https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-438

Point Cloud Based Approach to Stem Width Extraction of Sorghum

Jihui Jin and Avideh Zakhor
University of California, Berkeley

Abstract

A revolution in the field of genomics has produced vast amounts of data and furthered our understanding of the genotype-phenotype map, but is currently constrained by manually intensive or limited phenotype data collection. We propose an algorithm to estimate stem width, a key characteristic used for biomass potential evaluation, from 3D point cloud data collected by a robot equipped with a depth sensor in a single pass in a standard field. The algorithm applies a two step alignment to register point clouds in different frames, a Frangi filter to identify stem-like objects in the point cloud and an orientation based filter to segment out and refine individual stems for width estimation. Individually, detected stems which are split due to occlusions are merged and then registered with previously found stems in previous camera frames in order to track temporally. We then refine the estimates to produce an accurate histogram of width estimates per plot. Since the plants in each plot are genetically identical, distributions of the stem width per plot can be useful in identifying genetically superior sorghum for biofuels.

1. Introduction

Understanding the genotype-phenotype map is important to address growing challenges such as ensuring global food security and moving towards cleaner energy through biofuel production. By connecting specific genotypes to phenotypes, geneticists and breeders can select for more robust and higher yielding crops. Recent advances have led to cheaper and more accurate genotyping techniques providing a large variety of strains for breeders to select and screen. However, current phenotyping technologies represent a bottleneck. They are limited to manual measurements out in the field. Manual measurements are too labor intensive and time consuming and do not provide nearly enough data to supplement the genotypes available. Minervini *et. al.* [1] discusses the challenges in phenotyping using image analysis such as rapid plant growth, change in shape and size, or the non-static nature during data collection.

Various high throughput phenotyping platforms have recently been developed to automate the data collection of phenotypic traits such as plant width and height. These platforms integrate a variety of imaging sensors including thermal infrared imaging, fluorescence imaging, 3D imaging and tomographic imaging. Li *et. al.*[6] provide an extensive survey of the different types of sensors, their current applications, and their advantages and limitations towards phenotyping. These sensors can also be used for other tasks such as navigation through fields and segmentation of plants from ground as in Weiss's and Biber's[2] work.

One recent imaging technology that shows promise is the time-of-flight (ToF) sensor. ToF cameras provide 3D information of an entire scene at speeds of up to 160 frames per second. They are often spatially limited with lower resolution than traditional imaging techniques, but are rapidly improving. Klose *et. al.* [3] conducted a study of ToF sensors for automatic plant phenotyping and confirmed their usability in outdoor conditions.

Algorithms have been developed to extract phenotypic information from the 3D data collected by these imaging sensors. Terrestrial laser scanners have been used for stem extraction algorithms of trees[10][11], but trees tend to provide a larger surface with fewer branches. Other systems use multiple cameras or multiple views of the same plant to reconstruct complete 3D models which are then processed for phenotypic information [5][8][12]. Others still approach the problem using 2D data, such as Amean *et. al.* [7] who take advantage of the Frangi filter and Hough transform to segment the stems or Haug *et. al.* [9] who apply machine vision techniques.

Rather than segmenting individual plants, Bao *et. al.*'s[4] pipeline estimates the width on a series of homogeneous plants arranged in a plot. Their algorithm also allows for measurements to be taken in the field rather than greenhouses and laboratory settings, but it requires modifications such as increasing the spacing between plant rows to reduce occlusion .

Our goal is similar to Bao *et. al.*[4] in that we automate width estimation of crops in the field for sorghum as a source of biofuels. Our work focuses on automatic width estimation of sorghum plants using a point cloud generated by a ToF sensor mounted on a robot as it passes by the plant. Our algorithm is designed such that the field requires no modification to accommodate the robot. A typical field is divided into 3×3 meter plots. Each plot contains sorghum sharing the same genotype and is divided into four ranges as seen in Figure 1. Seeds are planted roughly 6 cm apart.

The robot traverses between the middle two ranges with the camera only mounted on one side facing the base of the stalks. The sorghum crops have little space between them so a single ToF camera frame often contains more than one stem which may or may not overlap with others as seen in Figure 8. The robot only sees one side of the crops so it also does not provide a 360° view of the plant which would simplify the task of isolating and estimating the width of stems.

Our approach is designed to take into account the above constraints. Specifically, we segment the various stems within a single frame, filter out extraneous structures such as leaves and debris, estimate the width from a partial incomplete view, and then register the various stems between frames. By tracking isolated stems, we can refine the width estimation and provide a more accurate histogram representing all stems detected within a typically homogeneous plot. Currently, most plots are measured manually with often less than five measurements per plot. Our algorithm

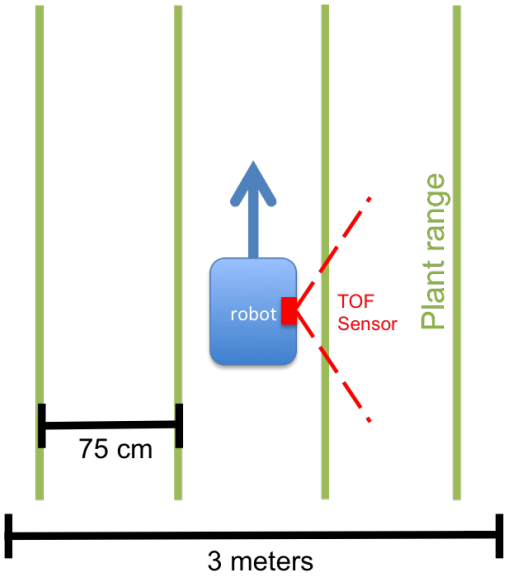


Figure 1. *A top down view of the plot. A single plot consists of 4 ranges of sorghum crops. A robotic platform with a ToF sensor mounted on the base traverses between two ranges.*

