

SCIENTIFIC REPORTS



OPEN

Hyperspectral Imaging for Determining Pigment Contents in Cucumber Leaves in Response to Angular Leaf Spot Disease

Yan-Ru Zhao^{1,*}, Xiaoli Li^{1,*}, Ke-Qiang Yu², Fan Cheng³ & Yong He¹

Received: 15 November 2015

Accepted: 25 May 2016

Published: 10 June 2016

Hyperspectral imaging technique was employed to determine spatial distributions of chlorophyll (*Chl*), and carotenoid (*Car*) contents in cucumber leaves in response to angular leaf spot (ALS). Altogether, 196 hyperspectral images of cucumber leaves with five infection severities of ALS were captured by a hyperspectral imaging system in the range of 380–1,030 nm covering 512 wavebands. Mean spectrum were extracted from regions of interest (ROIs) in the hyperspectral images. Partial least square regression (PLSR) models were used to develop quantitative analysis between the spectra and the pigment contents measured by biochemical analyses. In addition, regression coefficients (RCs) in PLSR models were employed to select important wavelengths (IWs) for modelling. It was found that the PLSR models developed by the IWs provided the optimal measurement results with correlation coefficient (*R*) of prediction of 0.871 and 0.876 for *Chl* and *Car* contents, respectively. Finally, *Chl* and *Car* distributions in cucumber leaves with the ALS infection were mapped by applying the optimal models pixel-wise to the hyperspectral images. The results proved the feasibility of hyperspectral imaging for visualizing the pigment distributions in cucumber leaves in response to ALS.

Cucumber (*Cucumis sativus*) has always been regarded as a widely cultivated plant in the world. FAO reported that the cucumber fruits production of the world reached to 60,895,225 tons in 2012¹. However, cucumbers often suffer from the invasion of various diseases, which affect the yield and quality of the cucumber fruits. Angular leaf spot (ALS) caused by the bacterium *Pseudomonas syringae* pv. *lachrymans* may severely affect cucumber growth by foliar infections. The bacterium firstly causes small, water-soaked lesions on the abaxial leaf surface; then, these infected areas turn brown or straw-colored². Mesophyll tissue affected by bacterial infection may show increased respiration, disintegration or collapse of cells, and degradation of pigments such as chlorophyll (*Chl*) and carotenoid (*Car*)³. The bacterial infection not only causes the reduction of pigments, but also results in spatial heterogeneity of leaf pigments. *Chl* is the most important pigment in plants and its concentration controls photosynthetic potential and primary production. *Car* is the second major group of plant pigments, and its concentration provides much complementary information on vegetation physiological status⁴. *Chl* and *Car* contents have been successfully used for diagnosing physiological state of plants under disease stresses^{5–7}. However, no research endeavors have been reported yet for characterizing the spatial distribution of pigments within the leaves infected by bacterium. Therefore, developing efficient methods for quantitative analysis of pigment contents and exploring pigment distributions in space are significant for the diagnosis of mild diseases. Detection of abnormal spots with lower *Chl* and *Car* values in leaves will be especially meaningful for detecting mild ALS diseases.

Ultraviolet spectrophotometry and high performance liquid chromatography have been used to determine *Chl* and *Car* contents through a series of operations including weighting, grinding, extracting supernatant, and measuring^{8,9}. However, these methods involve labor-intensive, time-consuming and tedious extraction procedures. The chlorophyll meter (SPAD-502), a simple and portable diagnostic tool, can measure the greenness or relative *Chl* content by acquiring the absorbance of leaves in the red and near infrared (NIR) regions at individual

¹College of Biosystems Engineering and Food Science, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China. ²College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China. ³Institute of Biotechnology, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, China. ⁴These authors contributed equally to this work. Correspondence and requests for materials should be addressed to Y.H. (email: yhe@zju.edu.cn)

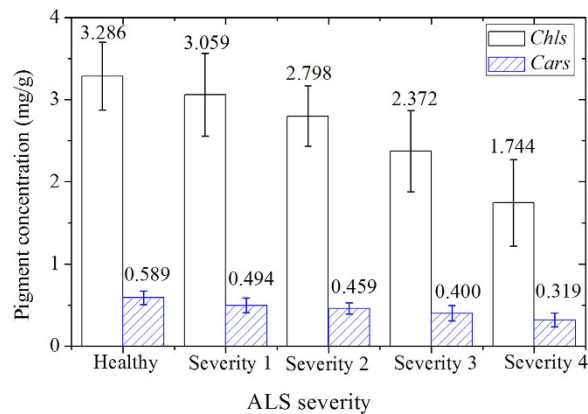


Figure 1. Mean measured *Chl*-BA and *Car*-BA contents in healthy leaves and ALS infected leaves. *Chl* and *Car* contents were determined by biochemical analysis (*Chl*-BA, *Car*-BA). Pigment “contents” values of the healthy and infected leaves were calculated and shown as 10 numbers to represent the variation trend of the pigments with the severity of the ALS disease.

points in a non-destructive way¹⁰. However, the chlorophyll meter method is an impractical tool for a large-scale real-time plants monitoring under field conditions. During the past few decades, spectroscopic techniques have been used extensively as powerful and effective analytical tools for estimating *Chl* and *Car* content in leaves^{11–13}. Previous literatures confirmed that spectral reflectance of vegetation in the visible region was primarily governed by pigments^{14–16}. Unfortunately, spectroscopic technique is often used to measure the averaged spectrum of the sample which fails to provide the spectral details at each pixel on gray-scale images of the targets. However, spectrum information with high spatial resolution is vital to detect the disease infection at early stage.

