Image-based 3D digitizing for plant architecture analysis and phenotyping

Thiago T. Santos, Alberto A. de Oliveira
Embrapa Agriculture Informatics
Brazilian Agricultural Research Corporation – Embrapa
Campinas, Brazil
{thiago, albertoao}@cnptia.embrapa.br

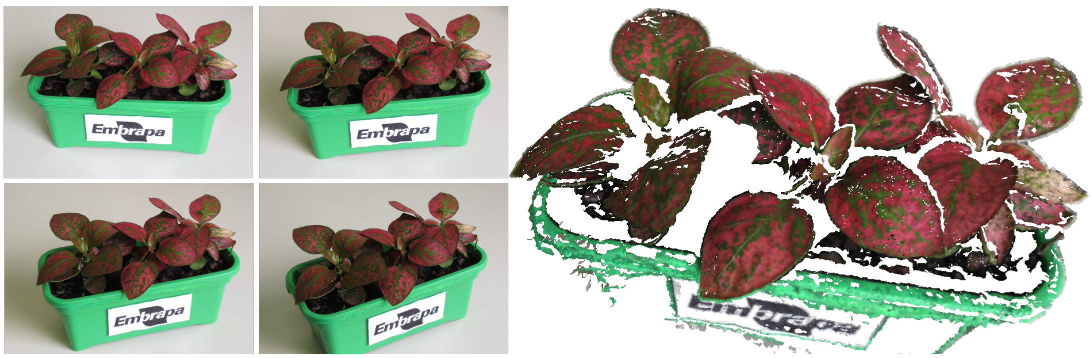


Fig. 1. 3D reconstruction using structure from motion and multiple view stereo. Only four images of *Ipoestia* were used as input (left) to produce the 3D model (right).

***Abstract*—Functional-structural plant modeling and plant phenotyping require the measurement of geometric features in specimens. This data acquisition is called *plant digitizing*. Actually, these measurements are performed manually, in invasive or even destructive ways, or using expensive laser scanning equipment. Computer vision based 3D reconstruction is an accurate and low cost alternative for the digitizing of plants not presenting a dense canopy. Sparse canopies are found in several important annual plants in agriculture as soybean and maize, at least in their early stages of development.**

This paper shows as the state of the art in structure from motion and multiple view stereo is able to produce accurate 3D models for specimens presenting sparse canopies. Three-dimensional triangular meshes are computed from a set of non-calibrated images, modeling a basil and an *Ixora* specimens and accurately representing their leaves and branches.

***Keywords*-plant phenotyping; plant architecture; structure from motion; multiple view stereo;**

I. Introduction

Three-dimensional scanning of plants refers to the construction of 3D models that allow measurement, analysis and simulation based on the specimens' geometric features. Currently, two areas are particularly concerned about this type of information: (i) functional-structural plant modeling and (ii) large scale plant phenotyping.

Functional-structural plant modeling aims to reproduce the pattern of growth and differentiation that determines plant architecture. The analysis of plant architecture assists on the understanding how the spatial distribution of physiological processes relates to plant morphogenesis [1]. According to Turnbull [2], manipulation of plant architecture has been one of the mainstays of plant improvement, one of the main responsible for the increase in agriculture productivity throughout history. Three-dimensional models can be applied on functional-structural plant simulations, such as radiation interception computation [3] or carbon source-sink relationships within the architectural framework of a specimen [4].

Platforms for large scale phenotyping have emerged in recent years with the goal of producing large amounts of phenotype data. Examples of such platforms include PHENOPSIS [5], used in studies on *Arabidopsis thaliana* performed at INRA and the TraitMill™ platform used on evaluation of transgenic rice (*Oryza sativa*) [6]. Such platforms should employ non-destructive techniques for data acquisition because they generate measurements at different moments throughout the plant development.

In functional-structural plant modeling, scanning of specimens has been performed using *contact methods* in which an operator moves a pointer over the plant surface [3], [7]. This process is tedious, labor intensive and requires long periods of time. Godin *et al.* [3] report that 24 working days were required for an operator, using a magnetic tracker (Polhemus Inc., Cochester, VT, USA), to scan eight specimens for their experiments on apple tree architecture. Rakocevic *et al.* [8] took up to 7 hours to scan clover specimens (*Trifolium repens L.*) on a small area of 100 cm^2 using a similar equipment.

