PLANT LEAF SEGMENTATION FOR ESTIMATING PHENOTYPIC TRAITS

Yuhao Chen[1], Javier Ribera[1], Christopher Boomsma[2] and Edward J. Delp[1]

[1]Video and Image Processing Laboratory (VIPER), Purdue University, West Lafayette, Indiana, USA
[2] Department of Agronomy, Purdue University, West Lafayette, Indiana, USA

ABSTRACT

In this paper we propose a method to segment individual leaves of crop plants from Unmanned Aerial Vehicle (UAV) imagery for the purposes of deriving phenotypic properties of the plant. The crop plant used in our study is sorghum [*Sorghum bicolor* (L.) Moench]. Phenotyping is a set of methodologies for analyzing and obtaining characteristic traits of a plant. In a phenotypic study, leaves are often used to estimate traits such as individual leaf area and Leaf Area Index (LAI). Our approach is to segment the leaves in polar coordinates using the plant center as the origin. The shape of each leaf is estimated by a shape model. Experimental results indicate that this approach can provide good estimates of leaf phenotypic properties.

Index Terms— phenotyping, segmentation, color image processing

1. INTRODUCTION

The process used by crop scientists to measure physical and chemical properties of a plant is known as phenotyping [1, 2]. It is a set of methodologies for recording, analyzing and measuring characteristic traits of a plant resulting from genetic and environmental factors. Many of these methods are invasive where the plant is damaged or disturbed and measurements are made. Imaging technologies provide a non-invasive and an effective way for data collection and plant trait estimation [2, 3, 4, 5, 6].

Previous work in image-based phenotyping can be categorized into two approaches. In one approach, individual plants are phenotyped in environments where the imaging conditions and the distribution of plants are controlled. Controlled environments can accurately predict plant growth but cannot be employed in large scale field studies [3]. In [7], 3D individual plants are constructed from depth images acquired from multiple view points. In [8, 9], leaves are destructively collected to obtain the leaf area by scaling the number of pixels in the segmentation mask. In [10], plants in an automated indoor facility are detected using color thresholding, then located using k-means, and finally segmented using active contours. The other approach to plant phenotyping consists in estimating plant traits from image data collected in outdoor environments (field-based). In [11], an orthomosaic of the field is created from UAV images and pixels are classified into crop, soil, and shadow using maximum likelihood estimates. The mask of the crop is used to estimate Normalized Difference Vegetation Index (NDVI) and correlate it with Leaf Area Index (LAI). LAI is defined as the total leaf surface per unit area of land [12]. In [13], field images from a hand-held device are acquired and segmented using histogram-based thresholding. The segmentation masks are then used to estimate LAI. Field-based approaches can greatly increase the efficiency of large scale phenotyping studies [3, 14]. Nonetheless, current phenotyping methods only provide an estimate of the average plant growth in the field due to the difficulty of measuring many individual plants [15].

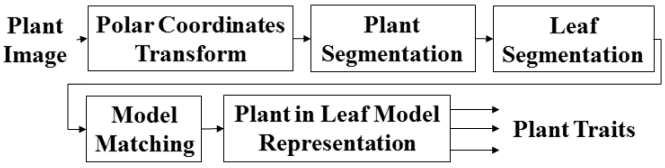

Fig. 1: Block diagram of our leaf modeling and plant traits estimation system.

Fig. 2: Image of a section of a sorghum field in West Lafayette, Indiana acquired from a UAV at an altitude of 40m.

In this paper, we estimate phenotypic traits at the plant level from UAV images. We describe a method to segment individual leaves. Figure 1 shows the block diagram of our approach. The image is converted into polar coordinates and a mask indicating plant material is obtained. Each leaf is detected by finding peaks in the mask in polar coordinates. Leaves are then reconstructed by a shape matching model that compensates for overlapping leaves. We limit the crop type to sorghum [*Sorghum bicolor* (L.) Moench] [16, 17].

