Multi-View Instance Matching for Plant Leaf Modeling

Takuma Doi[1] Fumio Okura[1,2] Yasuyuki Matsushita[1] Yasushi Yagi[1]
[1]Osaka University [2]JST PRESTO
{doi,okura,yagi}@am.sanken.osaka-u.ac.jp yasumat@ist.osaka-u.ac.jp

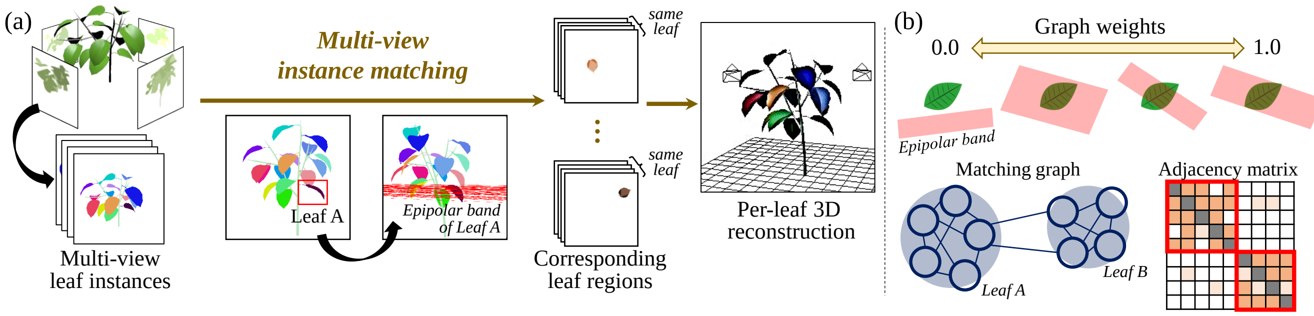

Figure 1: Overview of our method. (a) Proposed pipeline for per-leaf 3D reconstruction. (b) Multi-view instance matching using MRF optimization over a graph where edge weights are defined as the intersection of an epipolar *band* and a leaf instance. After spectral clustering, block diagonals indicate a set of corresponding leaf instances.

1. Introduction

This abstract introduces a work-in-progress report towards detailed modeling of plants from multi-view images. Plant shoot (*i.e.*, leaves and stems) modeling is actively studied among computer vision and graphics fields. Although the state-of-the-art study [2] achieves multi-view reconstruction of 3D branching patterns, detailed modeling of leaves, *i.e.*, estimating location, orientation, and shape of each leaf, still remains an open problem. We propose a multi-view per-leaf 3D reconstruction pipeline depicted in Figure 1(a). The proposed approach finds the correspondences among given leaf instances (*e.g.*, segmented by Mask R-CNN [1]) seen in multi-view images. A set of multi-view correspondences of leaf regions allows 3D reconstruction of each leaf by traditional 3D reconstruction approaches. This abstract summarizes the preliminary report with a focus on a multi-view instance matching method using an epipolar constraint and spectral clustering.

2. Multi-view instance matching

Given segmented instances in each image, finding correspondences among multi-view images is a challenging task. A straight-forward approach is to use an epipolar constraint based on the camera poses; however, the heavy occlusion of leaves causes matching ambiguity, *i.e.*, multiple leaves appear on the same epipolar line, as shown in Figure 1(a). To find the correct match among instances, we aggregate every two-view correspondence scores across multi-view images. We first compute epipolar lines from densely-sampled points on each leaf instance, which forms a *band* in another image. We then generate a matching graph, where each instance in each image forms a node, as shown in Figure 1(b). An edge weight in the graph is defined as the degree of intersection of an epipolar band and a leaf instance, which is analogous to the intersection of union (IoU) computation.

We solve the correspondence estimation as Markov random field (MRF) optimization over the weighted graph. Similar to the normalized-cut method [3], we convert the graph to an adjacency matrix and perform block diagonalization using a spectral clustering method. After spectral clustering, block diagonals indicate a set of corresponding leaf instances.

3. Experiments and discussion

To examine the effectiveness of the instance matching method, we performed a preliminary experiment using simulated plant models created in [2]. Assuming a correct instance segmentation is given, we simulated the instance images from 10 views by rendering each leaf in separate colors. The matching accuracies were 100.0 % and 94.3 % for plants with 4 and 16 leaves, respectively, showing that the proposed method successfully found the correct matches. The failure cases were caused by almost-fully occluded leaves, which we can easily ignore in practical applications by neglecting such small regions. We plan to unify our multi-view instance matching with a state-of-the-art instance segmentation methods and develop a new tool for a detailed plant phenotyping.

Acknowledgment: This work was supported in part by JST PRESTO Grant Number JPMJPR17O3.

References

[1] K. He *et al*. Mask R-CNN. In *ICCV'17*, 2017. 1

[2] T. Isokane *et al*. Probabilistic plant modeling via multi-view image-to-image translation. In *CVPR'18*, 2018. 1

[3] J. Shi and J. Malik. Normalized cuts and image segmentation. *TPAMI*, pages 888–905, 2000. 1