Hyperspectral imaging technique combines reflectance spectra with image processing. It has promoted recent progress in remote sensing and has been widely used for mapping the distribution of physiological characteristics in plants¹⁷. Yu *et al.*¹⁸ mapped spatial distribution of nitrogen in whole pepper plant with characteristic hyperspectral images. Shi *et al.*¹⁹ investigated chlorophyll concentration distribution to determine nitrogen deficiency in cucumber plants by using hyperspectral imaging combined with chemometrics. Zhang *et al.*²⁰ estimated nitrogen (N), phosphorus (P), and potassium (K) distributions in oilseed rape leaves by applying the optimal calibration models in each pixel of the reduced hyperspectral images. A majority of studies focusing on the disease detection rely on the spectral characteristics of the infected plants based on hyperspectral imaging^{21–23}. Hyperspectral imaging is capable of providing spectral details at each pixel in space, which makes it possible to map the pigment distributions in the samples. However, there is little research on mapping *Chl* and *Car* distributions of the plants in response to pathogen infection using hyperspectral imaging.

The aim of this work was to determine the *Chl* and *Car* distributions in the cucumber leaves with bacterial infection at five severities by using the hyperspectral imaging technique. The specific objectives were to: (1) extract important wavelengths (IW) from hyperspectral data for establishing the optimal quantitative relationships between spectra and pigment contents; (2) establish regression models for predicting *Chl* and *Car* contents and evaluate the performance of these quantitative analyses; (3) visualize *Chl* and *Car* concentration distributions in cucumber leaves in response to ALS based on the selected hyperspectral images.

Results and Discussion

Statistics of the pigments measured by biochemical analyses. A total of 194 cucumber leaves from healthy and infected plants with four severities of bacterial infection were collected for *Chl* and *Car* content biochemical analyses (*Chl*-BA, *Car*-BA). The statistical analysis results of the measured *Chl*-BA and *Car*-BA contents are shown in Fig. 1. The mean *Chl*-BA and *Car*-BA content were 3.286 mg/g and 0.589 mg/g for healthy plants, respectively. There was an obvious reduction trend of the averaged pigment contents along with the severity of disease, ranging from 3.059 mg/g to 1.744 mg/g for *Chl*-BA, and 0.494 mg/g to 0.319 mg/g for *Car*-BA. As expected, the mean values of *Chl*-BA and *Car*-BA contents displayed a descending tendency with the severities of the bacterial infection in leaves. The descending trends of *Chl*-BA and *Car*-BA content with disease stress in plants also confirmed previous results from the literature²⁴.

The SPXY algorithm²⁵ was implemented to separate all samples into a calibration set (130 samples) and a prediction set (64 samples) according to their differences in both x (spectra) and y (reference value) spaces. Figure 2 shows the statistical results of *Chl*-BA and *Car*-BA contents in calibration set and prediction set. There were evident variations of pigment contents in the calibration set, ranging from 0.944 mg/g to 4.503 mg/g for *Chl*-BA, and from 0.186 mg/g to 0.810 mg/g for *Car*-BA. The ranges of *Chl*-BA and *Car*-BA contents in the prediction set were from 1.352 mg/g to 3.857 mg/g and 0.252 mg/g to 0.701 mg/g, respectively. The calibration set has a lower mean value and a higher standard deviation (SD) than the prediction set. Generally, the calibration set with a wide concentration distribution of the chemical values is conducive to build a stable calibration model¹⁸.

Spectral features of leaves. The spectra of whole leaf were averaged as the representative spectrum of a sample, and 194 averaged spectra in the region of 380–1,030 nm were obtained altogether. The starting 18

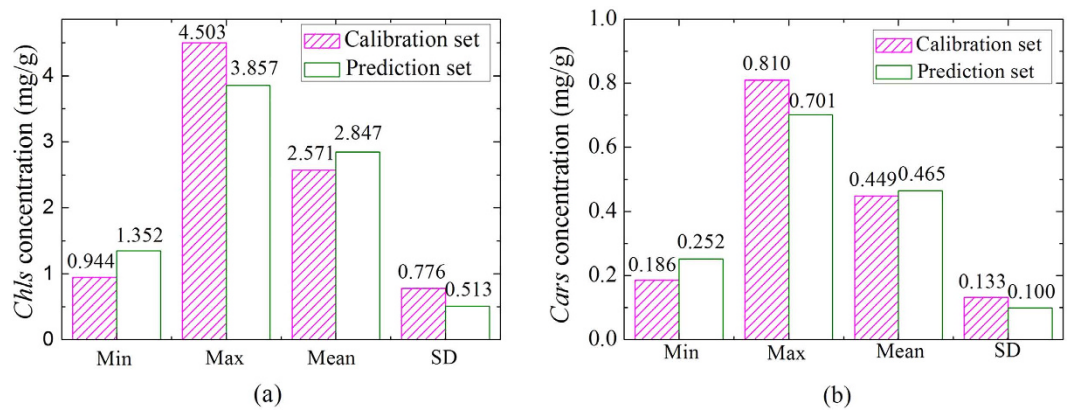


Figure 2. The measured *Chl*-BA (a) and *Car*-BA (b) content distributions in the calibration set and prediction set. Note: Min: Minimum; Max: Maximum; SD: Standard deviation. SPXY method was used to divide all samples into calibration set with 130 samples and prediction sets with 64 samples for modelling.

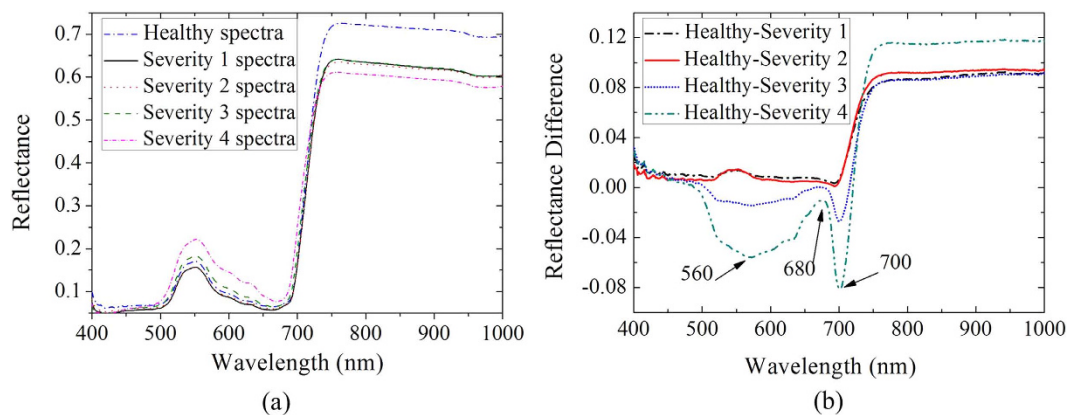


Figure 3. The average spectral reflectance curves of cucumber leaves. (a) Average spectral curves of cucumber leaves with five ALS disease severities; (b) Subtracted reflectance-differences between these average spectra. Reflectance of healthy samples and samples of four ALS disease severities ($n = 40, 39, 39, 37, 39$) were averaged respectively, resulting in five curves.

bands and ending 18 bands were removed to improve signal-noise-ratio (SNR), and wavebands in the range of 400–1,000 nm (476 wavelengths) were used for further processing.