2. LEAF SEGMENTATION

In this section, we describe how to find individual candidate leaves. As a sorghum plant grows the leaves spread out from the center of the plant. Analyzing individual leaves requires examining their width,

The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0000593. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Address all correspondence to Edward J. Delp, ace@ecn.purdue.edu

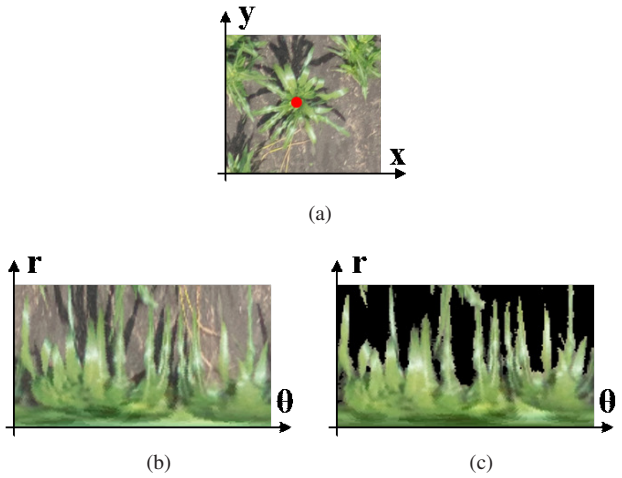

Fig. 3: a) A plant with its center labeled in red. b) Polar representation of a circular neighborhood around the plant. c) Segmented plant in polar coordinates.

length and orientation. To make the analysis easier we have all the leaves point in the same direction. This is done by transforming the image into polar coordinates. Let $I(x, y)$ denote the RGB values of the image at pixel coordinates (x, y), and let (C_x, C_y) be the coordinates of the plant center. The plant center is selected manually. We construct a circular neighborhood of radius R around (C_x, C_y) and convert it into a polar coordinate image $P(r, \theta)$:

$$P(r,\theta) = I\left(\left\lfloor C_x + r cos\left(\frac{\pi}{180}\theta\right)\right\rceil, \left\lfloor C_y + r sin\left(\frac{\pi}{180}\theta\right)\right\rceil\right), \tag{1}$$

where $\lfloor\cdot\rceil$ denotes rounding to the closest integer, and $0 \leq r < R$ and $0 \leq \theta < 360$ are integers. Figure 3a and Figure 3b show an example of the polar coordinate transformation. In polar coordinates, all leaves point along the positive r axis.

We assume the images contain only two classes: plant material (foreground) and soil (background). We use the color thresholding method described in [18] to obtain a foreground mask $F(r, \theta)$ such that

$$F(r,\theta) = \begin{cases} 1 & \text{if } P(r,\theta) \text{ contains plant material} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The background mask $B(r, \theta)$ is obtained as the complement of the foreground mask,

$$B(r,\theta) = 1 - F(r,\theta). \tag{3}$$

Figure 3c shows an example of a foreground mask. After thresholding, noise may be present as holes in the mask $F(r, \theta)$. In order to remove the holes without affecting the leaf morphology, we use connected components [19] on the background mask $B(r, \theta)$. The results of connected components is the number of components $\mathcal{C}$ and a labeled image $\Omega_B(r, \theta)$ with range $[1, \mathcal{C}]$ indicating the component label of the coordinate (r, θ). We obtain the set $\mathcal{H}$ of all labels whose component has less than T_A pixels:

$$\mathcal{H} = \{c \in [1,\mathcal{C}] : \sum_{r=0}^{R-1}\sum_{\theta=0}^{359} \delta\left(\Omega_B(r,\theta) - c\right) < T_A\}, \tag{4}$$

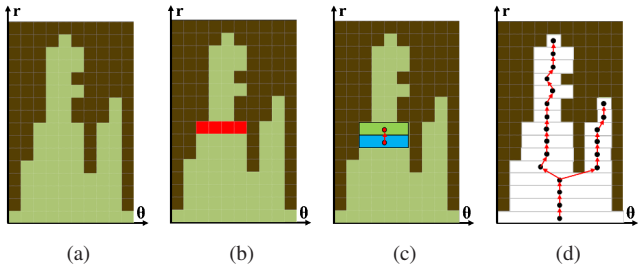

Fig. 4: a) Synthetic image with two leaves. b) A slice marked in red. c) Parent slice (blue) and child slice (green). d) Hierarchical slice representation.