If the manual acquisition of data is an obstacle to functional-structural modeling, in case of high-performance phenotyping it is a prohibitive impediment due to the need to process multiple individuals in a consistent way [9]. In phenotyping platforms, image processing techniques have been employed for measurement and analysis [9], [6], [10], which benefits from controlled environment and automated imaging. However, the systems used on these platforms are able to perform only simple measurements such as plant height and width, without obtaining any three-dimensional data [10]. Exceptions are the system for 3D reconstruction by laser scanning proposed by Kaminuma *et al.* [9], able to obtain three-dimensional surfaces for leaves and stems in *Arabidopsis*, and Biskup *et al.* [11] stereo system for leaf inclination computation.

Scanning techniques based on 3D laser scanning or stereo vision have emerged as non-intrusive and non-destructive alternatives to build three dimensional models with which plant measurements can be made automatically and consistently. Also, the employment of commodity digital cameras makes stereo vision a low cost alternative.

Contributions: This paper demonstrates that structure from motion and multiple view stereo can be used to produce accurate three-dimensional models of plants from a set of non-calibrated images. Using these models, structures as branches and leaves could be properly counted and measured automatically in a fast and consistent way.

II. Related work

Plant digitizing methods can be grouped into two classes: *contact* and *contactless*. The former methods, as presented before, require some kind of pointer touching the plant surface so that the three-dimensional model is obtained by recording its position. The latter methods employ laser scanners or cameras and build the plant model processing the data generated by these sensors.

A. Contact methods

The work of Lang [12] is recognized as the first to use a special apparatus to produce a discrete representation of a plant. It is a contact method employing a mechanical arm with potentiometers able to record its joints rotations. A disadvantage of such method is the equipment itself can change the canopy structure being modeled. Another problem is that different parts of the structure may be inaccessible by the pointer due to physical limitations.

Another contact method is the sonic digitizer employed by Sinoquet *et al.* [13] on experiments with maize cultivars. This method employs three ultrasonic sensors. The pointer is an ultrasound emitter and its position is determined by computing the distances to each pickup, calculated using the time intervals between emission and reception of sound in each sensor, assuming constant speed of sound through air. The method is sensitive to the wind and even the very structure of the vegetation could alter the propagation of sound, thus skewing the results.

Magnetic scanners are the contact method used more intensively in recent years [3], [8]. These devices produce a magnetic field that induces electric currents in coils inside the pointer, determining its position and orientation. Accuracy of a few millimeters can be obtained [8] and the magnetic field is not affected by plant structures, although metallic objects nearby may affect the measurement.

Contact techniques require an operator to move the pointer, sampling the plant surface. Although these techniques have the advantage of allowing the expert to register notes during the measurement, they are long and tedious procedures [3], so few specimens are scanned in each experiment. This is the main motivation for using non-contact techniques based on other forms of sensing, as described below.

B. Contactless methods: laser scanning

Three-dimensional models can be obtained with the use of Light Detection And Ranging (LiDAR) sensing systems. Such devices are capable of measuring the distance traveled by a laser beam to reach the surface of an object, which can be determined by performing a scanning with the beam.

Kaminuma *et al.* [9] use a laser scanner for constructing three-dimensional models for *Arabidopsis thaliana*. The models consists on the 3D surface of leaves and petioles, represented as polygonal meshes. The meshes are used to quantitatively determine two morphological attributes, the direction of the leaf blade and leaf epinasty, in order to characterize the phenotype of two different *Arabidopsis* ecotypes. The authors obtains a good sampling for surfaces due to peculiarities in the assembly of the experiment: the distance and the sample size allows a resolution of 0.045 mm per pixel, producing a dense 3D points cloud. However, different plants presenting larger dimensions such as trees, shrubs or even *Arabidopsis* specimens in more advanced stages of development can produce a sparse set of points.

Livny *et al.* [14] build 3D models from scattered points clouds obtained by a LiDAR scanner mounted on a car roof. The generated points correspond to trees near the vehicle's path. To produce a "skeleton" structure corresponding to each tree, the system considers criteria such as length, thickness, smoothness and density of the branches. This skeleton is a simplified model of the branched structure of the plant and its position in space, represented computationally as an directed acyclic graph. A first approximation of the skeleton is produced using the Dijkstra algorithm, taking the cloud of points as vertices of the graph. Using least squares based global optimization, the authors modify the position of the vertices in order to smooth the orientation of edges in the graph. The global optimization is more robust to noise and the non-uniformity in the density of points. The surface of the tree is determined by generalized cylinder centered on the edges of the graph and whose radius is determined by allometry. The results are visually compelling, the model looks like the imaged specimen, which makes the method suitable for applications in computer graphics. However, the fidelity of the model is not suitable for performing measurements as the needed in phenotyping applications. The method is unable to properly shape specimens having a high density canopy, not being able to operate in situations in which the structure of branches can not be sufficiently covered by the laser scanning process.

