Segmentation of thermal infrared images of cucumber leaves using K-means clustering for estimating leaf wetness duration

Dongmei Wen1, Aixin Ren1, Tao Ji1, Isabel Maria Flores-Parra4, Xinting Yang1, Ming Li1,2,4*
(1. Beijing Research Center for Information Technology in Agriculture/National Engineering Research Center for Information Technology in Agriculture/National Engineering Laboratory for Agri-product Quality Traceability/National Meteorological Service Center of Urban Agriculture, Beijing 100097, China; 2. College of Plant Protection, China Agricultural University, Beijing 100193, China; 3. Department of Sustainable Crop Production, Università Cattolica del Sacro Cuore, 29122 Piacenza, Italy; 4. Department of Informatics, University of Almeria, Almeria, CP04120, Spain)

Abstract: Leaf wetness duration (LWD) is a critical parameter used to predict plant disease, but its determination under actual field conditions is a major challenge. In this study, a method for determining LWD using thermal infrared imaging was developed and applied to cucumber plants grown in a solar greenhouse. Thermal images of the plant leaves were captured using an infrared scanning camera, and a leaf wetness area segmentation method consisting of two procedures was applied. First, a color space conversion was performed automatically by an image-processing algorithm. Then, the K-means clustering algorithm was applied to enable the segmentation of the wetness area on the thermal image. Subsequently, to enable overall thermal image analysis, an initial leaf wetness threshold (LWT) of 5% was defined (where wetness values higher than 5% indicated that the leaf was in a wet state). The results of comparative experiments conducted using thermal images of plant leaves captured using an infrared scanning camera and human visual observation indicated that the estimated LWD values were generally higher than the observed LWD values, because slight leaf wetness conditions were overlooked by the human eye but detected by the infrared scanning camera. While these differences were not found to be statistically significant in this study, the proposed method for determining LWD using thermal infrared imaging may provide a new LWD detection method for cucumber and other plants grown in solar greenhouses.

Keywords: thermal imaging, K-means clustering algorithm, leaf wetness duration, cucumber
DOI: 10.25165/ijabe.20201303.4301


1 Introduction

For many plant crops grown in solar greenhouses, ideal conditions in terms of temperature, humidity, and surface wetness are important contributors to the development of foliar fungal diseases[1]. Leaf wetness is particularly important because it provides the free water required by pathogens to infect foliar tissue[2]. Leaf wetness, defined as the visible presence of water on a leaf surface[3], can be caused by the roof or plastic film runoff, guttation, or fertilization in solar greenhouses[4,5]. Guttation, which occurs with adequate irrigation and high relative humidity, increases leaf wetness duration (LWD)[6]. When free water on plants exceeds a pathogen-specific length of time and temperatures are appropriate, pathogen spores can germinate and infect the host[7-8]. The incidence of downy mildew, caused by Pseudoperonospora cubensis, increases as LWD increases; studies conducted in growth chambers have confirmed the positive relationship between LWD and the incidence of this disease[9]. The infection model developed for predicting infection periods by fungal foliar pathogens usually use inputs based on subjective estimates of the cardinal temperatures and the leaf wetness duration requirement[10]. Thus, LWD is a key consideration in disease prediction models and for decision support systems used for foliar fungal plant pathogens[11]. LWD investigation is one of the major issues in plant disease research[12,13]. Numerous studies have been conducted on LWD and two main methods are currently used to determine LWD: (1) sensor measurements and (2) model estimations[4,14-17]. Different types of sensors based on static, mechanical, or electronic systems can be used to monitor leaf wetness[18]. These sensors can adequately respond to condensation, rain, and fog. However, there is no standard for the installation and maintenance of the sensors, which makes it difficult to obtain reliable measurements of LWD[19]. As an alternative to direct measurement of LWD using sensors, a number of different numerical models have been developed to quantitatively estimate LWD[20]. These models require a significant number of input parameters to achieve sufficient accuracies in the estimates, such as relative humidity, temperature, net radiation, and wind speed[21]. However, some of the parameters are particularly difficult to obtain[22,23]. Further, the considerable effort dedicated to developing a method for

