



2950 Niles Road, St. Joseph, MI 49085-9659, USA
269.429.0300 fax 269.429.3852 hq@asabe.org www.asabe.org

An ASABE Meeting Presentation

DOI: 10.13031/aim.20162462338

Paper Number: 162462338

Infield Biomass Sorghum Yield Component Traits Extraction Pipeline Using Stereo Vision

Yin Bao, Lie Tang, Patrick S. Schnable, Maria G. Salas-Fernandez

Iowa State University

**Written for presentation at the
2016 ASABE Annual International Meeting
Sponsored by ASABE
Orlando, Florida
July 17-20, 2016**

ABSTRACT. *Yield component traits such as plant height and stem diameter are dominant phenotypic data for biomass sorghum yield prediction. Extraction of these traits by machine vision during the growing season significantly reduces labor and time cost for large breeding programs. An automated 3D point cloud processing pipeline was developed to quantify different phenotypic variations in plant architecture of infield biomass sorghum. The input point cloud was generated by three side-view stereo camera heads placed vertically to capture extremely high plants. The features were extracted on a row plot basis instead of individual due to severe occlusion caused by densely populated leaves. Available features include plant height, plant width, vegetation volume index, and vegetation area index. Our strategy was to slice the point cloud along row direction into several equal volume slices and sum up the feature values with weights based on the point population and distribution in each volume slice. Therefore, the results were robust against empty space and abnormal individuals in the row plot. In addition, a semi-automated user interface was developed for users to measure stem diameters from the stereo images according to their specific sampling strategies. Users only need to zoom in on a stem segment and pick four corners of the rectangular segment. Metric measurement is then computed automatically based on image patch stereo matching using normalized cross correlation. The extracted stem diameters were compared to manual measurements in the field and a high correlation was obtained. The extracted features revealed great potential for automated field-based high-throughput phenotyping for plant architecture.*

Keywords. biomass sorghum, yield component traits, field phenotyping, stereo vision, 3D point cloud processing.

The authors are solely responsible for the content of this meeting presentation. The presentation does not necessarily reflect the official position of the American Society of Agricultural and Biological Engineers (ASABE), and its printing and distribution does not constitute an endorsement of views which may be expressed. Meeting presentations are not subject to the formal peer review process by ASABE editorial committees; therefore, they are not to be presented as refereed publications. Citation of this work should state that it is from an ASABE meeting paper. EXAMPLE: Author's Last Name, Initials. 2016. Title of presentation. ASABE Paper No. --. St. Joseph, MI: ASABE. For information about securing permission to reprint or reproduce a meeting presentation, please contact ASABE at <http://www.asabe.org/copyright> (2950 Niles Road, St. Joseph, MI 49085-9659 USA).

1. Introduction

Sorghum has promising potential to become one of the most productive bioenergy crops in the US because of its drought tolerance and high biomass yield potential (Rooney et al., 2007). Therefore, increasing biomass yield is a major breeding objective for biomass sorghum. It has been shown that plant architecture traits such as plant height (Lubberstedt et al., 1997; Salas Fernandez et al., 2009) and leaf angle (Morinaka et al., 2006) are highly correlated to sorghum biomass yield. For investigating genetic control of these traits, genome-wide association study (GWAS) is routinely used nowadays (Zhao et al., 2016). However, GWAS requires phenotypic data of a large amount of various genetic lines. Measuring phenotypic data by hand is extremely laborious and time-consuming.

Automated high-throughput field phenotyping is the key technology to scale up GWAS experiments and discover more genetic mechanisms of phenotypic variations. In the last decade, a few advanced automated high-throughput field phenotyping systems have emerged. BoniRob was an autonomous field robotic platform for early stage plant phenotyping (Ruckelshausen et al., 2009). It had four independently steerable drive wheels and adjustable chassis clearance from 0.4 m to 0.8 m. Color camera, 3D ToF camera, hyperspectral imaging system, laser distance sensor and light curtain were fused to measure plant population, crop density, inter-row spacing, plant height, and stem thickness (Klose et al., 2010). Another field-based high-throughput phenotyping platform was developed and evaluated at the University of Arizona (Andrade-Sanchez et al., 2014). This platform was modified based on an open rider sprayer, which had a maximum clearance of 1.93 m. A boom was used to support the sensors in the front, allowing measurement of four crop rows simultaneously. The sensors involved were sonar proximity sensor, infrared radiometer sensor and multi-spectral crop canopy sensor for canopy height, canopy temperature, and canopy reflectance respectively. Blue River Technology (CA, USA) developed Zea, a ground-based system for early season (V2-V5) corn plant phenotyping. It used a tunnel pulled by a tractor to block sunlight. 3D scanners were mounted inside the tunnel to collect RGB 3D point clouds. Nonetheless, none of the above systems are suitable for sorghum architecture traits characterization throughout the entire growing season because some sorghum genotypes can grow up to 13 feet. LemnaTec GmbH (Germany) provides a stationary gantry system which can move a sensor platform from 3 m to 6 m above the ground. The major limitations of such system include limited field size and high cost. Drone-based imaging systems have also been developed for high-throughput field phenotyping over the recent years. Regarding plant architecture traits, plant height, and leaf area index (LAI) can be extracted. However, aerial imaging suffers from occlusion caused by top canopies. Therefore, LAI from aerial imaging does not consider the leaf area below the visible canopies. On the other hand, side-view 3D LiDAR scanning of orchard trees has been adopted recently for individual tree identification (Underwood et al., 2015), tree trunk detection (Bargoti et al., 2015), and leaf area density estimation (Sanz et al., 2013). Side-view 3D profile proved to reveal more information about the plant geometry and canopy size at all levels.

Compared to LiDAR, the advantages of stereo camera include higher spatial resolution and lower price. It takes time for LiDAR system to scan a scene depending on the vehicle travel speed and LiDAR scan speed. Dynamic scene can cause inaccurate 3D point cloud or even blur the scene. Wind is a common field condition causing large movement of plant canopy. This situation is even worse for biomass sorghum where some genotypes have long and broad canopies in addition to extreme height. Stereo vision can capture a 3D scene instantaneously. It is more suitable for scanning biomass sorghum. Active 3D sensors such as Time-of-Flight (ToF) camera suffer more from sunlight than passive stereo camera does. ToF camera is better suited for night imaging.