The spectra of the samples at each bacterial infection severity were averaged to analyze spectra variations of all infected leaves and five spectra are shown in Fig. 3(a). All spectra had similar characteristics of the “green plants”^{26–30}. However, there were obvious reflectance value differences between the spectra of healthy and infected leaves in Fig. 3(a). In visible region (400–780 nm), infected leaves had higher reflectance values than healthy leaves, which was caused by the reduction of *Chl* content in the bacterial infected cucumber leaves²⁷. However, in NIR region (780–1,000 nm), reflectance values of infected leaves were lower than those of healthy leaves, which was mainly due to the internal structure damaged caused by the bacteria. Besides, red-edge swing descended and red edge position showed “blue movement” when the green plants suffered diseases stress³¹. Spectra showed an increasing trend in the visible region from severity 1 to severity 4 in infected leaves, which was consistent with the measured *Chl*-BA and *Car*-BA content variations shown in Fig. 1.

However, there were no evident difference in spectra among samples with severity 1, severity 2 and severity 3 ALS infections in Fig. 3(a), which also means that average spectrum of the whole leaf fails to provide spectral details of mild diseases. The obvious differences of these spectral curves were expressed through subtracting spectra of infected samples from healthy leaves spectra. The reflectance-differences (RDs) of the averaged spectra are displayed in Fig. 3(b). The spectra of the infected cucumber leaves were clearly distinguished by the RDs, especially around 560 nm, 680 nm, and 700 nm, which were closely related to the pigments of leaves³². The reflectance of pigments in visible region are the theoretical basis for establishing the mathematical correlation between spectral data and compositional contents.

Partial least squares regression models. As shown in Fig. 4(a), nine wavelengths are selected as the important wavelengths (IW_s) to predict the *Chl* using hyperspectral imaging (*Chl*-HSI) model. Among the selected wavelengths, seven wavelengths at 408, 484, 533, 582, 686, 708, and 750 nm were located in the visible

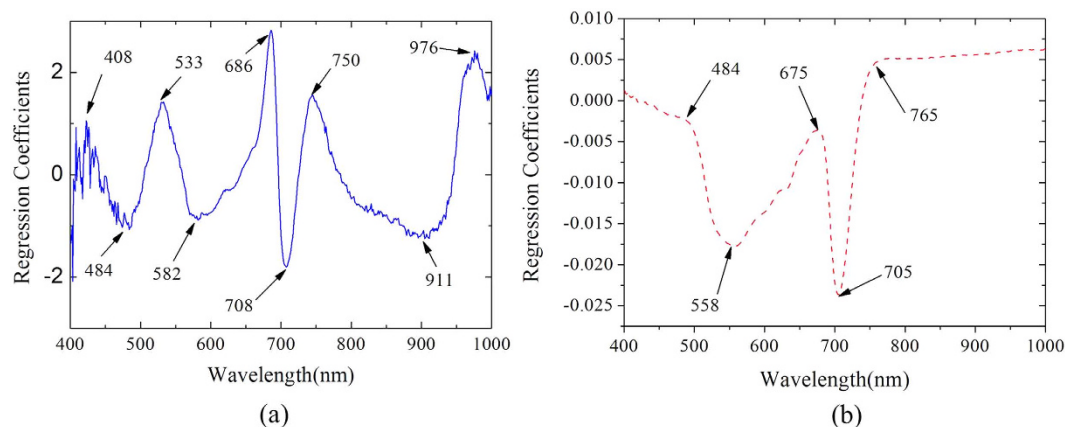


Figure 4. Regression coefficients (RCs) of each wavelength in PLSR based on the whole spectra. (a) RCs of each wavelength for *Chl*-HSI; (b) RCs of each wavelength for *Car*-HSI. The local extremum values of regression coefficient represent the importance of each wavelength. The selected important wavelengths which carried effective information were employed to establish the regression models to improve the speed of the modelling.

Pigments	Models	N. ^a /LVs ^b	Calibration		Validation		Prediction		Number
			Rc	RMSEC	Rcv	RMSECV	Rp	RMSEP	
<i>Chl</i>	W-PLSR ^c	476/7	0.913	0.315	0.894	0.347	0.87	0.255	Model 1
	RC-PLSR ^d	9/5	0.901	0.335	0.889	0.355	0.871	0.25	Model 2
<i>Car</i>	W-PLSR	476/2	0.796	0.08	0.787	0.081	0.886	0.049	Model 3
	RC-PLSR	5/1	0.803	0.079	0.793	0.08	0.876	0.05	Model 4

Table 1. Results of analysis by the PLSR models for estimating *Chl*-HSI and *Car*-HSI contents based on the IWs (RC-PLSR) and whole-wavelengths (W-PLSR). Note: N.^a: The number of the variables; LVs^b: Latent variables; W-PLSR^c: developed PLSR models based on whole wavelengths (476 variables); RC-PLSR^d: IWs selected by RCs were applied to the established PLSR models (9 and 5 variables).

region, which confirmed that pigments primarily governed the spectral reflectance of vegetation in the visible region¹⁶. In addition, 911 and 976 nm, which were close to the second overtones O-H stretching, were sensitive to water status of leaves. ALS destroys the cellular integrity of the cucumber tissue and may result in changes of the relative water content³³. Figure 4(b) shows the RCs of each wavelength and five wavelengths at 484, 558, 675, 705, and 765 nm were selected as the IWs for predicting *Car* by hyperspectral imaging (*Car*-HSI). There are some similar IWs for both *Chl* and *Car*, which explained by the similar optical properties of *Chl* and *Car*.

