Robotic Leaf Probing Via Segmentation of Range Data Into Surface Patches

G. Alenyà, B. Dellen, S. Foix, and C. Torras

Abstract— **We present a novel method for the robotized probing of plant leaves using Time-of-Flight (ToF) sensors. Plant images are segmented into surface patches by combining a segmentation of the infrared intensity image, provided by the ToF camera, with quadratic surface fitting using ToF depth data. Leaf models are fitted to the boundaries of the segments and used to determine probing points and to evaluate the suitability of leaves for being sampled. The robustness of the approach is evaluated by repeatedly placing an especially adapted, robot-mounted spad meter on the probing points which are extracted in an automatic manner. The number of successful chlorophyll measurements is counted, and the total time for processing the visual data and probing the plant with the robot is measured for each trial. In case of failure, the underlying causes are determined and reported, allowing a better assessment of the applicability of the method in real scenarios.**

I. Introduction

The automatic monitoring and maintenance of large botanic experimentation fields, e.g., for plant phenotyping, is a new research topic in agricultural robotics with many potential applications. For evaluating the state of a plant during growth, measurements and samples from leaves must be regularly taken and some pruning may need to be performed [1], [2]. Automation of these tasks with robots is however highly challenging due to the complex and deformable nature of plants, which pose problems for (i) recognition and localization of targets, i.e., leaves, given the varying appearances of plants, and (ii) the probing and handling of the plant under weakly constrained conditions in natural environments, among others.

Along this line, the European project GARNICS (Gardening with a Cognitive System) aims at 3D sensing of plant growth and building perceptual representations for learning the links to actions of a robot gardener. The project encompasses both the long-term learning of treatments to achieve specific goals (maximum leaf growth, homogeneous plant growth) as well as the short-term robot interaction with plants (for leaf surface measurement, addressing occlusion, probing). Since actions performed on plants (like watering or nutrient delivery) have strongly delayed effects, long-term learning is being addressed by extracting relevant cause-effects from weakly correlated inputs and outputs. In this paper we focus on the short-term robot activity required to collect such outputs (leaf measurements) along the experimentation period.

Authors are with Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens i Artigas 4-6, 08028 Barcelona, Spain; {galenya,bdellen,sfoix,torras}@iri.upc.edu

More precisely, we address in this work the problem of accurately placing a probing tool, i.e., a spad meter, on a leaf in order to measure the chlorophyll content. Challenges arise from the fact that leaves have to be segmented and recognized in a robust manner. Furthermore depth information needs to be acquired for placing the tool on the target. Our experimental platform consists of a robot arm equipped with a ToF sensor [3] and a spad meter, which are both mounted on the end-effector of the arm. The emphasis of this work is on *sensing-for-action* methods developed to segment leaves, fit quadratic surfaces to them, determine best candidates for probing, move the cameras to get a closer view, determine a suitable sampling point on the chosen leaf, and finally reach this point with a robot mounted tool. Intensity-based segmentation is complemented with depth data supplied by a ToF camera to delimit and fit surface patches to the leaves.

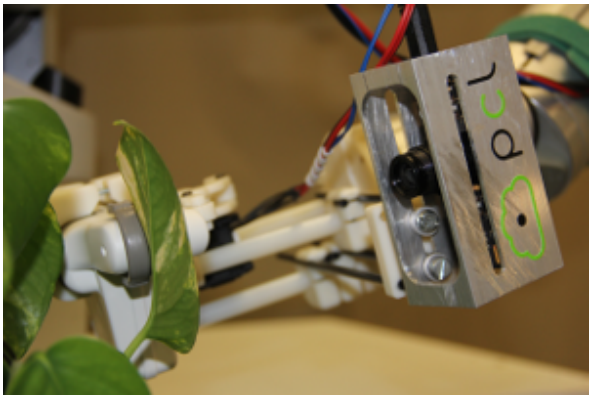
In this paper, the main focus is on the probing of the leaf. A next-best view algorithm for the system has been described in detail in [4].