C. Contactless methods: stereo vision

The work of Ivanov *et al.* [15] is possibly the first work in the literature using stereo vision to reconstruct the three-dimensional surface of a cultivar for measurement and analysis. The authors estimate the position and orientation of maize (*Zea mays L.*) and distribution of leaf area. The system uses a pair of cameras installed at 8.5 m from the ground on a cornfield presenting height of 2.5 m. Unfortunately, the difficulties imposed by the equipment available at the time (digital photography was not yet widespread) undermined the experiments. The segmentation of the leaves and determination of correspondences between images were performed manually, using photographic enlargements. Despite the limitations, this work was the forerunner of more recent systems, able to employ current advances in computing performance, digital imaging and computer vision.

Biskup *et al.* [11] developed a stereo vision system based on two digital cameras to create three-dimensional models of soybean plants foliage, with the aim of analyzing the angle of inclination of the leaves and its movement throughout the day. Given the importance of movement for the experiment, the system was able to process up to three images per second, recovering 3D information necessary to calculate the slope. Once the pair of cameras was fixed, the calibration was obtained using Zhang's algorithm [16] and a checkerboard pattern. The stereo system produces a dense set of points using image rectification and stereo triangulation [17]. To validate the method, the authors glued some soybean leaflets to a planar surface and, using an inclinometer, registered reference values for different inclinations. The average deviation between angles measured by stereo and reference angles was $1.9 \pm 0.3°$.

An alternative to automatic camera calibration and estimation of a sparse cloud of points in 3D is the use of structure from motion techniques (SfM), which simultaneously estimate the position of camera and the structure and position of objects in the scene.

SfM techniques usually produce a sparse set of points in 3D. Such points may be enough for the reconstruction of regular objects as buildings and furniture, generally well-defined by a set of vertices (corners). Complex objects, such as plants, require a dense set of 3D points able to sample their surfaces with enough resolution. Quan *et al.* [18] and Tan *et al.* [19] used as input for their systems a dense points cloud produced by the technique proposed by Lhuillier & Quan [20]. This technique uses SfM to estimate the position of cameras and an initial sparse set of points. Then, using a region growing algorithm [21], a dense set of points is produced. From this dense set points in 3D, Quan *et al.* [18] combine clustering, image segmentation and polygonal models in an interactive process in which the user helps the system to create a 3D model to the foliage. The user employs a graphical editor to draw branches to complete the model of small plants. Tan *et al.* [19] worked in the opposite direction: their system, based on the same algorithm [20], is able to recover a three-dimensional model for trunks and the main branches. However, smaller branches and leaves are artificially generated.

III. Technical background

The field of 3D reconstruction from multiple images (multi-view stereo – MVS) has achieved great progress in the last decade [22]. Recently, reconstruction in real time has been shown to be feasible on commodity hardware [23], [24], [25], enabling the construction of 3D models interactively using affordable consumer equipment. A frequently used approach is to recover the calibration of the cameras and a set of sparse three-dimensional points using SfM and thereafter using region growing techniques to produce a dense sampling of the objects' surfaces.

Real time SfM techniques from a free moving video camera are known as SLAM (simultaneous localization and mapping). In recent years, the robotics and augmented reality communities improved SLAM to build systems able to determine the camera position at multiple scales in real time robustly [26]. Newcombe & Davison [23] perform camera calibration using a SLAM system [26] and produce a 3D reconstruction for objects in an indoor office scene. Although the camera position is established in real time, seconds are needed for the system to produce the surfaces of the imaged objects. The method generates a dense cloud of points in 3D through the fusion of multiple depth maps produced between pairs of camera [27].

