Nondestructive Testing of Lettuce Nitrogen Stress Based on Multidimensional Image

Bao Guo Shen[1*], Jin Yue Dai[1], Xiao Dong Zhang[2], Zhao Hui Duan[2]

[1]Zhenjiang Key Laboratory of UAV Application Technology, Jiangsu Aviation Technical College, Zhenjiang, China

[2]Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education, Jiangsu University, Zhenjiang, China

*Correspondence：shenbg@jatc.edu.cn

Abstract

Visible light near infrared (VS-NIR) hyperspectral combined with three-dimensional laser scanning was applied to extract the VS-NIR features of lettuce nitrogen between 400-1700 nm and 3D morphological features of the plants. Such combination realizes the rapid quantitative detection of lettuce nitrogen. This study is based on the hyperspectral image data cube achieved from lettuce leaves with different nitrogen levels. Stepwise regression sensitive area was used and adaptive band selection method was combined to extract the characteristic spectrum and feature image of lettuce nitrogen and characterize the average image intensity. Also; the error caused by moisture variation content in lettuce nitrogen image features was compensated. Then a model of lettuce nitrogen hyperspectral image diagnosis was built. The reverse engineering software Geomagic Qualify was used to repair and smooth interference noise and discontinuous range which are based on the 3D laser scanning data of lettuce. Accordingly, the stem diameter, plant height, leaf area, and biomass features of different nitrogen levels of lettuce are obtained and the model of nitrogen detection about lettuce growth features was built based on reverse engineering and integral method. Multi-scale fusion lettuce nitrogen detection model is built by using the acquired hyperspectral images with growing features of lettuce nitrogen and adopting genetic algorithm combined with partial least squares regression. Results show the correlation coefficient R of the built model is 0.95; the model precision is much better than single feature of hyperspectral images and 3D laser scanning model. The feature extraction algorithm and the eigenvectors provide the reference for development of facilities for online monitoring system of crop growth information.

Keywords: Lettuce, Nitrogen, Hyperspectral images, 3D laser scanning, Information fusion

Introduction

China's facility area has reached 3.8 million hectares. Lettuce is one of the most important facilities for planting leafy vegetables in China. However, most of the planting facilities still use large-water and large-fertilizer irrigation and fertilization methods or adopt the set value regulation. However, such treatments do not consider or are unable to comprehensively and accurately monitor the fertilizer demand information and dynamic change of crops, and can hardly to achieve dynamic regulation according to the real demand of crops. The consequences are not only much waste and non-point source pollution, but also the direct effects on the taste, quality and economic benefits of lettuce. Therefore, it is urgent to accurately monitor the nutrition, and growth process of crops, judge the growth state of crops, and achieve precise regulation based on crop growth requirements.

During traditional facility nutrition management, nutrition testing is mainly based on expert experience and chemical determination. However, expert experience is greatly influenced by subjective factors, and despite its high detection accuracy, dynamic feedback control of nutrient solution is difficult due to low timeliness. In comparison, non-destructive detection methods such as hyperspectral remote sensing, visual image and 3D scanning do not harm crops; and gradually become a hot spot for non-destructive testing of nutrition because of their rapidity and timeliness [1]. Scholars have conducted relevant research and achieved some results [2-14]. Non-destructive diagnostic methods based on reflectance spectroscopy usually use point source sampling to reflect the comprehensive reflection characteristics in the sampling area. Although the multi-spectral

combination feature can reproduce the nutrient information of crops, it can-not reflect the light of the entire leaf area. In fact, the reflection characteristics and color texture of leaves and their distributions are subtly different. The visual sensor has a high resolution and a large range of field of view. With spectral methods via image segmentation, the influence of background and other factors can be removed to overcome the shortcomings of a small testing range and strict requirements on test sites which contribute to the acquisition of more crop information. However, the traditional vision sensor is usually limited by low spectral resolution, since it usually acquires a single color image in the visible range (400-700 nm); and can-not extract or analyze the image features of different bands. In comparison, the hyperspectral image technology can obtain hyperspectral image cube data composed of hundreds of wavelengths in the visible and near-infrared(VS-NIR) range of 400-1700 nm at a high resolution of 3-10 nm, which can acquire the reflection spectrum of plant leaves and leaf spectra in different periods of reflection intensity distribution for synchronous data acquisition and analysis. It has the advantages that are not possessed by the traditional spectrum and image technology. Therefore, in this study the hyperspectral image technique to extract and analyze the nitrogen characteristics at the crop leaf scales. Although this technology can analyze the characteristics of crop leaf size reflection intensity, color, and texture, there are differences in leaf area, plant height, stem diameter and other growth features caused by nitrogen stress. Synchronous and efficient acquisition of plant morphological features is difficult due to differences in data collection methods and multi-objective view field, blade stacking, and occlusion of targets.

Nitrogen abundance or deficiency of crops will directly lead to differences in leaf area, stem diameter and plant height, so nitrogen inversion of crops can also be used as an effective feature. However, the traditional visual image and contact measurement method cannot obtain the whole picture of the plant synchronously; and precision is not high. The 3D laser scanning can simultaneously acquire the overall morphological features of the plant; and can reach the micron level precision, so it can extract and analyze the differences in leaf area, plant height and stem length caused by nutritional stress. We will combine a specular like technology and 3D laser scanning, and make full use of the different lettuce leafs scale sensitive spectra of hyperspectral images, as well as the different levels of nitrogen stress caused by canopy-scale leaf area, plant height, stem diameter; and morphological differences, through multi-scale information fusion of nitrogen nutrition on lettuce inversion and quantitative analysis; We aim to improve the accuracy of non-destructive lettuce nitrogen nutrition detection; and provide a basis for accurate management of facility water and fertilizer based on crop growth information. There is no relevant reports worldwide.

Materials and Methods

Sample Cultivation

Experiment were carried out at the Venlo-type greenhouse of the Ministry of Modern Agricultural Equipment and Technology of Jiangsu University. For precise control of lettuce nitrogen, a soilless cultivation technique was used in sample cultivation. To ensure the balance of other nutrients, we precisely controlled nitrogen to obtain lettuce samples with different nitrogen stress levels.

The selected variety in the experiments was the annual heat-resistant lettuce from Italy. When the lettuce grew to the state of four leaves and one heart, the seedlings with similar growth were moved to hydroponic cultivation plates, and the nutrient solution was configured according to the Yamazaki formula [15]. The samples in groups A, B, C and D were treated at four nitrogen levels respectively. Nutrient solutions with 25%, 50%, 100% and 200% N in the standard formula (all in unit of mass fraction) were prepared. Totally 48 samples were collected from 12 plants at each level.

Test methods

Hyperspectral image acquisition and processing

The hyperspectral images of lettuce leaves were collected by a VS-NIR hyperspectral image acquisition system. This system mainly consists of VS-NIR cameras (VS:390.8-1050.1 nm, NIR :871.6-1766.3 nm), an ImspectorN17E

spectrograph, OLES30 lens, a DC adjustable light source of 150W tungsten halogen lamp, a glass fiber symmetrical line source, stages, a controller and a computer (Figure 1).

