262	3.724
t value		0.15	0.09	0.04	0.02	0.01

The F test result indicated that the variances of calculated values and manually measured values were homogeneous. Therefore, they can be compared by a t -test. When the significance level was 0.05, $t_{0.05} = 2.57$, and the t values at all speeds were all less than $t_{0.05}$. It proved that there were no significant differences between the calculated values using fuzzy ISODATA and the manually measured values. However, as shown in Table 4, the absolute error increased with the increase of sensor moving speed. The reason was that the faster the sensor moved, the less the amount of data it

acquired, which is demonstrated in Figure 4. When the speed reached 6 km/h, there were only about 40 points. Although the collected data were reduced, the distance value from the sensor to the crop canopy could still be obtained and relatively accurate after clustering.

The distribution of the absolute error between the manually measured values and the calculated values was checked using the MATLAB function `lillietest(x)` at different sensor moving speeds. This function outputs an h value and a p value. If $h = 0$, the inputted data follows a normal distribution. If $p > 0.05$, the null hypothesis can be accepted. The test result is shown in Table 5. It indicated that the absolute error was normally distributed at the 5% significance level.

Table 5. Normal distribution test result of absolute error.

Value	V = 0.5 km/h	V = 1 km/h	V = 2 km/h	V = 4 km/h	V = 6 km/h
h	0	0	0	0	0
p	0.5	0.3481	0.1853	0.2078	0.5

Note: when p is greater than the largest tabulate value in MATLAB, it is set 0.5.

Analysis of variance was applied to verify the influence of the sensor moving speed on the absolute error. The results are shown in Table 6. It indicated that the sensor moving speed has a significant influence on the absolute error at the 5% significance level.

Table 6. Variance analysis result of absolute error.

Source	Sum of Squares of Deviation	Degree of Freedom	Mean Squares of Deviation	F Value	Significance
Groups	14.91	4	3.73	3.16	*
Error	23.64	20	1.18		
Total	38.55	24			

Note: $F_{0.05}(4, 20) = 2.86$, "*" denotes significance.

3.4. Correlation between Calculated Values and Manually Measured Values

Regression analysis was used to describe the relationship between calculated values and manually measured values for all growth stages. The calculated distances using fuzzy ISODATA and manually measured distances from the sensor to the canopy were graphed on an XY scatter chart. After the data points were plotted, a trend line, equation, and R^2 were added to illustrate the relationship. The results are shown in Figure 5, which illustrated that the linear regression of calculated distances on manually measured distances had an R^2 of 0.88 with all growth stages included. The results were encouraging, since the calculated distances showed a strong linear correlation with the manually measured distances. The trend for increased correlation with time was similar to the absolute error described before. This trend in corn is intelligible considering the increased size of the plant leaf with advancing growth. Moreover, as the corn plant grew, the increased inclination angle of the leaf made it easier for the canopy to be detected by the sensor, therefore causing the correlation to increase.

The p -values of slope and intercept of the regression line were 1.53×10^{-19} and 0.00448 separately. They are far less than 0.01, which indicates that the two regression coefficients are highly significant. The residual distribution was tested using the MATLAB function `lillietest(x)`. The result suggested that it did not follow a normal distribution. We found two outliers, the standardized residuals of which are 3.145 and 3.574, respectively. Furthermore, the two data were from the 3-leaf growth stage. Since the leaves were very inclined at this stage, the ultrasonic echo perhaps did not return to the sensor after reflection. That led to the wrong measured data. Consequently, we eliminated the two outliers and performed the regression analysis one more. The new regression equation was $y = 0.876x + 10.587$, and the new R^2 value was 0.94. The p -values of the two coefficients were 1.38×10^{-24} and 1.43×10^{-6} separately. It indicates that they are more significant than those of the old regression equation. In addition, the residual was normally distributed after elimination of the two outliers. The standard

errors of regression coefficients were calculated for different growth stages. They were 0.56, 0.39, 0.14, 0.10 for the slope and 34.06, 23.96, 6.45, 3.79 for the intercept respectively from the 3-leaf to the 6-leaf growth stage. The smaller the error, the more accurate the coefficient. The errors decreased with the plant growth. In addition, we calculated the leverage of all points. The results demonstrated that there was not a high leverage point. None of points can strongly influence the slope of the regression line.