II. Related work

The problem of leaf segmentation has been addressed before by Quan *et al.* (2006), who proposed an image-based plant modeling system based on structure from motion, but which requires user interaction in the segmentation procedure to delineate some leaves [5]. In another related work, leaves were segmented from combined color images and stereo depth, and subsequently classified using the normalized centroid contour distance [6]. Different from these approaches, we extract leaves from ToF data and infrared-intensity images. Segmentation is fully automatic and based on a novel depth-segmentation algorithm which can be applied to sparse or noisy depth data and cope with curved surfaces. Another difference is that leaf models are fitted explicitly which allows localizing grasping points.

The proposed system for automated plant probing is related to vision-based robotic systems for fruit and vegetable picking, which have been proposed in the past for the automation of harvesting tasks [7], [8] (for a review and complete list of publications, see [9]). Commonly these systems first process and segment the data in order to identify and represent the target. Based on this representation, a robot action, i.e., cutting or grasping, is executed. Often the image processing task is eased by fixing the environment in a specific manner. For example, in a fruit-detachment system developed by Feng *et al.* (2008) strawberries were grown on a uniformly colored surface to simplify image segmentation

(A)



(B)

Fig. 1. A. WAM arm used in the experiments holding the ToF sensor and spad meter to measure the chlorophyll content of leaves. B. Probing of a Pothus leaf.

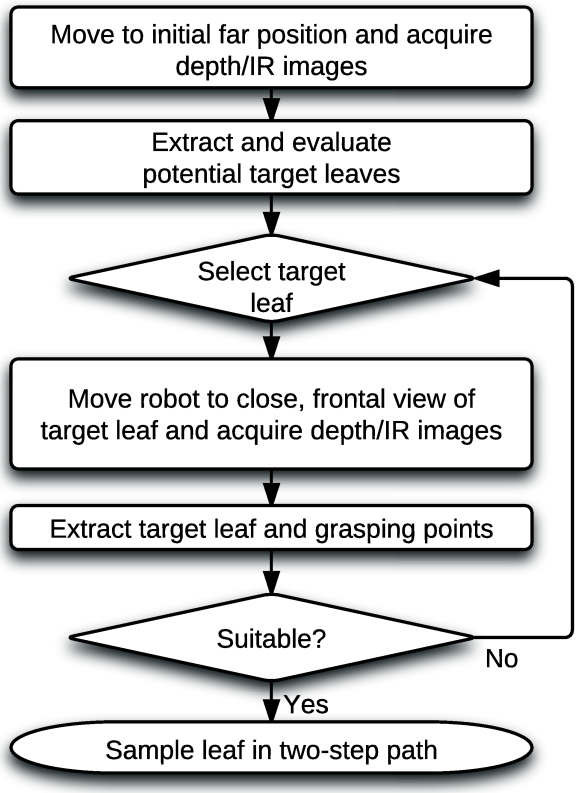


Fig. 2. Flow diagram of the suggested probing procedure (see Section III). In this paper, the main focus is on the probing procedure of the method.

[7]. In our system the environment is less constrained, and the proposed computer-vision system is thus more complex.

III. Overview of the method

Leaf probing is conducted following a two-stage approach (see Fig. 2). From a general view of the plant, the depth and infrared images are acquired with a ToF camera (Section IV), and are segmented into their composite surfaces as described in Section V. Leaf-model contours are fitted to the extracted segments, the validity of the fit and the graspability of the leaf are measured, and the segments are ranked (see Section VI). A target leaf is selected and the robot moves the camera to a closer, fronto-parallel view of it. If the leaf is still considered to be suitable for being sampled based on these criteria, the probing tool is placed onto the leaf following a two step path. If the target is considered to be non-suitable for probing, another target leaf (from the general view) is selected and the procedure is repeated.

IV. 3D image acquisition

Depth measurements are acquired by a ToF camera. This type of sensor has the main advantage of providing registered depth and infrared-intensity images of a scene at a high frame-rate. ToF cameras use the well-known time-of-flight principle to compute depth. The camera emits modulated infra-red light in order to measure the travelling time between the known emitted waves and the ones reflected back over the objects in the scene.

