

Curve-based Stereo Matching for 3D Modeling of Plants

Hamid Laga and Stanley J. Miklavcic^{ab}

^a*Phenomics and Bioinformatics Research Centre, University of South Australia*

^b*Australian Centre of Plant Functional Genomics (ACPGF)*

Email: {hamid.laga, stanley.miklavcic}@unisa.edu.au

Abstract: Realistic image-based 3D reconstruction of narrow plants, such as wheat and barley, is challenging because of plant's lack of a well defined surface, complex self-occlusions, and high degree of self-similarities. In this paper we present a new approach for image-based modeling of plants using a pair of cameras in a stereo setup. Our key idea is that while color and texture features are unlikely to match across images, local leaf shape, in particular the shape of its boundary curves, will. We propose a stereo matching algorithm that uses boundary curve features to build a dense depth map. We show that the local orientation of boundary curves can be used to recover the 2D leaf structure of plants, which in turn can be used for matching plant leaves across multiple views in a narrow baseline stereo setup. The proposed method is robust to slight movements of the leaves during imaging, which is an important concern in phenotyping narrow and fragile-leaf plants. We demonstrate results on a number of plants with varying complexity and discuss the performance and limitations of the proposed framework.

Keywords: *Plant phenotyping, orientation fields, 3D reconstruction, plant phenotyping*

1 INTRODUCTION

Plants are ubiquitous in nature and in urban environments. 3D plant models are used in many application areas including agriculture, plant biology, architecture, urban modeling, games and movies. Many methods have been developed for the creation of geometric plant models to enrich the realism of virtual environments (Deussen and Lintermann (2010)). In the past few years however, there is a growing interest in precise 3D modeling of plants from real-world data for biological studies. In such scientific fields, one is interested in measuring automatically, in a non-destructive manner, plant properties such as number of leaves, plant size, leaf shapes, and the 3D structure of the plant. One would like also to analyze and model the variation of these properties during the growth process and in response to various environmental factors, such as soil, weather and light conditions, and to the presence of obstacles.

Typical plant modeling techniques are procedural and are based on carefully designed growth rules (Prusinkiewicz and Lindermayer (1990); Deussen and Lintermann (2004); Talton et al. (2011)). Others, such as sketch-based techniques (Anastacio et al. (2006); Quan et al. (2006); Okabe et al. (2007); Chen et al. (2008)), are interactive, offering the user control and possibility to design the shape of the plant. Although the degree of realism of models created with these methods increasingly improved over the past few years, none of them is suitable for plant scientists who want to use for their studies accurate 3D reconstructions of actual plants. At the other end of the spectrum, image-based techniques have the best potential for producing realistic and accurate 3D plants, since they rely on images of real plants.

There exists a wide body of literature on image-based 3D modeling of real plants. Most of the published algorithms rely on either shape form silhouettes, known also as visual hull (Kutulakos and Seitz (2000); Yous et al. (2007)), or multi-view stereo matching (Seitz et al. (2006)). Visual hull techniques require accurate calibration and synchronization of the cameras. Rotating the plant is not feasible since the plant may undergo non-rigid motion (due to wind). In addition, given the high degree of occlusions in plants, visual hull techniques cannot generate accurate 3D models. Mirror-based systems (Kumar et al. (2012b)) overcome some of these limitations however one loses in the plant resolution. On the other hand, multiview stereo matching techniques rely on feature matching across multiple views of the same plant. However, observed color pixels cannot be matched efficiently across multiple viewpoints since plants exhibit high degree of self occlusions and self similarities. In addition, smoothness constraints along epipolar lines are not applicable when reconstructing plants, since plant leaves are very narrow and thus discontinuities are abundant in plant images.

