Proceedings of the 7th International Conference on Functional-Structural Plant Models, Saariselkä, Finland, 9 - 14 June 2013. Eds. Risto Sievänen, Eero Nikinmaa, Christophe Godin, Anna Lintunen & Pekka Nygren. http://www.metla.fi/fspm2013/proceedings. ISBN 978-951-651-408-9.

Automatic 3D plant reconstruction from photographies, segmentation and classification of leaves and internodes using clustering

Thiago Santos[1] and **Julio Ueda**[1]
[1]*Embrapa Agricultural Informatics, PO Box 6041, 13083-886 Campinas, Brazil*
**correspondence:* thiago.santos@embrapa.br

Highlights: A stereo approach for 3D plant modelling is presented. Using only a set of photographies, the method produces a dense 3D point cloud that samples the plant surface. Clustering automatically segments the plant structure into meaningful parts, which are classified as leaves or internodes. Measurements can be computed for each element, as area or surface normals.

Keywords: 3d plant models, image-based reconstruction, multiple view stereo, structure segmentation

INTRODUCTION

Non-invasive imaging and image analysis are novel technologies that have been employed to narrow the "phenotyping bottleneck". Three-dimensional plant models are traditionally acquired by invasive, slow and tedious manual measurements, aided by electromagnetic 3D tracking devices. Laser scanning figures as an alternative (Preuksakarn *et al.*, 2010; Delagrance and Rochon, 2011), but it presents some drawbacks as sensitivity to occlusion, lack of color and texture information and the high price of the laser scanning devices. A generated 3D point cloud must be properly segmented and classified for automatic measurement and characterization of the plant structure. Stereo-based techniques are emerging as a cheap and non-invasive alternative for 3D modelling of plants (Quan *et al.*, 2006; Biskup *et al.*, 2007; Santos and Oliveira, 2012). The present work extends our previous work (Santos and Oliveira, 2012) by (a) segmenting the plant models in significant structures using clustering in 3D space, (b) classifying the recovered segments into meaningful classes (leaves and internodes) and (c) performing area measurements in the model and comparing them against ground-truth data, validating the framework as an effective metrology tool.

MATERIAL

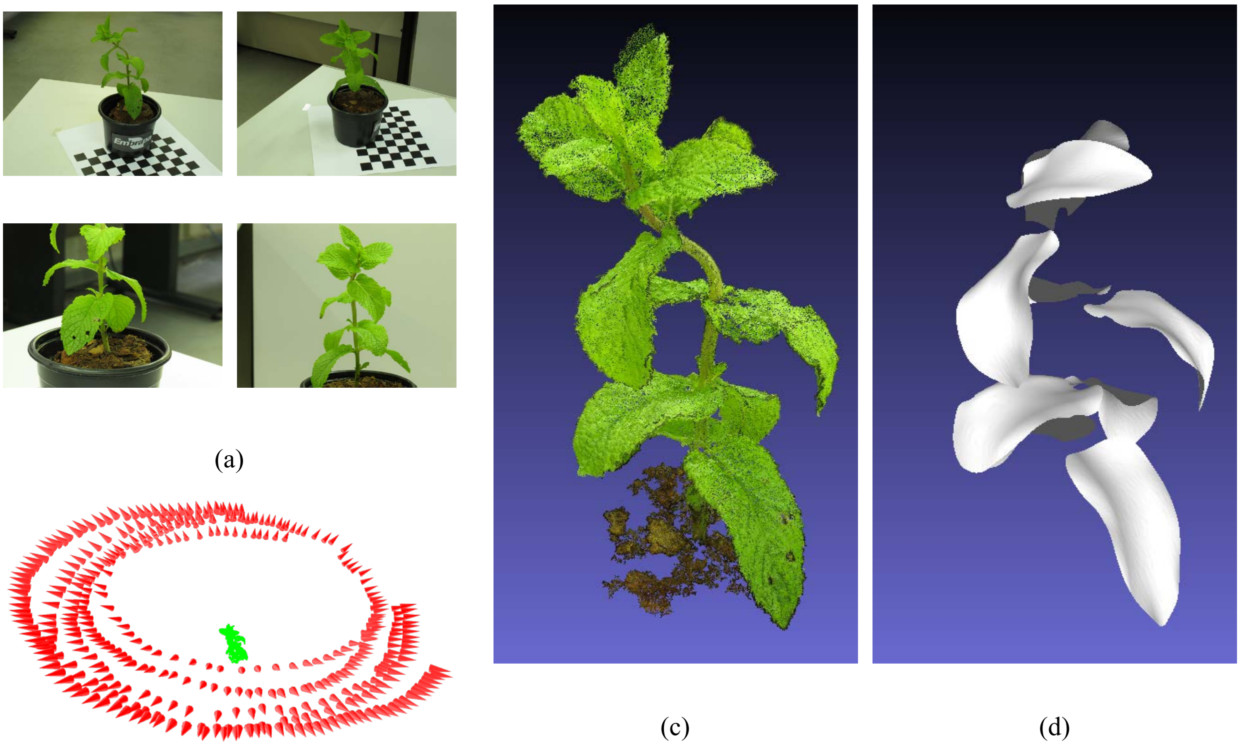
A set of 387 pictures of a mint specimen (*Mentha*) was acquired by a Canon Powershot G11 camera placed in different positions (see Fig. 1A and Fig. 1B). The potted specimen was photographed indoors, avoiding movements caused by wind. After the image acquisition step, the leaves were removed and placed in a table scanner to acquire the area measurements used as ground-truth.