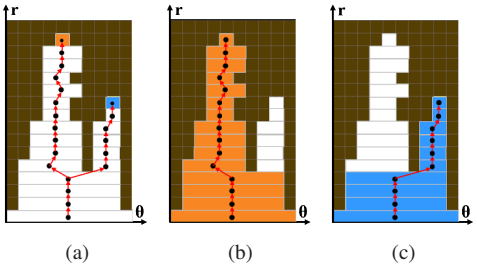

Fig. 5: a) Tip slices (orange and blue) in a hierarchical slice representation. b) Leaf segmented with an orange tip slice. c) Leaf segmented with a blue tip slice.

where $\delta(\cdot)$ is the Kronecker delta function. We fill the holes in the foreground mask $F(r, \theta)$, by combining all components in $\mathcal{H}$ and incorporating them into $F(r, \theta)$ as

$$F(r,\theta) \leftarrow F(r,\theta) + \sum_{c\in\mathcal{H}} \delta(\Omega_B(r,\theta) - c). \tag{5}$$

If there are leaves from neighboring plants, some additional independent components can appear in the foreground mask. We use connected components on the foreground mask $F(r, \theta)$ to retrieve all foreground objects. This gives us the labeled image $\Omega_F(r, \theta)$. We are only interested in the component that contains the plant center. The plant center, a single pixel in Cartesian Coordinates, transforms into a line spanning $r = 0$ in Polar Coordinates. Thus, the label of the connected component that corresponds to the plant is $\Omega_F(0, 0)$. The foreground mask is updated to contain only the plant component as

$$F(r,\theta) \leftarrow \delta(\Omega_F(r,\theta) - \Omega_F(0,0)). \tag{6}$$

We introduce a basic element, a slice, to analyze the leaf morphology. A slice is a connected set of coordinates in the foreground mask $F(r, \theta)$ with a constant r. Figure 4b shows an example of a slice marked in red.

By vertically stacking leaf slices, we obtain a leaf shape. We first define the parent-child relationship of two vertically connected slices. The slice with smaller r value is the parent of the slice with larger r value. Figure 4c shows the parent slice in blue and the child slice in green. With this relation, we construct a hierarchical representation of the slices (see Figure 4a and Figure 4d). In this hierarchical representation, each slice is considered as a node. The root node is the slice at $r = 0$ and corresponds to the plant center. There are three possible hierarchical relations between nodes. The simplest case occurs when a parent node has only one child node and the child

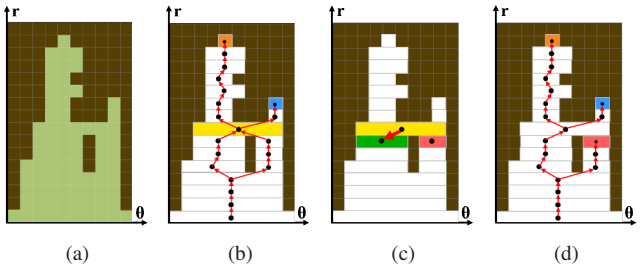

Fig. 6: a) Two syntethic leaves and a hole in the foreground mask. b) Hierarchical slice representation. The slice in yellow is where the two slices merge. The orange and blue slice are the two tip slices. c) The yellow slice is assigned to the closest slice (green). d) Hierarchical slice representation after the assignment. The three tip slices found are marked in orange, blue, and red slice.

also has only one parent. This one-to-one relation usually indicates that both the parent and child node are from a portion of a leaf. A different case occurs when a parent node has multiple child nodes and each child node has only one parent. In the last case, multiple parent nodes share the same child node. This relation is caused by holes in the foreground mask that remained after the noise removal step. When this occurs, we compute the distance between the slice centers of the parent nodes and the slice center of the child node. We assign the child node to its closest parent node according to these distances. We mark the other parent nodes as end-nodes of the hierarchy (see Figure 6).

We define the tip of a leaf as a node with no child. We reconstruct a leaf shape by combining the tip with all its ancestors (see Figure 5a, Figure 5b, and Figure 5c). We define set

$$\mathcal{A} = \{X_1, X_2, \ldots, X_N\}. \tag{7}$$

containing all N reconstructed leaf shapes $X_i, i = 1, \ldots, N$. We define an observed leaf shape $X \in \mathcal{A}$ as

$$X = \{x_1, x_2, \ldots, x_{L_X}\}. \tag{8}$$

where L_X is the number of slices in the leaf shape X. As each slice is one pixel thick, L_X is the length of the leaf shape X. Each slice $x_m, m = 1, \ldots, L_X$, is $\Theta(x_m)$ degrees wide.