Labatut *et al.* [28] use the SIFT framework [29] to detect points of interest and determine their correspondences between images. Projective constraint is employed to reduce the search space by establishing correspondences between pairs of pixels which are used to determine various points in 3D. Then, the set of points is used to build a Delaunay triangulation, which partitions the 3D space into a tetrahedron. An innovation of the method is to employ the fact that a Delaunay mesh partitions the space irregularly. Such a partition, compared to a regular set of voxels, has two advantages: (i) more efficient use of memory because large empty space regions are represented by few large tetrahedrons and (ii) a larger number of tetrahedron is used in small regions of the scene with great detail. The surface of the object is formulated as a binary classification problem. Each tetrahedron is classified as internal or external to the object. The surface is defined by a triangular mesh such that each triangle is the common face between an external and an internal tetrahedrons. This classification problem is formulated as a cut in a directed $s - t$ graph minimizing an energy function [30]. False positives obtained in the generation of the points have been filtered by this process of optimization. Three terms in the energy function were used in order to treat (i) occlusion, (ii) photo consistency and (iii) surface smoothness.

Furukawa & Ponce [31] employ the idea of *surfel*, small rectangular patches oriented in 3D space that are tangent to the objects's surface. Projecting the surfels on the input images, and using bilinear interpolation, the authors are able to evaluate the photo consistency between images without assuming fronto-parallel surfaces. This method is more robust to changes in pose of the object while keeping the problem tractable due to the simplification provided by surfels. A first sparse set of surfels is defined by the Harris detector and the difference of Gaussians (DoG). Correspondence between points were established employing epipolar constraint. Each pair of corresponding points is used for initialize a new surfel. The position and orientation of the surfel is refined using NCC photo consistency tests. After refinement, the surfel is added to the surfels set if it presents photo consistency for a number of images.

To produce a dense set of surfels, able to provide a good sampling for the object surface, Furukawa & Ponce's method add new surfels to the initial surfels set, iterating expansion and filtering steps. In the expansion step, new surfels neighboring the current ones are added to the set, covering unvisited regions of the object surface. In the filtering step, some of the new surfels are removed in a consistency check: they occlude or are occluded by existent surfels. According to comparative tests of Middlebury [22], [32], the method of Furukawa & Ponce is one of the best among the state of the art multiple view stereo methods.

IV. Plant digitizing using SfM and MVS

We start applying structure from motion for a input set of images covering a specimen. The images are generated slightly moving the camera, producing *short baseline* image pairs, as shown in Fig. 2 (in our plant digitizing experiments, the local feature matching produced poor results for wide baseline images). The first step of structure from motion framework is to employ local invariant feature detection and matching to produce a set of corresponding image points. We have tested SIFT [29] and SURF [33], local features invariant to image scale and rotation, usually employed in structure for motion due to their repeatability and accuracy [34].

The structure from motion system proposed by Snavely *et al.* [35] is used to recover the camera calibration data for each image and produce a sparse point cloud. This system employs an incremental method, adding one image at a time. It starts with an image pair presenting a large number of matches. The camera parameters for this pair are estimated using the five-points algorithm [36] and an initial set of 3D points is produced by triangulation. These initial camera parameters and 3D points are further refined by bundle adjustment [17]. Next, the system adds another image, that observes the largest number of the current 3D locations, and initializes this camera's extrinsic parameters using the direct linear transform (DLT) algorithm [17] inside a RANSAC procedure. New 3D points are added by triangulation and global bundle adjustment is applied again to refine the entire model[1]. This procedure is repeated until no remaining images observe enough reconstructed 3D points.

[1]Bundle adjustment is an essential procedure in image based 3D reconstruction [17]. This is why Snavely *et al.* call their system *Bundler*.

(a) All camera positions around a basil specimen

(b) Image 100 (c) Image 101 (d) Image 102

Fig. 2. A set of 143 images taken around a basil specimen (*Ocimum basilicum*). The short baseline between neighboring images helps on local feature matching.

Before building the plant surface mesh, we need to produce a *dense* point cloud using multiple view stereo. Similar to Tan *et al.* [19], we also employ a region growing strategy on the sparse point set. However, we choose Furukawa & Ponce patch-based algorithm [31] due to its recognized performance on object reconstruction [32].