In this study, three stereo camera heads were placed vertically for capturing side-view 3D point cloud of infiel biomass sorghum. An automated feature extraction processing pipeline was developed to extract several features which quantify different phenotypic variations in plant architecture. The available features include plant height, plant width, vegetation volume index, and vegetation area index. They are calculated based on a row plot basis instead of individual plant due to the severe occlusion caused by densely populated leaves. Our approach breaks the input point cloud into slices along the direction of plant row. Each volume slice is analyzed in terms of point population and spatial distribution to determine its weight for the final feature value. There are two reasons for the slicing approach. First, abnormal individual plant should not affect the overall plot-based feature. Sometimes in the same genotype row plot, a particular plant would grow taller than the average. Deciding the plant height from all the slices in a weighted way would reduce the bias towards the single tall plant. Second, empty space caused by non-emergent plants in the row plot should not contribute to the plot-based plant volume. Empty space can be identified by examining the shape of point cloud in each slice.

Stem diameter is also an important component trait for sorghum biomass yield. Sampling representative stems in a row plot requires decisions from a trained personnel. A semi-automated user interface was developed for users to measure stem diameter in the stereo images. The user would be asked to zoom in on a stem segment in the image. The projection of the stem segment is modeled by a convex quadrangle. The user only needs to pick four points on the stem edges, two points on the one side and the rest on the other side. The metric scale of the four corner points is recovered by stereo matching the image patch containing the quadrangle to its correspondence in the second stereo image. The matching cost is evaluated by Normalized Cross-Correlation (NCC) as it is robust against illumination difference between the stereo camera pair which is often expected under field lighting condition.

The extracted stem diameters were compared to manual measurements in the field. High correlations were obtained. The

automatically extracted features were demonstrated as well as the runtime performance, both showing great potential of the proposed methods for high-throughput field phenotyping for plant architecture.

2. Materials and Methods

2.1 Field Design

The field experiment was carried out in the summer of 2014 at Agricultural Engineering and Agronomy Research Farms of Iowa State University. 700 diverse sorghum lines were planted at two locations with each location having two repetitions. Commercial planting row spacing is 30 inches. However, the experimental trial in 2013 showed that at least 90-inch row spacing was necessary to avoid most long leaves of some sorghum genotypes blocking the stereo cameras. Therefore, 90-inch row spacing was used. Each row plot was 10 feet long, consisting of one genotype.

2.2 Data Acquisition System

In 2013, a robotic ground vehicle was developed for automatically collecting stereo images in the field (Bao et al., 2014). A compact utility tractor was retrofitted with a commercial auto-guidance system which enables the tractor to self-drive following a pre-recorded path by user. Therefore, the vehicle could traverse the entire field without human interaction. The auto-guidance system outputs GPS NMEA strings at 10 Hz. GPRMC string was used to log each imaging locations beforehand and trigger cameras during data acquisition. The imaging location map only needs to be created once for the entire growing season.

Point Grey Grasshopper GRAS-20S4C-C color camera was used to build the stereo camera (Fig. 1). This camera model allows multiple units to be daisy-chained via IEEE-1394b bus with a 125-microsecond synchronization accuracy. The imaging sensor has a resolution of 1624x1224. The image is approximately 2MB in 8-bit RAW format. The compatible lens of 6 mm focal length was used to obtain a view angle of 62.1 degrees.



Figure 1. The stereo camera heads and rotatable rig.

A rotatable camera rig attached in front of the tractor (Fig. 2) was designed to support the stereo camera heads as well as adjust the camera-to-plant distance. Moreover, stereo images of two rows were taken for one pass. An extension rig was attached to the rotatable rig when the plants grew taller than the field of view (FOV) provided by the middle-level stereo camera heads. This configuration allowed a vertical FOV of 9 feet and a horizontal FOV of 5 feet. Therefore, half of each 10-foot row plot was imaged.



Figure 2. Phenotyping robot without and with extension rig.

2.3 Automated Plot-based Feature Extraction Pipeline

The input 3D point cloud for our feature extraction pipeline was the fusion of three point clouds from the three stereo camera heads on the same side. For generating the point cloud from each stereo camera head, Semi-Global Matching (Hirschmüller et al., 2008) was adopted for its high efficiency and robust performance in practice. OpenCV, the most developed open-source computer vision library, provides a highly parallelized implementation, Semi-Global Block Matching (SGBM), to achieve near real-time performance by extensively using instruction level parallelism on modern CPU. Therefore, SGBM was used in our experiment with its default parameters. The stereo camera heads on the same side were calibrated and the point clouds from the middle and top stereo cameras were transformed into the coordinate system of the bottom stereo camera.

Our processing pipeline was developed by using Point Cloud Library (PCL) (Rusu et al., 2011). The input 3D point cloud is first downsampled with VoxelGrid (Rusu, 2010) filter. The input point cloud was partitioned by a 3D voxel grid (Fig. 3). A 3D voxel can be viewed as a 3D box. All the points in each voxel were approximated by their centroid. VoxelGrid filter serves as an effective way to regularize point cloud density and speed up subsequent processing. StatisticalOutlierRemoval (Rusu et al., 2008) filter was then applied to remove sparse outliers. StatisticalOutlierRemoval computes the distance from each point to all its neighbors and removes points whose mean distance to its neighbors are outside the interval chosen by a user-defined mean distance and standard deviation.

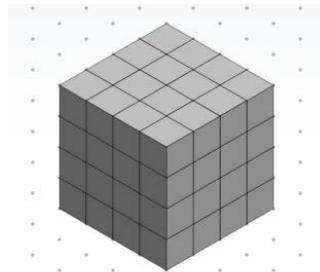


Figure 3. 3D voxel grid.