In this paper we address the image-based plant modeling problem that is fully automatic. We focus on plants with narrow leaves, such as wheat and barley, which exhibit high degree of self occlusions and self similarities. Since individual plant leaves and their features are extremely fine and self similar, they cannot be reliably matched across multiple views. Our key insight is that, in a narrow baseline setup, local orientations of leaves provide stable features across adjacent views. We use this property and propose a single-shot 3D plant modeling system that uses consumer digital cameras. Our reconstruction algorithm analyzes the captured images and computes orientation fields, which represent at each pixel the direction in which there is the strongest oriented plant feature. Using these fields, individual plant leaves are traced then reconstructed and refined in 3D using multi-view stereo. We show that the orientation field, which varies smoothly along leaf boundaries, can be used to automatically recover the plant structure from a single image. This will provide plant scientists with an efficient tool for analyzing the plant structure, counting the number of leaves and measuring their length and surface area. Finally, the advantage of performing stereo-matching using orientation features lies in their robustness against view-dependent appearance of plant leaves and in their uniqueness. We demonstrate this algorithm on plants with various levels of complexity.

2 METHODOLOGY

We use high-resolution RGB cameras arranged in a narrow baseline stereo setup and placed at roughly 1.5m distance from the plant, close enough to capture the plant surface. After background subtraction (Section 2.1), we compute an orientation field and a confidence map for each image (Section 2.2). We then recover the 2D structure of the plant on the images by an efficient integration of the orientation field (Section 2.3). Finally, we formulate the stereo correspondence problem as the problem of matching curves across the pair of images and propose an algorithm for solving it (Section 2.4). We present the experimental results in Section 3 and conclude in Section 4.

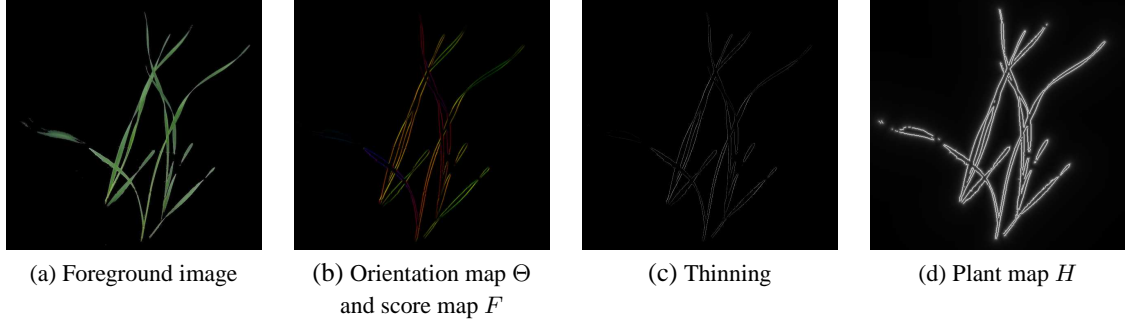


Figure 1. Plant image analysis process: (a) background subtraction, (b) orientation and score maps. Color indicates the orientation angle while intensity represents the magnitude of the maximum filter response, (c) non-maxima suppression of the score map F , (d) the plant map H . (The figure is best viewed in color).

2.1 Plant segmentation

The first stage of the processing is to segment the plant (the foreground) from the image (background). This problem known as background subtraction as been extensively investigated in the literature (Kumar et al. (2003)). In this paper, we learn in a supervised manner the foreground and background using Gaussian Mixture Models (GMM). We use few manually segmented images (typically one to three) as our training set. Let $\{FG, BG\}$ be the two states (foreground, background) of an image pixel \mathbf{s} . The distribution \mathbf{P} of \mathbf{s} may be modeled as a weighted sum of m Gaussians:

$$\mathbf{P}(\mathbf{s}|\Phi) = \sum_{k=1}^m \alpha_k \mathbf{P}(\mathbf{s}|k, \theta_k), \quad (1)$$

where $\Phi = \{\alpha_1, \dots, \alpha_m, \theta_1, \dots, \theta_m\}$, $\theta_k = \{\mu_k, \Sigma_k\}$ are the parameters of the k -th Gaussian, α_k are the mixing coefficients that sum to one, and m is the number of Gaussians in the model. In our experiments we found that one GMM f , with $m = 1$ for the foreground and one GMM g , with $m = 3$ for the background provide good segmentation results. We classify a pixel \mathbf{s} as foreground if the ratio $\frac{f(\mathbf{s}|\Phi_f)}{g(\mathbf{s}|\Phi_g)}$ is larger than one. Φ_f and Φ_g are the parameters of f and g , respectively.

