

Gaussian Process Shape Models for Bayesian Segmentation of Plant Leaves

Kyle Simek
ksimek@email.arizona.edu
Kobus Barnard
kobus@sista.arizona.edu

Interdisciplinary Visual Intelligence Lab
University of Arizona
Tucson, AZ

Abstract

We develop a novel probabilistic model for multi-part shapes based on Gaussian processes, which we apply to model rosette leaves of *Arabidopsis* plants. Our model incorporates domain knowledge of *Arabidopsis* leaves in two ways. First, leaves are modeled using two anatomical parts: a *blade* and a *petiole*. We model the two regions with separate Gaussian processes, with a smoothness constraint at the boundary. Second, we constrain all leaf petioles to initiate at the rosette center, which is also modeled. This Bayesian prior is combined with a simple likelihood function over foreground pixels to perform image segmentation by optimizing a posterior distribution. A simple data-driven approach is used to over-segment the image, then excess leaves are pruned using a Bayesian model selection criterion. We show that our approach is effective, even with minimal training data.

1 Introduction

The relationship between genetics and observable traits (phenotypes) is a central question in biological research. This interest has been driven by advances in inexpensive high-throughput genotyping methods, which has exposed a lack of phenotyping methods that are similarly efficient, inexpensive, and high-throughput. Rapid, inexpensive, and minimally invasive methods have been developed for collecting biological image data, and the need for efficient methods for analyzing this data is now paramount [9].

We present an approach for segmenting *Arabidopsis* rosette leaves using a new probabilistic 2D shape model based on Gaussian processes (GPs). Rosette segmentation using appearance alone is difficult in this scenario due to the subtlety of boundary cues and confounding leaf texture. Using a strong shape model can significantly improve segmentation when image evidence is weak.

Our model incorporates domain knowledge of *Arabidopsis* leaves in two ways. First, leaves are modeled as having using two anatomical parts: a *blade* and a *petiole*. We propose a piecewise Gaussian process that models the two regions as conditionally independent, given the transition points between them. This provides a better fit to the training data compared to a single GP, by allowing the blade region to vary more widely than the petiole. Second, we constrain all leaf petioles to initiate at the rosette center, which is also modeled. The result is a set of smooth leaves of varying shapes and sizes whose boundaries are c^0 continuous at the petiole tip and c^1 continuous everywhere else. We use this probabilistic shape prior in

conjunction with a simple foreground-based likelihood function to infer a set of rosette leaf shapes for a testing image, from which an image segmentation is obtained.

One problem inherent to segmentation is the question of how many segments are optimal? Some approaches use heuristic penalties on shape size or segment count to prevent over-segmentation, but these often do not generalize well to new datasets. Instead, we pose this as a problem of Bayesian model selection and propose an efficient method for comparing segmentations using the marginal model posterior. This requires evaluating an intractable integral; we propose an approximation that exploits our shape model’s Gaussian form, allowing the integral to be estimated analytically.

We first describe the Gaussian process leaf model, the foreground likelihood model, and the training process for both. We then outline a simple method for inference based on an initial over-segmentation, followed by a pruning step to maximize the marginal model posterior. The general framework we outline here is a proof of concept — the fact that it works well despite a weak likelihood function and simple inference algorithm shows the effectiveness of our shape model in the face of challenging circumstances.