Plant growth plane is defined as the plane that minimizes the distances from sorghum stems in the point cloud to itself. Our bottom stereo cameras were not installed parallel to the plant growth plane as shown in Fig. 4 where plant bases do not form a horizontal line in the image. For cropping out the plants of interest accurately, it is necessary to align roughly the plant growth direction with x -axis, plant row direction with y -axis with a predefined rotation matrix. The plants in background rows are removed based on z value. For our field design, any point whose z value is more than 1700 mm was removed.



Figure 4. Left images of three stereo cameras imaging the same row plot.

Next, Axis-Aligned Bounding Box (AABB) was extracted. Each edge of an AABB is aligned with one of the axes of the world coordinate system (Fig. 5). Therefore, an AABB can be defined by two of its vertices, P_{min} (X_{min} , Y_{min} , Z_{min}) and P_{max} (X_{max} , Y_{max} , Z_{max}). P_{min} is the vertex whose 3D coordinates minimum and P_{max} the vertex whose 3D coordinates are maximum. Given an input point cloud, P_{min} is determined by the minimum values in three dimension of all points and P_{max} the maximum values. The extracted AABB served as an initial bounding box. Following process would refine the AABB such that the absolute difference of X_{max} and X_{min} measures plot-based plant height and the absolute difference of Z_{max} and Z_{min} measures plot-based plant width.

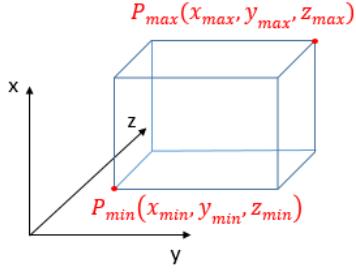


Figure 5. Axis-Aligned Bounding Box.

Soil ground is removed by changing X_{min} to a point that best separates plants from the ground. It is observed that the plant bases and the ground are mostly perpendicular to each other. Their intersection serves as a reasonable reference for measuring plant height. A template point cloud of right angle plate was placed roughly at the intersection of plant bases and ground (Fig 6).

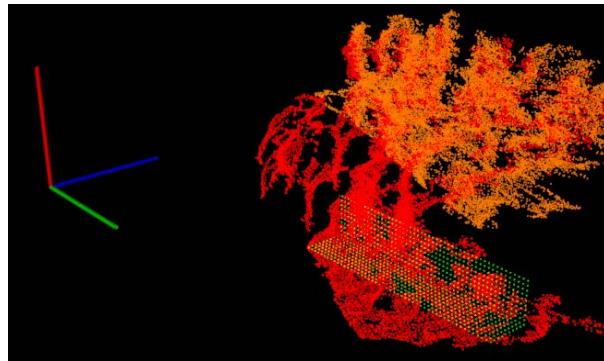


Figure 6. Right angle plate template point cloud (green).

Due to the unevenness of the ground, the auto-guidance system kept correcting the heading error, causing the camera-to-plant distance to vary within a 100 mm radius. A refinement is needed to align the template with target. Iterative Closest Point (ICP) (Zhang, 1994) algorithm was used for the refinement. ICP is an iterative process. In each iteration, for each point in the source point cloud, ICP finds the closest point in the target point cloud and estimates a rigid-body transformation to best align the each source point to its match. We limited the correspondence search radius to 300 mm to prevent the template from matching places other than the plant bases. Once the alignment was done, the points below the ground plane estimated by the template were removed. To account for the unevenness of soil ground and weed, the template ground plane was rotated around the base-ground intersection axis by 10 degrees such that the soil clumps and weed were below the plane. Then the points below the rotated ground plane were filtered. After this stage, it is assumed that the resultant point cloud only contains sorghum plants in the row plot of interest. Finally, X_{min} is set to the center point on the base-ground intersection edge (Fig. 7).

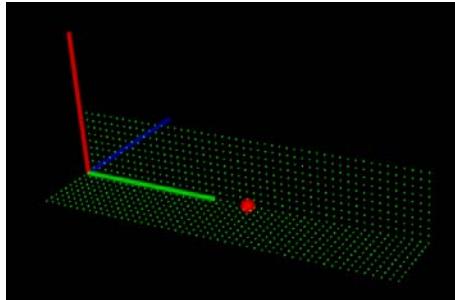


Figure 7. Template point cloud of right angle plate and the center point on the base-ground intersection edge (red sphere).

X_{max} also needs adjustment since it does not necessary represent the average plant height. Sometimes an abnormal plant would grow taller than the average. Our solution is to break the point cloud along the row direction (y-axis) into N_{slice} slices. In each slice i , the maximum x value of all points is denoted by x_{max}^i and the associated weight by w^i which is computed as the ratio of the slice population to the total population of the input point cloud. Weight w^i effectively reduces the contribution of the slice which contains empty space. The refined X_{max} equals the weighted median of x_{max}^i . Weighted median is obtained by first sorting x_{max}^i and then finding the first x_{max}^k satisfying $\sum_{i=1}^k w^i \geq 0.5$. Therefore, plant height is computed as

$$PH = X_{max} - X_{min}. \quad (1)$$

The same approach was used to refine Z_{min} and Z_{max} . Hence, plot-based plant width is

$$PW = Z_{max} - Z_{min}. \quad (2)$$

Note that the stereo cameras only see half side of the plants and the extracted plant width measures the half of the canopy span along the direction (z -axis) perpendicular to the plant growth plane.

Given the refined AABB, any point outside the AABB is removed. To quantify canopy volume, convex hull has been used to measure vegetation volume (Azzari et al., 2013). Given a set of 3D points, the convex hull is the smallest convex set that contains all points (Fig. 8). For any two points in a convex hull, the line segment connecting them is also contained in the convex hull. PCL uses Qhull (Barber et al., 1996) to compute convex hull.

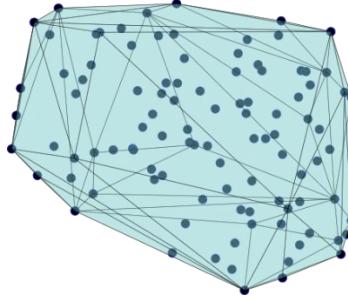


Figure 8. 3D convex hull.