The PLSR algorithm was employed to establish regression models for evaluating *Chl*-HSI and *Car*-HSI contents in the cucumber leaves based on the selected IWs and whole spectra, respectively. The results of the PLSR models are enumerated in Table 1. The best performance of Model 1 and Model 3 were achieved with *Rc* of 0.913, *Rp* of 0.870, RMSEC of 0.315, and RMSEP of 0.255 for *Chl*-HSI content estimation, *Rc* of 0.796, *Rp* of 0.886, RMSEC of 0.080, and RMSEP of 0.049 for *Car*-HSI contents prediction. However, Model 1 and Model 3 with 476 variables involved in the modeling were time-consuming and failed to offer the simplified regression functions for predicting *Chl*-HSI and *Car*-HSI contents. Model 2 and Model 4 also provided similar results with *Rc* of 0.901, *Rp* of 0.871, RMSEC of 0.355, and RMSEP of 0.250 for *Chl*-HSI content, *Rc* of 0.803, *Rp* of 0.876, RMSEC of 0.079, and RMSEP of 0.050 for *Car*-HSI content. Although 98.32% (476 vs. 9) of the variables in Model 2 were eliminated, *Rc* of Model 2 only showed a slight reduction from 0.913 to 0.901, and *Rp* had a similar performance with an increase from 0.870 in Model 1 to 0.871. Similar results obtained in Model 4. Moreover, Model 4 with only five wavelengths improved the modeling speed. Therefore, Model 2 and Model 4 were promising to predict *Chl*-HSI and *Car*-HSI contents of the cucumber leaves.

From Models 2 and 4, the quantitative relationships between the spectral reflectance and pigments contents are shown in function (1) and (2):

$$\begin{aligned}
 Y_{chls} = & -0.511 + 0.419X_{408nm} - 0.148X_{484nm} - 2.313X_{533nm} \\
 & - 2.315X_{582nm} - 0.871X_{686nm} - 3.252X_{708nm} \\
 & + 2.347X_{750nm} + 2.616X_{911nm} + 2.670X_{976nm}
 \end{aligned} \quad (1)$$

$$Y_{cars} = 0.170 - 0.007X_{484nm} - 0.716X_{558nm} - 0.121X_{675nm} - 0.953X_{705nm} + 1.017X_{765nm} \quad (2)$$

where, *Y* is the predicted pigment content, *X*_{*i*nm} is the reflectance spectra at the selected IWs.

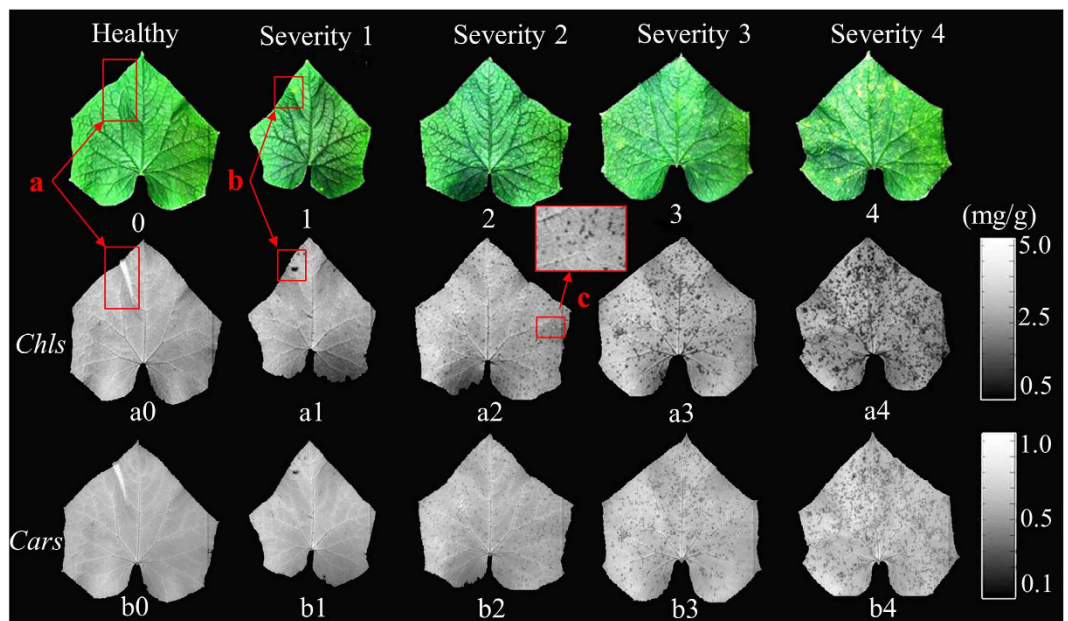


Figure 5. *Chl*-HSI and *Car*-HSI distributions in cucumber leaves with the ALS infection. *Chl*-HSI and *Car*-HSI of leaves in hyperspectral images were calculated based on the function (2) and (3), and the spatial distribution was executed by an image processing program. 0–4: RGB (R: 662 nm; G: 554 nm; B: 450 nm) images of healthy leaves and samples with disease severity 1, 2, 3, and 4, respectively; a0–a4: *Chl*-HSI content distributions of the samples; b0–b4: *Car*-HSI content visualization maps of the cucumber leaves; a: wrinkle part; b: disease spot; c: partially enlarged details of the disease spots.

As shown in Table 1, Model 2 and Model 4 were the optimal models to estimate *Chl*-HSI and *Car*-HSI contents. Therefore, the multi-linear functions (1) and (2) were capable to calculate the *Chl*-HSI and *Car*-HSI contents at each pixel point of the hyperspectral images of the samples.