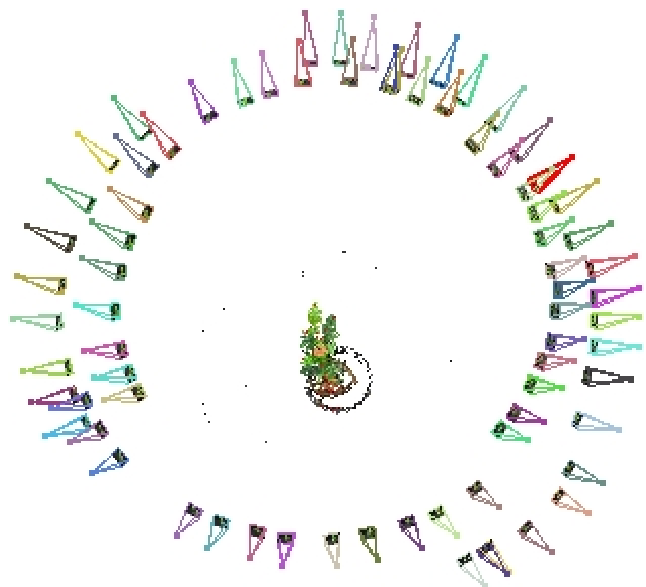
The patch-based multiple view stereo system returns a dense 3D point cloud. To produce a triangular mesh representing the plant surface we employ the ball pivoting algorithm [37].

V. Experiments

We present here two experiments[2] to demonstrate that structure from motion and multiple view stereo are able to digitize sparse plant canopies, as the ones found in cultivars like soybean, kidney beans and rice, or even maize and wheat in early development stages. In the first experiment, a set of 143 images taken around a basil specimen (*Ocimum basilicum*) was used as input, as shown in Fig. 2. In the second one, the input consisted in 77 images taken around an ornamental plant, a specimen of *Ixora coccinea*, as shown in Fig. 3.

SIFT feature detection and matching are computed using David Lowe's SIFT demo program, version 4 [39]. SURF feature detection and matching are computed using the C++

[2]Input data and results are available at the paper website [38].



(a) All camera positions around a *Ixora* specimen

(b) Image 50 (c) Image 51 (d) Image 52

Fig. 3. A set of 77 images taken around a ixora specimen (*Ixora coccinea*). Again, the short baseline between neighboring images helps on local feature matching.

implementation in the OpenCV library, version 2.4 [40]. Structure from motion is performed using Noah Snavely's *Bundler* [41] and patch-based multiple view stereo is computed using Yasutaka Furukawa's CMVS and PMVS2 packages [42], [43]. Finally, a triangular 3D mesh is generated using the ball-pivoting algorithm implementation in Meshlab [44].

VI. Results and Discussion

Fig. 4 shows the result for the *basil* images set. The original set of vertices corresponds to the PMVS2 output. Before running the ball pivoting algorithm, some dark vertices corresponding to the black background and the flower pot are filtered out by a color filter. The ball pivoting procedure produces a set of 3D meshes. Small meshes presenting less than 300 faces are removed, producing the final result. All leaves in the basil specimen are present in the 3D model.

A detailed view can be seen in Fig. 5. The vertices (dense 3D point cloud) produced by the patch-based multiple view stereo software can be seen in Fig. 5a while Fig. 5b shows the mesh produced by the ball pivoting algorithm. Removing the wireframe and showing just the faces, properly coloured, creates the realistic rendering observed in Fig. 5c. Note that even leaf veins can be observed in the final model, what could be further explored by 3D segmentation algorithms.

TABLE I
Performances results using SIFT keypoint detection.

	Basil	*Ixora*
Images	143	77
Keypoints detection	53m 60s	27m 45s
Keypoints matching	11m 50s	2m 28s
Structure from motion (*Bundler*)	23m 13s	2m 16s
Multiple view stereo (PMVS)	21m 56s	6m 14s

Quan *et al.* [18] also employ multiple view stereo for plant reconstruction, but the plant branches are not recovered in their experiments. Instead, branches are drawn by an user in an interactive procedure aided by a graphical interface. Our results demonstrate that SfM and multiple view stereo can accurately recover branches and other fine structures. However, as seen in the fragmented branches in Fig. 4c, it is necessary to get an input set containing views enough to solve the occlusions. For the basil set, some images in which the plant is seen from below would be necessary for a better branch reconstruction.

Fig. 6 shows the result for the *Ixora* images set. Again, dark vertices and small meshes are removed. All leaves found in the specimen are present in the 3D model and the entire plant is represented by one big mesh (no fragmentation in more than one connected component). However, the result presents many holes in leaves that do not correspond to real holes in the specimen. We believe they are caused by the specular highlights appearing on the shiny *Ixora*'s leaves.