Directly applying Qhull to our remaining vegetation point cloud is not accurate due to the possible empty space in the row plot. We use the same slicing strategy to identify slices that contain no plant. The AABB is broken into n equal sub-boxes along row direction. A convex hull is constructed for the points in each slice. The volume ratio of the convex hull to its corresponding sub-box indicates the vegetation occupancy in the sub-box. If the ratio is lower than a threshold α , the sub-box and the convex hull inside are marked as invalid. Then we obtain m valid sub-boxes. Denote β as the ratio of m to n . β is the correcting factor for effective ground area of the AABB. Thus, the effective ground area is calculated as

$$\text{Effective Ground Area (EGA)} = \beta(Y_{max} - Y_{min})PW \quad (3)$$

and the effective bounding box volume

$$\text{Effective Bounding Box Volume (EBBV)} = EGA \cdot PH. \quad (4)$$

We define total vegetation volume as the sum of the volumes of the m valid convex hulls

$$\text{Total Vegetation Volume (TVV)} = \sum_{i=1}^m \text{Volume}_{chull}^i. \quad (5)$$

Furthermore, we define vegetation volume index as the ratio of total vegetation volume to effective cubic volume

$$\text{Vegetation Volume Index (VVI)} = \frac{\text{Total Vegetation Volume}}{\text{Effective Bounding Box Volume}}. \quad (6)$$

Next, we quantify the vegetation surface area. To measure surface area, the point cloud needs to be converted to a surface presentation. Triangle mesh is widely used for its simplicity and efficiency (Fig. 9). PCL provides the GreedyProjectionTriangulation (Marton et al., 2009) algorithm for surface reconstruction. Before applying GreedyProjectionTriangulation, we use MovingLeastSquares (Rusu et al., 2011) filter in PCL to smooth out high-frequency components in our point cloud data since the canopies are normally smooth surface patches.

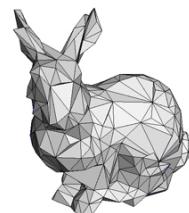


Figure 9. Triangle mesh surface representation.

Given the three vertices $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ of a triangle, the triangle area can be computed as

$$\text{Triangle Area (TA)} = \frac{|(\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_3 - \mathbf{p}_1)|}{2}. \quad (7)$$

The total vegetation area is approximated by the sum of areas of all triangles in the mesh

$$\text{Total Vegetation Area (TVA)} = \sum_{i=1}^N TA_i \quad (8)$$

where N denotes the number of triangles in the mesh. The total vegetation area includes canopies and stems. It includes panicles when the plants are mature. Moreover, we define vegetation area index as

$$\text{Vegetation Area Index (VAI)} = \frac{\text{Total Vegetation Area}}{\text{Effective Ground Area}}. \quad (9)$$

2.4 Semi-automated Stem Diameter Extraction

A user interface was developed for researchers to measure stem diameter. The user would be shown three images as in Fig. 4. Then the user can zoom in on a stem segment by roughly clicking on it. Only four reference points need to be picked to obtain the stem diameter (Fig. 10).

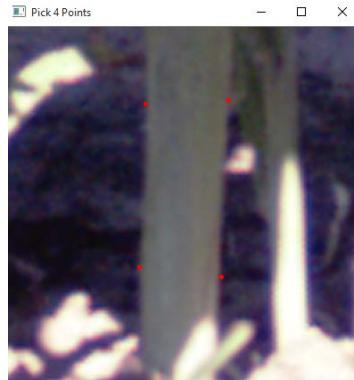


Figure 10. Four reference points on the stem edges chosen by the user.

The user may pick these points in any order as long as they can form a convex quadrangle. Next, the points are sorted such that $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4)$ corresponds to the top left corner, bottom left corner, top right corner and bottom right corner in the image coordinate system (Fig. 11).

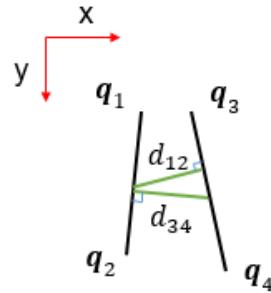


Figure 11. Sorted reference points and stem diameter estimation in the image coordinate system.

Let \mathbf{q}_{12} denote the middle point between \mathbf{q}_1 and \mathbf{q}_2 . Let \mathbf{q}_{34} be the middle point of \mathbf{q}_3 and \mathbf{q}_4 . The distance from \mathbf{q}_{12} to the line passing through \mathbf{q}_3 and \mathbf{q}_4 can be computed as

$$d_{12} = \frac{|(\mathbf{q}_3 - \mathbf{q}_4) \times (\mathbf{q}_4 - \mathbf{q}_{12})|}{|\mathbf{q}_3 - \mathbf{q}_4|}. \quad (10)$$

Similarly the distance from \mathbf{q}_{34} to the line passing through \mathbf{q}_1 and \mathbf{q}_2 is

$$d_{34} = \frac{|(\mathbf{q}_1 - \mathbf{q}_2) \times (\mathbf{q}_2 - \mathbf{q}_{34})|}{|\mathbf{q}_1 - \mathbf{q}_2|}. \quad (11)$$

The stem diameter in the image coordinate system is estimated by

$$d = \frac{d_{12} + d_{34}}{2}. \quad (12)$$

The stem diameter D in metric unit (for instance, millimeter) is calculated as

$$D = \frac{bd}{d_{shared}} \quad (13)$$

where b baseline between the two cameras forming the stereo camera head and d_{shared} the shared disparity of $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{q}_4)$. b can be obtained through camera calibration. d_{shared} . The triangulation principle is shown in Fig. 12.