METHOD

The proposed technique is composed by the following steps:

Multiple view stereo plant reconstruction – The method input is a set of several high-resolution photographies for each specimen. Camera position is automatically recovered by *structure from motion*. First, the SIFT algorithm (Lowe, 2004) is employed to detect and describe *image features* in each photography. The feature descriptors are used to find matches between features in different images. Projective reconstruction and robust estimation techniques (Hartley and Zisserman, 2003) are employed to define the relative position between images, i.e., the position of the camera at each image acquisition (Fig. 1B). Once each camera pose is defined, a *sparse* 3D point cloud for the plant surface is produced based on feature matching. Finally, a region growing multiple view stereo technique is employed to produce a *dense* 3D point cloud (Fig. 1C). Santos and Oliveira (2012) present a more detailed description of this 3D reconstruction step.

Segmentation by clustering of surface normals – The dense point cloud is segmented using a smoothness constraint, as proposed by Rabbani *et al.* (2006). First, the surface normals are estimated at each point p_i. This estimation is performed by finding a plane tangent to the surface by least-squares plane fitting, using the points in the neighbourhood of p_i. Then a region growing algorithm is applied: for a point p_i in a segment R, each neigh point p_j is added to R if the angle between the normal vectors of p_i and p_j is inferior to a threshold θ. This process is repeated until every point is assigned to a region. The detailed algorithm can be found in Rabbani *et al.* (2006).



(a) (b) (c) (d)

Fig. 1. Results for the *mint* dataset. (a) Input images acquired from different angles. (b) Result from the structure from motion step: camera poses (red cones) and a sparse point cloud for the plant (green). (c) Result from the multiple view stereo step, a dense point cloud sampling of the plant surface. (d) Smooth surfaces for the largest leaves, computed using the NURBS fitting procedure, after leaf segmentation and classification.

Classification using width/length ratio – Each segment is composed by a set of 3D points. Features can be computed from this set for segment characterization and further classification. For the specimen used in this work, the recovered segments correspond to leaves, internodes or spurious structures, as fragments from soil. These classes can be easily discriminated using their size and dimensions. Taking each point in the segment as a three-dimensional vector $p_i = (x_i, y_i, z_i)$, principal component analysis (PCA) was applied. In the transformed space, the ordered variances were used to describe the segments' dimensions in their main axes. A simple linear classifier was able to classify the leaves, if properly segmented.