2 Related work

Image segmentation continues to be a topic of great interest in computer vision, despite decades of research. One reason for this is that many segmentation tasks are based on domain-specific criteria, so no single segmentation algorithm can solve every problem. Many general-purpose approaches use Markov random fields (MRFs) to model interactions between neighboring regions and image evidence, and they use graph-theoretic algorithms for inference; for a survey, see Peng et al. [12]. Most of these techniques only incorporate very weak prior knowledge during segmentation, like a preference against small segments or in favor of smoothness within segments and boundaries at strong edges [1, 7, 14]. Tu and Zhu pose segmentation as Bayesian inference, learning appearance models for several different region types and Bayesian priors over boundary smoothness, region size, and number of regions [18]. Gaussian processes have been applied to Bayesian segmentation problems only recently. In one formulation, segments are modeled as level sets of multiple intersecting Gaussian process surfaces. This allows for nonlocal interactions and generalizes to scenes with varying complexity [3, 16]. Like us, the authors approximate the marginal posterior to compare segmentations of varying dimension, but they use a different approximation based on variational inference. Active shape models [2] model deformable shapes using a probabilistic framework. During training, a low-rank Gaussian model is learned from aligned training shapes with manually-chosen landmark points. The resulting model captures high dimensional modes of variation, which are used to deform a mean shape to fit image data. This model was applied to analyze leaf shape by Weight et al. [19], and Zhang et al. use a similar Gaussian model for 3D leaf petals [21]. Laga et al. propose a method for aligning leaf shapes that obviates the need for manually-chosen landmark points [8]. Our shape model also learns a Gaussian model from aligned shapes, but instead of estimating a full $N \times N$ covariance matrix, we estimate a covariance function that can represent a wide variety of leaf deformations with only a few parameters. As a result, our model is less prone to overfitting when training data is minimal and is robust to partial occlusion of the training shapes.

Automatic leaf segmentation has seen increasing interest in recent years. Minervini et al. propose an approach to the simpler problem of binary segmentation of *Arabidopsis* rosettes from background. They use an active contour model using a level-set representation but with a prior over foreground location, not shape [10]. Shape regularization is accom-

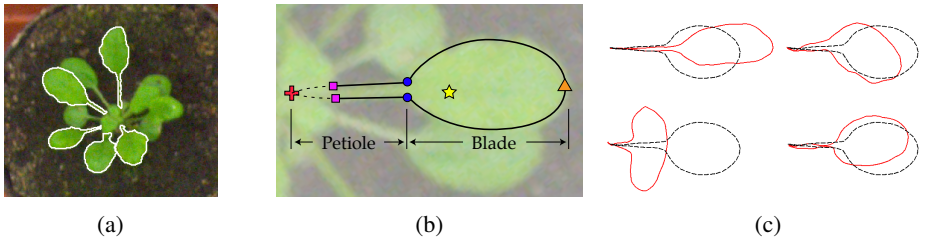


Fig. 1: (a) Example of training shapes (solid white). Only leaves with unoccluded blades were used (104 shapes); all shapes exhibit occlusion of the petiole. (b) Leaf shapes have a petiole and a blade region. Training shapes are aligned by rosette center (red plus) and rotated so the center of mass (yellow star) lies on the $+x$ axis. Shape contours are parameterized based on four landmark points: the petiole tip $z(-\pi) = z(\pi)$ (red cross), the blade tip $z(0)$ (orange triangle), and region transition points $z(-\pi/2)$ and $z(\pi/2)$ (blue dots). Occlusion boundaries (magenta squares) vary by training shape and are parameterized by proximity to the petiole tip. All other points are parameterized by linear interpolation. (c) Four samples from the trained shape prior (2) (solid red) with learned mean shape (dotted black).

plished by Gaussian smoothing in post-processing, and so it does not influence the optimization process. Cerutti et al. model leaf shape using a four-parameter polygon model and with inference based on active contours to segment images containing individual leaves. Regularization is imposed heuristically to avoid unnatural shapes [5]. Our approach is in a similar vein, but with a nonparametric probabilistic shape model, a regularizing prior learned from training shapes, and a task of multi-leaf segmentation, not single-leaf. Teng et al. segment multiple leaves using several closely-spaced views to recover 3D point clouds. Segmentation is then performed using both 2D and 3D data together, yielding high quality results [17]. Pape et al. perform 2D multi-leaf segmentation using a nonparametric color model for initial binary segmentation; leaf centers were estimated by local maxima of the foreground image’s distance transform [11]. However, for segmentation, that approach uses features of the foreground mask that are absent when leaves overlap significantly. We argue that stronger shape models will be needed to see further gains on this task.