ToF cameras have two main drawbacks: low resolution (e.g. 200×200 pixels for a PMD CamCube 3.0 camera) and noisy depth measurements due to systematic and non-systematic errors [10]. On one hand, low resolution can be a big problem for large environment applications, but it does not have such a negative impact when the camera is used at close ranges as it is our case. On the other hand, noisy depth measurements due to non-systematic errors get amplified by working in such a short range. Mainly the ones due to multiple light reception and light scattering. Systematic errors get highly reduced by calibration procedures and non-systematic ones can be reduced using filtering techniques [11].

Here we apply two filters to remove undesired wrongly estimated point depths and noise: a jump edge filter and an averaging filter [12]. Sometimes these false measurements are indicative of possible model misinterpretation or object occlusion and, therefore, their detection and 3D localization in the scene may provide valuable information for computing the next-best-view that can help to disambiguate or improve occluded leaf visibility and pose estimation [13].

V. Depth segmentation

In this section we describe an algorithm for segmenting the sparse and noisy depth data measured by the ToF camera

into surface patches in order to extract task relevant image regions, i.e., leaves. The method consists of the following steps:

1 Infrared-intensity segmentation. The infrared-intensity image (to which we will from now on refer as intensity image) is segmented into regions using a standard segmentation algorithm, yielding segments s_i. Details can be found in [14]. Segmentations are obtained for parameters $k = 150$, $\sigma = 0.5$ and min $= 100$.

2 Model fitting. For a given segment s_i we perform a minimization of the mean square distance

$$E_i = 1/N \sum_j (z_j - z_{j,m})^2 \tag{1}$$

of measured depth points $z_{j,m}$ from the estimated model depth $z_j = f_i(x_j, y_j)$, where $f_i(x_j, y_j) = ax_j^2 + by_j^2 + cx_j + dy_j + e$ is the data-model function and N is the number of measured depth points in the area of segment s_i. Surfaces are described by five parameters a, b, c, d, and e.

3 Building a segment graph. For each image segment, the boundary points are extracted. Two segments are considered neighbors if some of the respective boundary points are less than $d_{3D} = 5$ cm apart. The segments define the nodes V of the segment graph (V, e). We further assign a weight to each edge of the graph by measuring the pairwise dissimilarity e_d between the neighboring segments.

4 Segment dissimilarity. We define a dissimilarity measure between two segments s_i and s_j by computing the fitting errors

$$E_{i/j} = 1/n_i \sum_{p \in s_i} [f_j(x, y) - z(x, y)]^2 \quad , \tag{2}$$

and

$$E_{j/i} = 1/n_j \sum_{p \in s_j} [f_i(x, y) - z(x, y)]^2 \quad , \tag{3}$$

where $z(x, y)$ is the measured depth at (x, y), $f_j(x, y)$ and $f_i(x, y)$ are the estimated depth value using surface models at (x, y), and n_i and n_j are the number of points in segment s_i and s_j, respectively. The smaller error is selected and defines the edge dissimilarity e_d between the neighbors.

5 Graph based clustering. The pairwise dissimilarities between segments are used to sort the graph edges e_{ij} in order of increasing dissimilarity. Two segments are merged if their edge dissimilarity $e_d < d_{\text{merge}}$, where $d_{\text{merge}} = 15$ cm^2 is a parameter. After a merge, the edge dissimilarities and surface models for all neighbors of the affected segments are updated. Working consecutively along the ordered list, preference is given to merges of segments with large similarity, similar to Kruskal's algorithm for finding the minimum spanning tree of a graph [15].

The whole method requires currently about ≈ 2 s to segment an image, to fit surface models, and to computed probing points using an implementation integrated in ROS.