To learn the background model of Equation 1, we first select few images taken in similar conditions as in the system usage, manually segment them using interactive tools, such as Poisson Image Editing of Pérez et al. (2003), and fit a GMM of three Gaussians to the background and one Gaussian to the foreground. In practice, we used as few as three images for training. Note that the training stage should be executed whenever the acquisition conditions (background type, lighting conditions) change. Note also that the segmentation procedure may produce some background noise, which we eliminate using morphological operations.

2.2 2D orientation fields

An oriented filter kernel K_θ is a kernel that is designed to produce high response to features that are oriented along the direction θ when it is convolved with an image. Similar to recent work on hair modeling (Luo et al. (2012)), we define the orientation $\Theta(x, y)$ of an image I at pixel (x, y) as

$$\Theta(x, y) = \arg \max_{\theta} |K_\theta * I(x, y)|. \quad (2)$$

We define also the maximum response of the filter as

$$F(x, y) = \max_{\theta} |K_\theta * I(x, y)|. \quad (3)$$

In what follows, Θ is referred to as the orientation map and F as the score map. Next, we apply non-maxima suppression on the score map F ; the score of a pixel is set to zero unless it is the maximum score in direction orthogonal to θ . The resulting map is thresholded, obtaining a binary mask, and converted into a plant map H :

$$H(x, y) = \frac{1}{1 + d(x, y)}, \quad (4)$$

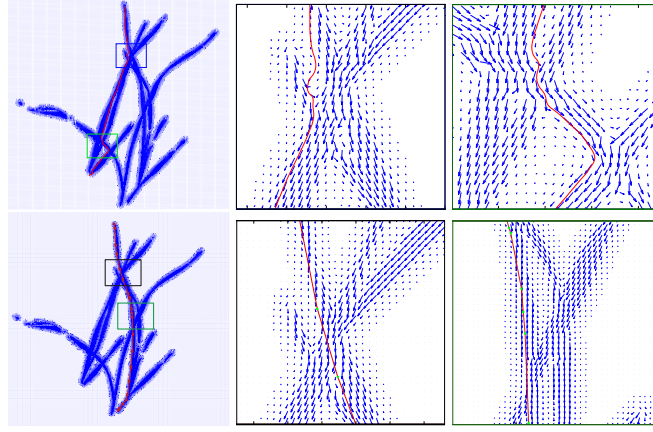


Figure 2. Tracing the vector field for recovering the plant structure using standard vector field integration (top row) and using the algorithm proposed in this paper (bottom row). Note that despite the overlap of several narrow leaves, our approach (bottom row) is able to track them correctly.

where $d(x, y)$ is the Euclidean distance to the closest foreground pixel in the binary mask. H takes a value of 1 where a line feature is present. It decays quickly when moving away from the image features.

We have experimented with several orientation filters, such as Gabor and Difference of Gaussian (DoG) filters. In practice we found that DoG filters provide better orientation fields. Thus, we define K_0 as:

$$K_0(x, y) = (G_\sigma(x) - G_{\sigma'}(x)) G_{\sigma'}(x), \quad (5)$$

where G_σ is a one-dimensional zero-mean Gaussian with standard deviation σ . Similar to Luo et al. (2012), we set $\sigma = 1$ and $\sigma' = 2$. The oriented filters K_θ are generated by rotating the original x -aligned filter K_0 with angles $\theta \in [0, \pi)$. In our experiments we use 180 different orientations, one every degree. The plant map H and orientation field Θ are important since they will be used to recover the structure of the plants from single images and for 3D reconstruction and modeling. Figure 1 summarizes this pipeline and shows examples of orientation and plant maps.