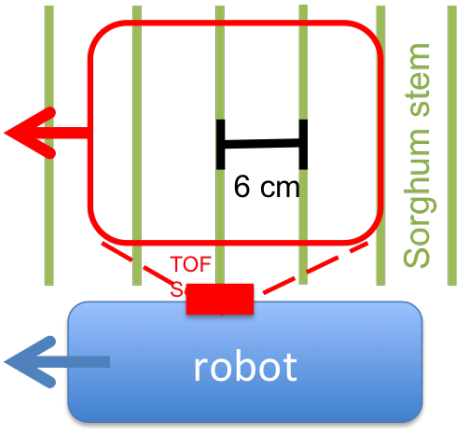


Figure 2. *A sideways view of the robot traversing past a single range. Sorghum crops are spaced 6 cm apart with the ToF sensor mounted directly on the side of the robot. The ToF sensor captures a limited FOV.*

measures more plots, provides more data points per plot, and allows for more measurements throughout the growing season with minimal human effort.

We organize the rest of the paper in the following way: First we discuss the details of our algorithm in Section 2. Next we explain the experimental setup designed to test the accuracy of our algorithms in Section 3. Then we analyze the results of our work in Section 4. Concluding remarks are in Section 5.

2. Tracking and Estimation Pipeline

Our proposed pipeline uses a recording of a single range of plants within a plot collected by the robot as seen in Figures 1 and 2. An unordered point cloud with each point represented by its (x,y,z) coordinates is extracted from each frame of the ToF camera and passed into the pipeline. As a pre-processing step, a Frangi filter[13] is applied to every point cloud generated from a frame. The Frangi parameters are designed to enhance vessel like structures within the point cloud, which typically preserves points corresponding to the stems. An adaptive threshold is applied to only keep points with a high response as seen in Figure 3. Orientation information generated from the Frangi filter is also kept for use later in the pipeline as seen in Figure 4. In summary, three forms of data are associated with frame i of the ToF camera: an unordered point cloud from the ToF sensor P_i, a Frangi cloud F_i, and the orientations O_i which point in the direction of the vessel of each point in the Frangi cloud .

Our pipeline consists of four major steps for each frame of the ToF camera: registration, stem extraction, tracking across frames, and width estimation are described in Sections 2.1, 2.2, 2.3 and 2.4 respectively. Every unique stem found during the extraction process is assigned an ID j. Every frame that contains stem j as determined by the temporal registration step has an associated estimated width $w_{i,j}$. As a post-processing step, every $w_{i,j}$ is reduced to a single estimate W_j and outputted as a histogram

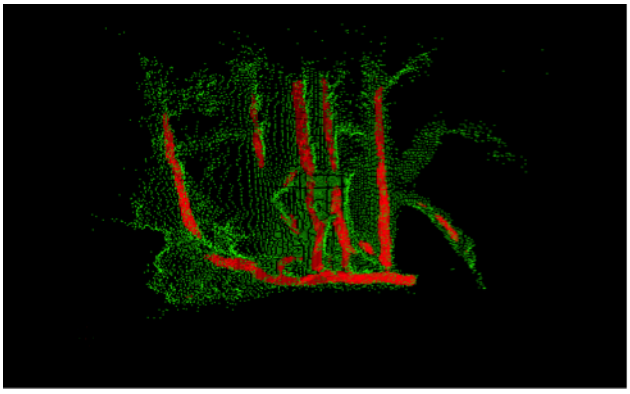


Figure 3. *Output of Frangi filter(red) overlaid on original point cloud(green) generated from a single frame of the ToF camera.*

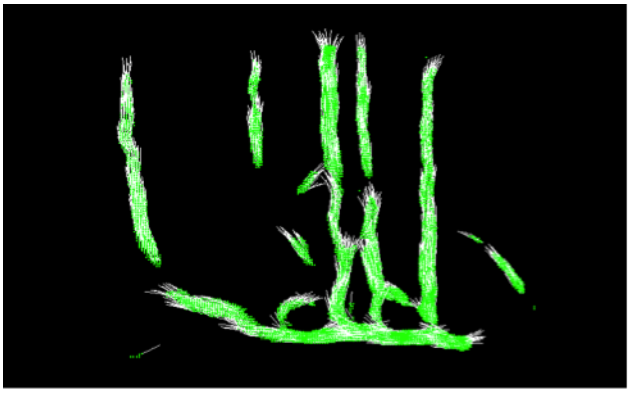


Figure 4. *Frangi point cloud(green) with vessel orientation(white) overlaid.*

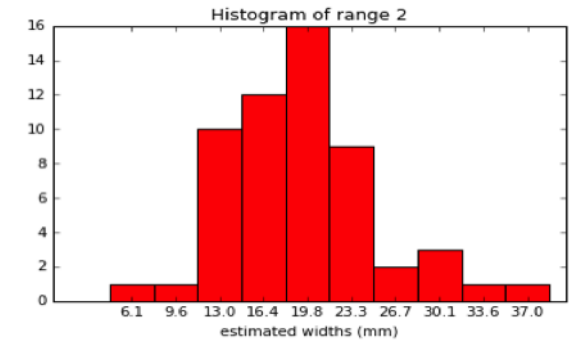


Figure 5. *An example histogram of width estimations generated from a single range.*

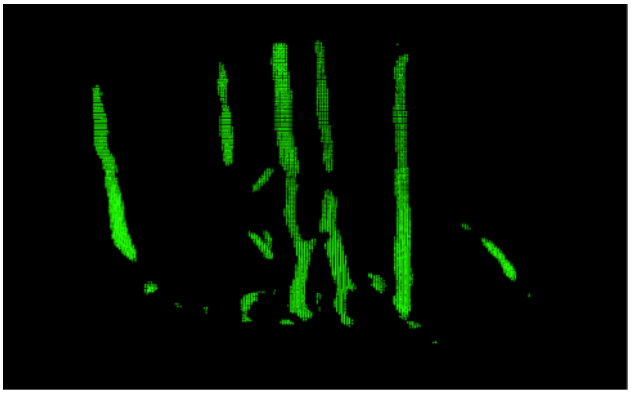


Figure 6. *Frangi cloud after points with horizontal orientations are removed.*

as seen in Figure 5. Tracking stems temporally prevents multiple measurements of the same stem from polluting the statistical information of a plot, so the histogram is an accurate representation of the biomass statistics.

