Branching Gaussian Processes
with Applications to Spatiotemporal Reconstruction of 3D Trees

Interdisciplinary Visual Intelligence Lab
http://ivilab.org

Kyle Simek* — kyle@matterport.com
Ravishankar Palanivelu† — rpalaniv@email.arizona.edu
Kobus Barnard‡ — kobus@cs.arizona.edu

*Matterport, Inc. †School of Plant Sciences, Univ. of Arizona ‡Computer Science, Univ. of Arizona

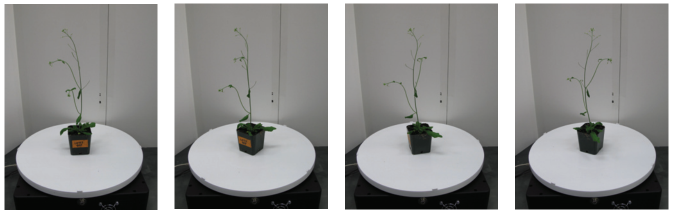
THE UNIVERSITY OF ARIZONA
matterport

Abstract

Problem Statement
- Multi-view 3D reconstruction of plant structure
- Plants undergo significant nonrigid motion during imaging
- Cameras have position & orientation error

Model: Branching Gaussian Process
- Novel probabilistic model for branching tree structure
- Can model smooth curves under random motion
- Guaranteed attachment between curves
- Robust to small camera error

Approach
- Approximate Bayesian inference with expectation propagation
- Bayesian model selection to explore different topologies

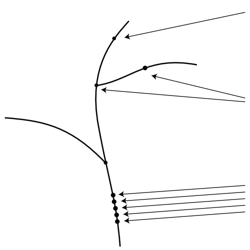


Four of eighteen input images. Plants exhibit nonrigid motion between images, complicating the reconstruction task.

Model

Representation: Curve tree

The *curve tree function* $f(c,t)$ maps point index (c, t) to 3D position.



Curve tree, f(c,t)

Each point has an index (c, t) denoting curve index c and position t.

The *parent function*, $\rho(c, t)$ maps each point (c, t) to its *branch point*, (c', t').

In practice, we model densely sampled points

$$\{(c_1, t_1), (c_2, t_2), \ldots, (c_N, t_N)\}$$

with 3D positions

$$z = \{z_1, \ldots, z_n\} = \{f(c_1, t_1), \ldots, f(c_N, t_N)\}.$$

To model motion, model one tree per time-step, τ : $Z = \{z_\tau\}_{1:T}$

~11k dimensions (200-300 points / plant, 18 time points, 3D points)

Branching Gaussian Process Prior

$p(Z) = \mathcal{GP}(0, k(c,t,\tau,c',t',\tau'))$ where

$$k(c,t,\tau,c',t',\tau') = \underbrace{k_r(c,t,c',t')}_{\text{"Mean tree"}} + \underbrace{k_\tau(\tau,\tau')k_r'(c,t,c',t')}_{\text{Temporal perturbations}}$$

- Spatial covariance (recursive)

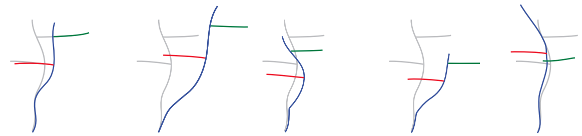
$$k_r(c,t,c',t') = \underbrace{\delta_{cc'}k_C(t,t')}_{\text{Smooth curves}} + \underbrace{k_r(\rho(c,t),c',t')}_{\text{Attachment}},$$

- Temporal covariance (Ornstein-Uhlenbeck process)

$$k_\tau(\tau,\tau') = \exp|\tau - \tau'|/\ell_\tau$$

Properties:
1. Curves are c^n continuous.
2. Curves are attached at branch points.
3. Subtrees are independent, conditioned on branch point.

Training: Maximum marginal likelihood from single specimen.



Random samples from trained prior (non-zero mean used for illustration)

Pixel Likelihood

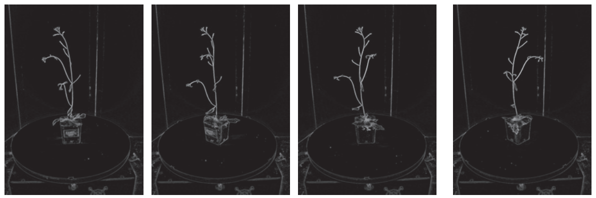
Trained random forest classifier to extract foreground likelihood maps.