3. LEAF MODELING

We use a shape modeling technique to remove noise and inter-plant occlusion. Figure 7 shows an example of the shape modeling of a leaf. Inter-plant occlusion is defined as leaves from neighboring plants occluding the peaks we want to seek as shown in Figure 7a. In addition, we define self-leaf as a leaf that belongs to the plant being analyzed and neighbor-leaf as a leaf belonging to neighboring plants. These two type of leaves vary in the distribution of their leaf widths. The width of the slices in self-leaves tends to decrease as r increases. On the contrary, the width of the slices in neighbor-leaves tends to increase as r increases. Inter-plant occlusion may create hourglass-like shapes (see Figure 7a). In order to estimate the true shape of each leaf in these scenarios, we define a finite set of K possible shape models:

$$\mathcal{D} = \{S_1, S_2, \ldots, S_K\}. \tag{9}$$

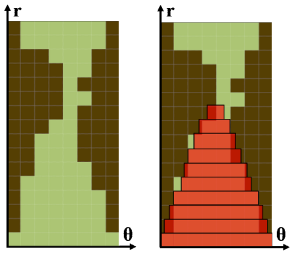

Fig. 7: a) Two synthetic leaves from different overlapping plants. b) A shape model (marked in red) to describe the observed leaf shape.

Each shape model $S \in \mathcal{D}$, consists of a set of slices:

$$S = \{y_1, y_2, \ldots, y_{L_S}\}, \tag{10}$$

where L_S is the number of slices in the shape model S.

The width $W_m, m = 1, \ldots, L_S$ of the m-th slice y_m in shape model S is modeled as random variable with normal distribution

$$p_{W_m}(w) = \mathcal{N}(\mu_m, \sigma_m^2), \tag{11}$$

where μ_m and σ_m^2 are the mean and variance, respectively. We call $p_{W_m}(\cdot)$ the slice matching probability. We set the value of σ_m^2 as

$$\sigma_m^2 = \max\{1, \tau\mu_m\}, \tag{12}$$

where τ is the tolerance rate. The higher the τ, the more disperse the width of a slice can be. We constrain σ_m^2 to be higher than 1 degree to tolerate a minimum amount of noise.

For an observed leaf shape $X \in \mathcal{A}$ and any possible shape model $S \in \mathcal{D}$, we define a shape model matching score that takes all slice matching probabilities into account:

$$f(X, S) = \sum_{m=1}^{L_S} \omega_{m,S} \;\; p_{W_m}(\Theta(x_m)), \tag{13}$$

where $\omega_{m,S}$ is the weight for the m-th slice of the shape model S. We assume that all the slices are equally weighted in the model, and

$$\sum_{m=1}^{L_S} \omega_{m,S} = 1\,. \tag{14}$$

We use the matching score to find the best shape model $\hat{S}$ to describe the observed leaf shape X as

$$\hat{S} = \underset{S \in \mathcal{D}}{\operatorname{argmax}} f(X, S)\,. \tag{15}$$

For each of all the observed leaf shapes $X_n, n = 1, \ldots, N$, we obtain $\hat{S}_n$. Some X_n may be matched to a shape model even if it is purely noise. To avoid this, we set a threshold T_B below which the match will be rejected, and the observation will not be considered. We use Otsu's method [20] to find the threshold T_B that separates strong and weak matching scores. We remove any observed leaf shape that has a matching score smaller than T_B. We denote the new leaf shape set as $\mathcal{A}'$ and the number of matched shapes as N'.

Finally, we transform the matched shape model back to Cartesian coordinates to estimate phenotypic traits. A horizontal line in polar coordinates is an arc in Cartesian coordinates. The width of

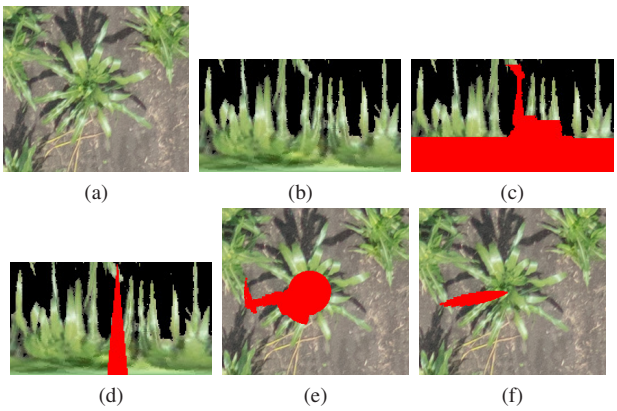

Fig. 8: a) The original image. b) Image segmented by color thresholding in polar coordinates. c) Segmentation of a leaf marked in red. d) The best shape model that describes the leaf is marked in red. e) Segmentation of the leaf transformed back into Cartesian coordinates. f) Shape model transformed back into Cartesian coordinates.