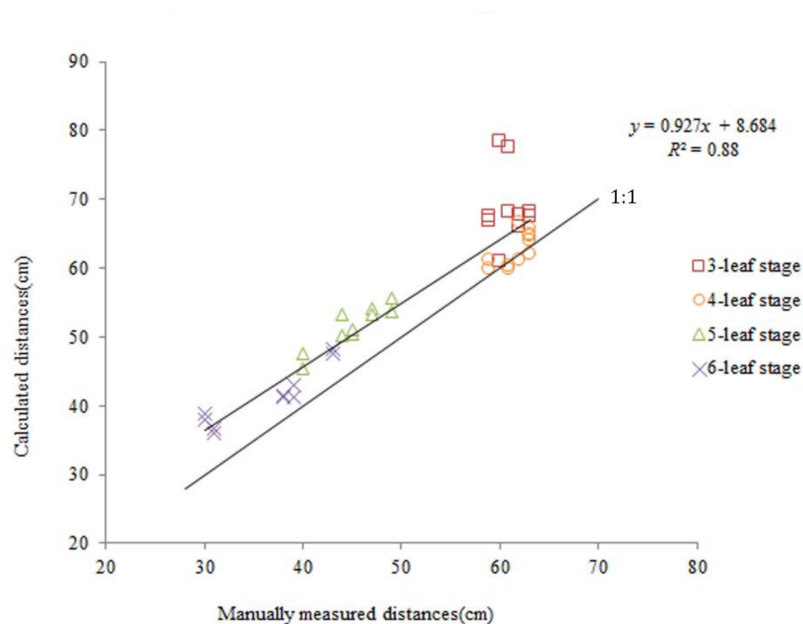


Figure 5. Calculated distances regressed onto manually measured distances from the sensor to the canopy for corn plants including all growth stages.

4. Conclusions

In this paper, ultrasonic ranging was combined with fuzzy ISODATA to obtain the distances from the sensor to the crop canopy for corn plants from the 3-leaf stage to the 6-leaf stage. The calculated values had a high correlation with the manually measured values. Applying the proposed method, we can acquire the distances from the sensor to the canopy and the ground at the same time for plants from the 3-leaf stage to the 5-leaf stage. After the 6-leaf stage, the distances to the ground cannot be obtained due to the increased blockage of sound waves by plant leaves. With the growth of plants, the error of calculated values decreased. The increased sensor moving speeds led to increased error. However, the errors were all less than 5 cm and acceptable in field application.

Our development can be used in the boom height control system of sprayers. We can install one ultrasonic sensor on each end of the boom to measure the distance from the boom to the crop canopy or the ground. The data were processed with a computer to judge if the distance was appropriate. If not, the computer would send a control signal to a hydraulic cylinder or an electrical motor to lift or lower either wing section of the boom. So, the boom can be maintained at a constant or near constant distance.

However, there are still some limitations in our work. The sampling rate of ultrasonic sensors is lower. So, the acquired data are fewer when the sensor moves at a higher speed. That will affect the clustering accuracy. The determination of parameters of ISODATA is difficult and requires full understanding of them and sufficient experience. Some outdoor conditions, such as the air temperature, humidity, and wind velocity, can influence the performance of ultrasonic sensors. When the sprayer works, uneven soil surface causes the vibration of boom, which may affect the distance measurement. Moreover, the mutual occlusion of leaves of different crops may result in more extraordinary data.

In the future, more research is needed to learn how the beam angle of ultrasonic sensors affects the measurement errors and calculation results. Some filter algorithms should be developed to eliminate the vibration effect and the abnormal data before clustering. The research subject should be extended to other crops, such as wheat, cotton, peanut, and so on. In addition to lab experiments, field experiments should also be conducted to verify the adaptability of our work. In order to achieve more data, laser sensors can be applied to take the place of ultrasonic sensors.

Author Contributions: F.L. was the principal investigator of the research and drafted the manuscript. X.B. and Y.L. provided guidance in developing the algorithm and performing the experiments. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Natural Science Foundation of Liaoning Province (No.2015020769).