3 Prior and likelihood

3.1 Prior

We have developed a probabilistic shape model based on random functions with Gaussian process priors. A random function $f(t) : T \rightarrow F$ is a random mapping from input space T to output space F . It may be interpreted as a set of random variables that are indexed by the input t . For continuous functions, the number of random variables is uncountably infinite, and we model them with a Gaussian process (GP) prior, which is an extension of the Gaussian distribution to uncountably infinite sets of random variables. The defining property of the Gaussian processes is that any finite subset of random variables is Gaussian distributed. A GP is defined by two functions: a mean function $\mu(t)$ that gives the mean of a random variable $z(t)$ for any $t \in T$; and a covariance function $k(t, t')$ that gives the covariance between any two random variables $z(t)$, and $z(t')$.

We represent shape by a vector-valued random function, $z : (-\pi, \pi] \rightarrow \mathbb{R}^2$, where the two output dimensions correspond to x and y coordinates of the 2D shape boundary. It is

convenient to represent the vector-valued shape function as a scalar-valued function with an extra input representing the output dimension of interest, $z : (-\pi, \pi] \times \{1, 2\} \rightarrow \mathbb{R}$, where $z(t, 1)$ represents the x dimension and $z(t, 2)$ represents the y dimension.

We place a Gaussian process prior over z ,

$$z(t, d) \sim \mathcal{GP}(\mu(t, d), k(t, d, t', d')), \quad (1)$$

where $k(\cdot)$ is a covariance function defined in Sec. 3.2 and the mean $\mu(\cdot)$ is learned from data in Sec. 4. Modeling a continuous random function explicitly is impractical, so we select a finite subset of the input space $\mathbf{t} = (t_1, \dots, t_n)$ and model the corresponding random variables

$$\mathbf{z} = (z(t_1, 1), \dots, z(t_n, 1), z(t_1, 2), \dots, z(t_n, 2))^T.$$

These random variables represent the (x, y) coordinates of sampled points on the shape boundary. We choose \mathbf{t} to be 100 values uniformly spaced in $(-\pi, \pi]$, which is dense enough to capture fine detail in leaf shapes. By properties of Gaussian processes, the resulting 200-dimensional random variable is Gaussian distributed

$$\mathbf{z} \sim \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} K_x & \mathbf{0} \\ \mathbf{0} & K_y \end{bmatrix}\right), \quad (2)$$

where

$$\mu_{x,i} = \mu_x(t_i, 1), \quad \mu_{y,i} = \mu_y(t_i, 2), \quad (3)$$

$$K_{x,ij} = k(t_i, 1, t_j, 1), \quad K_{y,ij} = k(t_i, 2, t_j, 2). \quad (4)$$

The shape model also has an angle parameter ϕ , which has uniform distribution $p(\phi) = \mathcal{U}(0, 2\pi)$. Rotation is not well modeled by Gaussian processes, so the GP is learned using shapes aligned to a default orientation, and the angle parameter models rotation of the aligned shape around the rosette center.

Multiple leaves are treated as independent. Let $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$ be a set of n independent leaf shapes \mathbf{z}_i with angles $\Phi = \{\phi_1, \dots, \phi_n\}$. The full shape prior is then

$$p(\mathbf{Z}, \Phi) = p(\mathbf{Z})p(\Phi) = \prod_i p(\mathbf{z}_i)p(\phi_i). \quad (5)$$

The model assumes all leaves initiate at the rosette center, $c \in \mathbb{R}^2$, which is uniformly distributed over image dimensions $W \times H$,

$$p(c) = 1/(WH) \quad c_1 \in [0, W), c_2 \in [0, H). \quad (6)$$

The rosette center variable acts as a global translation transformation when rendering the shapes in the likelihood function. The full prior is then,

$$p(c, \mathbf{Z}, \Phi) = p(c)p(\mathbf{Z})p(\Phi). \quad (7)$$

3.2 Multi-region covariance function

The GP covariance function controls properties of a random shape's variation from the mean. We tailor a covariance function specifically to *Arabidopsis* leaves by applying standard covariance functions in four ways: (1) summing covariance functions; (2) applying different covariance functions in different shape regions; (3) adding constraints to specific points; and

(4) transforming the covariance function’s input space. These techniques allow our model to generate a wide variety of realistic leaf shapes with relatively few training parameters. Because training shapes are aligned with the x -axis, the x and y dimension exhibit different random characteristics, which we model independently. Below we construct a covariance function that is applied to both dimensions but with different trained parameters.