2.3 Recovering the plant structure in 2D

Recovering the structure of plants, especially narrow-leaf ones, from 2D images is very challenging because of the high degree of self-occlusions and self similarities. We propose in this section an algorithm that traces the plant leaves, starting ideally from the tips, and recovers an estimate of their 2D structure from a single image. Tracing is based on the orientation field Θ and the plant map H .

First we select a point $p_s(x_s, y_s)$ with $H(p_s) = 1$ as a seed point and grow it in a direction that is determined by the orientation map Θ . Naive integration of Θ does not recover correctly the plant structure as shown in Figure 2 (top row). Instead, we follow an approach that is similar to Beeler et al. (2012) for hair modeling. Let $p(x, y)$ be the last point on the reconstructed curve. We define a growth cone $\Delta(r, \gamma)$ with growth resolution (cone's height) r and opening angle 2γ . The cone's apex is placed at $p(x, y)$ and its principal axis of symmetry is oriented in the direction $\Theta(x, y)$. Pixels falling inside the cone define potential growth candidates. For each potential growth candidate $p_i(x_i, y_i)$ we compute a score:

$$\zeta(p_i) = \left(1 - \frac{|d\theta|}{2\gamma}\right) H(p_i), \quad (6)$$

where $d\theta$ is the angular deviation between the vectors $v = (\cos \theta, \sin \theta)^t$ and $v_i = (\cos \theta_i, \sin \theta_i)^t$, $\theta = \Theta(x, y)$ and $\theta_i = \Theta(x_i, y_i)$. At each iteration we add to the curve the pixel with the highest score and repeat the procedure starting from the newly added pixel. We iterate until the maximum score of candidate pixels is lower than a threshold value, which we set to 0.1 in our experiments.

The proposed procedure depends on several control parameters. The opening angle γ controls the strength of curvature that is tolerated. In our experiments we found that $\gamma = \pi/12$ achieves good results on the plant