2.1 Registration

Every point cloud produced from a single frame of the ToF camera represents a small slice of the physical plant range. The Frangi cloud F_i exists in the same frame as point cloud P_i based on the physical location of the camera during data collection. However, the clouds F_i and P_i, do not share the same coordinate system as P_j and F_j where $j \neq i$. A physical stem exists in a subset of the point clouds generated from the frames. In order to reason about the same stem across the mulitple frames and their corresponding point clouds, we must first register the point clouds to a single coordinate system.

In order to register the clouds P_i to P_{i-1}, we perform two transformations $T_{1,i}$ and $T_{2,i}$. The first transformation $T_{1,i}$ is a translation determined by the vector $\overrightarrow{C_i C_{i-1}}$ where C_i denotes the centroid of point cloud P_i. The robot in which the camera is mounted moves at a low enough velocity such that frames i and $i-1$ typically have large overlaps. Thus, the translation of P_i to $T_{1,i}(P_i)$ as determined by the centroids is sufficient to provide an inital aligment with P_{i-1}.

We then determine $T_{2,i}$ using Iterative Closest Point(ICP) to refine the alignment of $T_{1,i}(P_i)$ to P_{i-1}. Transformations $T_{1,i}$ and $T_{2,i}$ are then applied to F_i, which exists in the same frame as P_i, to register it as well with the global coordinate frame as defined by the first frame. Thus $T_{1,1}$ and $T_{2,2}$ are both merely the identity matrix.

2.2 Stem Extraction

The Frangi filter retains all vessel like points. This includes the stems as well as unnecessary parts such as leaves. Every point in F_i has a corresponding orientation in O_i. We take advantage of the vertical orientation of stems by removing points in F_i whose orientations are close to horizontal. Specifically, all points in F_i whose corresponding orientations in O_i lie closer to the horizontal plane than the vertical axis can be removed as seen in Figure 6.

Gaps between clusters of points in F_i are introduced as a result of removing points with horizontal orientations. We then use a flood-fill algorithm to segment the various clusters in F_i. We define a cluster as a group of points where every point is at least within a certain distance of another point in the cluster. Clusters that consist of less than a minimal number of points are removed as noise.

It is important to make the distinction between clusters and stems. A stem is an actual physical plant. A cluster is a collection of points in point clouds F_i or P_i that corresponds to the stem. We design the radius of the flood fill algorithm such that each cluster only corresponds to a single stem, but a single stem may consist of multiple clusters.

Each remaining cluster often still contains points associated with leaves which were not removed in the horizontal points removing filter. However, a majority of the points within a cluster have roughly the same orientation. We use this to refine a cluster further. Specifically, we construct a 2D histogram of the orientation vectors by first projecting them onto the ground plane and then binning them. We then locate the bin with the highest count and keep all points within its small neighborhood in the cluster.

Inspecting Figure 3, we note that the width of the Frangi points in F_i corresponding to a single stem does not necessarily match up with the width of the points in P_i corresponding to the same stem. This is because the Frangi filter is only designed to preserve vessel like structures, not biomass information. Thus, we need to use the cluster of points from F_i to generate a corresponding cluster of points in P_i in order to accurately estimate the width. To do this, we perform a radial based search of all points in P_i centered around every point in each cluster in F_i to select relevant points in P_i. The resulting clusters can be seen in Figure 7. Note that clusters can correspond to the same stem as seen by the yellow and purple clusters. We address this issue in the next step.

2.3 Tracking Across Frames

Often stray leaves or other horizontal occluding objects cut a stem into multiple parts. An unordered set based merging algorithm combines all clusters from the same stem. Each cluster begins as its own set. All clusters that are determined to come from the same stem by our match criteria have their sets merged. We represent every cluster k_α^i in frame i of the ToF camera as a line ℓ_α^i determined by its centroid c_α^i, and a vector representing the average orientation o_α^i in order to determine whether two clusters belong to the same stem as seen in Figure 9. For the remainder of this section, we assume all clusters belong to frame i and thus drop the superscript.

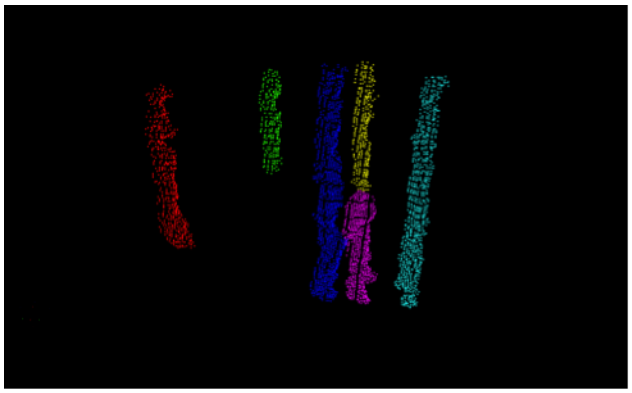


Figure 7. *Frangi cloud clustered and refined into stem candidates are then used to extract corresponding stem points from the original point cloud.*