Before data acquisition, a pre-sampling experiment was carried out. The exposure time of hyperspectral image formation was set as 15 ms and the scanning speed as 1.25 mm/s, the peak reflection intensity of the pre-sampling image of lettuce was 3000, which ensured the clarity and non-distortion of the image; After that, black-white field calibration was carried out, and the range of reflection intensity was set at 0-4000. According to the noise intensity and distribution, a second-order butter-worth filter was selected to remove noise.
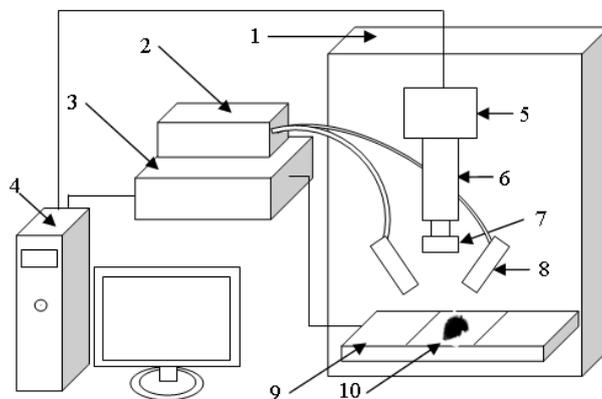

Figure 1. Hyperspectral imaging system. 1: is light box ,2: light source,3: displacement console, 4: computer, 5:VS-NIR camera, 6: imaging spectrometer, 7: lens, 8: glass fiber optic light, 9: displacement table, 10: lettuce leaves.

The collection of hyperspectral image was based on Spectral Cube. The collected VS-NIR spectra ranges from 390.8 to 1766.3 nm, and the spatial resolution is 31.25 um. Within the sampling range, hyperspectral image data cubes composed of 768 images of different spectral segments were obtained at intervals of 1.3nm (390.8 - 1050.1 nm) and 3.5 nm (871.6 - 1766.3nm).

Lettuce growth information scanning and data acquisition

A handheld self-positioning 3D laser scanner was applied to collect morphological data of the whole lettuce. This 3D laser scanner, mainly consists of a computer, Handyscan3D (EXA scan), FireWire adapter, FireWire cables, power supply and other components (Figure 2). The specification of the instrument was: measuring speed at 25.000 measures per second, resolution of 0.050 mm; Accuracy is 0.040 mm; shooting distance is300mm; field depth between 150 and minus 150; laser crossover area is 210 mm plus 210 mm.
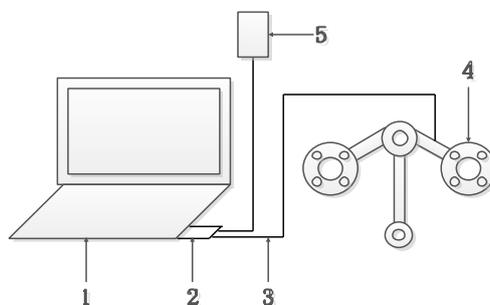
Figure 2. Handheld self-position 3D laser scanner. 1: Computer, 2: firewire adapter, 3: firewire cable, 4: handyscan3d (EXA scan), 5: power supply.

During the scanning, the high-reflection target point with a 6-mm-diameter black contour was firstly pasted on the lettuce leaves and cultivation plates to be scanned. Due to the curved surface of lettuce leaves, the minimum distance between two target points should be 20 mm when pasting the reflected target points. When the surface

of a cultivation plate was flat, the distance between the control target points was 100 mm. After that, the 3D laser scanner was used to measure the calibration plate to correct the sensor parameters and ensure data acquisition accuracy. Before data acquisition by the 3D laser scanner, the laser power, shutter time and resolution of the acquisition software of the scanner sensor must be determined to ensure the clarity of the 3D model. After analysis and comparison, the final laser power was set at 65%, the shutter time was 7.2 ms, and the resolution was 0.50 mm. Finally, through hand-held scanning, 3D data of all the lettuce samples were sequentially obtained.

Lettuce growth information scanning and data acquisition

The weighing method was applied to measure fresh biomass. Specifically, firstly, the total weight of the cultivation board together with the lettuce was controlled for 3 minutes; Secondly, the weight of the cultivation board and the planting basket was subtracted to obtain the fresh biomass of the lettuce. Each test sample was analyzed by an AA3 continuous flow analyzer, and its nitrogen content was calculated by the formula (1) [16].

$$N = \frac{c}{m \times (1-w)} \times 100\% \quad (1)$$

where: c -- observed value of sample solution (mg); m --test sample weight (mg); w -- water content of the sample (%). Laboratory chemical analysis showed, the nitrogen content of all samples were distributed between 0.81% and 4.62%.

Results and analysis

Feature Extraction and Analysis of Nitrogen Hyperspectral Images of Lettuces

Image background segmentation

Background segmentation was firstly carried out to obtain the hyperspectral target image of lettuce leaves. Therefore, we cut the image with the circumscribed rectangle of the blade on ENVI. After that, through threshold analysis of the target image and background, the 480 nm image was finally selected for threshold segmentation based on the two-peak method [17], at the threshold of 157. Then, gray level inversion was carried out on the segmented binary target image, and image processing was further carried out based on mathematical morphology operation to fill in the residue and remove isolated noise points. Finally, the hyperspectral sequence target image of lettuce leaves was obtained by multiplying the original hyperspectral image with the processed binarization target image.

Nitrogen characteristic spectrum extraction

The hyperspectral image data cube consists of 768 images within 390.8 to 1766.3 nm. Each image corresponds to a reflection intensity distribution (or gray scale distribution) of the target image at the corresponding wavelength, and the information amount is extremely rich compared to the conventional color image. As reported, the sensitivity range of lettuce nitrogen is mainly 390- 450 nm (including 49 bands), 520 - 570 nm (40), 600 - 650nm (39) and the red-edge region of 680 - 750 nm (44) [18,19]. To improve the effectiveness of feature extraction, we adopted sensitive interval stepwise regression [20,21] to screen nitrogen features. The main idea of stepwise regression is to introduce the equation one by one in all the variables considered n the order from large to small significance degree to nitrogen. F test was carried out at each step to ensure that before introduction of new variables, only the variables with significant influence on the dependent variable were included in the regression equation, while the insignificant variables were eliminated.

Specifically, with nitrogen in lettuce green mountain area as example, each wavelength was selected every 1.3 nm (forming a specular system at the resolution of 1.3 nm) between 520 and 570 nm. Then a total of 40 wavelengths in the average grey value of an image as a variable, and the size of partial correlation coefficients between variables in a multivariable equation, were introduced the spectral variable xi one by one, and total nitrogen content of y AA3 was measured values to establish the regression model:

$$y_i = b_o + \sum_{i=1}^{40} b_i x_i + \delta \quad (2)$$

where: b_o is the constant term of the model; b_i (i = 1, 2... is the partial regression coefficient of selected variables); δ is the residual; x_i is the average gray level or reflection intensity of feature images of selected bands; y_i is the measured total nitrogen content.