Distribution maps of pigments in cucumber leaves. The optimal simplified models (functions (1) and (2)) were employed to estimate *Chl*-HSI and *Car*-HSI values at each pixel of the hyperspectral images. Then, *Chl*-HSI and *Car*-HSI distribution maps in the cucumber leaves were generated by a developed image processing program. Figure 5 shows the *Chl*-HSI and *Car*-HSI content distributions in the cucumber leaves in response to ALS. Two gray-scale color bars (0–255 levels) were generated with different *Chl*-HSI and *Car*-HSI values from small to large shown in different color from black (0 level) to white (255 level).

RGB images of the cucumber leaves with five severity stages are shown in Fig. 5(0–4). There were no obvious symptoms appeared on the upper side of the leaves in Fig. 5(0–2), therefore it was difficult to diagnose the bacterial early infection according to the RGB images. However, several yellow points appeared on the upper side of the cucumber leaves with the bacterial infection at severity 3 and severity 4, such as Fig. 5(3,4). The color difference between the infected area and the non-infected region was not evident. Overall, it was not easy to estimate the epidemic situation according to the RGB images of the samples. Figure 5(a0–a4, b0–b4) show the *Chl*-HSI and *Car*-HSI distribution maps in the cucumber leaves. More and more disease spots with low *Chl*-HSI and *Car*-HSI contents appeared with the increasing severity of the bacterial infection. *Chl*-HSI values of the whole healthy leaf were homogeneous, and there was no abnormal point in Fig. 5(a0) except for the (a) region, which was caused by the wrinkle part. The points with lower *Chl*-HSI content than other parts were clearly displayed at (b) point in Fig. 5(a1). However, this point is hard to be detected at (b) region in Fig. 5(1). Furthermore, several small disease spots are displayed in Fig. 5(a2), as illustrated by the partially enlarged details of the disease spots in (c) region. However, there was no obvious disease spots appeared in Fig. 5(2). It could be concluded that the disease spot with lower *Chl*-HSI content was easily detected in the *Chl*-HSI distribution maps at early stage. More disease spots with low *Chl*-HSI content are shown in Fig. 5(a3, a4). Meanwhile, Fig. 5(b0–b5) show the *Car*-HSI content distributions in the cucumber leaves. It could be found that *Car*-HSI content had similar variations with the *Chl*-HSI content in the cucumber leaves in response to the ALS infection. *Chl* and *Cars* are both light-harvesting pigments, but *Chl* is the most abundant and critical pigment for photosynthesis. Therefore, *Chl*-HSI distributions had better performance than *Car*-HSI distributions for detecting ALS severities.

Disease symptoms commonly reflect the interaction between plant and pathogen. Once bacterial pathogen successfully infects the leaves, a series of physiological changes would turn up during the process. Infected points firstly would appear water soaked symptom, pigments break would be then broken down and the infected area would become straw colored. Finally, the pathogen would cause wilting, browning, and death of leaves³. When the cucumber leaves were attacked by the *Pseudomonas syringae* pv. *lachrymans*, water soaked small disease spots firstly appeared on the underside of leaves. However, these small disease spots can't be easily detected from the upper surface by naked eyes. More disease spots appeared on the upper side of the leaves with the increasing of

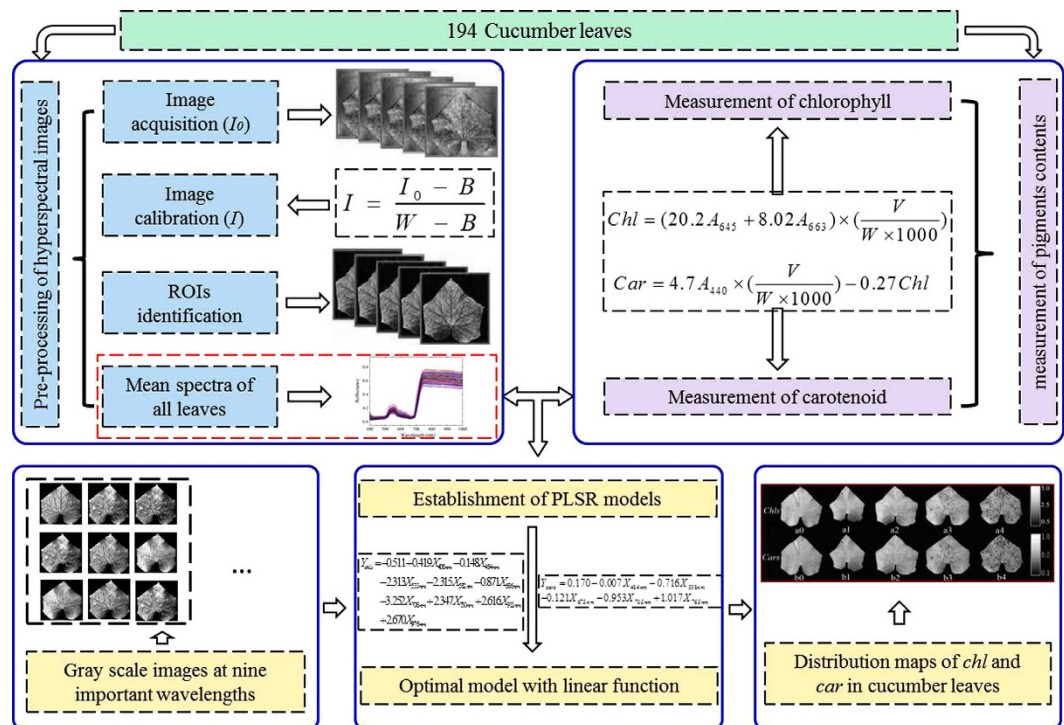


Figure 6. Main steps of determining of Chl-HSI and Car-HSI values in cucumber leaves infected with ALS by hyperspectral imaging. There were three main steps to map the Chl-HSI and Car-HSI spatial distribution. (1) After hyperspectral images acquisition and correction, reflectance values were extracted from the regions of interest (ROIs). (2) Then Chl and Car contents were determined by biochemical analysis (Chl-BA, Car-BA). Averaged spectra of samples and contents of Chl-BA, Car-BA were used to select the important wavelengths (IW) for developing the partial least square regression (PLSR) models. (3) The established PLSR models were used to calculate the Chl-HSI and Car-HSI values at each pixel on the hyperspectral images. Finally, Chl-HSI and Car-HSI spatial distribution in cucumber leaves infected with different severities of *Pseudomonas syringae* pv. *lachrymans* infection were displayed with the help of MATLAB software.

disease severity. The variation of pigments in infected cells is one of the most important characteristics associated with the diseases severity of plants. As shown in Fig. 5, the visualization of Chl and Car spatial distribution are helpful and meaningful to observe the ALS disease severities.