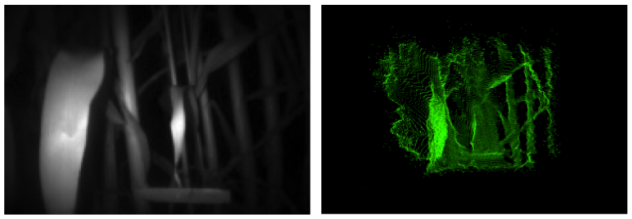


Figure 8. *(a) Infrared image from the ToF camera representing the field of view with minor occlusion. (b) Unordered point cloud generated by the ToF camera.*

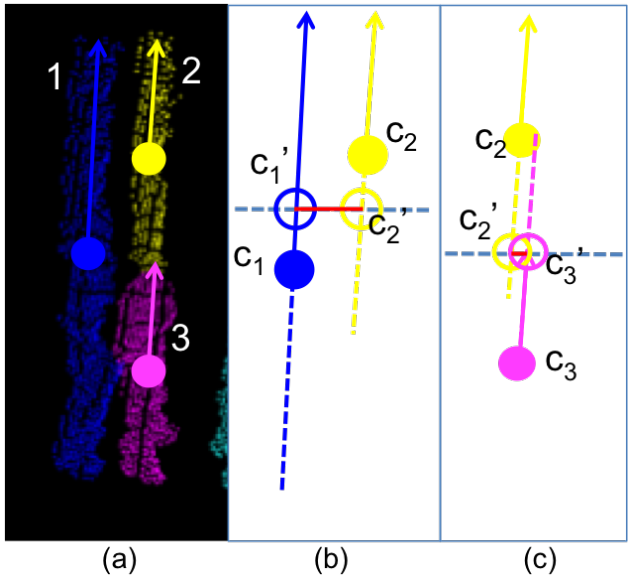


Figure 9. *(a) Three potential clusters undergoing the merging algorithm. A line representing the vector as determined by their centroids and average orientation vectors is overlaid. (b) A 2D example where cluster 1 (blue) and cluster 2 (yellow) do not satisfy the distance criteria. Although the orientation of the vectors align, the distance between the points c_1' and c_2' on the plane of comparison is too large. (c) An example of a successful match between cluster 2 (yellow) and cluster 3 (pink). The distance between the points c_2' and c_3' on the plane of comparison satisfies the distance criteria and thus clusters 2 and 3 are determined to lie on the same physical stem.*

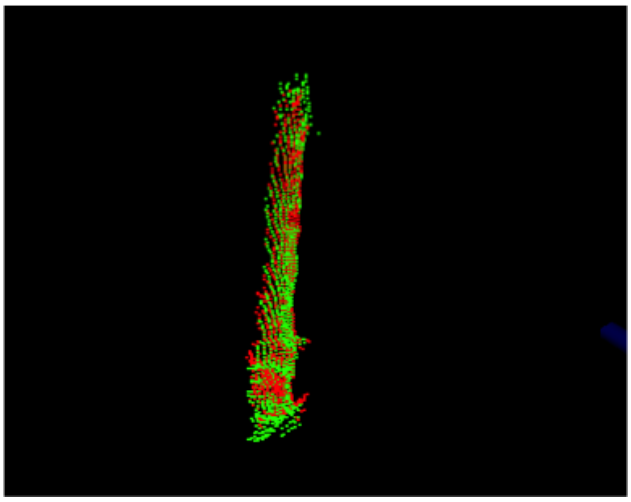


Figure 10. *Two corresponding stems from different frames matched and assigned the same stem ID.*

Two clusters k_α and k_β existing in 3D space that belong to the same physical stem must lie on or close to the same line. We compare the direction of ℓ_α and ℓ_β, the lines representing k_α and k_β respectively, using the average orientations. Specifically, the angle between o_α and o_β must be below a certain threshold. In order to determine whether ℓ_α and ℓ_β occupy the approximately same position in space, we first determine a plane for comparison. We find the points c'_α and c'_β that lie on ℓ_α and ℓ_β respectively and intersect a plane parallel to the ground offset at the distance $\frac{c_{\alpha,y}+c_{\beta,y}}{2}$ where $c_{\alpha,y}$ and $c_{\beta,y}$ represent the distance of the centroids of k_α and k_β from the ground respectively. The offset distance of the plane is determined in order to minimize errors introduced by inaccuracies from the sensor and Frangi filter. We then determine the distance between c'_α and c'_β as the error metric. If the error in distance between c'_α and c'_β exceeds a certain threshold, the clusters k_α and k_β occupy a different region in the point cloud and thus do not belong to the same physical stem. Two clusters k_α and k_β must satisfy the orientation and distance criteria in order to be considered a match. An example of a successful and a failed match based on the distance metric is shown in Figure 9. For the case of the clusters shown in Figure 7, only the yellow and purple clusters are considered a match and thus merged.

Once merged, stem candidates in frame i must then be registered with stems found in previous frames. An example of two corresponding stems can be seen in Figure 10. Each physical stem found thus far has a unique ID assigned to it. We only consider a subset of these stems active if they are likely to exist in the current frame i. We compare the stem candidates in frame i to the list of all active stems collected from frames 1 to $i-1$. Correspondence is determined with the same match criteria as the merging algorithm using the line determined by the average orientation and centroid. Correspondence is determined through a greedy matching algorithm comparing c^i_α and o^i_α with c^j_β and o^j_β and assigning ID j to the cluster k^i_α on successful matches. Any remaining clusters are determined to be new stem candidates and assigned a unique stem ID accordingly.