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

References

1. Qi, L.; Fu, Z. Experimental study on spray deposition uniformity. *Trans. Csaе* **1999**, *15*, 113–117.
2. Chen, Z.; Wu, C.; Yang, X. Spray distribution uniformity of boom spraying. *J. Jiangsu Univ.* **2008**, *29*, 465–468.
3. Jeon, H.; Womac, A.; Gunn, J. Sprayer boom dynamic effects on application uniformity. *Trans. ASAE* **2004**, *47*, 647–658. [[CrossRef](#)]
4. Andreas, H.; Hans, O.; Hartje, S. A novel method for testing automatic systems for controlling the spray boom height. *Biosyst. Eng.* **2018**, *174*, 115–125.
5. Ramon, H.; Missotten, B.; Baerdemaeker, J. Spray boom motions and spray distribution: Part 2, Experimental validation of the mathematical relation and simulation results. *J. Agri. Eng. Res.* **1997**, *66*, 31–39. [[CrossRef](#)]
6. Paolo, B.; Emilio, G. Field-crop-sprayer potential drift measured using test bench: Effects of boom height and nozzle type. *Biosyst. Eng.* **2017**, *154*, 3–13.
7. Cui, L.; Xue, X.; Ding, S. Analysis and test of dynamic characteristic of large spray in boom and pendulum suspension damping system. *Trans. CSAE* **2017**, *33*, 61–68.
8. Qiu, B.; Yan, R.; Ma, J. Research progress analysis of variable rate sprayer technology. *Trans. CSAM* **2015**, *46*, 59–72.
9. Dionisio, A.; Martin, W.; Roland, G. An Ultrasonic System for Weed Detection in Cereal Crops. *Sensors* **2012**, *12*, 17343–17357.
10. Thomas, F.; Felix, R.; Michal, W. Assessment of forage mass from grassland swards by height measurement using an ultrasonic sensor. *Comput. Electron. Agric.* **2011**, *79*, 142–152.
11. Shrestha, D.S.; Steward, B.L.; Birrell, S.J. Corn plant height estimation using two sensing systems. In Proceedings of the 2002 ASAE Annual International Meeting/ CIGR XVth World Congress, Chicago, IL, USA, 28–31 July 2002; p. 021197.
12. Kataoka, T.; Okamoto, H.; Kaneko, T.; Hata, S. Performance of corn height sensing using ultrasonic sensor and laser beam sensor. In Proceedings of the 2002 ASAE Annual International Meeting/ CIGR XVth World Congress, Chicago, IL, USA, 28–31 July 2002; p. 021184.
13. Chang, Y.K.; Zaman, Q.U.; Rehman, T.U.; Farooque, A.A.; Esau, T.; Jameel, M.W. A real-time ultrasonic system to measure wild blueberry plant height during harvesting. *Biosyst. Eng.* **2017**, *157*, 35–44. [[CrossRef](#)]
14. Liu, X.; Li, Y.; Li, M. Design and test of smart-targeting spraying system on boom sprayer. *Trans. CSAM* **2016**, *47*, 37–44.
15. Wei, X.; Shao, J.; Miao, D. Online control system of spray boom height and balance. *Trans. CSAM* **2015**, *46*, 66–71.
16. Wang, S.; Zhao, C.; Wang, X. Design and experiments on boom height adjusting system. *J. Agri. Mech. Res.* **2014**, *36*, 161–164.
17. Jin, X.; Zhang, H.; Zhou, Y. Agricultural land gradation based on the fuzzy ISODATA clustering method. *Trans. CSAE* **2008**, *24*, 82–85.
18. Park, H.; Jun, C. A simple and fast algorithm for K-means clustering. *Expert Syst. Appl.* **2009**, *36*, 3336–3341. [[CrossRef](#)]
19. Jin, H.; Shi, D.; Chen, Z. Evaluation indicators of cultivated layer soil quality for red soil slope farmland based on cluster and PCA analysis. *Trans. CSAE* **2018**, *34*, 155–164.

20. Pan, R.; Liu, J.; Song, D. Agricultural land classification based on fuzzy comprehensive analysis. *Trans. CSAE* **2014**, *30*, 257–265.
21. Zhao, R.; Huang, X.; Zhong, T.; Xu, H. Application of clustering analysis to land use zoning of coastal region in Jiangsu Province. *Trans. CSAE* **2010**, *26*, 310–314.
22. Bezdek, J.C. A physical interpretation of fuzzy ISODATA. *IEEE Trans. SMC* **1976**, *6*, 387–390.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).