VI. Extraction of grasping points

We assume that the procedure described above delivers segments that correspond to leaves of the plant. This assumption may not always hold, but it is a good enough working hypothesis. The grasping point should lie within the part of the leaf that points away from the stem of the plant. This way, the risk of collisions with the stem and other leaf parts can be reduced. We further want to approach the leaf from the side to maximize the touched leaf area. For this purpose, a leaf-specific contour needs to be fitted to the leaf segment boundary in order to map leaf-specific grasping points along the segment boundary.

We extract the outer 2D boundary C_i of segment i, consisting of a set of points $\{x, y, z\}$. For each plant type, we have extracted the leaf boundary which is characteristic for the specific plant. We smooth the boundary points with a Gaussian function. The resulting values provide a set of weighted boundary points $\{x, y, w\}_m$, defining our model boundary C_m. The z-value is constant and can be ignored, since the viewing direction of the camera is assumed to be aligned with the z-axis. The distance of the transformed boundary to a model boundary for a set of four transformation parameters (scaling, rotation, translation) provides a minimization criteria. Using a Nelder-Mead simplex search, the transformation parameters are determined.

Once the segment contour has been fitted to the model contour, we can identify grasping points. We assume that predefined grasping points are provided together with the leaf contour model, as illustrated in Figure 3, lower panel. For each model grasping point, we find the point on the segment contour C_i^{trans} that has the smallest distance to the model grasping point. Together with the resulting grasping point $\mathbf{x}_g = (x_g, y_g, z_g)_i$, we also provide the validity measure of the fit v_i.

We further compute the surface normal of the leaf using a principal components analysis. From this, the orientation of the tool relative to the leaf both for moving to a close view and for aligning the tool is determined.

VII. Experimental setup

The experimental setup includes a PMD CamBoard Time-of-Flight camera rigidly attached to the last link of a Barrett WAM arm (Fig. 1). As can be observed, the camera is displaced from the robot end-effector position to leave room for a spad meter we have adapted to take samples of some selected leaves.

We have opted for a configuration where the tool is outside the field of view of the camera. This implies that, during the robot motion from the close view of the leaf to the placement of the tool, the leaf is not in the camera field of view, and the motion is then performed in open loop. Implicitly we are

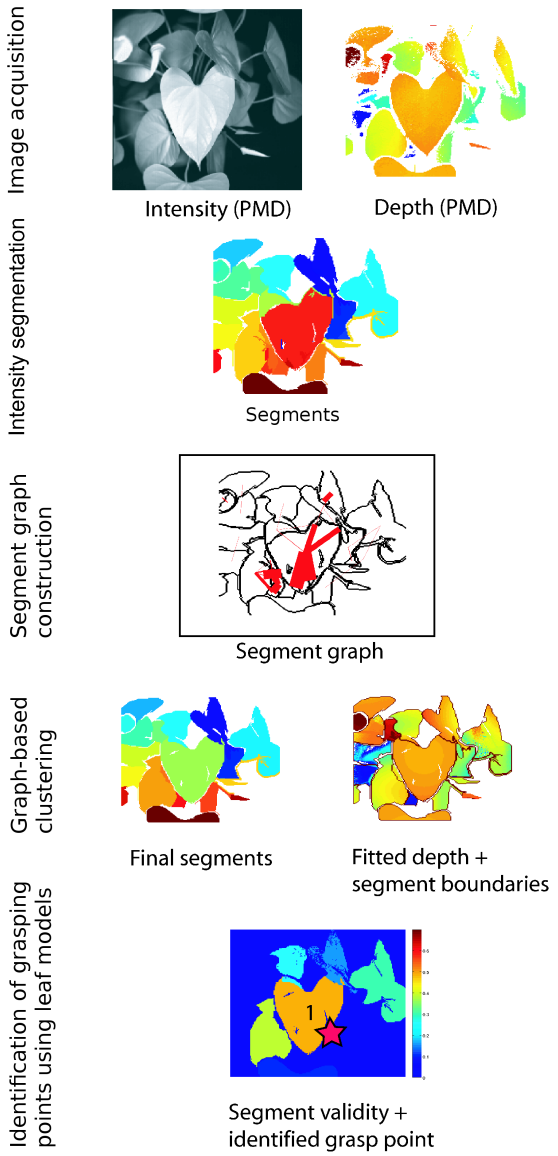


Fig. 3. Schematic of the leaf-extraction algorithm. ToF data (depth and infrared intensity) are acquired and the infrared-intensity image is segmented and surface models are fitted to the segments. A segment graph is constructed and a graph-based segment merging procedure is employed having the goal of finding the segments and surface models which fit the data best. Final segment contours are fitted to predefined model contours and grasping points are determined.