Received data: 2018-04-09  Accepted date: 2020-04-07
Dongmei Wen, PhD, research interest: Early detection and prediction model of greenhouse vegetable disease, Email: wendm@nercita.org.cn; Aixin Ren, Master candidate, research interest: Prediction model of greenhouse vegetable disease, Email: 1787031145@qq.com; Tao Ji, PhD candidate, research interest: Prediction model of plant disease, Email: a513653089@qq.com; Isabel Maria Flores-Parra, Master, Professor, research interest: image processing, Email: iflores@ual.es; Xinting Yang, PhD, Professor, research interest: information technology in agriculture, Email: yangxt@nercita.org.cn; Aixin Ren, PhD, Professor, research interest: image processing, Email: a513653089@qq.com; Ming Li, PhD, Associate Professor, research interest: warning system of greenhouse vegetable disease. Mailing address: Room 302, Building 15, Beijing Academy of Agriculture and Forestry Science, 98 Shuguang Huayan Middle Road, Haidian District, Beijing 100097, China. Tel: +86-10-51503912, Email: lim@nercita.org.cn.

K-means clustering for estimating leaf wetness duration.
determining LWD suggests a need for yet another alternative approach.

As an emerging tool in modern agriculture, machine vision technology can simulate human eye functions with the added capability of continuous acquisition, storage, and analysis of data. To date, machine vision research in agricultural fields has primarily focused on leaf wetness caused by spray. Derksen and Jiang used machine vision technology to characterize the sedimentary features of the spray droplets. Ramalingam et al. used multispectral image technology to study the characteristics of a high-capacity spray on a leaf. To date, very few studies have used machine vision technology to determine leaf wetness conditions have been performed.

The K-means algorithm is an unsupervised learning technique for clustering proposed by James MacQueen in the 1960s and widely used for image segmentation in medicine and biology. Mohd et al. apply K-means clustering detecting the hot spot of thermal infrared images, and prove this algorithm can be used for the hot spot detection on thermal infrared images of electronic boards. Etehadtavakol et al. apply K-means and fuzzy c-means for color segmentation of thermal infrared breast images to separate total colors to detect different temperature regions on the human body surface. In this study, we apply K-means clustering for the color segmentation of collected thermal infrared images.

In conventional digital images, water droplets on a leaf are difficult to discern because of their transparency, low volume, and uneven distribution across the leaf surface, and the veins in the leaf may also obscure the water droplets. In recent years, thermal infrared imaging technology has been widely applied in biological and abiotic stress testing. Thermal imaging can reflect the temperature distribution across the leaf surface, which is affected by leaf wetness. As such, thermal imaging can theoretically be used to monitor changes in the leaf surface as conditions change from dry to wet.

In this study, we developed a method for determining LWD using thermal infrared imaging and applied it to cucumber plants grown in a solar greenhouse. Our objectives were (1) segment the wet part of leaves on the thermal infrared image accurately; (2) develop and validate a new method for estimating LWD.

2 Materials and methods
2.1 Plant preparation
The cucumber (Cucumis sativus L.) seedlings were planted on September 28, 2016, and transplanted on October 15, 2016. The cucumber plants were grown in a greenhouse at Beijing Xiaotangshan Precision Agricultural Experimental Base in Changping District, Beijing, China (116.47°E, 40.18°N). Each greenhouse had an area of 50×7 m², was constructed of metal arches covered with polyethylene film, and was oriented in a north-south direction. A brick wall formed the eastern part of the greenhouse, and a glass window was on the western end. The substrate used was a 2:1 mixture of peat and vermiculite with 500 g chicken manure and 10 g N15-P10-K15 compound fertilizer. Cucumber plants were watered with tap water every 5-8 d, 8-10 m³ of chicken manure and 10 g N15-P10-K15 compound fertilizer. The substrate used was a 2:1 mixture of peat and vermiculite with 500 g chicken manure and 10 g N15-P10-K15 compound fertilizer. Cucumber plants were watered with tap water every 5-8 d, 8-10 m³ of chicken manure and 10 g N15-P10-K15 compound fertilizer.

In this study, cucumber samples for the thermography measurement were selected at the sixth true leaf stage. For each sample, to avoid variation in the LWD caused by wind from the door, we randomly selected nine plants in the middle of the greenhouse for analysis. Every leaf selected was complete one and each plant selected 3 leaves from the upper, middle and lower parts of the plant.