Table I presents the time spent at each step on a Intel Core i7® processor. This table reports the results using the SIFT keypoint detection and matching computed using Lowe's demo program, which took most of the processing time. However, GPU based SIFT implementations can process hundreds of images in few minutes.

A. Limitation

As mentioned before, the method is limited to not too dense plant canopies[3]. Fortunately, several cultivars in early stages of development present the necessary sparsity to good reconstructions.

The image acquisition stage in this work was performed manually, moving a tripod around the specimens. Custom user interfaces [25] could be developed to assist on this acquisition step, helping the user to produce a suited input images set. In automated phenotyping platforms, robotic arms or alternative engines, as the Stanford Spherical Gantry [45], could be employed to move the camera around the specimens. Ideally, 6DOF would be desirable to produce views able to solve occlusions for different specimens and species.

VII. Conclusion

In this paper, we demonstrated as structure from motion and multiple view stereo can be employed as a powerful alternative

[3]It is important to note that the digitizing of dense canopies can be prohibitive even for manual procedures. Also, all the current alternatives would be *invasive*.

(a) A general view. (b) Top view. (c) Side view.

(d) Image 1 (e) Image 21 (f) Image 41 (g) Image 61 (h) Image 81 (i) Image 101 (j) Image 141

Fig. 4. Basil reconstruction.

(a) Vertices (b) Wireframe (c) Flat faces

Fig. 5. Basil reconstruction (detail).

for non-invasive, non-destructive plant digitizing. The models presented are detailed and accurate and could be used on the computation of several useful information as number of leaves, leaf area, leaf angle, plant height and plant topology.

The structure from motion system used in this work takes several minutes to produce the camera calibration and the sparse point cloud. The patch-based multiple view stereo algorithm can take hours in standard commodity hardware, so relegated to batch processing. However, recent advances in the visual SLAM field made possible *real-time* image based reconstruction of objects [23], [46]. Real-time plant reconstruction in 3D would mean not just a fast way for plant digitizing, but also a formidable tool for plant measuring, augmented reality aided instrumentation and applied robotics on precision agriculture. This is a topic for further investigation.

Acknowledgment

The authors would like to thank William R. Schwartz for his revision of the manuscript. This work was supported by Brazilian Agricultural Research Corporation (Embrapa) under grant 03.11.07.007.00.00.

References

[1] C. Godin, E. Costes, and H. Sinoquet, "Plant architecture modelling - virtual plants and complex systems," in *Plant Architecture and its Manipulation*, ser. Annual plant reviews, C. Turnbull, Ed. Blackwell, 2005, vol. 17, pp. 238–287. [Online]. Available: http://www-sop.inria.fr/virtualplants/Publications/2005/GCS05

(a) A general view. (b) Top view. (c) Side view.

(d) Image 1 (e) Image 11 (f) Image 21 (g) Image 31 (h) Image 41 (i) Image 51 (j) Image 61 (k) Image 71

Fig. 6. *Ixora* reconstruction.

(a) Vertices (b) Wireframe (c) Flat faces

Fig. 7. *Ixora* reconstruction (detail).

[2] C. Turnbull, *Plant architecture and its manipulation*, ser. Annual plant reviews. Blackwell, 2005, no. v. 17. [Online]. Available: http://books.google.com/books?id=O-NDVp9FYQgC

[3] C. Godin, E. Costes, and H. Sinoquet, "A Method for Describing Plant Architecture which Integrates Topology and Geometry," *Annals of Botany*, vol. 84, no. 3, pp. 343–357, Sep. 1999. [Online]. Available: http://dx.doi.org/10.1006/anbo.1999.0923

[4] G. Lopez, R. R. Favreau, C. Smith, E. Costes, P. Prusinkiewicz, and T. M. DeJong, "Integrating simulation of architectural development and sourcesink behaviour of peach trees by incorporating Markov chains and physiological organ function submodels into L-PEACH," *Functional Plant Biology*, vol. 35, no. 10, pp. 761–771, Nov. 2008.