Besides the bacterial infection, pigment disorder of leaves also could be caused by fungal, viral infections, insect, and nutrient deficiencies. For example, nitrogen deficiency could cause a low chlorophyll density distribution on leaves. In this case, however, the symptom is usually stabilized, and would never spread over infected time. Also, some insects could cause significant damage by eating mesophyll³⁴. Nevertheless, the destroyed area always express as a worm eaten path or wormholes which contain no pigment. These characteristics of both stresses were mostly different from that of diseases. Leaf spot is regarded as one of the most common fungal and bacterial plant symptoms. Generally, shapes, sizes, and colors of leaf spots caused by various pathogens are different³⁵, and these physical features would change with the severity degree or incubation stage of infection. In addition, many microbial diseases would spread over a large area in groves and plantations through accidental introduction of vectors or through infected plant materials³⁶. The specifics of various symptoms change the spectral and spatial information of plants. Thereby, hyperspectral imaging has the advantage for real-time detection of plant disease at early stage by mapping the pigment spatial distribution or revealing the physical specificities of diseases. In future work, more kinds of diseases with different severities or infection incubation should be carried out to develop more adequate models on plants' pigment images production. More attention should be paid on the statistical analysis of physical specifics (shape, infection size and others) of various diseases in future investigation.

Conclusions

In summary, this study shows that the hyperspectral imaging coupled with chemometrics is an effective method to measure Chl-HSI and Car-HSI contents and to generate distribution maps of pigments in cucumber leaves with the ALS infection. The PLSR models with nine wavelengths at 408, 484, 533, 582, 686, 708, 750, 911, 976 nm and five wavelengths at 484, 558, 675, 705, and 765 nm could provide better accuracy than the whole-PLSR models for detecting Chl-HSI and Car-HSI contents of cucumber leaves. In addition, ALS disease spots were shown clearly in the Chl-HSI and Car-HSI distribution maps. Applying important wavelengths to detect mild disease from mapping pigment spatial distribution is critical for sensor's real time assessment.

Materials and Methods

Samples. A total of 100 cucumber plants (*Cucumis sativus* L, Jingyu, Tianjin, China) were grown in a greenhouse at Zhejiang University, Hangzhou, China. Twenty uninoculated plants kept as a control group with non-inoculated. *Pseudomonas syringae* pv. *lachrymans*, obtained from Institute of Biotechnology, was inoculated onto cucumber plants at 6-leaf stage as described in the literature³⁷. The samples were picked two weeks later after inoculation. The disease severity was determined by the severity and number of disease spots on a cucumber leaf by agriculture experts, and the following five ALS severity groups were defined:

- (1) non-diseased plants in which no symptoms were observed on the leaves (40 samples);
- (2) infected plants at severity 1 in which initial symptoms were seen with less than 2 disease spots on one leaf (39 samples);
- (3) infected plants at severity 2 in which symptoms were a little more severe than stage 1 with 2~4 disease spots on the leaf (39 samples);
- (4) infected plants at severity 3 with 5~9 disease spots on one leaf (37 samples);
- (5) infected plants at severity 4 with more than 10 disease spots on one leaf (39 samples).

In this study, a total of 194 samples with five different disease severities were picked in an hour. The tested samples were immediately enclosed into individual plastic bags to keep fresh in incubator with crushed ice around to maintain the constant temperature of 2–4 °C and transported to the laboratory within 2 hours for capturing hyperspectral images.

Equipment. The hyperspectral images of cucumber leaves were captured by a line-scanning hyperspectral imaging system (380–1030 nm) in reflectance mode. The system is composed of a spectrograph (ImSpectorV10, Spectral Imaging Ltd., Finland), a CCD camera (C8484-05G, Hamamatsu, Hamamatsu) coupled with a lens (OLES23; Specim, Spectral Imaging Ltd., Oulu, Finland), an illumination system with two 150W quartz tungsten halogen lamps (Fiber-Lite DC950 Illuminator, Dolan Jenner Industries Inc., USA) adjusted at an angle of 45° to illuminate the camera's field of view (FOV), a conveyor operated by a stepper motor (IRCP0076, Isuzu Optics Crop, Taiwan), and a computer with data acquisition and preprocessing software (Spectral-Cube data acquisition V10 software). The spectral resolution is 2.8 nm in 380–1030 nm, and CCD camera has a resolution of 672 × 512 pixels. The distance between samples and the lens was 485 mm, the speed of conveyor was set as 3.8 mm/s, and the exposure time was 0.09s during the image acquisition. To acquire a three-dimension hypercube, each sample was placed on a black background which had a very low reference. The upper side of the cucumber leaf was faced to the camera.

Image processing. Raw hyperspectral image (I_0) was calibrated by white (W) and dark (B) reference images. The white image was acquired from a standard Teflon tile (~99.9% reflectance); the dark image (~0% reflectance) was obtained by turning off the light source, completely covering the camera lens with its opaque cap and recording the completely response³⁸. The calibrated image (I) was calculated by the following equation:

$$I = \frac{I_0 - B}{W - B} \times 100 \quad (3)$$

A whole cucumber leaf was selected as a region of interest (ROI) of a sample. A spectral matrix of 196 samples × 512 bands was generated. There are many approaches to execute the target image segmentation from hyperspectral images, such as “Manual Mask” and “Automatic Mask”^{39,40}. A simple threshold segmentation based on a single band grayscale image was finished with the aid of Environment Visualizing Images (ENVI) softwares (ITT Visual Information, Solutins, USA). In addition, this method is also effective in MATLAB R2009a (The MathWorks, Inc., Natick, MA, USA).