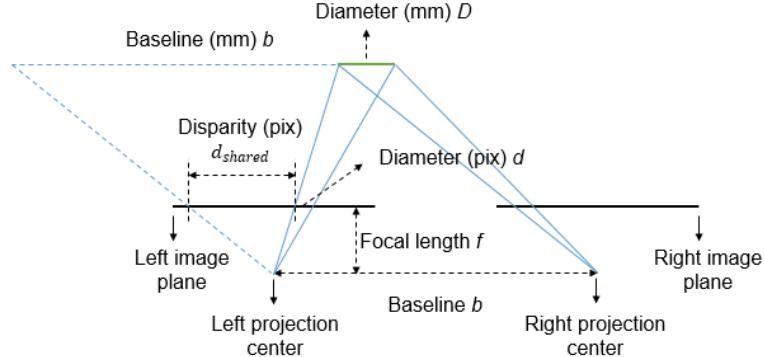


Figure 12. Triangulation principle for measuring stem diameter.

The shared disparity d_{shared} is calculated by matching the image patch containing the stem segment in the reference (left) image of the stereo image pair to its correspondence in the target (right) image. Since the stem segment is relatively small compared to the full image size and there is not enough resolution to reconstruct the curved surface on the stem segment, it is reasonable to assume that $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ and \mathbf{q}_4 lie on the same the plane fronto-parallel to the camera projection plane. In other words, the depth values of all four reference points to the camera pair are equal. The stem segment patch matching is evaluated by using Normalized Cross-Correlation for it is well-known for robustness against radiometric differences in real-world images (Hirschmüller et al., 2007). The desired shared disparity is the one of minimum aggregated matching costs among all possible values:

$$d_{shared} = \underset{d \in \varepsilon}{\operatorname{argmax}} m(P, d) \quad (14)$$

where ε denotes the possible disparity range which can be obtained through stereo camera calibration and P the image patch formed by the quadrangle of stem segment. The aggregated costs for matching the patch P given a disparity d are computed as

$$m(P, d) = \frac{1}{n} \sum_{q \in P} \frac{(rp(q) - \bar{rp})(tp(q, d) - \bar{tp})}{\sigma_r \sigma_t} \quad (15)$$

where n denotes the number of pixels in image patch P , q denotes a pixel in P . $rp(q)$ computes the intensity of pixel q in the reference P . \bar{rp} and σ_r are the average intensity and standard deviation of P . $tp(q, d)$ computes the intensity of the pixel determined by the spatial position of q and the horizontal offset disparity d . \bar{tp} and σ_t are the average intensity and standard deviation of the target image patch P' .

3. Results and Discussion

The results were computed on a desktop workstation with a 3.5GHz Xeon HexaCore CPU. The stereo images were first downsampled by a factor of 0.5 with the final resolution of 812x612. Our voxel size of VoxelGrid downsampling was 0.01x0.01x0.01 m. The default settings in PCL were used for StatisticalOutlierRemoval filter.

3.1 Plant Height and Plant Width

The plot-based plant height (PH) and plant width (PW) were extracted with the volume slice occupancy weighted median approach. The only parameter is the number of volume slices, N_{slice} . We demonstrate how varying N_{slice} affects the result and how to decide a reasonable value. If N_{slice} equals 1, the resultant PH measures the maximum height of the whole point cloud, which may not be representative. Increasing N_{slice} linearly decreases the thickness of the each volume slice. With the volume slice becoming thinner, the probability of including lower canopies increases and the estimated PH shrinks. The PH tends to show a slow decrease with N_{slice} increasing when N_{slice} is larger than 3 (Fig. 12).

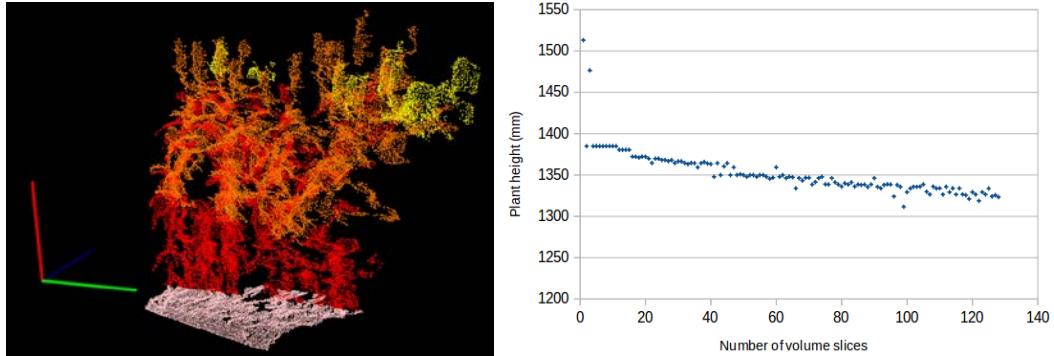


Figure 12. Sample 3D point cloud (left) and plant height vs number of volume slices (right).

Choosing a reasonable N_{slice} depends on plant grow stage and intra row spacing. Before flouring, PH is normally measured from the ground to the whorl. The final three to four leaves are mostly taller than the whorl. A large N_{slice} would help to reduce the difference between actual PH and our estimated PH. After flouring, the sorghum head is taller than the leaves. Ideally there should be one plant in each volume slice. Then our method would measure the exact PH of each plant. The ratio of row plot length to average intra row spacing is a reasonable reference for N_{slice} . However, choosing different values of N_{slice} during the growing season is difficult because different sorghum genetic lines may have various timings of flouring. Therefore, a fixed N_{slice} was used to process our data of all growth stages. Taking into account both early stage and mature stage, the number of volume slices is calculated as

$$N_{slice} = \left\lceil 2 \frac{\text{Row Plot Length}}{\text{Average Intrarow Spacing}} \right\rceil \quad (15)$$

where $\lceil \cdot \rceil$ denotes ceiling function. The goal is to capture half plant in each volume slice. Since our row plot in the image was 1.5 m and average intrarow spacing about 0.15 m, we used $N_{slice} = 20$ for our dataset. Fig. 13 and Fig. 14 demonstrates the PH and PW estimation as the height and width of AABB for early stage and mature stage. They also shows that our estimation was robust against empty space and abnormal individual plant in the row plot.