2.2 Thermal image acquisition
Thermal images of each leaf were obtained using an infrared scanning camera (FLIR A615, FLIR Systems, USA) with a spectral sensitivity of 7.5-14 μm and a geometric resolution of 0.69 mm (640×480 pixels) in the greenhouse. For each leaf, thermal images were captured every 60 min from 18:00 when leaves were totally dry until the leaf fully wet at night, and from sunrise, until the leaf was fully dry in the morning for four nights (November 1-5). The temperature is 23±2°C in the daytime and 15±2°C at night. The humidity is 40±3% in the daytime and 35±2% at night.

Each day, total thermal images in JPEG format were used in this study to estimate LWD on the cucumber plants. The associated estimates of LWD (the time between when wetness first appeared and disappeared) were concurrently calculated. Concomitantly, the existence of wetness on the leaves was visually assessed by humans on select days (November 1-5). During each day, beginning at 18:00 and ending at 11:00 or 12:00 the next day, the wetness for at least 12 min was defined based on the presence of water that was visible to the human eye. The observed LWD as the standard value.

2.3 Image pre-processing
In order to improve the image quality, image-processing techniques were used to remove background noise and enhance important features of the samples before computation analysis. Median filtering, which is commonly used as an image-processing tool to remove background noise effects and improve image quality before segmentation, was used. The RGB color space image was transformed into a gray image, filtering the gray image from R, G and B channels respectively, and then three filtered images were synthesized into RGB images.

2.4 Conversion of RGB images into \(L^*a^*b^*\)
Several different color space alternatives exist. In this study, we selected the International Commission of Illumination (CIE) \(L^*a^*b^*\) color space, which is most commonly used because of its uniform distribution of colors and proximity to human color perception. The \(L^*a^*b^*\) color space mathematically describes colors in three dimensions: lightness (\(L^*\)) and color opponents of green-red (\(a^*\)) and blue-yellow (\(b^*\)).

Selection of the \(L^*a^*b^*\) color space required transformation from the red-green-blue (RGB) color space. The transformation required two steps: (1) transforming RGB to XYZ coordinates and (2) transforming XYZ to \(L^*a^*b^*\) coordinates. Equations (1)-(5) support these two required transformations, respectively:

\[
\begin{align*}
X &= 0.2417 Y &= 1.5010 Z &= 1.1170 \\
Y &= 1.0000 Y &= 0.0565 Z &= 5.5943 \ \\
L &= \left\{ \begin{array}{ll}
116 f\left(\frac{Y}{Y_0}\right)^{1/3}, & \text{if } \frac{Y}{Y_0} \leq 0.008856 \\
903.3 \frac{Y}{Y_0}, & \text{if } \frac{Y}{Y_0} > 0.008856 \\
\end{array} \right.
\end{align*}
\]

\[a = 500 \left[ f\left(\frac{X}{X_0}\right)^{\gamma} - f\left(\frac{Y}{Y_0}\right)^{\gamma} \right] \]


$$b = 200 \left[ f \left( \frac{Y}{\lambda_0} \right)^3 - f \left( \frac{Z}{\lambda_0} \right)^3 \right]$$

(4)

$$f = \left\{ \begin{array}{ll}
7.787 t^{1.16}, & t > 0.008856 \\
0, & \text{otherwise}
\end{array} \right.$$

(5)

where, $X_0$, $Y_0$, and $Z_0$ are the initial reference values, and $t$ is the ratio between $X$, $Y$, and $Z$ and $X_0$, $Y_0$, and $Z_0$.

### 2.5 Segmentation of leaf wetness area

Image segmentation was used to divide each image into related regions as segmentation of the leaf wetness area in the thermal image is a key step in LWD estimation. After the color space was fully defined, the K-means clustering algorithm was next applied to support the segmentation of the thermal image. This algorithm is commonly used in computer vision applications for image segmentation[36].

A digital image can be viewed as a data point set, $X=\{X_1, X_2, \cdots, X_n\}$ with $n$-dimensional vectors. The K-means clustering algorithm segments this image (defined by $X$) into $K$ clusters by defining a subset, $Z=\{C_1, C_2, \cdots, C_K\}$, of $X$ that minimizes the objective function

$$J = \sum_{i=1}^{K} \sum_{X \in C_i} (X - C_i)^2$$

where the objective function is minimized, and

$$\frac{\partial J}{\partial C_i} = \sum_{X \in C_i} (X - C_i)$$

The objective function can be alternatively expressed as follows:

$$J = \sum_{i=1}^{K} \sum_{X \in C_i} (X_i - C_i)^2$$

(6)

when the objective function is minimized, and

$$\frac{\partial J}{\partial C_i} = \sum_{X \in C_i} (X_i - C_i)$$

The objective function can be alternatively expressed as follows:

$$J = \sum_{i=1}^{K} \sum_{X \in C_i} (X_i - C_i)^2$$

(7)

where, $n_i$ is the number of data points in cluster $i$.