Leaves are modeled with two semantic regions, the *blade* and the *petiole* (see Fig. 1b). Shape functions are parameterized so $z(0)$ is the extreme tip of the blade, $z(\pi)$ is the tip of the petiole, and the transitions between regions are $z(-\pi/2)$ and $z(\pi/2)$. We model the two regions with different covariance functions that are joined smoothly at the transition points.

We start by defining the petiole and blade covariance functions,

$$k_p(t, t') = \lambda_p^2 k_m(t, t', \sigma_p), \quad (8)$$

$$k_b(t, t') = \lambda_{b_1}^2 k_m(t, t', \sigma_{b_1}) + \lambda_{b_2}^2 k_m(t, t', \sigma_{b_2}), \quad (9)$$

where k_m is the Matern kernel with $\nu = 3/2$ [20],

$$k_m(t, t', \sigma) = \left(1 + \frac{\sqrt{3}|t - t'|}{\sigma} \right) \exp\left(\frac{-\sqrt{3}|t - t'|}{\sigma} \right). \quad (10)$$

The Matern kernel generates functions that are c^1 continuous, and the parameter σ controls the smoothness scale. The blade region’s covariance (9) is the sum of two Matern kernels, allowing for large- and small-scale variation to occur simultaneously. Only one Matern term is used for the simpler petiole region.

To combine the two region covariances, we use the “Markov link kernel” of Reece et al. [13], constraining the random functions to be c^1 smooth at the transition point. Define

$$\partial k(t, t') = \frac{\partial}{\partial t'} k(t, t') \quad \text{and} \quad \partial \partial k(t, t') = \frac{\partial \partial}{\partial t \partial t'} k(t, t'). \quad (11)$$

Let $\mathbf{t}_0 = (-\pi/2, \pi/2)^\top$ to be the indices of the transition between regions. The Markov link kernel for the two regions is

$$k_0(t, t') = \begin{cases} k_b(t, t') + g_b(t) (K_c - M_b) g_b(t')^\top & t, t' \in (-\pi/2, \pi/2) \\ k_p(t, t') + g_p(t) (K_c - M_p) g_p(t')^\top & t, t' \notin (-\pi/2, \pi/2) \\ g_b(t) K_c g_p(t')^\top & t \in (-\pi/2, \pi/2], t' \notin (-\pi/2, \pi/2) \\ g_p(t) K_c g_b(t')^\top & \text{otherwise,} \end{cases} \quad (12)$$

where

$$g_\times(t) = [k_\times(t, \mathbf{t}_0) \quad \partial k_\times(t, \mathbf{t}_0)] M_\times^{-1}, \quad M_\times = \begin{bmatrix} k_\times(\mathbf{t}_0, \mathbf{t}_0) & \partial k_\times(\mathbf{t}_0, \mathbf{t}_0) \\ \partial k_\times(\mathbf{t}_0, \mathbf{t}_0) & \partial \partial k_\times(\mathbf{t}_0, \mathbf{t}_0) \end{bmatrix},$$

and $K_c = \text{diag}([\sigma_0^2 \quad \sigma_0^2 \quad \sigma_1^2 \quad \sigma_1^2])$. Here, σ_0^2 is the variance of the transition points and σ_1^2 is the variance of their first derivative.

The petiole always initiates at the rosette center, so we constrain the initial petiole point and its derivative to have zero variance,

$$k_1(t, t') = k_0(t, t') - g_0(t) M_0^{-1} g_0(t')^\top. \quad (13)$$

We also want the shape to scale and bend from tip to tip. For this we add two linear terms with transformed inputs,

$$k_2(t, t', d) = k_1(t, t') + \lambda_{lm} k_L(\mu(t, t', d), \mu(t, t', d)) + \lambda_{lc} k_L(\cos(t/2), \cos(t'/2)) \quad (14)$$

where $k_L(t, t') = tt'$ is the standard linear covariance function. The middle term transforms the inputs of k_L by the mean function $\mu(\cdot)$, which generates random offsets proportional to the mean shape. The second term transforms k_L by a sinusoid, which models variation in the blade tip while the petiole tip remains stationary.