the arc is called cord length. For a model shape $S \in \mathcal{D}$, we obtain the cord length c_m of the m-th slice y_m in S as

$$c_m = 2r_m \sin\left(\frac{\mu_{m,r}}{2}\right), \tag{16}$$

where r_m is the r coordinate of slice y_m, and $\mu_{m,r}$ is the mean μ_m in radian. We estimate the total area of S as

$$\begin{aligned} A(S) &= \left(\sum_{m=1}^{L_S} c_m\right) + \frac{r_m^2}{2}(\mu_{L_S,r} - \sin(\mu_{L_S,r})) \\ &= \left(\sum_{m=1}^{L_S} 2r_m \sin(\frac{\mu_{m,r}}{2})\right) + \frac{r_m^2}{2}(\mu_{L_S,r} - \sin(\mu_{L_S,r})), \end{aligned} \tag{17}$$

where the second term corresponds to the tip of the leaf, a small circular segment area. LAI is obtained as

$$\text{LAI} = \frac{\sum_{n=1}^{N'} A(\hat{S}_n)}{A_u}, \tag{18}$$

where A_u is the unit area the plant occupies.

4. EXPERIMENTAL RESULTS

We evaluated our method with the image shown in Figure 2. This image was acquired from a UAV flying at an altitude of 40 meters with a speed of 8m/s on June 30, 2016 in West Lafayette, Indiana. The centers of 10 plants were manually selected. R, the radius of the circular neighborhood around each plant center, was set to 70 pixels. We set the area threshold T_A for rejecting background holes to 150. Shape models were generated as triangles in polar coordinates. The widths μ of the triangles spanned between 12 degrees and 100 degrees, and the length was between 20 and 70 pixels. A total of 4,539 shape models were constructed. The tolerance rate τ was set to 0.3 empirically. A_u, the unit plant area was set to 19,600 pixels.

We thank Professor Ayman Habib and the Digital Photogrammetry Research Group (DPRG) from the School of Civil Engineering at Purdue University for providing the images used in this paper.

Fig. 9: a) Different leaves are segmented in different colors. b) The models describing each leaf are marked in different colors.

The leaf segmentation was evaluated by visually comparing the original images (such as Figure 8a) with the segmented images (such as Figure 9a). The method was evaluated using the 10 selected plants. The number of leaves in the selected plants was 82. Out of these 82 leaves, a total of $TP = 57$ true positive leaves were correctly detected. A total of $FP = 18$ false positive leaves were detected, meaning that we incorrectly detected 18 noisy segments of the image as leaves. Out of these 82 leaves, the method produced a total of $FN = 25$ false negative samples, meaning that 25 leaves could not be detected. Precision and recall was used to evaluated the performance of our method [21, 22]. Precision is defined as

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{19}$$

Recall is defined as

$$\text{Recall} = \frac{TP}{TP + FN} \tag{20}$$

We obtained a precision and recall of 69.5% and 76.0%, respectively. Since the number of false positives and the number of false negatives is similar, the precision and recall are similar. There is no bias towards missing leaves or overdetecting leaves. Hence, the average number of detected leaves is very close to the true value. Based on the shape models, phenotypic traits were estimated. Traits included the area (Equation (17)) and length (L_S) of each leaf, and the leaf count (N') and LAI (Equation (18)) of individual plants . Figure 8 shows the segmentation result of a leaf for one plant. The length and area of the leaf shown was estimated to be 98 pixels and 1,000 square pixels, respectively. The final segmentation result of an entire plant is shown in Figure 9a and 9b. We correctly detected and segmented 8 out of 14 leaves and the LAI of the plant was estimated to be 0.21.

5. CONCLUSIONS

This paper presented a method to segment leaves of sorghum plants from UAV imagery. We can obtain phenotypic traits such as the area and length of each leaf and the leaf count and LAI of each plant. Future work will refine our shape models and investigate optimal weights for the shape model matching. We are collecting much more imagery and plan on ground truthing it with the help of agronomists so that we can also investigate the use of deep learning methods.

6. REFERENCES

[1] F. Fiorani and U. Schurr, "Future scenarios for plant phenotyping," *Annual Review of Plant Biology*, vol. 64, pp. 267–291, February 2013.