assuming that the leaf will not move and that the robot has enough precision along this small motion.

The robot and plant initial relative configuration assures that the plant's region of interest is reachable by the robot's tool. In the same way that plant position is guaranteed to be inside the field of view of the camera's initial pose. In the close view, the camera is placed in a frontal configuration at 30 centimeters of the localized leaf.

Since the camera system is attached to the end-effector of the robot, the current position of the camera and, consequently, its measurements can be determined from the robot configuration. We use an implementation of the inverse

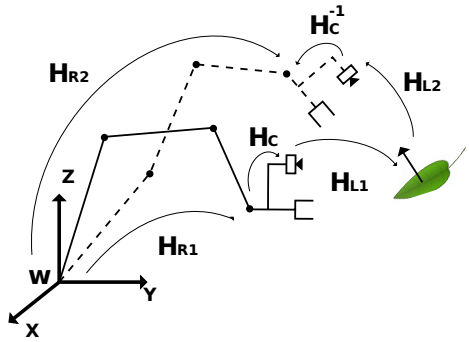


Fig. 4. Robot motion. Transformations involved in the computation of the desired robot position H_{R2} from the current position H_{R1}, the hand-eye calibration H_C, the observed leaf position from the camera H_{L1} and the desired target position H_{L2}.

kinematics of the manipulator using the KDL library to obtain the cartesian position of the tool center point (TCP) $\mathbf{H_{R1}}$ in the space of the robot. The hand-eye calibration of the camera with respect to the TCP $\mathbf{H_C}$ has been computed off-line. By concatenating transformation matrices, we can express every point cloud in the robot's coordinate frame. Hence, once the plant leaves have been segmented and properly localized, the robot tool can be correctly placed over the desired leaf to be sampled. We use the same approach to compute the desired position of the robot in the two different tasks: next view and probing. For illustration, see Fig. 4.

We have found that the best orientation for next view sensing is to place the camera system perpendicular to the leaf surface. Additionally, with this position we aim to minimize the difference in depth of the point cloud, as ToF cameras exhibit a depth error that depends on the current distance [16].

VIII. Leaf Probing and Chlorophyll Measurements

The main focus of this work is the assessment of the robustness of the probing procedure. For this we assume that the robot has been moved such that the camera is in a fronto-parallel position and at a distance of 30 cm with respect to the target leaf surface. Moving to a closer, fronto-parallel view of a leaf allows better verification of suitable leaves for probing and thus also a better determination of grasping points. For this purpose the validity of selected target segments for different view points using both a model leaf (cut from paper board) and real leaf was evaluated. We observed that the validity decreases with increasing angle, i.e., the further we move away from the fronto-parallel position, the more difficult it becomes to recognize the leaf due to view-dependent shape distortions and other visibility impairments (see Fig. 5). This also implies that the grasping point cannot be accurately determined past some angle, because the model-leaf contour together with the associated grasping point will fit the segment boundary only very poorly. A close view for probing is desirable since it increases the amount of data that can be gathered about a