After that, the total sum of squares SST and coefficient of determination R^2 of the regression model were derived. The bass for measuring the contributions of hyperspectral image feature variables to the nitrogen model is the decision coefficient of the model and the significance of the F test. That is, when the model R^2 is larger than the pre-set R^2 and F_i is larger than the test level F_α, then x_i has a significant impact on nitrogen content, and the introduction is successful; otherwise, the variable is eliminated. Following this, we selected the next variable. The optimal characteristic variables and regression models were obtained until the variables were no longer introduced and eliminated.

The following stepwise regression criteria adopted: when the independent variable enters, the equation F is larger than 2.35, and when the selected variable makes F smaller than 1.97, the variable is removed, and R^2 is larger than 0.5, and the number of variables in each group is less than or equal to 5. Based on these conditions, the regression equation of each interval was obtained as follows:

$$N_{390-450nm} = 7.27 + 9.83 AG_{402} + 37.52 AG_{418} - 52.13 AG_{429} - 17.86 AG_{446} \quad (3)$$

$$N_{520-570nm} = -5.21 + 31.07 AG_{522} - 24.33 AG_{540} - 31.65 AG_{556} + 41.06 AG_{569} \quad (4)$$

$$N_{600-650nm} = 3.28 - 16.39 AG_{615} + 41.56 AG_{636} - 17.91 AG_{650} \quad (5)$$

$$N_{680-750nm} = 22.58 + 31.63 AG_{685} - 12.47 AG_{699} + 79.46 AG_{706} - 5.25 AG_{741} \quad (6)$$

Where A_{Gi} (i theinterval: 390 – 750 nm) is the gray-scale variable of hyperspectral image; N_j (j is 390 – 450 nm, 520 – 570 nm, 600 – 650 nm, 680 – 750 nm) is the predicted value of nitrogen content in hyperspectral sensitive range j.

The significance F value of the N390-450nm, N520-570nm, N600-650nm and N680-750nm equations was 31.29, 37.06, 19.51, and 21.89 respectively, R^2 was 0.59, 0.65, 0.55, 0.61 respectively, and the standard deviation was 0.62, 0.55, 2.51, 1.85 respectively.

The sensitive intervals of the 15 characteristic bands obtained above were set in advance according to previous research. To overcome the interband orrelation in sensitive intervals and the multi-collinearity of bands with different sensitive intervals, reduce the redundant information in hyperspectral image data and improve the effectiveness of data processing, we adopted the adaptive band selection (ABS) method.

ABS fully considers the spatial and spectral correlations of various bands. The relevant mathematical models are as follows:

$$I_i = \frac{\sigma_i}{(R_{i-1,i} + R_{i,i+1})/2} \quad (7)$$

$$\sigma_i = \left[\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (f(x,y) - \bar{f}_i)^2 \right]^{\frac{1}{2}} \quad (8)$$

$$R_{i,j} = \frac{E\{(f_i(x,y) - \bar{f}_i)(f_j(x,y) - \bar{f}_j)\}}{\sqrt{E\{(f_i(x,y) - \bar{f}_i)\}^2} \sqrt{E\{(f_j(x,y) - \bar{f}_j)\}^2}} \quad (9)$$

Where: σ_i is standard deviation of the i-th band; E{}is the--mathematical expectation; $R_{i-1,i}$ and $R_{i,i+1}$ are the correlation coefficient between the i-th band and its two adjacent bands; I_i is the index factor of the i-th image; M, N are image line, column number of pixels; $f_i(x, y)$ is the i-th image; f_i is the pixel average of the i-th image.

The index obtained by ABS fully considers the enrichment of image information and the similarity between adjacent bands. A larger li means, more information is contained in the image, and the multicollinearity and other band images are less relevant.

Table1. Index and its corresponding band number

Wavelength /nm	index	Band number	Wavelength /nm	index	Band number
402	985.37	11	615	575.31	182
418	689.96	24	636	906.55	198
429	727.39	33	650	873.21	209
446	941.04	47	685	796.35	236
522	823.57	108	699	991.87	247
540	877. 49	122	706	998.22	253
556	1088.58	135	741	370.43	280
569	723.89	145			

According to Eq. (7), the exponential size of each image was calculated. Table 1 shows the index list for the sensitive bands using ABS. Clearly the 402, 446, 556, 636, 699, and 706 nm images have the larger index and are more representative.

(a)Group D   (b)Group B   (c)Group C   (d)qroup A
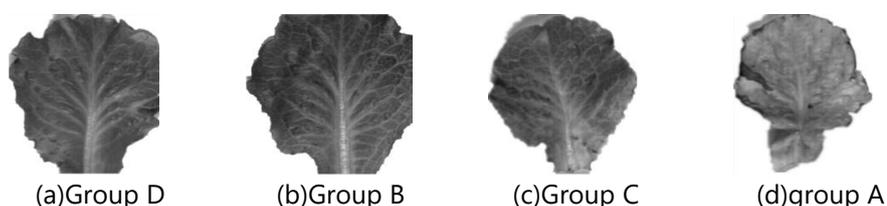
Figure 3 (a), (b), (c), (d) are the characteristics of lettuce leaves at 556 nm under different nitrogen stresses

Expression and quantification of nitrogen features

Figure 3 shows the lettuce leaves at four nitrogen levels and the characteristic wavelength of 556 nm. Clearly, the gray scale gradient is more obvious with the increase of nitrogen level, and the texture of veins is also different, but the gray value of the point on the image corresponds to the spectral intensity of the point at the specific wavelength. Therefore, the gray value (or reflection intensity) of the feature image and the texture feature can be extracted to characterize the nitrogen content of the lettuce sample.

At a particular wavelength, the average gray level of the image can represent the intensity distribution of the lettuce sample area. Thus, the average gray level (AG) on the feature image can be used to characterize the nutrients s and extrinsic of lettuce leaves. The gray mean feature of lettuce leaf images at sensitive wavelength, AG, is calculated as follows:

$$AG = \frac{1}{N}\sum_{i=1}^{N} f_i(x,y) \qquad (10)$$

Where: N is the number of pixel points in the leaf region in the whole image (i=1,2… , N); f(x,y) represents the gray value of each pixel point (x,y).

Leaf veins and other tissue texture structures of lettuce leaves with different nitrogen levels may differ to some extent. Therefore, texture features can be used to invert lettuce nitrogen. However, canopy-scale crop growth monitoring can be hardly conducted directly under the greenhouse conditions, and thus cannot be directly applied to the later automatic monitoring equipment as an effective feature. Therefore, we only extracted and quantified the grayscale features of lettuce nitrogen images, which was used as a hyperspectral image feature vector for modeling and diagnosis.

Feature Compensation Method

Previous studies on crop nitrogen diagnosis based on hyperspectral images usually only used the nitrogen characteristic images acquired in the visible spectrum between 400 and 700nm to conduct modeling and analysis, but did not consider the influence of crop water difference on the nitrogen hyperspectral characteristic images. Under different water contents, the nitrogen characteristic images of the same spectral segment will differ, and the difference changes with the specific spectral segment. In this study, samples with different water contents were taken from group C under the same nitrogen level. Under the water content conditions of 71%, 86% and 91%, the hyperspectral image grayscale of leaves with the same nitrogen level at 556 nm nitrogen are different to some degrees (Figure 4), which is also an important factor causing model errors and affecting prediction accuracy.