[2] R. T. Furbank and M. Tester, "Phenomics-technologies to relieve the phenotyping bottleneck," *Trends in Plant Science*, vol. 16, pp. 635–644, November 2011.

[3] J. W. White, P. Andrade-Sanchez, M. A. Gore, K. F. Bronson, T. A. Coffelt, M. M. Conley, K. A. Feldmann, A. N. French, J. T. Heun, D. J. Hunsaker, M. A. Jenks, B. A. Kimball, R. L. Roth, R. J. Strand, K. R. Thorp, G. W. Wall, and G. Wang, "Field-based phenomics for plant genetics research," *Field Crops Research*, vol. 133, pp. 101–112, July 2012.

[4] G. A. Johnson, G. P. Cofer, S. L. Gewalt, and L. W. Hedlund, "An engineering approach to image-based phenotyping," *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 381–383, July 2002, Washington, DC.

[5] A. Mutka and R. Bart, "Image-based phenotyping of plant disease symptoms," *Frontiers in Plant Science*, vol. 5, no. 734, 2015.

[6] A. Bucksch, J. Burridge, L. M. York, A. Das, E. Nord, J. S. Weitz, and J. P. Lynch, "Image-based high-throughput field phenotyping of crop roots," *Plant Phisiology*, vol. 166, no. 2, pp. 470–486, October 2014.

[7] R. F. McCormick, S. K. Truong, and J. E. Mullet, "3D sorghum reconstructions from depth images identify QTL regulating shoot architecture," *Plant Physiology*, vol. 172, pp. 823–834, August 2016.

[8] P. Chaudhary, S. Godara, A. N. Cheeran, and A. K. Chaudhari, "Fast and accurate method for leaf area measurement," *International Journal of Computer Applications*, vol. 49, pp. 22–25, July 2012.

[9] C. Lü, H. Ren, Y. Zhang, and Y. Shen, "Leaf area measurement based on image processing," *Proceedings of the International Conference on Measuring Technology and Mechatronics Automation*, pp. 580–582, March 2010.

[10] M. Minervini, M. M. Abdelsamea, and S. A. Tsaftaris, "Image-based plant phenotyping with incremental learning and active contours," *Ecological Informatics*, vol. 23, pp. 35–48, September 2014.

[11] F. Agüera, F. Carvajal, M. Pérez, and F. Orgaz, "Multitemporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop," *Biosystems Engineering*, vol. 132, pp. 19–27, January 2015.

[12] D. J. Watson, "Comparative physiological studies on the growth of field crops: I. variation in net assimilation rate and leaf area between species and varieties, and within and between years," *Annals of Botany*, vol. 11, pp. 41–76, January 1947.

[13] J. Liu and E. Pattey, "Retrieval of leaf area index from top-of-canopy digital photography over agricultural crops," *Agricultural and Forest Meteorology*, vol. 150, pp. 1485–1490, October 2010.

[14] J. L. Araus and J. E. Cairns, "Field high-throughput phenotyping: the new crop breeding frontier," *Trends in Plant Science*, vol. 19, no. 1, pp. 52–61, 2014.

[15] D. Kelly, A. Vatsa, W. Mayham, L. Ngô, A. Thompson, and T. KazicDerek, "An opinion on imaging challenges in phenotyping field crops," *Machine Vision and Applications*, pp. 1–14, December 2015.

[16] National Research Council, "Sorghum," in *Lost Crops of Africa. Volume I: Grains*, vol. 1, pp. 127–144. The National Academies Press, Washington, DC, 1996.

[17] S. K. Panguluri and A. A. Kumar, "Phenotyping in sorghum," in *Phenotyping for Plant Breeding: Applications of Phenotyping Methods for Crop Improvement*, vol. 1, pp. 73–110. Springer New York, New York, NY, 2013.

[18] J. Ribera, F. He, Y. Chen, A. F. Habib, and E. J. Delp, "Estimating phenotypic traits from UAV based RGB imagery," *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Workshop on Data Science for Food, Energy, and Water*, August 2016, San Francisco, CA (to appear).

[19] E. R. Davies, "Binary shape analysis," in *Machine Vision: Theory, Algorithms, Practicalities connected components*, pp. 161–167. Morgan Kaufmann, San Francisco, CA, 2004.

[20] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Janurary 1979.

[21] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, June 2006.

[22] D. M. W. Powers, "Evaluation: From precision, recall and f-factor to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, December 2011.