During the iterative minimization of the objective function, the cluster centers of K-means were constantly updated, and data points in the same cluster increased in similarity, while data points in different clusters decreased in similarity. When the objective function was minimized and the cluster centers were no longer updated, the image segmentation process was complete.

Figure 1 shows a flowchart of the complete thermal image segmentation process for the cucumber plants considered in this study. After the image pre-processing, color images were converted to binary images and the appropriate image segmentation threshold selected. Accordingly, target pixels comprising the thermal image were labeled 1 (white), while all other pixels were classified as background pixels and labeled 0 (black). Finally, the target pixels (wet surface area of a leaf) comprising the segmented images from the thermal images were extracted and automatically represented in RGB.

### 2.6 Leaf wetness duration estimation

A predefined leaf wetness threshold (LWT) can be used to identify a leaf’s wet state and subsequently estimate LWD with some level of accuracy. In this study, an LWT was defined based on the target image’s pixel[37]. Specifically, the LWT was determined as follows:

$$LWT = \frac{N}{M} \times 100\%$$

(8)

where $N$ and $M$ are the pixels for the target image and single blade image, respectively. LWT>5% was used to indicate leaf wetness in the study. The times when leaf wetness first appeared (LWT>5%) and disappeared (LWT<5%) were recorded as part of this study.

### 2.7 Statistical analysis of LWD estimates

To evaluate our proposed method for determining LWD, we compared the estimated LWD results obtained from the thermal image analysis with human visual observation. Differences between the two sets of results were characterized using a Student’s t-test (SPSS 16.0, SPSS Inc., Chicago, IL, USA) and the root means square error (RMSE)[37]. The RMSE was determined as follows:

$$RMSE = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2}$$

(9)

where $d_i$ and $\bar{d}$ are the estimated and observed LWD results, respectively; and $n$ is the sample number.

### 3 Results and discussion

#### 3.1 Differences between thermal infrared images of the wet and dry leaf surface

Thermal infrared imaging facilitated the differentiation between wet and dry leaves (Figure 2). In conventional digital images, water droplets on a leaf are difficult to discern because of their transparency, low volume, and uneven distribution across the leaf surface[30]. In contrast, the thermal infrared image store temperature rather than color data in each pixel. Furthermore, other studies have confirmed a correlation between leaf temperature and wetness state, given that water is the primary source of infrared absorption[38]. A previous study explored the possibility of using thermal imaging as a tool to identify water stress in plants, and this could be used in irrigation scheduling[39, 40]. Hence, from this study, thermal images have great potential because leaf wetness causes a change in leaf temperature and color. Therefore, we thought it possible to segment the water droplets in

![Figure 1](image-url)
an infrared image.

In this study, appreciable differences in color were apparent between the wet and dry leaf surface in one leaf, due to the leaf temperature changes. Therefore, we fully utilized the color difference between the wet (blue) and dry (yellow) of the thermal infrared image, in Figure 2b. This will enable the effective segmentation of the wet areas in thermal images.

### 3.2 Image processing results

Figure 3 shows sample thermal images of a cucumber leaf before (Figure 3a) and after denoising (Figure 3e). Image quality was improved after removing background noise.

Figure 4b shows a sample of the RGB color space that was transformed into the $L^*a^*b^*$ color space (Figure 4b). Figure 4c shows the wetness region of thermal images. The K-means clustering algorithm can successfully segment the wet part of the leaves.

Using thermal imaging methods for determining LWD requires associated methods based on computer machine vision[39]. Hence, wet area segmentation of thermal images is a key process in the estimation of LWD. Our results in this study clearly show that $K$-means clustering provides a method for wet area segmentation. This is because of the difference between the background color and an interest region[41]. In the process of image acquisition, the temperature in the greenhouse was changing from day till night. But our image acquisition was mainly at night, with small temperature fluctuation. We think the impact of temperature on the thermal images can be ignored.