(a)91%   (b)86%   (c)71%
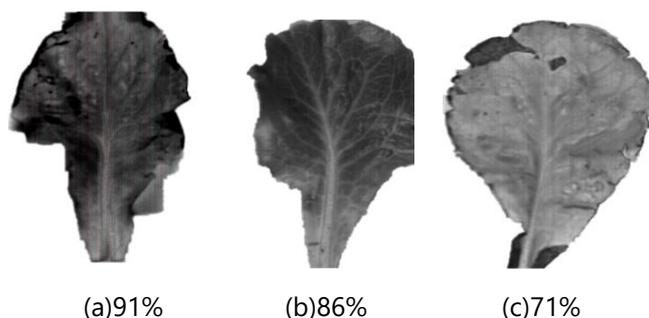
Figure 4 The characteristic diagrams of group C lettuce leaves at different moisture contents at 556 nm are shown in (a), (b), (c), respectively

To overcome the influence of water stress on nitrogen characteristic images, we obtained near-infrared hyperspectral images of lettuce leaves from 390.8 to 1766.3 nm simultaneously, and extracted hyperspectral water images of lettuce at 775, 960 and 1420 nm by the method described in section 2.1.2. Based on the water content characteristics of crop nitrogen images in different spectra, the water content response models of lettuce leaves in different spectra were established to compensate the nitrogen contents of lettuce. Specifically, partial least squares regression on the gray (strength) variable of the moisture content image and the measured moisture content W[21, 22] was firstly performed to establish the prediction model of lettuce moisture content:

$$W = 65.09 + 43.82AG_{775} + 12.65AG_{960} - 117.72AG_{1420} \qquad (11)$$

where: AG_{775}, AG_{960}, AG_{1420} represent the mean gray values of lettuce leaf images at sensitive wavelengths of 775, 960 and 1420 nm, respectively.

The correlation coefficient of the lettuce water content prediction model is 0.906, and the mean square error is 1.02. Due to the existence of errors, if a modified model is established by directly using the samples of the predicted moisture content and nitrogen content at the same level; and if linear error compensation is carried out directly, the compensation effect will be unsatisfactory. Therefore, the moisture content of samples was predicted by the detection model, and the nitrogen features were compensated based on the difference at reflection response of different water content levels. In this case of estimated water content the variation rate $\triangle Wi$ of the nitrogen hyperspectral image characteristic variable AG_i(i=1,2...6)with the water content level can be determined by combining the total nitrogen content obtained by AA3 chemical detection and the reflectance value at the characteristic wavelength of nitrogen. Accordingly, the correction coefficient ΔAG_i of the nitrogen hyperspectral feature variable AG_i at different water content can be calculated. AG_i was corrected according to equation (12) (Table 1).

Table2. Hyperspectral characteristic variables of lettuce nitrogen in moisture content rate of $\triangle Wi$ and correction coefficient of $\triangle AG_i$

Feature image	Rate of change	$\triangle AG_i$

(nm)	($\triangle Wi$)	$W \geq 80\%$	$70\% \leq W \leq 80\%$	$60\% \leq W \leq 70\%$	$W \leq 60\%$
402	2.08×10-1	3.18×10-3	-4.95×10-2	-5.72×10-2	-8.06×10-2
446	3.74×10-1	1.85×10-3	-3.66×10-2	-4.05×10-2	-7.29×10-2
556	8.52×10-1	6.76×10-2	-7.13×10-2	-8.05×10-2	-5.97×10-2
636	11.25×10-1	8.28×10-2	-9.71×10-2	-4.38×10-2	-8.58×10-2
699	9.76×10-1	2.55×10-2	-9.64×10-2	-8.75×10-2	-8.07×10-2
706	9.55×10-1	2.69×10-2	-8.87×10-2	-8.49×10-2	-8.52×10-2

Based on the hyperspectral features of lettuce nitrogen obtained from different spectral segments, 24 of 48 rosette lettuce samples were used to detect nitrogen by PLSR.

$$N = 23.39 + 6.14AG_{402} + 25.66AG_{446} - 31.52AG_{556}$$

$$+ 66.85AG_{636} + 45.65AG_{699} - 56.76AG_{706} \qquad (12)$$

The remaining 24 samples were used for prediction analysis, Results showed that the correlation coefficient of the model without correction of moisture content error was only 0.88 and the root-mean-square error (RMSE) was 1.03; In comparison, the correlation of the prediction model reached 0.92 and the RMSE was 0.46, which greatly improved the model accuracy.
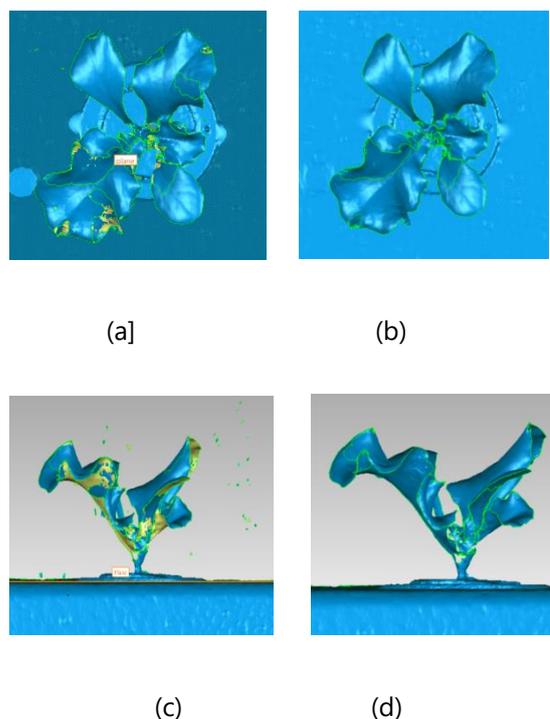
Extraction and analysis of nitrogen morphological characteristics of lettuce

Three-dimensional scan data preprocessing

The model acquired by 3D laser scanning will generate more noise due to the reflection interference of objects around the lettuce sample. However, the uneven structure of lettuce surface can cause problems such as scanning vulnerability, so no accurate scanning model can be built. For this reason, the model was repaired to be an ideal lettuce model on the reverse engineering software Geomagic .

During the data repair, the obtained lettuce 3D data was first imported into the Geomagic qualifier to convert the lettuce model composed of triangles into a point cloud, and redundant noise points were eliminated by the software. Then the 3D point cloud was transformed into a triangular surface model by encapsulation. After that, the parts of lettuce surface with holes were filled. Finally, the lettuce model was smoothed.

Figure 5(a) and (c) are the original top view and dominant view of 3D data, respectively, while Figure 5(b) and (d) are the 3D data after the model repair. After the data repair, noise was eliminated and the 3D data was kept continuous. The restored lettuce 3D model was transformed into a point cloud model and the file was saved in the format of igs. The digital design module of CATIA was used to save the imported point cloud in the format of ASC, which contains the coordinate data of the point cloud.