Both the x and y dimensions are modeled using covariance function of the form (14), substituting 1 and 2 for the d input, respectively. Let $k_{2,x}$ be the shape covariance in the x dimension, with parameters

$$(\sigma_{0,x}, \sigma_{1,x}, \sigma_{p,x}, \lambda_{p,x}, \sigma_{b1,x}, \lambda_{b1,x}, \sigma_{b2,x}, \lambda_{b2,x}, \lambda_{lm,x}, \lambda_{lc,x}).$$

Similarly, $k_{2,y}$ is the shape covariance in the y dimensions, with parameters

$$(\sigma_{0,y}, \sigma_{1,y}, \sigma_{p,y}, \lambda_{p,y}, \sigma_{b1,y}, \lambda_{b1,y}, \sigma_{b2,y}, \lambda_{b2,y}, \lambda_{lm,y}, \lambda_{lc,y}).$$

Because variation in the x dimension is simpler than in the y dimension, in practice we set $\lambda_{b2,x}$, $\sigma_{b2,x}$, and $\sigma_{1,x}$ to zero. The resulting full shape covariance function has 17 parameters, which are learned from training data in Section 4.

3.3 Likelihood

If image evidence is sufficiently informative, it can be difficult to assess the prior’s effectiveness, because the likelihood dominates. Our goal is to demonstrate a strong shape prior, so we choose a likelihood function that uses only known foreground masks as evidence. Such masks are provided by the 2015 CVPPP Leaf Segmentation Challenge dataset. Let $D = \{d_i\}$ be the binary foreground mask with N pixels $d_i \in \{0, 1\}$, where 1 denotes foreground. Let P be the space of closed 2D polygons. We define a rendering function $\rho : \mathcal{P}^M \rightarrow \{0, 1\}^N$ that maps M shape polygons $\mathbf{z}_i \in P$ to a rasterized binary map by setting all pixels lying on the interior of any polygon to one and all other pixels to zero. The likelihood function is a product of i.i.d. terms over all pixels,

$$p(D|\mathbf{Z}) = \prod_i p(d_i|\text{fg})^{\rho_i(\mathbf{Z})} p(d_i|\text{bg})^{1-\rho_i(\mathbf{Z})} \quad (15)$$

where $p(d_i|\text{fg})$ and $p(d_i|\text{bg})$ are Bernoulli distributions with parameters μ_{fg} and μ_{bg} , respectively. During inference, this likelihood gives strong preference to leaf shapes whose rendering overlaps maximally with white pixels in the foreground mask and minimally with black pixels. Our general approach permits more sophisticated likelihood functions (e.g. using color or texture), but the foreground likelihood is simple to train and is sufficient to illustrate the effectiveness of our shape model.

4 Training

For training, we selected 104 example shapes across 28 rosettes. Data was obtained from the CVPPP Leaf Segmentation Challenge dataset A1 [15], which provides individual leaf segmentation for 128 rosettes. Leaf segments without occluded blades were manually chosen in each of 28 training rosettes. Polygon shapes were obtained from boundary pixels on each selected leaf segment. Some example training shapes are shown in Fig. 1a. The location of the rosette center was labeled in each image and transition between petiole and blade regions were labeled for each training shape.

Shapes were rotated and translated so the rosette center lies at the origin and the shape’s center of mass lies on the positive x -axis. The left part of the petiole is occluded in most

training shapes, so we discard the leftmost 10 points so training shapes do not include occlusion boundaries. The result is an open polygon with a start and end point, with points ordered counter-clockwise.

The points of each training polygon were parameterized as described in Fig. 1b, so the petiole tip, blade tip, and bottom and top transition points correspond to $z(\pi)$, $z(0)$, $z(-\pi/2)$, and $z(\pi/2)$, respectively. Endpoints of the observed shape polygon (magenta squares in Fig. 1b) were assigned indices as follows. Let d_1 (resp. d_2) be the distance from the first (last) polygon point to the origin, and let L_1 (L_2) be its Euclidean distance from the blade tip, $z(0)$. The first and last polygon points are assigned indices $-(1 - d_1/(d_1 + L_1))\pi$ and $(1 - d_2/(d_2 + L_2))\pi$, respectively. All other points are assigned indices by linearly interpolating between these six landmark points.

