Computers and Electronics in Agriculture 150 (2018) 41–51

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

# Estimation of soybean leaf area, edge, and defoliation using color image analysis

Wei-zhen Liang, Kendall R. Kirk*, Jeremy K. Greene

*Department of Plant and Environmental Science, Edisto Research and Education Center, Clemson University, 64 Research Rd, Blackville, SC 29817, United States*

ARTICLE INFO

*Keywords:*
Soybean
Mahalanobis distance
Canny edge detection algorithm
Defoliation
Statistical regression

ABSTRACT

Soybeans are an extremely important crop in the United States, providing the largest source of animal protein feed and a main source of vegetable oil in the country. The goal of this research is to identify methodologies to estimate percent defoliation of the soybean canopy and leaves using RGB images taken in the field. The Mahalanobis distance classification method was used to process sets of images and calculate leaf area (number of pixels) corresponding to two classes (leaf and background) with eight different color groups. The Canny edge detection algorithm provided an efficient method for detecting leaf edges, and threshold $t_2 = 20$ was found to be the optimal value for estimating soybean leaf edge. The segmentation results showed a performance of 96% for soybean leaves using Mahalanobis distance classification. Two statistical regression models (polynomial and logistic regression) for defoliation of soybean were developed based on individual images of trifoliate leaves taken from the field. The models both utilized leaf area and edge to provide estimates of soybean defoliation; however, a logistic equation has potential to provide greater understanding and more accurate estimates of defoliation with variations, especially at low defoliation. The $R^2$ and root mean square error (RMSE) of estimated and observed defoliation of trifoliate leaves were 0.90 and 6.16%, respectively. The validation of soybean canopy defoliation and its corresponding trifoliate leaves defoliation also provided reasonable correlation ($R^2 = 0.96$ and RMSE = 1.85%). This approach could lead to use of remotely sensed imagery for estimating defoliation in soybeans and timely intervention with integrated pest management strategies.

## 1. Introduction

Soybeans are an extremely important crop in the United States, providing the largest source of animal protein feed and the second largest source of vegetable oil in the country. Domestic production of soybeans in 2016 was estimated by USDA to be $40.9 billion (USDA, 2016). In addition to being a major economic row crop grown in the United States, soybeans also serve as an important rotational crop that "feeds" subsequent crops by improving the fertility of the soil. Numerous yield-limiting factors can undermine the value of soybeans as an agricultural commodity in the southeastern United States. Abiotic factors, such as inadequate rainfall, irrigation, and temperatures, and biotic factors, such as weeds, diseases, and insects, can negatively impact profitable production of soybeans in the region. Insects can exert pressure on the crop, and signs of their impact are dynamic during a given growing season, with symptoms of feeding injury changing from subtle to marked in little time. The primary pests that reduce soybean yield in the southeastern United States are stink bug, corn earworm, velvetbean caterpillar, soybean looper, kudzu bug (Clemson University Extension; North Carolina State University Extension).

Defoliation is the most common type of insect injury observed by soybean producers, and it can occur from emergence to maturity. The current technique to estimate defoliation is to randomly remove 10 leaflets from the top, middle, and lower part of the canopy and visually determine the amount of defoliation (University of Minnesota; University of Nebraska, Lincoln; Clemson University). Estimating the amount of defoliation in soybean caused by insects can be a frustrating and time consuming process. Soybean leaf defoliation is typically overestimated because human eyes tend to focus on damage (Clemson University; University of Minnesota; Kogan and Herzog, 1980). Therefore, accurate estimation of defoliation will help save producers money and protect the environment from unnecessary applications of insecticide. Prior studies measuring leaf area and defoliation have used visual estimates (Stotz et al., 2000), desktop scanners (O'Neal et al., 2002), leaf-area meters (O'Neal et al., 2002; Malone et al., 2002), hand tracings of defoliated leaves (Hoy and Hall, 1993), or comparisons of treated leaves to appropriate controls (Jensen et al., 1977). Manual measurement of leaf area and defoliation is technically complex,

* Corresponding author.
*E-mail address:* kirk2@clemson.edu (K.R. Kirk).

https://doi.org/10.1016/j.compag.2018.03.021
Received 13 November 2017; Received in revised form 7 March 2018; Accepted 17 March 2018

expensive, and labor intensive. Furthermore, these measurements estimate defoliation from single plants or individual leaves (one prediction value for each plant or each leaf), which may not be representative of the plant canopy across the field. Therefore, it would be beneficial to develop and apply image processing to assess features of the soybean canopy for improving measurement and estimation of defoliation.

Image processing has been used as an effective tool for analyses in various crops and applications. In recent years, several studies have used image processing to assess features of crop canopies for general purposes, such as defining fertility requirements, and for specific applications like disease detection, smart spraying, and yield estimation (Diago et al., 2012; Hitimana and Gwun, 2014; Masood et al., 2016). However, there are no studies that have evaluated soybean defoliation using image processing. Our work aims to develop a fast, robust, and inexpensive methodology for straightforward image processing and interpretation of Red-Green-Blue (RGB) imagery taken in the field, allowing accurate definition of leaf area and edges and estimation of defoliation in soybeans.