### 3.3 Leaf wetness duration estimation

We chose one leaf from the broader dataset as an example to show the estimated leaf wetness state over time based on the LWT (Figure 5). The results indicated that the selected cucumber leaf was in a wet state (LWT>5%) between 23:00-24:00 and 07:00-10:00 on November 1-2 and 4-5 (Figures 5a and 5d), and between 10:00-11:00 on November 2 and 3 (Figures 5a and 5b).

To assess the accuracy of our proposed method for determining LWD using thermal images, we compared the estimated LWD averaged results obtained from the thermal image analysis with human visual observation (Table 1). The estimated LWD averaged from thermal images was determined to be 9.93, 9.81, 10.56, and 9.33 h on November 1-2, 2-3, 3-4, and 4-5, respectively, while the observed LWD averaged was 9.47, 9.52, 10.15, and 9.00. The estimated and observed LWD results were not significantly different (Student’s $t$-test, $p>0.05$) for any of the observation periods considered (November 1-2, 2-3, 3-4, and 4-5).

<table>
<thead>
<tr>
<th>Date of observation</th>
<th>Mean LWD ($\pm$standard error) [h]</th>
<th>Test statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated</td>
<td>Observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>November 1-2</td>
<td>9.93±0.02</td>
<td>$t=0.262$</td>
<td>$p=0.611$</td>
</tr>
<tr>
<td>November 2-3</td>
<td>9.81±0.21</td>
<td>$t=0.133$</td>
<td>$p=0.717$</td>
</tr>
<tr>
<td>November 3-4</td>
<td>10.56±0.21</td>
<td>$t=0.172$</td>
<td>$p=0.680$</td>
</tr>
<tr>
<td>November 4-5</td>
<td>10.33±0.02</td>
<td>$t=0.156$</td>
<td>$p=0.695$</td>
</tr>
</tbody>
</table>

To further compare the estimated and observed LWD results, we constructed scatter plots and calculated RMSE values. Figure 6 shows the correlation between the estimated and observed LWD results for each of the observation periods considered in this study. The results indicated that the difference between estimated LWD and overall observation period was 1 h, and lower RMSE values indicated higher LWD estimate accuracy. These results also indicated that over 90% of the estimated LWD values were equal to or higher than the observed LWD values.

One limitation with leaf surface wetness duration is the selection of the wetness threshold (LWT). Predefined LWTs are generally required for the estimation of LWD[37], but are rarely determined through image analysis. In this study, we selected an LWT of 5% based on target image pixels. We found that the differences between the estimated and observed LWD results were not statistically significant, and the estimated LWD values were generally higher than the observed LWD values, which are both sufficient leaf wetness time for an infection to occur[37]. The
The difference between the estimated and observed LWD was generally <0.5 h (Table 1). We theorize that slight leaf wetness droplets may be overlooked by the human eye but detected by the infrared scanning camera.

Figure 5  Estimated leaf wetness over time based on the LWT

Figure 6  Correlation between estimated and observed LWD results
4 Conclusions

This study shows that thermal infrared images are very suitable for detecting changes in wetness on cucumber leaves. In our proposed method, the images of the cucumber leaves were pre-processed, after which the wet area was segmented automatically from the thermal image using an image analysis algorithm and K-means clustering. This was accomplished based on the differences in color between the wet and dry leaf surfaces. The good agreement between the estimated and observed LWD shows that the proposed method can be used to predict LWD with acceptable accuracy. The LWD estimation method proposed in this study is rapid, accurate, and non-destructive. Furthermore, it can be integrated into greenhouse disease forecast models and warning systems as an aid to suppress the development and propagation of infectious plant diseases. Thus far, the use of thermal image processing and computer vision for LWD estimation is still new, which means that there are still many issues to be explored, such as the application of the proposed method in an online manner in greenhouses.

Acknowledgements

This study was supported by the China National Key R&D Program (2017YFE0122503), National Natural Science Foundation of China (31401683), Beijing Municipal Excellent Talents Project (2016000057592G260), and Postdoctoral Science Foundation of Beijing Academy of Agriculture and Forestry (968). The authors are also grateful to the research centre ‘CIMEDES’ at University of Almeria, Spain.

[References]