[5] C. Granier, L. Aguirrezabal, K. Chenu, S. J. Cookson, M. Dauzat, P. Hamard, J.-J. Thioux, G. Rolland, S. Bouchier-Combaud, A. Lebaudy, B. Muller, T. Simonneau, and F. Tardieu, "Phenopsis, an automated platform for reproducible phenotyping of plant responses to soil water deficit in arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit," *New Phytologist*, vol. 169, no. 3, pp. 623–635, 2006. [Online]. Available: http://dx.doi.org/10.1111/j.1469-8137.2005.01609.x

[6] C. Reuzeau, V. Frankard, Y. Hatzfeld, A. Sanz, W. Van Camp, P. Lejeune, C. De Wilde, K. Lievens, J. de Wolf, E. Vranken, R. Peerbolte, and W. Broekaert, "Traitmill?: a functional genomics platform for the phenotypic analysis of cereals," *Plant Genetic Resources*, vol. 4, no. 01, pp. 20–24, Apr. 2006. [Online]. Available: http://dx.doi.org/10.1079/PGR2005104

[7] E. Costes, H. Sinoquet, J. J. Kelner, and C. Godin, "Exploring Within-tree Architectural Development of Two Apple Tree Cultivars

Over 6 Years," *Annals of Botany*, vol. 91, no. 1, pp. 91–104, Jan. 2003. [Online]. Available: http://dx.doi.org/10.1093/aob/mcg010

[8] M. Rakocevic, H. Sinoquet, A. Christophe, and C. Varlet-Grancher, "Assessing the Geometric Structure of a White Clover (Trifolium repens L.) Canopy using3-D Digitising," *Annals of Botany*, vol. 86, no. 3, pp. 519–526, Sep. 2000. [Online]. Available: http://dx.doi.org/10.1006/anbo.2000.1209

[9] E. Kaminuma, N. Heida, Y. Tsumoto, N. Yamamoto, N. Goto, N. Okamoto, A. Konagaya, M. Matsui, and T. Toyoda, "Automatic quantification of morphological traits via three-dimensional measurement of Arabidopsis," *The Plant Journal*, vol. 38, no. 2, pp. 358–365, 2004. [Online]. Available: http://dx.doi.org/10.1111/j.1365-313X.2004.02042.x

[10] A. Hartmann, T. Czauderna, R. Hoffmann, N. Stein, and F. Schreiber, "HTPheno: An image analysis pipeline for high-throughput plant phenotyping," *BMC Bioinformatics*, vol. 12, no. 1, pp. 148+, 2011. [Online]. Available: http://dx.doi.org/10.1186/1471-2105-12-148

[11] B. Biskup, H. Scharr, U. Schurr, and U. Rascher, "A stereo imaging system for measuring structural parameters of plant canopies," *Plant, Cell & Environment*, vol. 30, no. 10, pp. 1299–1308, 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1365-3040.2007.01702.x

[12] A. R. G. Lang, "Leaf orientation of a cotton plant," *Agricultural Meteorology*, vol. 11, no. 0, pp. 37 – 51, 1973.

[13] H. Sinoquet, B. Moulia, and R. Bonhomme, "Estimating the three-dimensional geometry of a maize crop as an input of radiation models: comparison between three-dimensional digitizing and plant profiles," *Agricultural and Forest Meteorology*, vol. 55, no. 3-4, pp. 233 – 249, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016819239190064W

[14] Y. Livny, F. Yan, M. Olson, B. Chen, H. Zhang, and J. E. Sana, "Automatic reconstruction of tree skeletal structures from point clouds," in *ACM SIGGRAPH Asia 2010 papers*, ser. SIGGRAPH ASIA '10. New York, NY, USA: ACM, 2010. [Online]. Available: http://dx.doi.org/10.1145/1866158.1866177

[15] N. Ivanov, P. Boissard, M. Chapron, and B. Andrieu, "Computer stereo plotting for 3-d reconstruction of a maize canopy," *Agricultural and Forest Meteorology*, vol. 75, no. 1-3, pp. 85 – 102, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016819239402204W

[16] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000. [Online]. Available: http://dx.doi.org/10.1109/34.888718

[17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, Apr. 2004. [Online]. Available: http://www.worldcat.org/isbn/0521540518

[18] L. Quan, P. Tan, G. Zeng, L. Yuan, J. Wang, and S. B. Kang, "Image-based plant modeling," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 599–604, Jul. 2006. [Online]. Available: http://dx.doi.org/10.1145/1141911.1141929

[19] P. Tan, G. Zeng, J. Wang, S. B. Kang, and L. Quan, "Image-based tree modeling," *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007. [Online]. Available: http://dx.doi.org/10.1145/1276377.1276486