(a)   (b)

(c)   (d)

Figure 5. (a), (b) are lettuce top view 3D scan data, (c), (d) are lettuce main view 3D scan data

Extraction of nitrogen morphological features of lettuce

The difficulty in calculating lettuce volume (which can be converted into biomass), leaf area, plant height and stem diameter lies in the irregular morphology of lettuce. The lettuce shape based on 3D laser scanning is composed of a series of point cloud data. Therefore, we extracted the growth features of lettuce, (e.g. volume, leaf area, plant height and stem diameter), by establishing a spatial geometric model of the point cloud data.

(1) Volume calculation

N tiny cubes were used to fit the lettuce shape. Specifically, lettuce is composed of n cubes in size of *a* to the third power, and its volume is the sum of the effective cubes (Figure 6). This algorithm does not consider the lettuce shape, but only needs to judge the size and effectiveness of each cube element. The valid cubs represent the solid volume part of the lettuce, while the invalid cubes are the part outside the volume. At the same time, 3D point cloud data were segmented and projected onto the normal plane perpendicular to the plant height and transformed into 2D data points. The method is as follows:

① The lettuce point cloud data is equally divided in the height direction (Z axis) in step lengths of *a* (step length *a* is much smaller than the leaf thickness) to form *n*-layer lettuce segments. When *a* tends to be infinite and *n* tends to infinite, the lettuce volume can be considered to as an irregular pattern of n-layer bottom area S_k and height *a*;

② The cross-sectional area S_k of each layer of lettuce is calculated. The point cloud data of each layer is projected onto the X-Y plane perpendicular to the height direction, and equidistantly divided in steps of *a* along the X and Y axes to generate i by j pixel cells. The pixel cells are judged one by one according to the point cloud data projected by each piece of lettuce into each pixel cell. When the pixel cloud contains the point cloud of the lettuce projection, it is a valid pixel and is marked as 1; otherwise, it is marked as 0. The number M of effective pixel cells is counted, and the cross-sectional area S_k of the lettuce is calculated as the product of M and the area of the unit pixel. The volume *V* and the S_k of lettuce are calculated as follows:

$$S_k = aaM \qquad (13)$$

$$V = \sum_{i=1}^{n} S_k = \sum_{i=1}^{n} aaM \qquad (14)$$

where: *a* is the step size.

Based on the calculated raw vegetable volume and the measured fresh weight of lettuce, a biomass detection model based on 3D scan data was established:

$B_m=0.13+0.91V$ (15)

where: B_m is the biomass of lettuce. The correlation coefficient of the model is 0.98, and the RMSE is 0.26. An accurate inversion of biomass features based 3D morphology can be realized by using the lettuce plant volume and the biomass model.
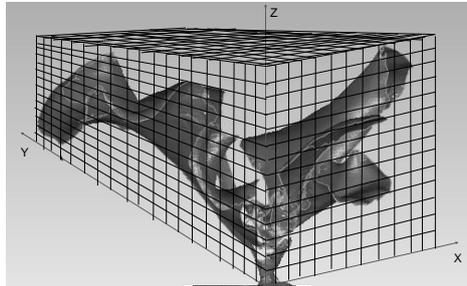
Figure 6. The 3D grid model of lettuce

(2) Leaf area calculation

The point cloud data is interpolated to form an irregular triangular grid, the area S_i of each triangle is calculated, and the sum of them can represent the leaf area S_c. The lettuce leaf area is computed as follows:

$B_m=0.13+0.91V$ (16)

(3)Plant height calculation

Let the coordinates of any point of the cloud data be f(x, y, z), and we only need to find the maximum value z_{max} and minimum value z_{min} along the Z-axis direction. The coordinates of z_{max} and z_{min} are f(x1, y1, z1); and f(x2, y2, z2) respectively. The plant height P_h can be obtained by calculating the distance between the two as follows:

$P_h=z_{max}-z_{min}=z_1-z_2$ (17)

(4) Stem diameter calculation

Starting from the bottom of the planting basket, the section of lettuce stem was cut off every 3.3mm along the height direction, and 3 cross sections were cut off. The diameter of each cross section was calculated, and the average value was calculated to determine the stem thickness of lettuce. The cross-section image of lettuce stem consists of a layer of approximately circular point clouds. The maximum x_{max} and minimum x_{min} along the X axis and the maximum y_{max} and minimum y_{min} along the Y axis are calculated in the x-y plane. The diameter of this cross-section is calculated. The stem diameter of lettuce, L_a, is calculated as follows (Figure 7):

$L_a=\sum_{i=1}^{3}[(x_{i\ max}-x_{i\ min})+(y_{i\ max}-y_{i\ min})]/6$ (18)

where:x_{imax} and x_{imin} are the maximum and minimum values along the X-axis in the cross-section of the i-th layer (i=1, 2, 3); y_{imax} and y_{imin} are the maximum and minimum values along the Y-axis on the i-th cross-section.
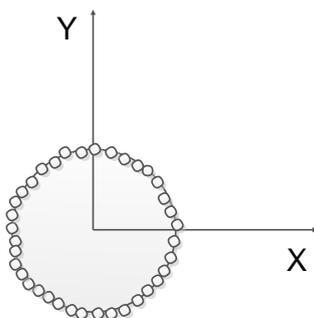
**Figure 7. Stem diameter coordinates of lettuce**

The volume, leaf area, plant height and stem thickness of lettuce can be calculated by importing the point cloud data in format of ASC into Matlab through programming.

**Morphological characteristics analysis of nitrogen in lettuce**

The leaf area of lettuce not only reflects the growth, but is closely related to the final yield. Figure 8 shows the morphological features of lettuce over time at different nitrogen levels. With the rise of nitrogen level, leaf area, plant height, stem diameter and biomass of groups A-C all increased to some extent. Among them, the standard formula group C had the largest eigenvalue increment, and, this trend became more apparent over time. However, the features value of the 200% excessive nitrogen group did not increase significantly with the increase of nitrogen application rate, but was lower than group C. These results indicate that deficient nitrogen and excessive nitrogen will both inhibit lettuce growth to some extent, and that the use of lettuce features at different growth stages can be used as an effective feature to estimate its nitrogen levels.
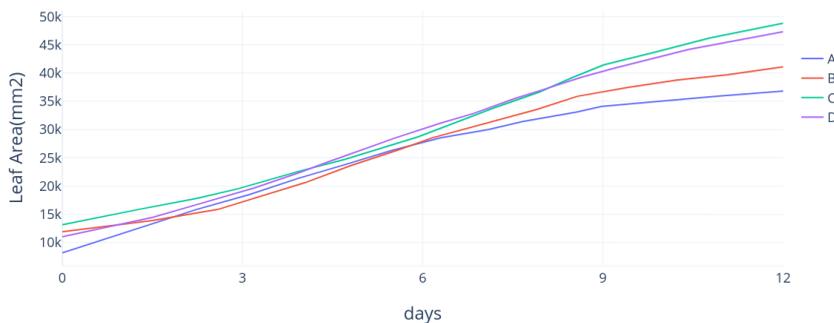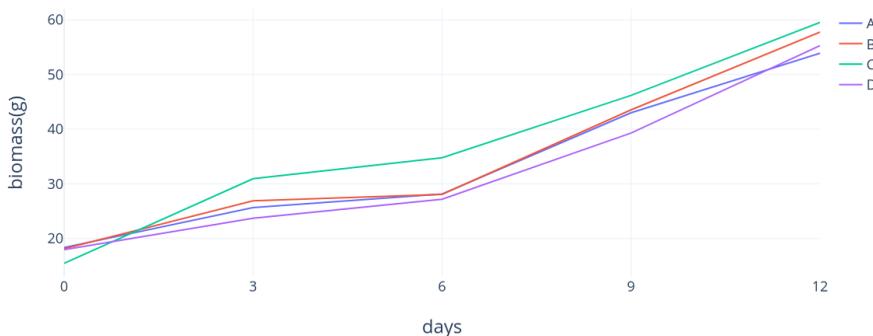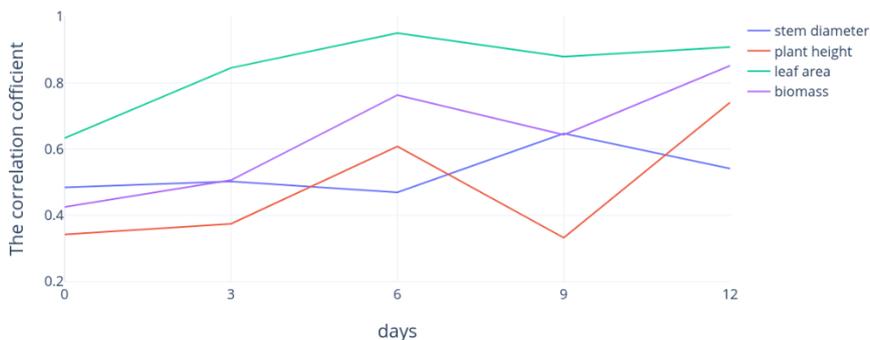