Biochemical analyses. After the collection of hyperspectral images, samples were immediately cut into pieces and weighted to execute chemical analysis. Each measurement was performed in three replications. Each sample with 0.1 g fresh weight leaf tissue without leaf vein was immersed using a mixture (20 ml) of 80% acetone and 100% ethanol (1:1) for 24 hours in the dark to extract the pigments. *Chl*-BA and *Car*-BA contents were measured by 752UV/Vis spectrophotometry (Inesa Instrument, Shanghai, China) and calculated per unit fresh mass basis with the published methods^{41,42}:

$$Chl = (20.2A_{645} + 8.02A_{663}) \times \left(\frac{V}{W \times 1000} \right) \quad (4)$$

$$Car = (4.7A_{440} - 0.27 \times (20.2A_{645} + 8.02A_{663})) \times \left(\frac{V}{W \times 1000} \right) \quad (5)$$

where, *Chl* content was the sum of contents of *Chl* a and *Chl* b; V was the volume of the mixed solvent (ml); W was the fresh weight of the leaf that was measured (g); A_{440} was the absorbance value of the extract solution at 440 nm, A_{645} was the absorbance value of the extract solution at 645 nm, A_{663} was the absorbance value of the extract solution at 663 nm.

Data Analysis. Regression coefficients (RCs) of partial least squares regression (PLSR) can be used to measure the association between each variable (X) and the response (Y)⁴³. Generally, variables with small absolute

value of RCs can be eliminated and RCs is an effective method to select important wavelengths. PLSR is a statistic method based on multiple linear and stepwise regressions and especially suitable for modeling when the number of variables is greater than the number of samples, and when there is colinearity between variables⁴⁴. In addition, the number of latent variables (LVs) of the PLSR model was based on the lowest value of predicted residual errors sum of squares to avoid over-fitting or under-fitting problems. PLSR was applied to construct the calibration models with the aid of Unscrambler X10.1 (CAMO Process As, Oslo, Norway). It combines the most useful information from hundreds of bands into the first several factors to develop the calibration model. Leave-one-out cross-validation was applied to validate the performance and evaluate over-fitting of the calibration models. The prediction performance of the models were evaluated by correlation coefficients (R) of calibration (R_C), cross-validation (R_{CV}), prediction (R_p), root mean square error of calibration (RMSEC), cross-validation (RMSECV), and prediction (RMSEP)¹⁸.

Spatial distribution of leaf pigments. Observation of the *Chl* and *Car* content variances in cucumber leaves is helpful to measure the growth conditions of the plants. An optimal calibration model is firstly established with the mean spectra of the samples and their corresponding chemical values³⁷. Then *Chl*-HSI and *Car*-HSI contents of each pixel were predicted by inputting the spectrum of each pixel into the optimal regression models. The distribution maps of *Chl*-HSI and *Car*-HSI contents were generated according to the predicted values in the corresponding spatial position. This visualization process was carried out in MATLAB R2009a. Figure. 6 illustrates the main steps of mapping of *Chl*-HSI and *Car*-HSI spatial distribution in cucumber leaves infected with ALS by hyperspectral imaging technique.

References

1. FAOSTAT. Food and agricultural commodities production. (2012) Available at: <http://faostat.fao.org/site/339/default.aspx>. (Accessed: 5th August 2015).
2. Kuźniak, E. *et al.* Involvement of ascorbate, glutathione, protein S-thiolation and salicylic acid in benzothiadiazole-inducible defence response of cucumber against *Pseudomonas syringae* pv *lachrymans*. *Physiol. Mol. Plant. P.* **86**, 89–97 (2014).
3. Agrios, G. N. *Plant Pathology* 5th ed (ed. Agrios, G.) (Elsevier Academic Press 2005).
4. Blackburn, G. A. Wavelet decomposition of hyperspectral reflectance data for quantifying photosynthetic pigment concentrations in vegetation. *Proceedings of the XXth ISPRS Congress; Commission. 7*, 12–23 (2007).
5. Cherif, J. *et al.* Analysis of *in vivo* chlorophyll fluorescence spectra to monitor physiological state of tomato plants growing under zinc stress. *J. Photoch. Photobio. B.* **101**, 332–339 (2010).
6. Matile, P., Hortensteiner, S., Thomas, H. & Krautler, B. Chlorophyll breakdown in senescent leaves. *Plant physiol.* **112** (1996).
7. Guo, D.-P. *et al.* Photosynthetic rate and chlorophyll fluorescence in leaves of stem mustard (*Brassica juncea* var. *tsatsai*) after turnip mosaic virus infection. *Plant Sci.* **168**, 57–63 (2005).
8. Braumann, T. & Grimme, L. H. Reversed-phase high-performance liquid chromatography of chlorophylls and carotenoids. *BBA-Bioenergetics* **637**, 8–17 (1981).
9. Wellburn, A. R. The Spectral determination of chlorophylls *a* and *b*, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J. Plant Physiol.* **144**, 307–313 (1994).
10. Uddling, J., Gelang-Alfredsson, J., Piikki, K. & Pleijel, H. Evaluating the relationship between leaf chlorophyll concentration and SPAD-502 chlorophyll meter readings. *Photosynth. Res.* **91**, 37–46 (2007).
11. Yoder, B. J. & Pettigrew-Crosby, R. E. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–9,500nm) at leaf and canopy scales. *Remote Sens. Environ.* **53**, 199–211 (1995).
12. Gitelson, A. A., Gritz, Y. & Merzlyak, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **160**, 271–282 (2003).
13. Datt, B. Remote sensing of chlorophyll *a*, chlorophyll *b*, chlorophyll *a + b*, and total carotenoid content in eucalyptus leaves. *Remote Sens. Environ.* **66**, 111–121 (1998).
14. Daughtry, C. S. T., Walthall, C. L., Kim, M. S., Brown de Colstoun, E. & McMurtrey, J. E. III. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* **74**, 229–239 (2000).
15. Gitelson, A. The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. *Int. J. Remote Sens.* **13**, 3367–3373 (1992).
16. Peñuelas, J. & Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* **3**, 151–155 (1998).
17. Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G. & Frias, J. M. Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends Food Sci. Tech.* **18**, 590–598 (2007).
18. Yu, K. Q. *et al.* Hyperspectral imaging for mapping of total nitrogen spatial distribution in pepper plant. *PLoS one* **9**, doi: 10.1371/journal.pone.0116205 (2014).
19. Shi, J.-Y. *et al.* Nondestructive diagnostics of nitrogen deficiency by cucumber leaf chlorophyll distribution map based on near infrared hyperspectral imaging. *Sci. Hortic-Amsterdam.* **138**, 190–197 (2012).
20. Zhang, X., Liu, F., He, Y. & Gong, X. Detecting macronutrients content and distribution in oilseed rape leaves based on hyperspectral imaging. *Biosyst. Eng.* **115**, 56–65 (2013).
21. Lelong, C. C. D., Pinet, P. C. & Poilvé, H. Hyperspectral imaging and stress mapping in agriculture: a case study on wheat in beauce (France). *Remote Sens. Environ.* **66**, 179–191 (1998).
22. Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U. & Herppich, W. B. Early detection of *Fusarium* infection in wheat using hyperspectral imaging. *Comput. Electron. Agr.* **75**, 304–312 (2011).
23. Rumpf, T. *et al.* Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Comput. Electron. Agr.* **74**, 91–99 (2010).
24. Scarpari, L. M. *et al.* Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis pernicioso*. *J. Exp. Bot.* **56**, 865–877 (2005).
25. Galvão, R. K. H. *et al.* A method for calibration and validation subset partitioning. *Talanta* **67**, 736–740 (2005).
26. Wu, C., Niu, Z., Tang, Q. & Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agr. Forest Meteorol.* **148**, 1230–1241 (2008).
27. Broge, N. H. & Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* **76**, 156–172 (2000).
28. Filella, I. & Peñuelas, J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *Int. J. Remote. Sens.* **15**, 1459–1470 (1994).
29. Boochs, F., Kupfer, G., Dockter, K. & KÜHbauch, W. Shape of the red edge as vitality indicator for plants. *Int. J. Remote. Sens.* **11**, 1741–1753 (1990).