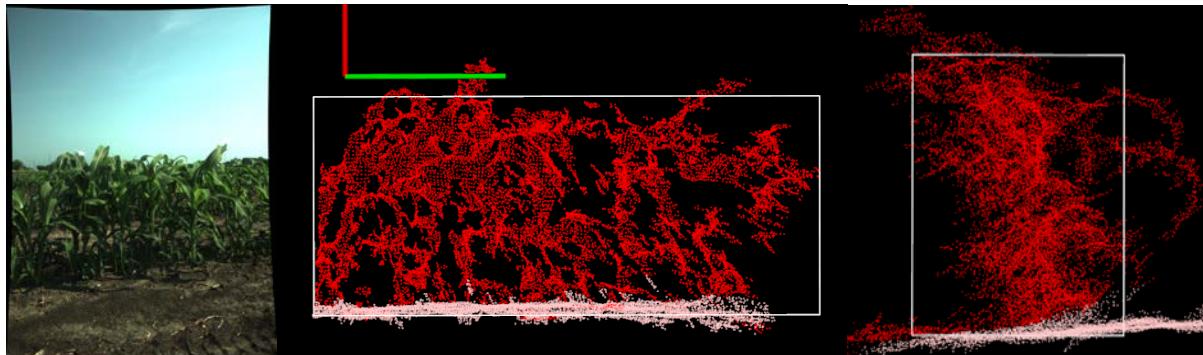


Figure 13. Plant height and width estimation at an early stage. Left: RGB image. Middle: Front view of the 3D point cloud. Right: Side view of the 3D point cloud.

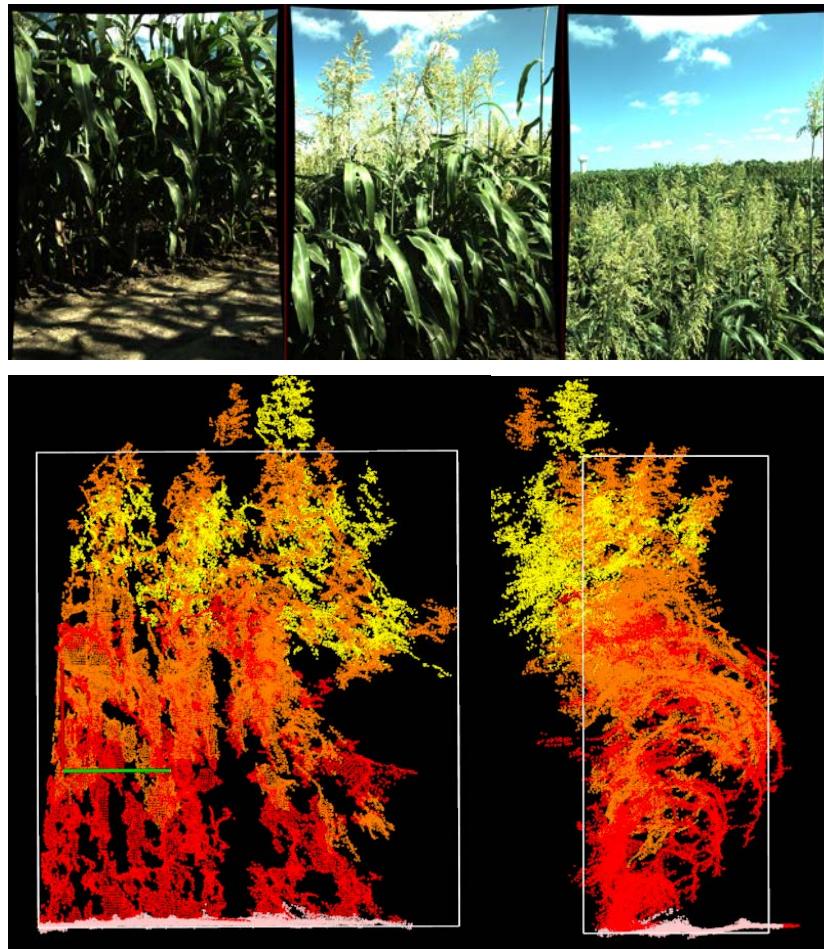


Figure 14. Plant height and width estimation at a mature stage. Top: RGB image. Bottom left: Front view of the 3D point cloud. Bottom right: Side view of the 3D point cloud.

The runtime performance for *PH* and *PW* extraction is illustrated in Table 1. The algorithm visits each point once for volume slicing and once for searching height/width in each volume slice. Weighted median computation is linear to N_{slice} which is far less than the total number of points. Therefore, the algorithm complexity is $O(n)$ where n is the point cloud size.

Table 1. PH and PW extraction runtime for Fig. 13 and Fig. 14.

| Point cloud ID | Points | Time |
|----------------|--------|-----------|
| Fig. 13 | 15026 | 0.017 sec |
| Fig. 14 | 75905 | 0.082 sec |

3.2 Vegetation Volume Index

Our *VVI* requires two parameters, N_{slice} and volume ratio threshold α . The same N_{slice} in *PH* and *PW* estimation was used in our *VVI*. During experiment, it was found that 0.3 is a proper value for volume ratio threshold α , meaning that a valid volume slice should contain a set of points whose convex hull volume is larger than 30 percent of the slice volume. Fig. 15 shows the valid convex hulls extracted from an input point cloud. The empty space in the point cloud was successfully detected showing no convex hull in the corresponding volume slices.