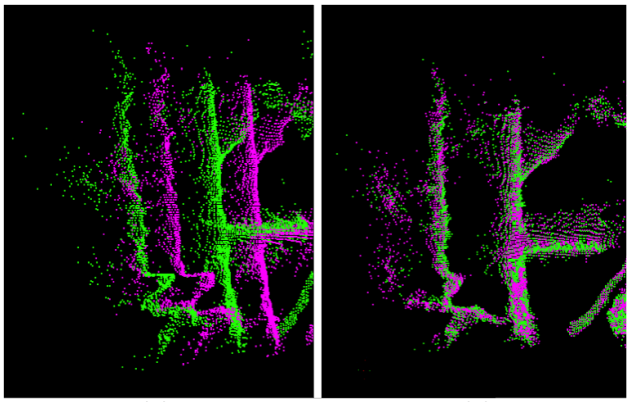


Figure 11. *Example of the registration step. (a) The pink and green represent two point clouds, one registered to the global frame and the other extracted from the ToF sensor. (b) The point clouds after undergoing our registration process*

2.4 Width Estimation

Width is defined as the maximum distance between points in the horizontal direction. Each cluster is aligned with the vertical axis based on the average orientation of the associated Frangi cluster. The width is then estimated at multiple increments along the vertical axis. Outliers are removed and remaining estimates are averaged and stored according to the stem ID.

3. Experimental Set Up

We collected data from a series of sorghum plots located at the University of Illinois at Urbana-Champaign. Plots are arranged as described in Section 1 and shown in Figures 1 and 2. A robot platform mounted with the PMD pico sensor, a commercial time-of-flight camera, traversed between ranges of a plot while recording at 45 frames per second. The sensor is placed approximately 15-20 cm away from the plants. Every frame yields an infrared image and a point clouds around the base of the plants as seen in Figure 8. The ToF sensor is limited to a resolution of 160×120 with a small FOV of roughly $82^\circ \times 66^\circ$, so only a small portion of the plant is visible.

To validate the width estimation accuracy, we tagged and manually measured the width of 100 plants distributed across 10 different plots to use as the ground truth. Ten separate recordings were taken of the plots with tags were placed within the FOV of the ToF camera so that correspondence between estimated values and ground truth measurements could be established.

We ran our algorithm on all 10 videos. The algorithm automatically detects all stems within a video and outputs an associated estimated width. We then manually determined which estimated measurement generated by our algorithms corresponded to which tagged plants with ground truth measurements using the infrared images and point clouds in order to evaluate accuracy.

4. Results

We present examples of the intermediate steps of the algorithm as well as the final results when compared to ground truth. In Figure 11, we show an example of the registration process. The cloud C_{i-1} in green in Figure 11a represents a the previous cloud registered to the global coordinate frame. The pink cloud C_i is the point cloud extracted from the current frame. Figure 11b displays the results of our two-step registration process.

In Figure 12, we show two examples of the stem extraction process on a single frame. Figures 12a and 12d represent the infrared images seen by the ToF sensor. Figures 12b and 12e represent the points clouds C_i and Frangi clouds F_i. Figures 12c and 12f represent the result of our stem extraction step prior to the cluster merging algorithm. The registration step is done on a per cluster basis, or individual stems, rather than whole frames. An example of the correspondence between two extracted clusters can be seen in Figure 10.

In order to evaluate the accuracy of the width estimation step, we use the tagged stems with available ground truth measurements. Of the 100 tagged stems, only 95 were found within the FOV of the ToF camera when inspecting the infrared images manually. Missing tags were either occluded by large leaves or placed too low/high. Of the 95 tags found, 88 were automatically detected throughout the tracking process. Undetected stems failed to show up during the stem extraction process and thus had no associated estimated width.

Percent error is defined as the ground truth minus estimated width divided by ground truth. Note that this accounts for both under and over estimation as seen in Figure 13. Absolute percent error is the absolute value of the error from ground truth. Estimated measurements from the 88 detected plants achieved an average of 0.50% error and 11.5% absolute error when compared to the manual measurements as ground truth.

Typically we capture around 20-30 cm out of the 3 m of the range in the plot within the FOV of the ToF camera. Each plant stem occupies roughly 2-3 cm of that range. The generated point cloud from a single frame has anywhere between 8,000 to 15,000 points. Each stem extracted by our algorithm from a single frame corresponds to between 200 to 1,000 points.

Individual estimates and errors can be seen in Figure 14. Our estimated width tends to follow the general trend of manual measurements such as around tags 69-75. It seems to do more poorly with particularly narrow stems as seen around tags 8, 19, and 25 which have a width of around 1.5 cm instead of the average 2 cm. This could be partially due to the low resolution of the ToF camera being unable to capture the narrow stems or overestimations due to the pipeline combining touching neighboring stems. The average absolute width error is only 2.38 mm. The PMD pico is rated at a depth resolution of less than 3 mm at a distance of 50 cm, so we are also approaching the maximum depth resolution. An example of the histogram outputted by our algorithm for a single range can be seen in Figure 5.

The variation of estimates across frames in the form of box and whisker plots can be found in Figure 15, which is very close to the final estimate used by our algorithm. The thinner lines, or whiskers, denote the minimum and maximum width estimated. The red line overlaid on the box and whisker plots shows the corresponding ground truth measurements. We note that certain plants have large variation across frames such as numbers 7, 34, 76 and 88. However, there does not seem to be a correlation between error and large variation as plant numbers 7, 34 and 88 show less than 10% error as seen in Figure 13. Only plant number 76 shows significant error with our algorithm heavily overestimating its width. On the flip side, there are instances of very