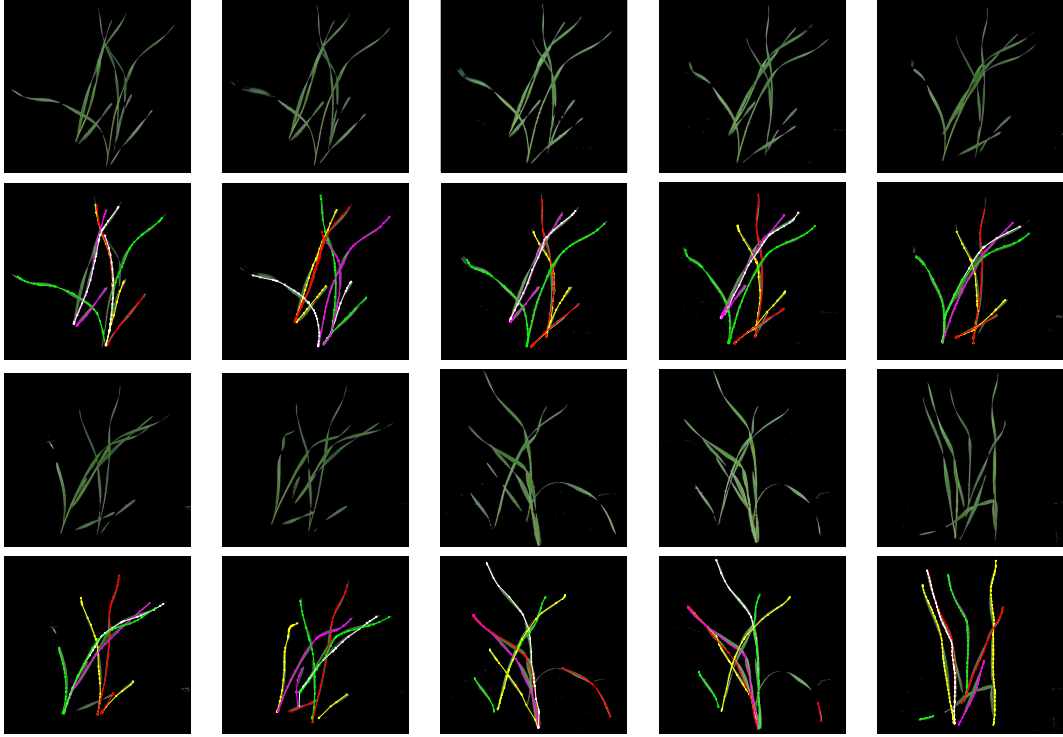


Figure 3. Recovered 2D structures from a single image. First and third rows show the input images after background subtraction. The second and fourth rows shows the recovered plant structure.

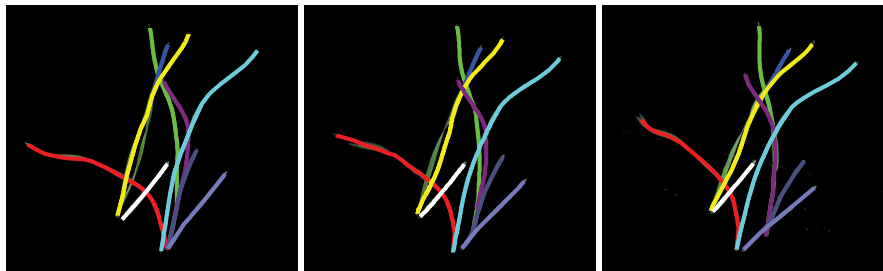
images we have experimented in this paper. The growth resolution r depends both on the thickness of the plant and on the image resolution. The smaller the value of r is, the less robust is the tracing procedure to self-occlusions. Large values, on the other hand, may lead to errors in the tracking. In our acquisition setup and with images of size 764×1148 we found that $r = 50$ provides good results.

The selection of the initial point p_s (the seed point) is very important. In our implementation we choose them to be the leaf tips, which can be efficiently detected using tip detection procedures such as the one proposed by Kumar et al. (2012a). Curves at the tips however have high curvature and thus they should be treated with a special care. In our implementation, in the first five iterations of the algorithm we set the opening angle γ to $\pi/6$. After the five iterations, γ is reset to $\pi/12$. This enables us to automatically seed the integration procedure, while producing plausible results in most of our experiments, see Figure 3 for some examples.

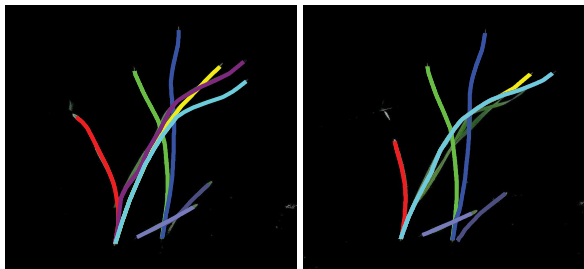
2.4 Curve-based stereo matching

In this section, we focus on recovering the 3D structure of a plant using a stereo matching algorithm that is based on matching curve segments. Our key insight is that in the presence of high degree of self-occlusions and self-similarities, color-based stereo matching would fail to recover accurate 3D structure of the plant. We instead rely on the local orientations of plant features, which provide reliable matches in a narrow baseline setup. Below, we elaborate further on this idea.