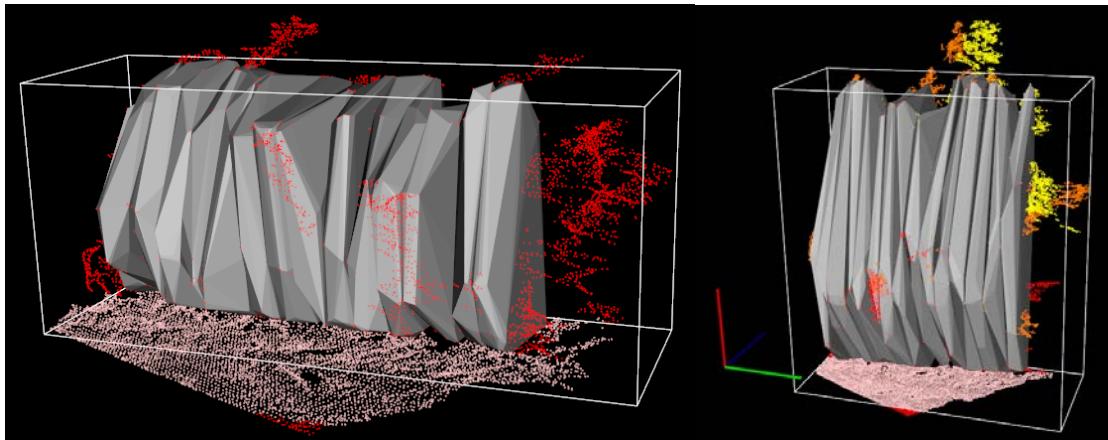


Figure 15. Valid convex hulls for VVI calculation. Left: Fig. 13. Right: Fig. 14.

The empty space is caused by insufficient points not able to form a plant shape. First, there could be no actual plant in the volume slice. The points belong to the leaves of nearby plants. Second, there is a plant in the volume slice. However, due to occlusion, the base of the plant stem is not visible in the view. Our VVI computation discards those invalid volume slices such that the resultant index is invariant to irregular population density in the row plot.

Table 2 shows the runtime performance for VVI computation.

Table 2. VVI extraction runtime performance for Fig. 13 and Fig. 14.

| Point cloud ID | Points | Time |
|----------------|--------|-----------|
| Fig. 13 | 15026 | 0.007 sec |
| Fig. 14 | 75905 | 0.019 sec |

3.3 Vegetation Area Index

Our VAI computation uses MovingLeastSquares and GreedyProjectionTriangulation for smoothing and triangle mesh generation. The default parameter settings in PCL were used. Fig. 16 shows the reconstructed triangle mesh for Fig.13 and Fig. 14.

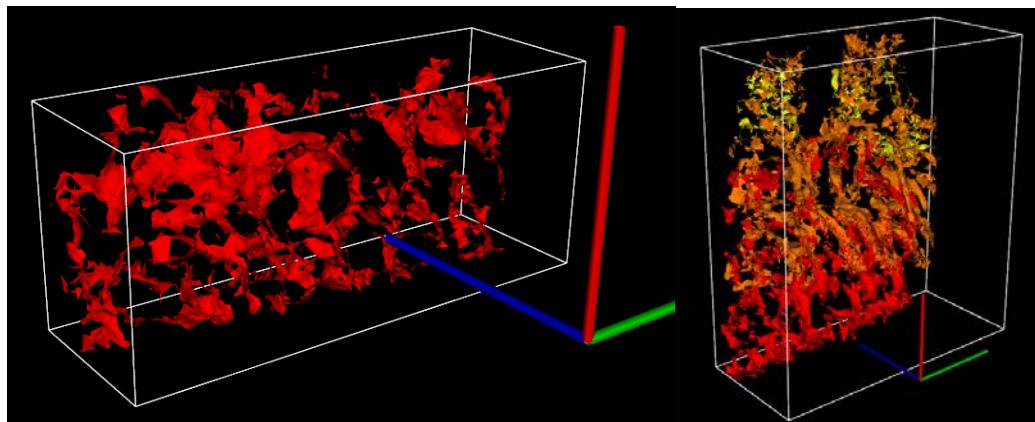


Figure 14. Triangle mesh for Fig. 13 and Fig. 14.

MovingLeastSquares and GreedyProjectionTriangulation are the most time-consuming subroutines in our processing pipeline. Table 3 illustrates the runtime performance of the subroutines in VAI computation.

Table . VAI subroutine runtime performance for different data size.

| Point cloud ID | Points | MovingLeastSquares | Triangulation | Total leaf area |
|----------------|--------|--------------------|---------------|-----------------|
| Fig. 13 | 15026 | 0.157 sec | 0.424 sec | 0.030 sec |
| Fig. 14 | 75905 | 0.847 sec | 2.234 sec | 0.153 sec |

3.4 Stem Diameter

There were 39 stem diameters with different sizes that were manually measured with a caliper in the field and marked with red ribbons such that they could be identified in the image. Since the cross section of a sorghum stem is not circular, when measuring the stem diameter in the field with a caliper, we tried to measure the diameter in the row direction as the cameras would capture. Then we used the semi-automatic software to estimate each stem diameter four times and average the results. Fig. 17 shows the comparison between caliper-based and imaging-based approach. A correlation coefficient of

0.958 was achieved.

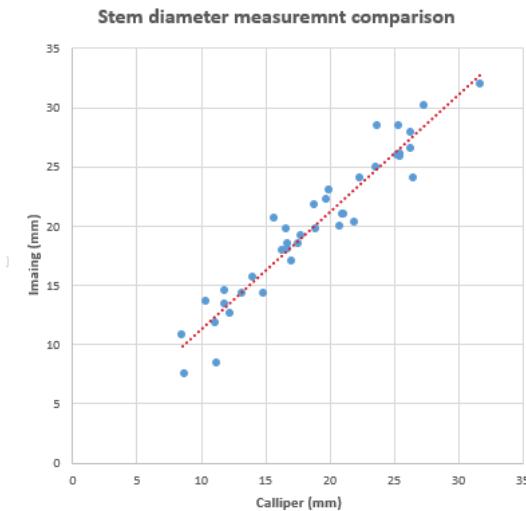


Figure 17. Stem diameter measurement comparison: Imaging (y-axis) vs Caliper (x-axis).

The trend line also shows that the imaging-based approach tended to give larger value than the caliper-based approach did. The intersection of the trend line and y-axis is about 1 mm. One possible explanation would be the disparity error caused by patch-based correspondence matching. We simplified the curved stem surface with a fronto-parallel plane such that the obtained disparity can be shared for any pixel in the image patch. However, the disparity close to the center of the stem segment would have a larger value than that close to the edge because disparity is inversely proportional to depth. The simplified model would result in a smaller disparity. According to Equation 13, the diameter D would become larger. For our experiment, the error was within the required range. One could use more complex models such as slanted plane or even B-spline surfaces. But that would require estimation of more parameters and robustness might be reduced.