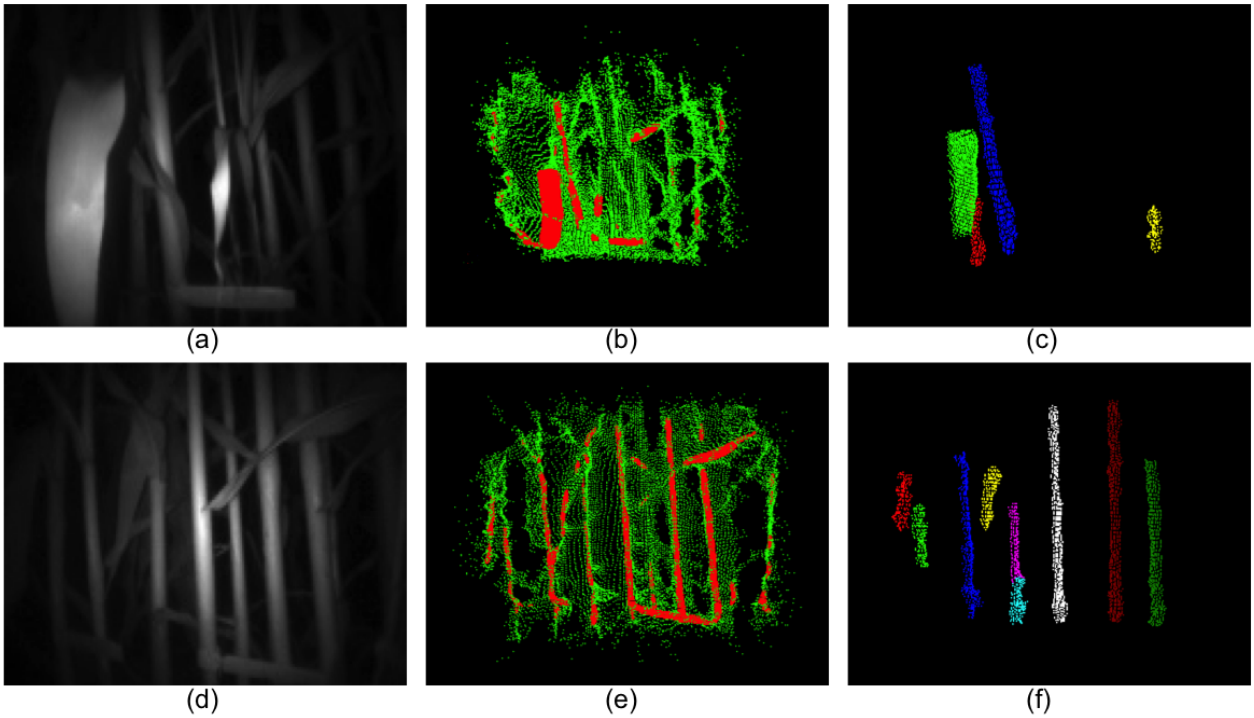


Figure 12. *(a)-(c) and (d)-(f) represent two examples of the stem extraction process. (a), (d): infrared images generated by ToF camera; (b), (e): corresponding point cloud and Frangi cloud; (c), (f) the extracted clusters prior to the merging algorithm.*

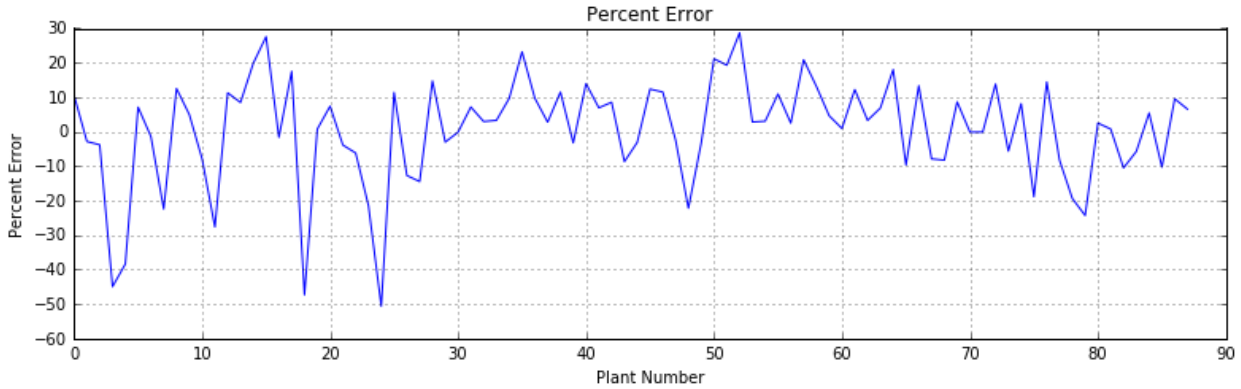


Figure 13. *Graph of the percent error of the estimated width of 88 tagged plants when compared to manual measurements.*