Image rectification. Rectification is the process of projecting the stereo images onto a common image plane in such a way that the corresponding points between the two images have the same row coordinates. This process is particularly useful for stereo matching, because the 2-D stereo correspondence problem is reduced to 1-D problem: given a point in the left image, its corresponding point on the right image can be found by searching left of his location along the same line. There are several techniques that can be used for rectification. When camera parameters are available, computed in a calibration step for instance, one can use the algorithm of Fusiello et al. (2000). In the experiments presented in this paper we use the non-calibrated version of the algorithm proposed in Fusiello and Irsara (2011). In both the calibrated and uncalibrated cases, the Matlab source codes are available for download at the authors website.



(a) Leaf correspondences across three consecutive frames.



(b) Correspondences across two consecutive frames.

Figure 4. Leaf correspondences computed with our matching algorithm. Correspondences are color-coded.

Curve matching. Once we have rectified our pair of images, search for correspondences is restricted along the horizontal raster lines. There are various techniques to recover a dense disparity map of the object of interest. The most common approach is to use dense pixel-based stereo matching. However, our objects of interest are narrow plants, such as wheat and barley, which are composed of narrow textureless leaves, with high degree of self-occlusions and self similarities. Thus, standard pixel-based stereo matching does not provide efficient 3D reconstructions. Instead, we propose a curve-based technique that recovers the 3D structure of the plant.

We match all pairs of curves after pruning pairs whose raster line intervals do not overlap. We then assign for each pair of curves c_i and c_j , c_i from the left image and c_j from the right image, a matching score s_{ij} equal to the number of pixels where their local orientations are similar. Using these scores, we compute the three-nearest neighbors for every curve and retain the best reciprocal match. Similar to Sinha et al. (2012), since the images are rectified, pixels where matching curves meet each raster line must correspond. Thus, the computed set of curve correspondences specifies a set of pixel correspondences.

3 RESULTS

Figure 3 shows few results of plant structure recovery from 2D images. In these examples, the images are of the same pot that contains three wheat plants. Each plant has three leaves. Thus, in total we have nine leaves with relatively high level of intersection and overlap. The images are taken around the plant by rotating every time the pot with approximately ten degrees. Thus the plant leaves may undergo some non-rigid deformations due to interaction with wind. As we can see in Figure 3, our algorithm is able to trace the leaves in most of the cases despite the high degree of overlap and self-occlusion.

Figures 4-(a) and (b) show results of our curve matching procedure. In Figure 4-(a), we use the first three frames of Figure 3, while in Figure 4-(b) we use the fifth and sixth frames. We can see that our algorithm can match correctly the leaves despite the high degree of occlusions and non-rigid motion of the leaves. Figure 4-(b) shows also some limitations of our algorithm. In fact, when two leaves merge together and then split, our algorithm fails to track them correctly. We plan to address this issue in future works.

4 CONCLUSIONS

We proposed in this paper an algorithm for recovering the structure of narrow-leaf plants and for matching plant leaves across multiple views. We found that, in absence of color or texture features, local orientations of plant leaves provide reliable features for computing stereo correspondences, which is the first step towards full 3D reconstruction. We plan to extend our approach to handle complex plants at their late growth stages, where

the plant can have many leaves, and explore multi-cue matching as in Kumar et al. (2007). We will explore other applications of the orientation fields, such as non-photo realistic rendering (Ning et al. (2010, 2011)).