[20] M. Lhuillier and L. Quan, "A quasi-dense approach to surface reconstruction from uncalibrated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 418–433, Mar. 2005. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2005.44

[21] ——, "Match propagation for image-based modeling and rendering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1140–1146, Aug. 2002. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2002.1023810

[22] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," in *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1. Washington, DC, USA: IEEE Computer Society, Jun. 2006, pp. 519–528. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2006.19

[23] R. A. Newcombe and A. J. Davison, "Live dense reconstruction with a single moving camera," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, Jun. 2010, pp. 1498–1505. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2010.5539794

[24] G. Vogiatzis and C. Hernndez, "Video-based, Real-Time Multi View Stereo," *Image and Vision Computing*, Feb. 2011. [Online]. Available: http://dx.doi.org/10.1016/j.imavis.2011.01.006

[25] H. Du, P. Henry, X. Ren, M. Cheng, D. B. Goldman, S. M. Seitz, and D. Fox, "Interactive 3D Modeling of Indoor Environments with a Consumer Depth Camera," in *13th International Conference on Ubiquitous Computing (Ubicomp 2011*, Sep. 2011.

[26] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," in *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, Nov. 2007, pp. 1–10. [Online]. Available: http://dx.doi.org/10.1109/ISMAR.2007.4538852

[27] R. Szeliski, *Computer Vision: Algorithms and Applications (Texts in Computer Science)*, 1st ed. Springer, Nov. 2010. [Online]. Available: http://www.worldcat.org/isbn/1848829345

[28] P. Labatut, J.-P. Pons, and R. Keriven, "Efficient Multi-View Reconstruction of Large-Scale Scenes using Interest Points, Delaunay Triangulation and Graph Cuts," in *Computer Vision, 2007. Proceedings ICCV 2007. IEEE Computer Society Conference on*, Oct. 2007, pp. 1–8. [Online]. Available: http://dx.doi.org/10.1109/ICCV.2007.4408892

[29] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004. [Online]. Available: http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94

[30] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, 2001. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=969114

[31] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2009.161

[32] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "Middlebury Multi-view Stereo Bebchmark," http://vision.middlebury.edu/mview/, 2011.

[33] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, Jun. 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1077314207001555

[34] T. Tuytelaars and K. Mikolajczyk, "Local Invariant Feature Detectors: A Survey," *Foundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2007. [Online]. Available: http://www.nowpublishers.com/product.aspx?product=CGV&doi=0600000017

[35] N. Snavely, S. Seitz, and R. Szeliski, "Modeling the world from internet photo collections," *International Journal of Computer Vision*, vol. 80, pp. 189–210, 2008, 10.1007/s11263-007-0107-3. [Online]. Available: http://dx.doi.org/10.1007/s11263-007-0107-3

[36] D. Nistér, "An efficient solution to the five-point relative pose problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 756–777, Jun. 2004. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2004.17

[37] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin, "The ball-pivoting algorithm for surface reconstruction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 349–359, 1999. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=817351

[38] Embrapa Agriculture Informatics, "PlantScan Project Website – WGARI 2012." [Online]. Available: http://www.agropediabrasilis.cnptia.embrapa.br/web/plantscan/wgari2012

[39] D. Lowe, "Demo software: SIFT keypoint detector," July 2005. [Online]. Available: http://www.cs.ubc.ca/~lowe/keypoints/

[40] Willow Garage, "OpenCV (Open Source Computer Vision)." [Online]. Available: http://opencv.willowgarage.com

[41] N. Snavely, "Bundler: Structure from Motion (SfM) for Unordered Image Collections," April 2010. [Online]. Available: http://phototour.cs.washington.edu/bundler/

[42] Y. Furukawa, "Clustering Views for Multi-view Stereo (CMVS)." [Online]. Available: http://grail.cs.washington.edu/software/cmvs/

[43] ——, "Patch-based Multi-view Stereo Software (PMVS - Version 2)." [Online]. Available: http://grail.cs.washington.edu/software/pmvs/

[44] "MeshLab." [Online]. Available: http://meshlab.sourceforge.net/

[45] M. Levoy, "Stanford Spherical Gantry," February 2002. [Online]. Available: http://graphics.stanford.edu/projects/gantry/

[46] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison, "DTAM: Dense tracking and mapping in real-time," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, Nov. 2011, pp. 2320–2327. [Online]. Available: http://dx.doi.org/10.1109/ICCV.2011.6126513