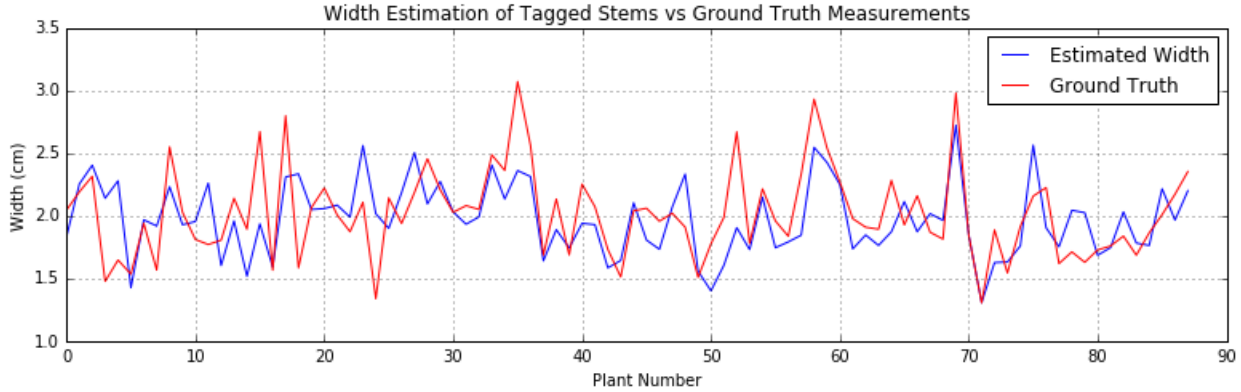
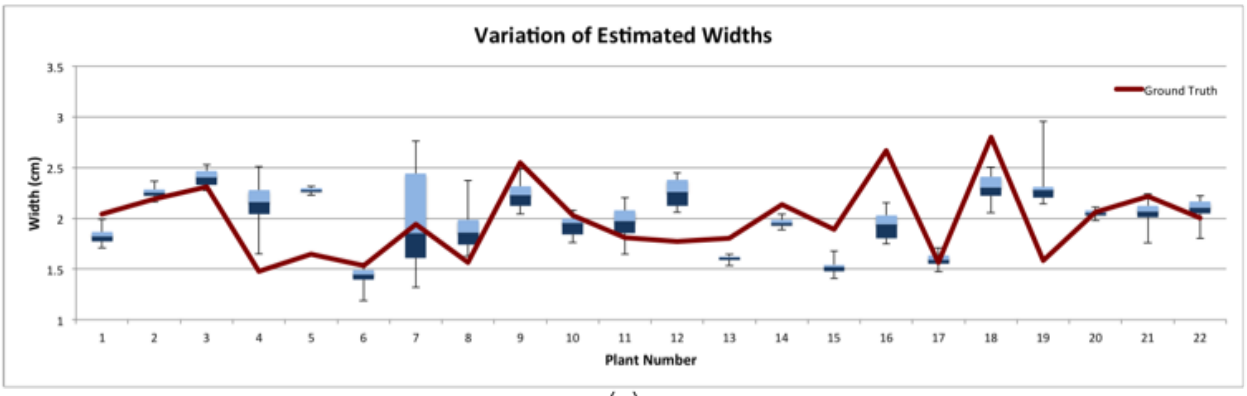
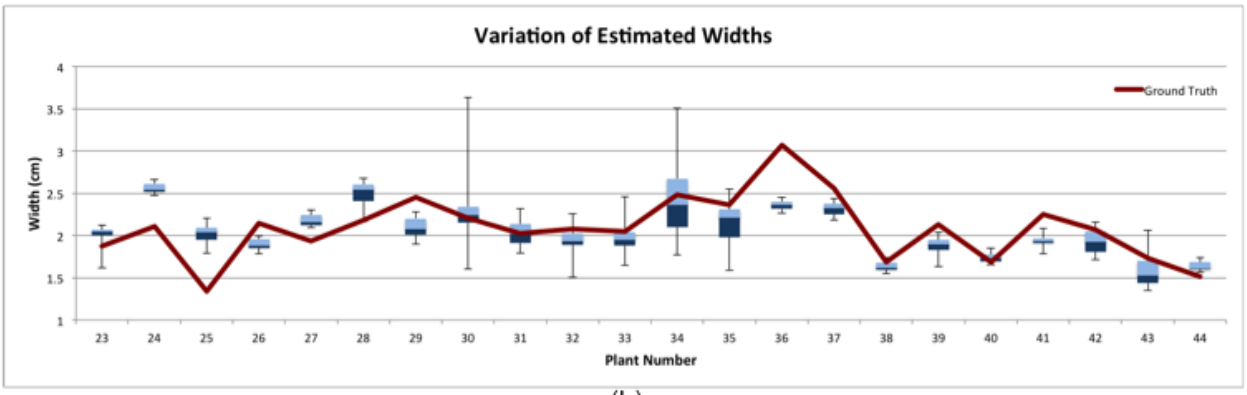


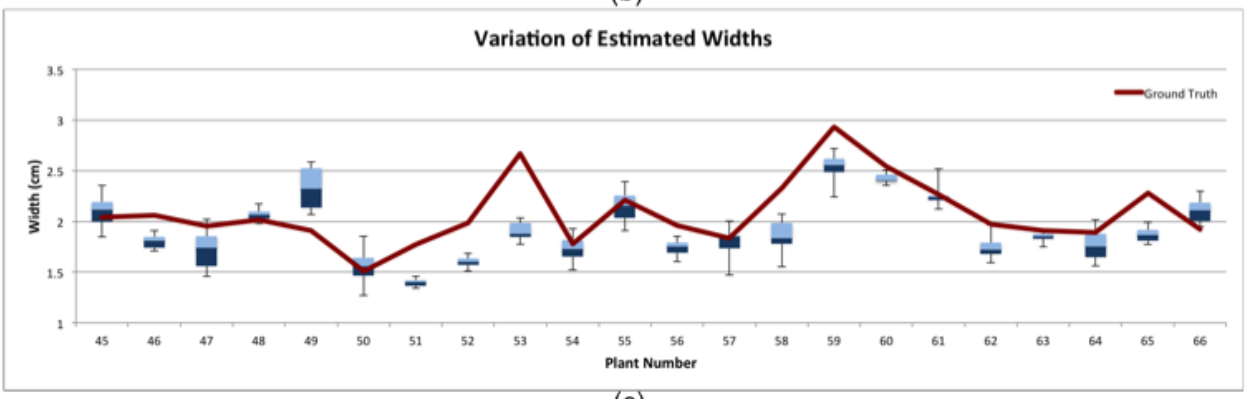
Figure 14. *Estimated width compared against the ground truth of 88 tagged plants.*