30. Peñuelas, J., Filella, I., Biel, C., Serrano, L. & Savé, R. The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote. Sens.* **14**, 1887–1905 (1993).
31. Moshou, D. *et al.* Intelligent multi-sensor system for the detection and treatment of fungal diseases in arable crops. *Biosyst. Eng.* **108**, 311–321 (2011).
32. Ustin, S. L. *et al.* Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. *Remote Sens. Environ.* **113**, S67–S77 (2009).
33. Mahlein, A.-K., Oerke, E.-C., Steiner, U. & Dehne, H.-W. Recent advances in sensing plant diseases for precision crop protection. *Eur. J. Plant Pathol.* **133**, 197–209 (2012).
34. Zhao, Y., He, Y. & Xu, N. A novel algorithm for damage recognition on pest-infested oilseed rape leaves. *Comput. Electron. Agr.* **89**, 41–50 (2012).
35. Patil, J. & Kumar, R. Advance in image processing for detection of plant diseases. *Journal of Advanced Bioinformatics Applications and Research* **2**, 135–141 (2011).
36. Sindhuja, S., Ashish, M., Reza, E. & Cristina, D. A review of advanced techniques for detecting plant diseases. *Comput. Electron. Agr.* **72**, 1–13 (2010).
37. Wei, G., Klopper, J. W. & Tuzun S. Induced systemic resistance to cucumber diseases and increased plant growth by plant growth-promoting rhizobacteria under field conditions. *Phytopathology* **86**, 221–224 (1996).
38. Zhu, F., Zhang, D., He, Y., Liu, F. & Sun, D.-W. Application of visible and near infrared hyperspectral imaging to differentiate between fresh and frozen-thawed fish fillets. *Food Bioprocess Tech.* **6**, 2931–2937 (2013).
39. Wu, D. *et al.* Rapid prediction of moisture content of dehydrated prawns using online hyperspectral imaging system. *Anal. Chim. Acta.* **726**, 57–66 (2012).
40. ElMasry, G., Sun, D.-W. & Allen, P. Non-destructive determination of water-holding capacity in fresh beef by using NIR hyperspectral imaging. *Food Res. Int.* **44**, 2624–2633 (2011).
41. Arnon, D. Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. *Plant Physiol. Biochem.* **21**, 1–15 (1949).
42. Yang, C. W. *et al.* Comparative effects of salt-stress and alkali-stress on the growth, photosynthesis, solute accumulation, and ion balance of barley plants. *Photosynthetica* **47**, 79–86 (2009).
43. Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. A review of variable selection methods in Partial Least Squares Regression. *Chemometr. Intell. Lab.* **118**, 62–69 (2012).
44. González-Fernández, A. B., Rodríguez-Pérez, J. R., Marabel, M. & Álvarez-Taboada, F. Spectroscopic estimation of leaf water content in commercial vineyards using continuum removal and partial least squares regression. *Sci. Hortic-Amsterdam* **188**, 15–22 (2015).

Acknowledgements

This research was supported by the National Natural Science Foundation of China (31471417), Specialized Research Fund for the Doctoral Program of Higher Education of China (20130101110104).

Author Contributions

Y.-R.Z. carried out the experiment and wrote this manuscript. X.L. designed the experiment and revised the manuscript. K.-Q.Y. managed the data processing and writing of the manuscript. F.C. carried out the experiment. Y.H. reviewed the initial design of the experiments and made guidance for the writing of the manuscript. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zhao, Y.-R. *et al.* Hyperspectral Imaging for Determining Pigment Contents in Cucumber Leaves in Response to Angular Leaf Spot Disease. *Sci. Rep.* **6**, 27790; doi: 10.1038/srep27790 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>