Leaf surface fitting using NURBS – Leaves surfaces were approximated by NURBS fitting (Piegl, 1991), getting a smooth and regularized 3D mesh representing the surface (Fig. 1D).

RESULTS

Fig. 1B shows the recovered camera poses and sparse 3D model produced by structure from motion. The chessboard observed in Fig. 1A is generally used in computer vision for camera calibration, but in the present experiment it is employed just to define the scale factor for the final model – the leaves' textures and edges provide all the image features needed for the camera pose estimation. Fig. 1C shows the 3D dense point cloud produced by the multiple view stereo step. After cloud segmentation and classification, the largest leaves were successfully identified. Small leaves were sub-sampled in the point cloud, resulting in over-segmentation and misclassification. A smoothed surface was produced for each one of the correct leaves by NURBS fitting. Table 1 shows the estimated area vs. the ground-truth data.

Table 1. The 3D model as a measurement tool. Individual leaf area computed on the the smooth NURBS surfaces vs. ground-truth produced using a common scanner.

Leaf/Area (mm²)	Ground-truth	Computed on the 3D model	Difference
1	1434.48	1510.60	5.31%
2	1464.68	1500.91	2.47%
3	1144.59	1167.69	2.02%
4	1157.50	1177.38	1.72%
5	1007.02	1005.48	-0.15%
6	791.39	735.33	-7.08%
7	899.51	956.64	6.35%
8	954.69	1079.98	13.12%
9	660.58	660.68	0.02%

CONCLUSION

In the proposed methodology, a free moving camera is able to capture the plant structure from different views, differently of Biskup *et al.* (2007) fixed-camera approach. The structure is recovered only from image data, without human intervention as the branches sketches employed by Quan *et al.* (2006). The large number of input images should not be a problem because video acquisition is able to provide thousands of video frames that could be automatically selected. Further steps under development are (i) more extensive tests on different species, (ii) an interactive tool to help human operators to perform computer-aided video acquisition with feedback and (iii) an egomotion system based on structure from motion for robot path planning in automatized platforms.

FUNDING

This work was supported by Brazilian Agricultural Research Corporation (Embrapa) under grant [03.11.07.007.00.00].

LITERATURE CITED

Biskup B, Scharr H, Schurr U, Rascher U. 2007. A stereo imaging system for measuring structural parameters of plant canopies. *Plant, Cell & Environment*, **30**:1299–1308.

Delagrange S, Rochon P. 2011. Reconstruction and analysis of a deciduous sapling using digital photographs or terrestrial-LiDAR technology. *Annals of botany,* **108**:991–1000.

Hartley R, Zisserman A. 2003. *Multiple View Geometry in Computer Vision*, 2th edn. Cambridge: Cambridge University Press.

Lowe DG. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60:**91-110.

Piegl L. 1991. On NURBS: a survey. *IEEE Computer Graphics and Applications,* **11**(1):55-71.

Preuksakarn C, Boudon F, Ferraro P, Durand J-B, Nikinmaa E, Godin, C. 2010. Reconstructing plant architecture from 3D laser scanner data . In *Proc. of the 6th Intl. Workshop on Functional-Structural Plant Models*, 14-16.

Quan L, Tan P, Zeng G, Yuan L, Wang J, Kang SB. 2006. Image-based plant modelling. *ACM Transactions on Graphics*, **25**: 599–604.

Rabbani T, van den Heuvel F, Vosselmann G. 2006. Segmentation of point clouds using smoothness constraint, In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* **36**(5):248–253.

Santos TT, Oliveira AA. 2012. Image-based 3D digitizing for plant architecture analysis and phenotyping. In *Workshop on Industry Applications (WGARI) in SIBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images),* 21-28. http://www.decom.ufop.br/sibgrapi2012/index.php/call/wgari.