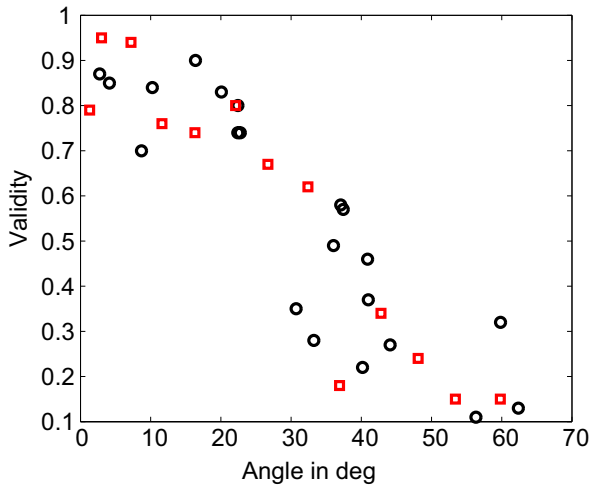


Fig. 5. Combined validity results for the artificial (red square) and the real leaf (black circle). Line fitting to the data points suggests an average error of the measured validity of $\approx \pm 0.1$ for a nearly planar leaf. The validity measures the correlation between the measured and transformed 2D contour of a segment and a 2D model-leaf contour.

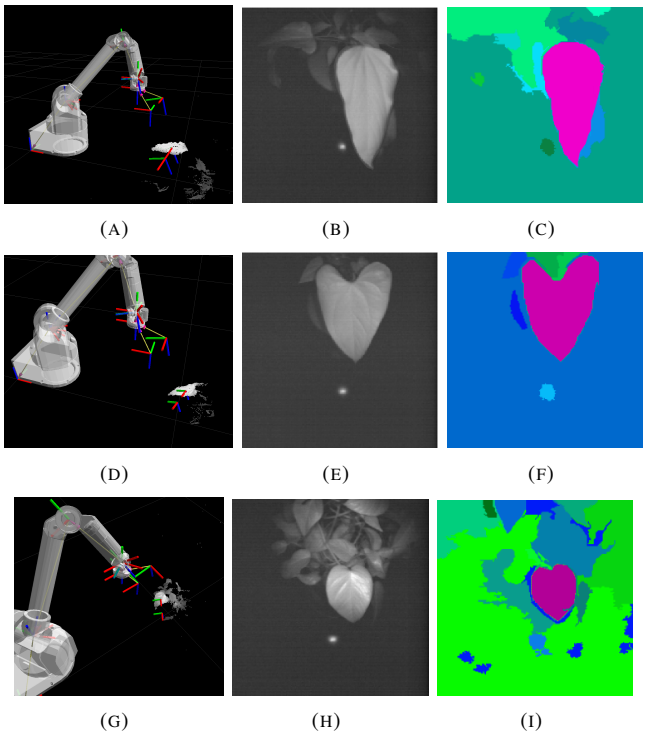


Fig. 6. Left to right. Robot model, PMD intensity image, and the segmented target leaves for Anthurium White (A-C), Anthurium Red (D-F), and Pothos (G-I).

leaf (the resolution).

Once the robot arm is brought into the desired position, the reliability of the probing procedure, including target selection, probing point extraction, and path planning can be evaluated. This is done by automatically measuring the chlorophyll level of the target leaf at the probing point repeatedly with a spad meter mounted on the robot arm. Three plants are investigated: Anthurium White, Anthurium Red, and Pothus.

First, a target leaf is found from the ToF data and the infrared-intensity image in the close view. The poses of the robot, the camera, the tool, and the target leaf are computed for path planning. In Fig. 6 examples of the computed poses in the robot model (left panels), the PMD infrared intensity images (middle panels), and the segmented target leaves (right panels) are shown for the different plants. Because the camera is already in the close, fronto-parallel view with respect to the leaf, the segmentation problem is eased compared to the far view (not shown). For the selected target leaf, a grasping point is extracted and the probing movement is executed (see video at `www.iri.upc.edu/groups/perception/LeafProbing`). If the probing is successful, the measured chlorophyll level of the leaf is reported. The time for computing the probing point and execution of the probing is measured. The results of experiments conducted for the different plants are summarized in the tables 1-3. The most common failure reason are errors in the leaf-model fitting, leading to a wrong estimation of the grasping point (failure reason 1). Other less frequent reasons are bad normal estimation (failure reason 2), probing a point of a leaf patch with low chlorophyll content (failure reason 3), or an overly long path chosen by the kinematics (failure reason 4).