(a)
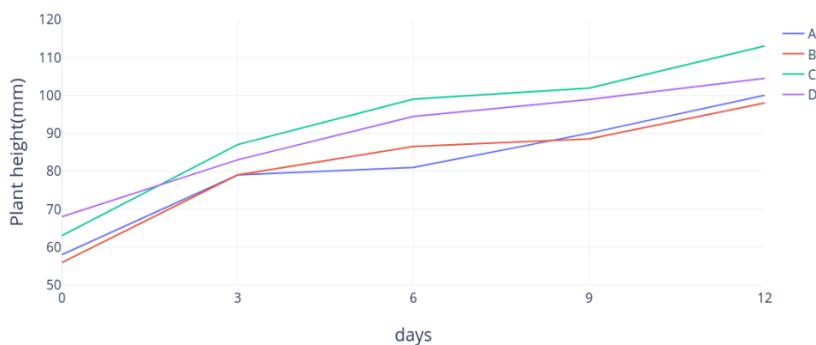
(b)
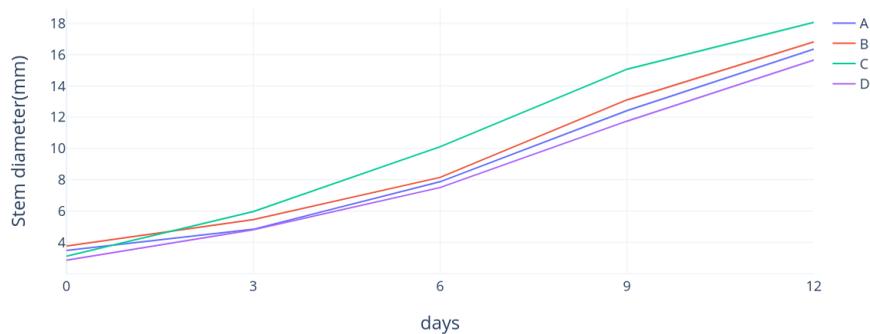
(c)

(d)

Figure 8. Morphological features of lettuce rosette stage in different nitrogen stresses are shown in (a), (b), (c), (d)

Correlation analysis between growth features and nitrogen in different growth stages

Based on the features of lettuce stem size, plant height, leaf area and biomass obtained at different growth stages, the correlation analysis was conducted with the nitrogen content of lettuce at different stages (the measured value of SPAD was substituted for that of AA3).
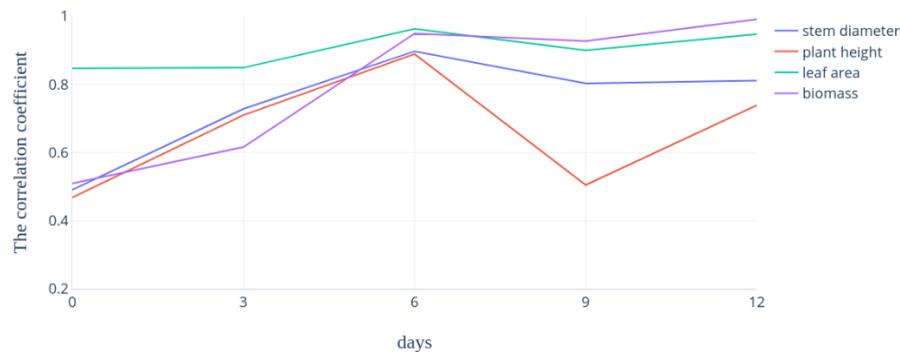

Figure 9. Correlation analysis of lettuce rosette stage in growing and nitrogen

The growth features of lettuce at the early rosette stage are less correlated with nitrogen (Figure 9). As the plant grows, the correlation between plant growth and nitrogen is gradually increasing. The growth features and nitrogen features both maximize at the rosette stage. Among them, nitrogen stress is most correlated with plant

biomass and leaf area, but less correlated with stem diameter. The main reason is that the changes of stem thickness are not obvious in the middle and late stages of plant growth.

The sample of D, has a certain negative correlation with the morphological features in the middle and late stages. Thus, to obtain better model features, we analyzed the correlation between growth features and lettuce nitrogen after the removal of group D.

Figure 10. Correlation analysis of lettuce rosette stage in growing and nitrogen (excluding the group D)

The correlation between crop growth and nitrogen is intensified significantly (Figure 10). It is indicated that the negative correlation between the late D-group and the growth features (or negative correlation with certain growth features) leads to a decrease in the correlation of the growth features in the overall correlation analysis. The main reason is that there is a certain degree of feature overlap between group D and the stress samples.

Based on the growth features of lettuce nitrogen, a nitrogen detection model was established by using 24 of the 48 rosette lettuce samples. The growth feature detection model including group D is:

$$N = 13.26 - 0.24L_a + 0.15P_h + 7.1 \times 10^{-6}S_c + 0.03B_m \qquad (19)$$

The correlation coefficient of the model is 0.90 and the RMSE is 0.87.

The growth feature detection model excluding the group D is:

$$N = 13.24 - 0.54L_a + 0.13P_h + 1.02 \times 10^{-4}S_c + 0.06B_m \qquad (20)$$

The correlation coefficient is 0.93 and the RMSE is 0.62.

Results show that crop growth features can be used for diagnosis of lettuce nitrogen. However, because nitrogen is only one of the main influence factors on crop nitrogen, the growth rate-based nitrogen diagnosis effect is poor in the early growth stage, and when nitrogen is excessive, the growth rate features are negatively correlated to some extent due to overlapping features. Therefore, it is difficult to achieve the desired nitrogen diagnosis effect by relying on a single growth feature.