To assess the repeatability, when conducting the four measurements for each stem, we intentionally varied the positions of the four reference points on the stem edges to observe the standard deviation (SD) of using this method. The results have shown that the SD of four measurements is less than 1 mm for most of the stems (Fig. 18).

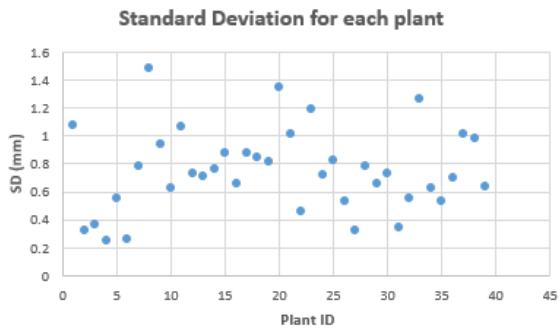


Figure 18. Standard deviation of four stem diameter measurements taken on each stem of a total of 39 plants.

The runtime mostly depended on the image patch stereo matching. Therefore, it is linear to image patch size. For our dataset, the time needed to perform patch matching and stem diameter computing never exceeded 0.1 seconds. This lag was hardly noticed by users.

4. Conclusion and Future Work

In this study, we developed an automated plot-based feature extraction pipeline to characterize some phenotypic variations in plant architecture for infiel biomass sorghum based on stereo vision. We also developed a user-interactive software to measure stem diameter from the stereo images. The extracted features are well suited for high-throughput field-based phenotyping. Furthermore, they are not limited to sorghum. The same approach could be applied to maize and orchard trees.

Deep convolutional neural network has gained a success in objection detection and semantic segmentation in recent years. For future work, we would like to explore the possibility of automated stem diameter extraction with deep ConvNet.

References

- Andrade-Sanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmo-Silva, A. E., French, A. N., ... & White, J. W. (2014). Development and evaluation of a field-based high-throughput phenotyping platform. *Functional Plant Biology*, 41(1), 68-79.
- Azzari, G., Goulden, M. L., & Rusu, R. B. (2013). Rapid characterization of vegetation structure with a microsoft kinect sensor. *Sensors*, 13(2), 2384-2398.
- Bao, Y., Nakami, A. D., & Tang, L. (2014). Development of a Field Robotic Phenotyping System for Sorghum Biomass Yield Component Traits Characterization. In 2014 Montreal, Quebec Canada July 13–July 16, 2014 (p. 1). American Society of Agricultural and Biological Engineers.
- Barber, C. B., Dobkin, D. P., & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4), 469-483.
- Bargoti, S., Underwood, J. P., Nieto, J. I., & Sukkarieh, S. (2015). A Pipeline for Trunk Localisation Using LiDAR in Trellis Structured Orchards. In *Field and Service Robotics* (pp. 455-468). Springer International Publishing.
- Fernandez, M. G. S., Becroft, P. W., Yin, Y., & Lübbertedt, T. (2009). From dwarves to giants? Plant height manipulation for biomass yield. *Trends in plant science*, 14(8), 454-461.
- Hirschmüller, H., & Scharstein, D. (2007, June). Evaluation of cost functions for stereo matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on* (pp. 1-8). IEEE.
- Hirschmüller, H. (2008). Stereo processing by semiglobal matching and mutual information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2), 328-341.
- Klose, R., Möller, K., Vielstädte, C., & Ruckelshausen, A. (2010, March). Modular system architecture for individual plant phenotyping with an autonomous field robot. In *Proceedings of the 2nd International Conference of Machine Control & Guidance* (pp. 299-307).
- Lübbertedt, T., Melchinger, A. E., Schön, C. C., Utz, H. F., & Klein, D. (1997). QTL mapping in testcrosses of European flint lines of maize: I. Comparison of different testers for forage yield traits. *Crop science*, 37(3), 921-931.
- Marton, Z. C., Rusu, R. B., & Beetz, M. (2009, May). On fast surface reconstruction methods for large and noisy point clouds. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on* (pp. 3218-3223). IEEE.
- Morinaka, Y., Sakamoto, T., Inukai, Y., Agetsuma, M., Kitano, H., Ashikari, M., & Matsuoka, M. (2006). Morphological alteration caused by brassinosteroid insensitivity increases the biomass and grain production of rice. *Plant Physiology*, 141(3), 924-931.
- Rooney, W. L., Blumenthal, J., Bean, B., & Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*, 1(2), 147-157.
- Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., & Weiss, U. (2009). BoniRob—an autonomous field robot platform for individual plant phenotyping. *Precision agriculture*, 9(841), 1.
- Rusu, R. B., Marton, Z. C., Blodow, N., Dolha, M., & Beetz, M. (2008). Towards 3D point cloud based object maps for household environments. *Robotics and Autonomous Systems*, 56(11), 927-941.
- Rusu, R. B. (2010). Semantic 3D object maps for everyday manipulation in human living environments. *KI-Künstliche Intelligenz*, 24(4), 345-348.
- Rusu, R. B., & Cousins, S. (2011, May). 3d is here: Point cloud library (pcl). In *Robotics and Automation (ICRA), 2011 IEEE International Conference on* (pp. 1-4). IEEE.
- Sanz, R., Rosell, J. R., Llorens, J., Gil, E., & Planas, S. (2013). Relationship between tree row LIDAR-volume and leaf area density for fruit orchards and vineyards obtained with a LIDAR 3D Dynamic Measurement System. *Agricultural and forest meteorology*, 171, 153-162.
- Underwood, J. P., Jagbrant, G., Nieto, J. I., & Sukkarieh, S. (2015). Lidar-Based Tree Recognition and Platform Localization in Orchards. *Journal of Field Robotics*, 32(8), 1056-1074.
- Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. *International journal of computer vision*, 13(2), 119-152.
- Zhao, J., Mantilla Perez, M. B., Hu, J., & Salas Fernandez, M. G. (2016). Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum. *The Plant Genome*.