Color classification techniques in the RGB color spectrum can be used to distinguish canopy from background (e.g. soil and shadow) in images. Several statistical measurements of similarity between groups, in terms of multiple characteristics, have been proposed, such as Kolmogorov's variation distance, Bhattacharyya distance, and Mahalanobis distance (Devroye et al., 1996). Mahalanobis distance classification is widely used for pattern recognition and data analyses when groups have different means but similar standard deviations (Devroye et al., 1996) and is most suitable in image processing for precision agriculture (Chen et al., 2010; Diago et al., 2012). Chen et al. (2010) extracted 28 color features from corn imagery using Mahalanobis distance for identifying five China corn varieties at a success rate of 90%. Diago et al. (2012) extracted 40 colors in 7 groups and used Mahalanobis distance to determine each pixel belongs to what color group and characterize grapevines, leaves, and background. The results showed a performance of 92% effectiveness for leaves and 98% effectiveness for grapes. The goal of leaf color classification is to separate the leaf color from background using Mahalanobis distance and determine the pixel numbers to characterize the leaf.

Edge detection is another important technique for image processing that allows for defining leaves. The goal of leaf edge detection in our research is to extract the leaf edge or margin and determines the boundary pixels number to characterize the leaf shape. The concept of edge detection is to find the region where the sharp intensity or colors change. Edge detection employs operators to approximate the first derivative of gray level gradient of an image in the spatial domain. The locations of local maximum of the first derivative are considered to be edge points. Prewitt and Sobel (Sobel, 1978) operators are examples of gradient-based edge detectors. Marr and Hildreth (1980) proposed the Laplacian of Gaussian operator for edge detection using a Gaussian filter for smoothing images and calculating the second derivative, with zero-crossing points considered as edge points. Canny (1986) also presented an optimal edge detector that gives the edge information of both intensity and direction.

The Sobel operator (Langner, 2001; Vairalkar and Nimbhorkar, 2012; Amlekar et al., 2015) and Canny edge detection algorithm (Kumar, 2012; Amlekar et al., 2015) are convenient and popular methods for detecting edges of leaves. Most studies (Masood et al., 2016) have estimated defoliation, incidence of plant diseases, and infection based on individual images of leaves instead of using images of a section of plant canopy, which may not effectively represent the condition of the crop unless replicated meticulously; care must be taken to ensure that the samples analyzed are representative of the population.

Our research aims to develop reliable and user-friendly software and methodology for accurately estimating leaf area, leaf edge, and color information to be used for estimation of defoliation of soybean leaves from RGB imagery taken in the field. The objectives of this paper are to: (1) compare Mahalanobis distance algorithm and without Mahalanobis distance for calculating leaf area pixel number, (2) compare Sobel and Canny edge detection methods for calculating leaf edge pixel number, (3) develop defoliation model to estimate leaf defoliation using leaf area and edge pixel number, (4) validate the proposed algorithms and defoliation model using other soybean images and compare the results.

## 2. Materials and methods

In our developed software, the area of interest of an image (area to be analyzed) can be drawn manually as a rectangle by the user. The software will georeference and digitize images for use in soybean management and assessment. The image digitization process involves scanning each pixel in the image taken from field, outputting pixel brightness, red, green, and blue color space values, leaf area (A) (pixel), and leaf edge (E) (pixel), and defoliation (%). The pixel and color information (e.g. brightness) for a particular area of interest are exported as ".csv" files, which can then be imported into MATLAB, Python software for further color analysis.

To be able to correlate the estimated leaf area, leaf edge and defoliation of canopy with real plant measurements, a detailed experimental setup for the acquisition of images was developed based on soybean canopy photos taken from the field and its corresponding 6 images of trifoliate leaves (2 from top, 2 from middle, and 2 from bottom of

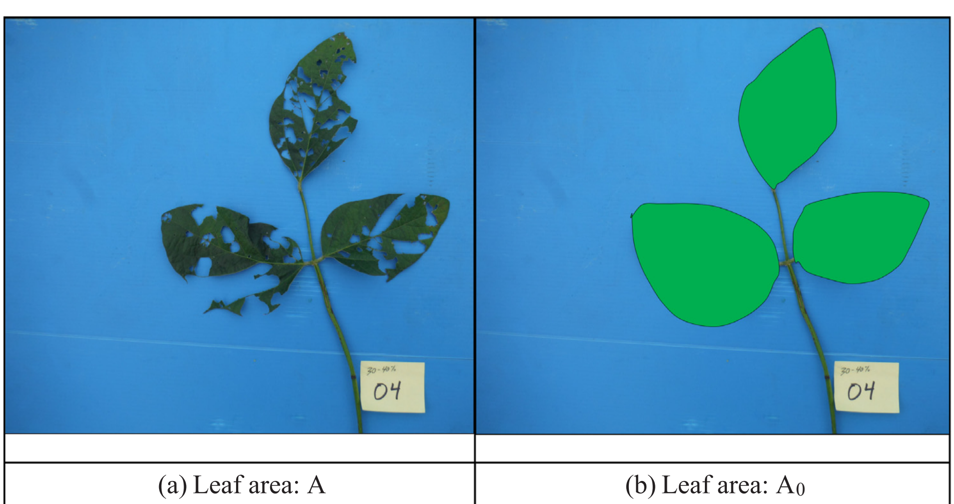


**Fig. 1.** Example performed over a soybean image of a defoliated leaf. (a) Original image (A). (b) Processed image to obtain complete leaf area ($A_0$).