REFERENCES

- Anastacio, F., M. C. Sousa, F. Samavati, and J. A. Jorge (2006). Modeling plant structures using concept sketches. In *Proceedings of the 4th international symposium on Non-photorealistic animation and rendering*, pp. 105–113. ACM.
- Beeler, T., B. Bickel, G. Noris, P. Beardsley, S. Marschner, R. W. Sumner, and M. Gross (2012, July). Coupled 3d reconstruction of sparse facial hair and skin. *ACM Trans. Graph.* 31(4), 117:1–117:10.
- Chen, X., B. Neubert, Y.-Q. Xu, O. Deussen, and S. B. Kang (2008, December). Sketch-based tree modeling using markov random field. *ACM Trans. Graph.* 27(5), 109:1–109:9.
- Deussen, O. and B. Lintermann (2004). *Digital Design of Nature: Computer Generated Plants and Organics*. Springer-Verlag.
- Deussen, O. and B. Lintermann (2010). *Digital Design of Nature: Computer Generated Plants and Organics* (1st ed.). Springer Publishing Company, Incorporated.
- Fusiello, A. and L. Irsara (2011, July). Quasi-euclidean epipolar rectification of uncalibrated images. *Mach. Vision Appl.* 22(4), 663–670.
- Fusiello, A., E. Trucco, and A. Verri (2000, July). A compact algorithm for rectification of stereo pairs. *Mach. Vision Appl.* 12(1), 16–22.
- Kumar, P., M. Brooks, and A. Dick (2007). Adaptive multiple object tracking using colour and segmentation cues. In *Computer Vision: ACCV 2007*, Volume 4843 of *Lecture Notes in Computer Science*, pp. 853–863. Springer Berlin Heidelberg.
- Kumar, P., J. Cai, and S. Miklavcic (2012a). Automated detection of root crowns using gaussian mixture model and bayes classification. In *International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pp. 1–7.
- Kumar, P., J. Cai, and S. Miklavcic (2012b). High-throughput 3d modelling of plants for phenotypic analysis. In *Proceedings of the 27th ACM Conference on Image and Vision Computing New Zealand, IVCNZ '12*, pp. 301–306.
- Kumar, P., S. Ranganath, and W. Huang (2003). Queue based fast background modelling and fast hysteresis thresholding for better foreground segmentation. In *Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia*, Volume 2, pp. 743–747.
- Kutulakos, K. N. and S. M. Seitz (2000, July). A theory of shape by space carving. *Int. J. Comput. Vision* 38(3), 199–218.
- Luo, L., H. Li, S. Paris, T. Weise, M. Pauly, and S. Rusinkiewicz (2012). Multi-view hair capture using orientation fields. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1490–1497. IEEE Computer Society.
- Ning, X., H. Laga, S. Saito, and M. Nakajima (2010). IR2s: interactive real photo to Sumi-e. In *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering*, pp. 63–71. ACM.
- Ning, X., H. Laga, S. Saito, and M. Nakajima (2011). Contour-driven sumi-e rendering of real photos. *Computers and Graphics* 35(1), 122 – 134.
- Okabe, M., S. Owada, and T. Igarashi (2007). Interactive design of botanical trees using freehand sketches and example-based editing. In *ACM SIGGRAPH 2007 courses*, SIGGRAPH '07, New York, NY, USA. ACM.
- Pérez, P., M. Gangnet, and A. Blake (2003, July). Poisson image editing. *ACM Trans. Graph.* 22(3), 313–318.
- Prusinkiewicz, P. and A. Lindermayer (1990). *The algorithmic beauty of plants*. Springer-Verlag.
- Quan, L., P. Tan, G. Zeng, L. Yuan, J. Wang, and S. B. Kang (2006, July). Image-based plant modeling. *ACM Trans. Graph.* 25(3), 599–604.
- Seitz, S., B. Curless, J. Diebel, D. Scharstein, and R. Szeliski (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. In *IEEE Conference on Computer Vision and Pattern Recognition*, Volume 1, pp. 519–528.
- Sinha, S. N., K. Ramnath, and R. Szeliski (2012). Detecting and reconstructing 3d mirror symmetric objects. In *European conference on Computer Vision (ECCV)*, pp. 586–600. Springer-Verlag.
- Talton, J. O., Y. Lou, S. Lesser, J. Duke, R. Měch, and V. Koltun (2011, April). Metropolis procedural modeling. *ACM Trans. Graph.* 30(2), 11:1–11:14.
- Yous, S., H. Laga, M. Kidode, and K. Chihara (2007). GPU-based shape from silhouettes. In *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, pp. 71–77.