Establishment of multi - characteristic nitrogen fusion model

The above analyses show that the hyperspectral image features and growth features of lettuce nitrogen are moderately correlated with plant nitrogen. However, the model involving a single feature is inaccurate, and cannot be directly applied to online monitoring of lettuce nitrogen. To make fully use the complementary advantages of multiple features and realize high-precision online nitrogen monitoring of lettuce, we used PLSR [22] for information fusion and establish a multi-feature lettuce nitrogen detection model based on the acquired hyperspectral image features of 402, 446, 556, 636, 699 and 706 nm as well as the growth features of stem diameter, plant height, leaf area and biomass.

In the multi-information fusion, the commonly-used modeling methods include multiple linear regression (MLR), PLSR and self-learning black box models based on back-propagation neural network and support vector machine. Generally, the accuracy of black box models is greatly improved in comparison with the explicit model.

In the construction of the online monitoring system, however, due to a large number of training samples and the selectivity difference of parameters such as kernel function, the real-time performance of the model is poor and uncertain, which complicates the system implementation. Therefore, we established an explicit model to facilitate the later construction of the system.

It is hoped that the significance of the regression equation can be maintained while the goodness of fit is improved. PLSR is the integration of multiple MLR, canonical correlation analysis and principal component analysis. The extracted components not only summarize the information of the feature variable system, but also well explain the nitrogen of lettuce, and reduce the noise interference in the system. Therefore, we used PLSR to establish a lettuce nitrogen prediction model.

Due to the data disparity between hyperspectral images and growth features variables, and to improve the balance and convergence speed of the model and eliminate the model errors caused therefrom, two types of feature variables are normalized as follows [23]:

$$x_i^{'} = (x_i - x_{min})/(x_{max} - x_{min}) \qquad (21)$$

where: x is the eigenvalue of eigenvector; i is the feature sequence (i=1,2,3...); x_{min}, x_{max} are the minimum and maximum of the eigenvalues of samples in the eigenvector respectively.

$N = -4.72 + 12.34AG_{402} - 8.52AG_{446} + 34.71AG_{556}$

$-26.73AG_{636} + 10.94AG_{699} - 15.62AG_{706} \qquad (22)$

Results show that the correlation coefficient between the predicted value and the measured value is 0.97, and the RMSE is 0.39. The prediction accuracy of the model combining the hyperspectral image and growth features is improved significantly.

The above model involving six hyperspectral image features and four growth features can better predict lettuce nitrogen in the rosette period. However, in addition to a reduction of computational efficiency and real-time performance, the involvement of too many feature variables also greatly increases the complexity and cost of the crop growth information line monitoring equipment, and decreases the reliability due to the overly complex detection links. To construct low-cost growth monitoring equipment while meeting the accuracy requirements, we used the genetic algorithm (GA) to further optimize the image and growth features of the rosette lettuce based on the above features.

GA performs self-learning and optimization by simulating the principle of survival of the fittest and survival of the fittest in natural evolution, and has certain advantages in solving complex optimization problems such as large space and nonlinearity. Therefore, we combine GA and PLSR to optimize the combination of features with higher prediction accuracy, and thereby established an optimal prediction model for lettuce nitrogen.

GA adopts a binary coding scheme, and the string length is the number of features, which is 10. When the i-th bit is 1, the i-th feature is selected, and otherwise the feature is masked. Thus, each individual represents a different subset of features, a candidate solution. The selected population size is 50, the crossover probability P_c is 0.9, the mutation probability P_m is 0.1, and the termination condition is that the genetic algebra reaches 50.

The goal of feature subset selection is to use fewer features to achieve the same or better modeling effect. Therefore, the fitness evaluation consists of two parts: prediction accuracy and the number of features used. 1.Firstly the selected feature subset is subjected to PLSR. Then, the test samples are used to predict, and the correlation coefficient R between the predicted value and the measured value is determined as the prediction accuracy. 2. Each feature subset contains a certain number of features. If the prediction precision of two feature subsets is the same, then the subset containing fewer features is selected. Among these two parts of prediction accuracy should be considered first. Therefore, the fitness function is determined as follows:

$$fitness = R \times 100 + k \times d \qquad (23)$$

where: *R* is the model prediction accuracy, *d* is the number of features in the selected feature subset. Here, the weight of the prediction accuracy is set to be 100 to increase its importance. A high prediction accuracy means the *fitness* is large; and this feature has the opportunity to win in the competition. The optimal feature

combination obtained by the operation is 0111011001, *fitness*=94.18. The selected features are the gray mean value variables, leaf area and biomass growth feature variables of the feature images at 446, 556, 636 and 706 nm. PLSR is performed on the selected feature variables to obtain the PLSR model:

$$N = -18.51 - 23.75 AG_{446} + 37.28 AG_{556} - 19.23 AG_{636} - 9.57 AG_{706} + 15.32 L_a + 29.65 B_m \qquad (24)$$
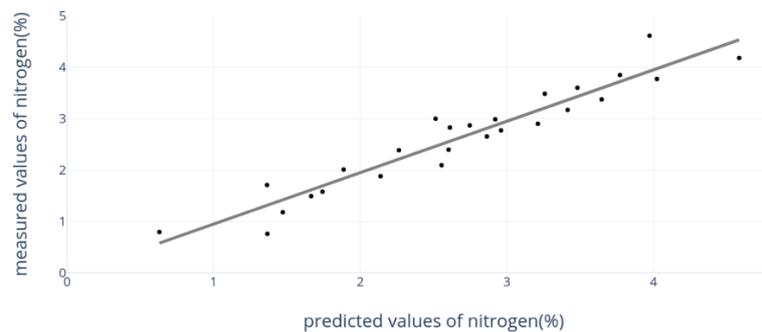
Figure 11. Lettuce nitrogen of predicted values and measured values in rosette stage

The model is verified and the predicted results are shown in Figure 11. Results show the correlation coefficient between the predicted and measured values of nitrogen is 0.95, and the RMSE is 0.57. The prediction accuracy of the model is slightly lower than the original PLSR model involving 10 features. Nevertheless, as the number of features is reduced by 40%, it is beneficial to improving the operating efficiency of the system, and effectively reduce the research and development cost of the crop growth information monitoring equipment.

Conclusion

Visible-near-infrared high-spectral imaging was combined with 3D laser scanning for rapid and non-destructive detection of nitrogen in lettuce, and the characteristic images of nitrogen as well as the response rules of growth parameters of rosette lettuce under different nitrogen conditions were acquired by establishing a spatial geometric mode.

(1) The 480 nm band image was segmented by a 157 gray scale threshold to obtain the target image sequence. Based on the 3D laser scanning system, the 3D morphological features of lettuce with different nitrogen levels were obtained. Morphological parameters such as leaf area, stem diameter, plant height and biomass were extracted.

(2) To overcome the influence of the interaction between lettuce nitrogen and water on nitrogen detection, we firstly used the 775, 960, 1420 nm lettuce moisture content model to estimate the moisture content of nitrogen samples. The nitrogen features were modified by using different nitrogen image feature variables with the water content change rate, which improved the detection accuracy of the hyperspectral image model of lettuce nitrogen.