Probing succeeded in 90% of the cases for Anthurium White, 85% of the cases for Anthurium Red, and 70% of the cases for Pothus. For the Pothus plant, the worst result was obtained. This might be due to the small size of the leaves, posing limits to the depth estimation and thus probing point estimation. Overall, experiments were successful in 82% of the cases.

IX. Conclusions

We presented a method for modeling, monitoring, and sampling plant leaves using infrared-intensity images and depth maps acquired with a time-of-flight (ToF) camera. Since quadratic surface models are used to guide the segmentation of the infrared-intensity image, sparse or noisy depth data can be used, which often poses a problem to approaches working in the depth space directly. Target leaves are selected using leaf models, and this way a probing point for chlorophyll measurements with a robot-mounted spad meter can be selected. We tested the approach with a real robot arm equipped with the spad meter and the ToF camera. We tested the robustness of the method by repeatedly probing a leaf. This was done for several plants. On average, probing succeeded in 82% of the cases. The most frequent failure reason was a wrong estimation of the probing point, caused by poor model fitting or segmentation errors.

Acknowledgements

This research is partially funded by the EU GARNICS project FP7-247947, by the Spanish Ministry of Science and Innovation under projects PAU+ and MIPRCV Consolider

Anthurium Andreanum (White)				
Exp.	Success	Time [s]	Chlorophyll	Failure Reason
1	Y	21.05	44.2	x
2	N	26.44	-	1
3	Y	24.04	43.8	x
4	Y	25.06	43.7	x
5	Y	21.48	44.3	x
6	N	24.62	-	1
7	Y	24.32	44.4	x
8	Y	25.60	43.9	x
9	Y	23.45	44.4	x
10	Y	23.23	44.6	x
11	Y	24.91	43.9	x
12	Y	25.87	44.2	x
13	Y	22.53	43.8	x
14	Y	21.71	44.4	x
15	Y	25.03	43.5	x
16	Y	27.95	43.8	x
17	Y	21.76	43.5	x
18	Y	22.04	43.1	x
19	Y	23.95	44.0	x
20	Y	22.64	44.6	x

TABLE I

Anthurium Andreanum (Red)				
Exp.	Success	Time [s]	Chlorophyll	Failure Reason
1	Y	18.87	49.6	x
2	Y	19.31	49.0	x
3	Y	19.54	50.2	x
4	Y	20.88	49.6	x
5	Y	21.18	49.1	x
6	Y	20.77	50.2	x
7	N	18.68	-	2
8	Y	19.59	48.3	x
9	Y	19,22	49.9	x
10	Y	18.77	48.3	x
11	Y	21.15	49.9	x
12	Y	19.35	50.3	x
13	N	21.11	-	1
14	Y	18.97	49.5	x
15	N	19.29	-	1
16	Y	18.11	32.8	1
17	Y	18.67	50.1	x
18	Y	19.64	48.6	x
19	Y	19.98	46.6	x
20	Y	18.48	49.3	x

TABLE II

Ingenio CSD2007-00018, and the Catalan Research Commission. B. Dellen acknowledges support from the Spanish Ministry for Science and Innovation via a Ramon y Cajal program. S. Foix is supported by PhD fellowship from CSIC's JAE program.

Epipremnum Aureum (Pothos)				
Exp.	Success	Time [s]	Chlorophyll	Failure reason
1	Y	19.61	50.9	x
2	Y	22.22	51.9	x
3	Y	19.23	51.8	x
4	N	19.86	-	1
5	Y	18.52	47.6	x
6	Y	19.97	52.3	x
7	Y	21.39	47.7	x
8	N	22.92	-	1
9	Y	18.10	50.5	x
10	Y	19.75	51.4	x
11	Y	19.59	49.5	x
12	Y	18.67	34.6	3
13	N	24.26	-	1
14	Y	22.10	48.5	x
15	N	31.53	-	1
16	Y	29.81	50.5	4
17	N	19.78	-	1
18	Y	18.49	50.9	x
19	Y	19.65	51.9	x
20	N	21.04	-	1

TABLE III

REFERENCES

[1] T. Fourcaud, X. Zhang, A. Stokes, H. Lambers and Ch. Köner: "Plant Growth Modelling and Applications: The Increasing Importance of Plant Architecture in Growth Models", *Annals of Botany* 101: 1053-1063, 2008.