For each image:
1. Render tree z_τ into image from time τ as binary foreground image
2. Evaluate each binary pixel $\gamma_{\tau j}$ against likelihood map, independently.

$$p(D_\tau|\boldsymbol{z}_\tau) = \prod_i p(d_i|\mathtt{fg})^{\gamma_{\tau i}} p(d_i|\mathtt{bg})^{1-\gamma_{\tau i}}.$$

Full likelihood:

$$L(Z) = p(Z|D_{1:T}) = \prod_{\tau=1}^{T} p(D_\tau|\boldsymbol{z}_\tau)$$



Foreground log-likelihood maps. Classifier was learned from a single image.

Inference

Goal: minimize $p(Z|D_{1:T}) \propto p(Z)L(Z)$
- Intractable: $L(Z)$ is non-convex, hundreds of dimensions.
- Approximate as block-diagonal Gaussian, $\hat{L}(Z)$ (3x3 blocks)
- Approximate posterior is $q(Z) \propto p(Z)\hat{L}(Z)$
- Minimize KL divergence between $q(Z)$ and $p(Z|D)$

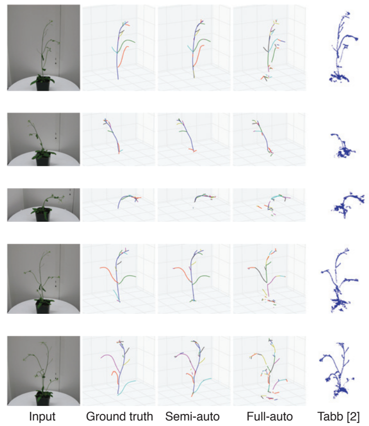
Expectation propagation
1. Bootstrap initial guess for 3D model (see paper for details)
2. For each image D_i:
   For each point z_{ij}:
   - Estimate point posterior $q(z_{ij})$ by importance sampling
   - Extract Gaussian likelihood, $\hat{L}(z_{ij}) \propto q(z_{ij})/p(z_{ij})$

   Construct full likelihood, $\hat{L}(\boldsymbol{z}_i) = \text{blockdiag}_j(\hat{L}(z_{ij}))$
3. Update $q(Z)$ with new $\hat{L}(\boldsymbol{z}_i)$ (Kalman filter, RTS smoother)
4. Repeat 2, 3 until convergence
5. Extract $p(D)$ for model selection

Model selection:

Propose birth/death of stems to optimize p(D) (see paper).

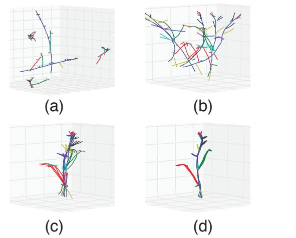
Results



Reconstruction results, compared with Tabb's Shape-from-Silhouette Probability Maps algorithm.

Results using 9 and 18 views. Metrics: Intersection-over-union (IoU); tree-structure consistency (TSC, see paper); Diadem neuron metric.

	IoU	TSC	Diadem [1]
Semi (9)	0.63 ± 0.14	0.49 ± 0.09	0.81 ± 0.09
Semi (18)	0.70 ± 0.08	0.56 ± 0.06	0.84 ± 0.06
Auto (9)	0.50 ± 0.11	0.41 ± 0.07	0.41 ± 0.18
Auto (18)	0.43 ± 0.08	0.59 ± 0.11	0.44 ± 0.18



Samples from the posterior, after conditioning on (a) zero (prior), (b) one, (c) three and (d) 18 views. Note: image (a) not to scale.

References:

[1] Gillette, T.A., Brown, K.M., Ascoli, G.A.: The DIADEM Metric: Comparing Multiple Reconstructions of the Same Neuron. Neuroinformatics 9(2-3) (April 2011)

[2] Tabb, A.: Shape from Silhouette Probability Maps: Reconstruction of Thin Objects in the Presence of Silhouette Extraction and Calibration Error. (CVPR 2013)

This work was supported by the iPlant Collaborative under NSF Award Numbers DBI-0735191 and DBI-1265383 and the Department of Education's GAANN Fellowship.
Special thanks to Dr. Amy Tabb for sharing her shape from silhouette probability maps code for our evaluation.