(3) Genetic algorithm (GA) and partial least squares regression (PLSR) were used to establish the multi-feature fusion detection model of lettuce nitrogen. The accuracy of the model reached 0.95. The extracted features and model can provide **references** for the development of crop growth information systems.

Under natural conditions, the accurate acquisition of crop nutrition and growth information is the key problem that restricts the acquisition and application of crop growth information. Although this study uses multi-feature information fusion foe quick detection of crop nitrogen, it is still based on the laboratory environment. Therefore, how to accurately obtain the nutrition and growth features of crops under natural conditions; and how to build and develop a low-cost online monitoring system for crop growth information are the focus of the next step.

References

1. Bei MR，Luo XH，Yang HZ. Simultaneous determination of nitrogen，phosphorus and potassium in rubber leaf samples by AA3 continuous flow analyzer (CFA) [J]. Chinese Journal of Tropical Crops, 2011, 32(7):1258-1264.(in Chinese). DOI:10.3969/j.issn.1000-2561.2011.07.015

2. Gu Q, Deng JS, Lu C, Shi YY, Wang K, Shen ZQ. Diagnosis of Rice Nitrogen Nutrition Based on Spectral and Shape Characteristics of Scanning Leaves[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 43(8):170- 174. (In Chinese). DOI:10. 6041/j. issn. 1000-1298. 2012. 08. 031

3. Giacomo D, Stefania D. A multivariate regression model for detection of fumonisins content in maize from near infrared spectra[J].Food Chemistry,2013, 141(4):4289-4294.DOI:10.1016/j.foodchem.2013.07.021

4. Hoyos-Villegas V, Fritschi FB, 2013. Relationships among vegetation indices derives from aerial photographs and soybean growth and yield[J]. Crop Science, 53(6): 2631-2642. DOI: 10.2135/cropsci2013.02.0126

5. Hao Y. Chen B. Quantitative determination of low amino acid contents in tea by using near-infrared spectroscopy [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(6): 216- 220. DOI: 10.6041 /j.issn.1000-1298.2014.06.033

6. Kamruzzaman M, Sun DW, EIMasry G. Fast detection and visualization of minced lamb meat adulteration using NIR hyperspectral imaging and multivariate image analysis[J]. Talanta, 2014, 21(103): 130-136.

7. Knox NM, Skidmore AK. Remote sensing of forage nutrients: combining ecological and spectral absorption feature data ［J］.ISPRS Journal of Photogrammetry and Remote Sensing, 2012 72: 27- 35.

8. Lorente D, Aleixos N, GmezSanchis J. Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment[J]. Food Bioprocess Technol, 2015, 5(4): 1121-1142. DOI: 10.1007/s11947-011-0725-1

9. Li F. The principle of regression analysis method and the actual operation of SPSS [M]. Beijing: China Financial Publishing House, 2014(3).

10. Li XW, Lu X, Zhang Z, Chen J, Shi HG ,Tian M. Diagnosis of Nutriton and Recommended Model of Topdressing For Cotton ［J］.Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12) :209- 214. DOI: 10.6041 /j.issn.1000-1298.2014.12.031

11. Nativ R, Zeev S, Y Cohen, Victor A, Ran E, Timea I, Clara Shendereya, Arnon D, Uri Y. Estimating olive leaf nitrogen concentration using visible and near-infrared spectral reflectance [J]. Bio systems Engineering, 2012, 114(2013):426-434. DOI: 10.1016/j.biosystemseng.2012.09.005

12. [Ouyang AG, Xie XQ, Liu YD. Selection of NIR variables for online detecting soluble solids content of apple ［J. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(4): 220-225. DOI: 10.6041 /j.issn.1000-1298.2014.04.035

13. Pu RL, Gong P. Hyperspectral Remote Sensing and Its Application[M]. Beijing: Higher Education Press, 2000(8).

14. Qin JW, Thomas FB, Mark AR. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence[J]. Journal of Food Engineering, 2009, 93(2):183-191.DOI: 10.1016/j.jfoodeng.2009.01.014

15. Ramoelo A, Skidmore AK, Schlerf M. Savanna grass nitrogen to phosphorous ratio estimation using field spectroscopy and the potential for estimation with imaging spectroscopy[J]. Journal of Applied Earth Observation and Geoinformation, 2013, 23:334 -343. DOI: 10.1016/j.jag.2012.10.009

16. Shetty N, Gislum R. Quantification of frusta concentration in grasses using NIR spectroscopy and PLSR[J]. Field Crops Research, 2011, 120(1): 31-37. DOI: 10.1016/j.fcr.2010.08.008

17. Su JM, Fu RH, Zhou JB. Practical Guide to Statistical Software SPSS for Windows [M]. Beijing: Publishing House of Electronics Industry, 2000(6).

18. Tong QX. Analysis of typical spectral features and their characteristics in China[M]. Beijing: Science Press, 1990(8).

19. Wang ZG. Nutrition and Quality of Vegetables [M]. Beijing: Science Press, 2009: 36-37.

20. WANG YY, HE P, WEI T, LI SS. A Research of an Infrared Image Segmentation Algorithm Based on the Two-dimensional Entropy[J]. Journal of Air Force Engineering University (Natural Science Edition), 2015, 16(1):77-80. DOI:10.3969/j.issn.1009-2015.01.017

21. Yang Wei, Li MZ, Sun H, Zheng LH. De-noising Algorithm of Multispectral Images and Nonlinear Estimation of Nitrogen Content of Cucumber Leaves in Greenhouse[J]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(7) :216- 220. DOI: 10.6041 /j.issn.1000-1298.2013.07.038

22. Zhang ZA, Luo B. New approach for image retrieval based on color and spatial features [J]. Journal of xidian university, 2008, 35(2): 577-581.

23. Zhang XL, Liu F, He Y. Detecting macronutrients content and distribution in oilseed rape leaves based on hyperspectral imaging[J]. Biosystems Engineering, 2013, 15(1): 56-65.

Authors	Name	Bibliography	Photo
1st Author *Corresponding author	Bao Guo Shen	Zhenjiang Key Laboratory of UAV Application Technology, Jiangsu Aviation Technical College, Zhenjiang 212134, China; PHD candidate, Associate Professor, Research interest: detection of crop growth information. Address: 88 Ruicheng Road, Zhenjiang 212134, China. Tel.:+86-511-87052353, Email: shenbg@jatc.edu.cn	
2nd Author	Jin Yue Dai	School of Engineering, Jiangsu Aviation Technical College, Zhenjiang 212134, China;Post-Graduate, Professor, Research interest: machine design and manufactur.Address:88 Ruicheng Road, Zhenjiang 212134, China. Tel: +86-511-87056622, Email address: daijy@jatc.edu.cn.	
3rd Author	Xiao Dong Zhang	PhD, Professor, Research interest: hyperspectral and machine vision technology research facility in crop growth information, detection and applications. Address: 301 Xuefu Road, Zhenjiang 212013, China. Tel: +86-511-88797338, Email: zxd700227@126.com	
4th Author	Zhao Hui Duan	MSc student, engaged in the research of nutritional information detection in crops. 370404480@qq.com	