[2] E.J. Van Henten, B.A.J. Van Tuijl, G.-J. Hoogakker, M.J. Van Der Weerd, J. Hemming, J.G. Kornet, and J. Bontsema: "An Autonomous Robot for De-leafing Cucumber Plants grown in a High-wire Cultivation System", *Biosystems Engineering* 94(3), pp. 317-323.

[3] R. Klose, J. Penlington and A. Ruckelshausen: "Usability study of 3D Time-of-Flight cameras for automatic plant phenotyping", Workshop on Computer Image Analysis in Agriculture, pp. 93-105, August 2009.

[4] G. Alenyà, B. Dellen, and C. Torras: "3D modelling of leaves from color and ToF data for robotized plant measuring". *Proc. IEEE Intl. Conf. on Robotics and Automation (ICRA'11)*, Shanghai, May 2011.

[5] L. Quan, P. Tan, G. Zeng, L. Yuan, J. Wang and S.B. Kang: "Image-based Plant Modelling", ACM Siggraph, pp. 599-604, 2006.

[6] Ch.-H. Teng, Y.-T. Kuo, and Y.-S. Chen: "Leaf segmentation, classification, and three-dimensional recovery from a few images with close viewpoints". *Optical Engineering* 50(3), doi:10.1117/1.3549927, 2011.

[7] G. Feng, C. Qizin, and M. Masateru: "Fruit detachment and classification method for strawberry harvesting robot", *International Journal of Advanced Robotic Systems*, vol 5, pp. 41-48, 2008.

[8] S. Kitamura and K. Oka: "Recognition and cutting system of sweet pepper for picking robot in greenhouse horticulture", *Mechatronics and Automation, 2005 IEEE International Conference*, vol. 4, pp. 1807-1812, 2005.

[9] T. Grift, Q. Zhang, N. Kondo, K.C. Ting: "A review of automation and robotics for the bio-industry", *Journal of Biomechatronics Engineering*, vol. 1, pp. 37-54, 2008.

[10] S. Foix, G. Alenyà, and C. Torras: "Lock-in Time-of-Flight (ToF) cameras: a survey", *IEEE Sensors J.*, vol. 11(9), pp. 1917-1926, 2011.

[11] S. Fuchs and G. Hirzinger: "Extrinsic and depth calibration of ToF-cameras", *Proc. 22nd IEEE Conf. Comput. Vision Pattern Recog.*, vol. 1-12, Anchorage, June 2008, pp. 3777–3782.

[12] S. Fuchs and S. May: "Calibration and registration for precise surface reconstruction with time of flight cameras," *Int. J. Int. Syst. Tech. App.*, vol. 5, no. 3-4, pp. 274–284, 2008.

[13] S. Foix, G. Alenyà, and C. Torras: "Towards plant monitoring through next best view", *Proc. 14th Int. Conf. Catalan Assoc. Artificial Intell.*, Lleida, Oct. 2011.

[14] P.F. Felzenszwalb and D.P. Huttenlocher: "Efficient graph-based image segmentation". *International Journal of Computer Vision* 59, pp. 2004, 2004.

[15] J.B. Kruskal: "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem", *Proceedings of the American Mathematical Society*, vol. 7, no. 1, pp. 4850, 1956.

[16] S. Foix, G. Alenyà, J. Andrade-Cetto and C. Torras: "Object modeling using a ToF camera under an uncertainty reduction approach", *IEEE Intl. Conf. on Robotics and Automation (ICRA'10)*, Anchorage, pp. 1306-1312, 2010.