4.1 Training the prior

The prior training procedure is detailed in the supplemental material. Briefly, the GP prior mean was obtained by averaging the training polygons. The 17 prior covariance parameters were optimized using maximum marginal likelihood [20]; the x and y dimensions were trained independently, and we developed a method to update parameters in blocks to mitigate local minima (see supplemental material).

4.2 Training the likelihood

The Bernoulli likelihood parameters μ_{fg} and μ_{bg} are trained from ground-truth foreground segmentations using maximum likelihood. We set μ_{fg} is set to the percentage of training foreground pixels overlapping with foreground pixels in the corresponding original mask, after dilating the training mask by two pixels. Similarly, μ_{bg} to the percentage of background training pixels that overlap foreground pixels in the original mask, after eroding the training mask by two pixels. Since the training data is effectively noise-free, this dilation and erosion adds some tolerance for imperfect fitting around the boundaries. The choice of two pixels of dilation is made empirically; using a larger value allows the inferred shape to stray farther into regions lacking evidence, while using a smaller value forces inferred shapes to be nearly exact, which tends to overfit.

5 Inference

We use a data-driven approach to hypothesize leaf shapes, and use local minimization to find the maximum posterior configuration of the hypothesized shapes. Later, we derive a Bayesian model selection criterion and use it to prune spurious shapes. We estimate the rosette center c by taking the morphological skeleton of the foreground mask, converting it to an adjacency graph, and choosing the vertex with minimal eccentricity. We then measure the distance from each foreground point to c and select extreme points $\{p_i\}$ whose radius is a local maximum. For each p_i , we construct a shape \mathbf{z}_i and rotation θ_i so the blade's tip $z(0)$ lies on p_i and the petiole's tip $z(\pi)$ lies on c . Given this initial model (\mathbf{Z}, Φ, c) , we estimate the maximum posterior,

$$\arg \max_{(\mathbf{Z}, \Phi)} p(\mathbf{Z}, \Phi | D) \propto p(D | \mathbf{Z}, \Phi, c) \prod_i p(\mathbf{z}_i). \quad (16)$$

Given N hypothesis shapes with 100 points each, this is an optimization problem in $200N + 1$ dimensions, but the effective dimension is much lower. Let $K = UDU^T$ be the eigendecom-

Method	Dice (%)	Symm. Dice (%)	Count Err.	Abs. Count Err.
Ours	86.5 ± 7.4	68.6 ± 8.1	-4.3 ± 1.8	4.3 ± 1.8
Pape et al. [11]	N/A	74.2 ± 7.7	-1.9 ± 2.5	2.6 ± 1.8

Tab. 1: Segmentation results. The Dice metric measures overlap between the testing and ground truth segmentations, averaged over all testing images. ‘‘Symm. Dice’’ is the symmetric best Dice metric. For details, see Scharr et al. [15]. Note that Pepe et al. is evaluated on 128 training images, whereas we evaluate on 90 held-out images. ‘‘Count Err.’’ and ‘‘Abs. Count Err.’’ are equal, because all count errors are negative.

position of the shape prior covariance matrix from Eq. (2), which defines a shape-space $U^\top \mathbf{z}$ whose dimensions are uncorrelated. Eigenvalues in D less than 0.25 have less than one pixel of random variation, so they are discarded along with the corresponding columns of U ; in our model the resulting space has 25 dimensions per shape. The optimization (16) converges much faster when run in this reduced space. To further improve optimization speed, each shape is optimized independently in a first pass, before optimizing all shapes jointly.