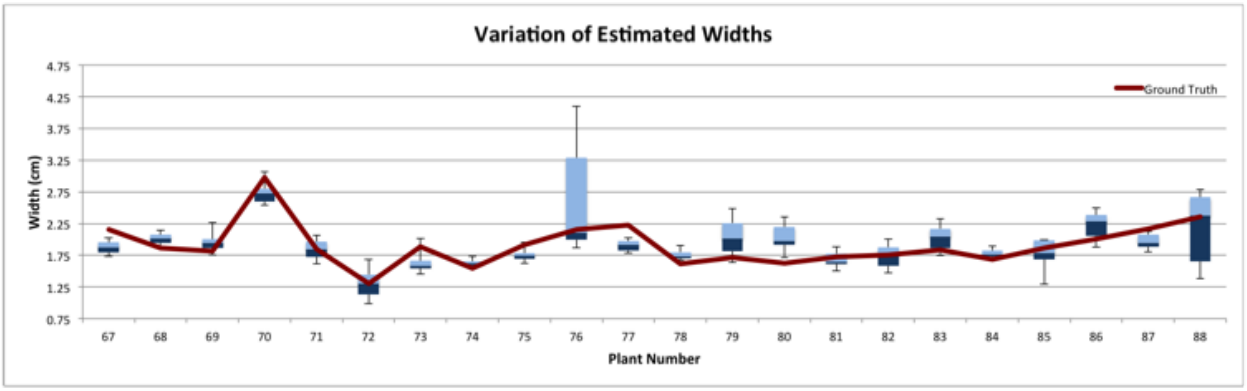
(a)



(b)



(c)



(d)

Figure 15. *A series of box and whisker plots generated from all estimated widths of every tagged plant with ground truth overlaid. The bottom edge of the dark blue box and the top edge of the light blue box denote the 25th percentile and 75th percentile respectively. The shared edge between the dark and lighter shade of blue denotes the median of all estimates.*

low variation and high % error such as plant numbers 5, 25, 36, 65, 67, and 73.

5. Conclusion

We have shown a pipeline for segmenting, estimating and then tracking stems within a plot. The pipeline is fully automatic, requiring no manual effort. The width estimation has been shown to achieve under 1% signed error and under 12% absolute error. Each plot yields over 50 estimates, a significant improvement on the sub 5 measurements typically taken manually.

Our algorithm is designed to run automatically without modifications to existing fields minimizing the obstacles for large scale adoption by breeders. By producing significantly more data throughout the growth cycle, our pipeline will improve our understanding of the genomes key to high-yield biomass sorghum crops to aid in the transition towards biofuels. This work can easily be extended towards other crops such as corn.

Acknowledgments

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-AR0000598. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

[1] Massimo Minervini, Hanno Scharr, Sotirios A. Tsaftaris, Image analysis: the new bottleneck in plant phenotyping [applications corner], IEEE Signal Processing Magazine 32, no. 4, pg 126-131 (2015). Harvard

[2] Ulrich Weiss, Peter Biber, Plant detection and mapping for agricultural robots using a 3D LIDAR sensor., Robotics and autonomous systems, 59.5, 265-273 (2011).

[3] Ralph Klose, Jaime Penlington, Arno Ruckelshausen, Usability study of 3D time-of-flight cameras for automatic plant phenotyping, Bornimer Agrartechnische Berichte 69.93-105, 12 (2009).

[4] Yin Bao, Lie Tang, Patrick S. Schnable, Maria G. Salas Fernandez, Infield Biomass Sorghum Yield Component Traits Extraction Pipeline Using Stereo Vision, 2016 ASABE Annual International Meeting, American Society of Agricultural and Biological Engineers, pg 1. (2016). APA

[5] Anthony Paproki, Xavier Sirault, Scott Berry, Robert Furbank, Jurgen Fripp, A novel mesh processing based technique for 3D plant analysis, BMC Plant Biology, 12.1, 63 (2012).

[6] Lei Li, Qin Zhang, Danfeng Huang, A review of imaging techniques for plant phenotyping, Sensors, 14.11, 20078-20111 (2014).

[7] Z. Mohammed Amean, Tobias Low, Cheryl McCarthy, Nigel Hancock, Automatic plant branch segmentation and classification using vesselness measure., Proceedings of the Australasian Conference on Robotics and Automation (ACRA 2013). Australasian Robotics and Automation Association (2013).

[8] Gerie van der Heijden, Yu Song, Graham Horgan, Gerrit Polder, Anja Dieleman, Marco Bink, Alain Palloix, Fred van Eeuwijk, Chris Glasbey, SPICY: towards automated phenotyping of large pepper plants in the greenhouse, Functional Plant Biology, 39.11, 870-877 (2012).

[9] Sebastian Haug, Peter Biber, Andreas Michaels, Jrn Ostermann, Plant Stem Detection and Position Estimation using Machine Vision.

[10] Pyare Pueschel, Glenn Newnham, Gilles Rock, Thomas Udelhoven, Willy Werner, Joachim Hill, The influence of scan mode and circle fitting on tree stem detection, stem diameter and volume extraction from terrestrial laser scans, ISPRS journal of photogrammetry and remote sensing, 77, 44-56 (2013).

[11] Gabor Brolly, Gza Kirly, Algorithms for stem mapping by means of terrestrial laser scanning, Acta Silvatica et Lignaria Hungarica, 5, 119-130 (2009).

[12] J. C. Noordam, J. Hemming, C. Van Heerde, F. B. T. F. Golbach, R. Van Soest, E. Wekking, Automated rose cutting in greenhouses with 3D vision and robotics: analysis of 3D vision techniques for stem detection, In International Conference on Sustainable Greenhouse Systems-Greensys2004 691, pp. 885-892 (2004).

[13] Alejandro F. Frangi, Wiro J. Niessen, Koen L. Vincken, Max A. Viergever, Multiscale vessel enhancement filtering, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer Berlin Heidelberg, 1998, pg. 130-137.

Author Biography

Jihui Jin received his BS in Electrical Engineering Computer Science at the University of California Berkeley (2016). He is currently a research scientist at Signetron Inc, focusing on 3D signal processing. In the past, he has worked in the Media Integrated Communication Lab at Osaka University and in the Swarm Lab at Berkeley. His research interests are augmented reality, 3D sensors, and computer vision.