The initial model estimate is likely to have false positive leaves, so we next apply pruning. Bayesian posterior densities of differing dimension cannot be compared directly, so applying a maximum posterior criterion for pruning is invalid. Instead, we compare the marginal posteriors of the different models, after integrating out all continuous parameters. The model is characterized by the number of segments $N = |\mathcal{Z}|$. The marginal posterior of N is

$$p(N|D) = \int p(N, \mathbf{Z}, \Phi|D) d(\mathbf{Z}, \Phi) \quad (17)$$

$$= p(N) \int p(\mathbf{Z}, \Phi|N) p(D|\mathbf{Z}, \Phi) d(\mathbf{Z}, \Phi). \quad (18)$$

For simplicity, we model the prior $p(N)$ as uniform with infinite support and omit it from the computation. The integral (18) is intractable in general, but for a fully linear Gaussian model it can be computed analytically. Our prior is Gaussian, but the likelihood is nonlinear, so exact analytical evaluation is not possible. Instead, we apply the Laplace approximation [4],

$$p(N|D) = p(N) p(\tilde{\mathbf{Z}}, \tilde{\Phi}|N) p(D|\tilde{\mathbf{Z}}, \tilde{\Phi}) (2\pi)^{d/2} \prod_{i=1}^N |K^{-1} + \Psi_i^{-1}|^{1/2}, \quad (19)$$

where $(\tilde{\mathbf{Z}}, \tilde{\Phi})$ are the maximum posterior parameters from (16), and Ψ_i is the Hessian of the likelihood (15) w.r.t. (\mathbf{z}_i, θ_i) , evaluated at $(\tilde{\mathbf{Z}}, \tilde{\Phi})$. The Laplace approximation gives the exact solution to (18) under an approximation of the likelihood as Gaussian. We use numerical differentiation to approximate each Ψ_i as a diagonal matrix.

To remove false positives, we propose pruning each leaf in the model sequentially, accepting the proposal if the marginal model posterior (19) improves. This simple inference procedure works reasonably well in practice, and it illustrates a general framework that can be applied to much more sophisticated inference strategies. For example, adding birth proposals could improve the false-negative rate, and split and merge moves could fix under- or over-segmented regions, respectively.

6 Experiments

We tested our model on the A1 training set of the CVPPP 2015 Leaf Segmentation Challenge (LSC) dataset [15]. The dataset has 128 images with ground-truth segmentations. We used



Fig. 2: Example segmentation output by our algorithm (best viewed in color). **Left:** Raw segmentation. **Middle:** Segmentation after applying foreground mask. **Right:** Ground truth segmentation. Most errors result from mistakes made by our initialization heuristic, which could be fixed by additional trans-dimensional moves. For example, few interior leaves are found, which could be corrected by adding random birth moves.

the first 28 images to train our model. The next 10 images were used to tune the optimization procedure. We evaluated our algorithm on the last 90 images.

To evaluate, a segmentation is generated from the shape model as follows. Beginning with a $W \times H$ map of zeros, the shapes are ordered by decreasing area and render shapes in order using the painter’s algorithm [6]. Each shape is drawn with a pixel intensity equal to its ordinal number in the sorted list. As a result, small leaves appear in front of larger leaves, as they tend to appear in *Arabidopsis* rosettes.

Segmentation performance is evaluated using the three Leaf Segmentation Challenge metrics [15]. Results are summarized in Table 1, and an example segmentation is shown in Fig. 2. Errors arise from two main causes. First, segment count is consistently underestimated, because the leaf proposal strategy targets only exterior leaves, missing interior leaves. Results could be improved by using additional cues for proposals or just randomly sampling shapes from the prior. Errors also occur when the rosette center is estimated wrongly, which could be improved by optimizing it jointly with the shape polygons.

Recall that our method uses foreground masks as weak evidence and a rudimentary bottom-up method for proposing shapes. This forces segmentations to be informed primarily by the prior shape model. Results show that our shape model is effective at estimating segmentations despite weak evidence, showing its promise as a general purpose leaf prior. Although Pape et al. achieve better results, they do so using a strong appearance model, more sophisticated mechanism for proposing segments. Our shape model could be incorporated into approaches like theirs to identify unnaturally shaped segments and to estimate shape boundaries when edge evidence is absent.

7 Conclusion

We have described a novel probabilistic model for multi-part shapes using piecewise Gaussian processes. We use prior knowledge to add constraints to our model, which contributes robustness to partial occlusions during training and testing and helps the model generalize despite relatively few training examples. We apply this model in a simple Bayesian framework for model-based image segmentation and show that it is effective at distinguishing individual leaves, even with partial occlusion and a weak evidence model. In addition to leaf segmentation, our model also semantically segments the leaf’s anatomical parts (the blade and petiole) and can provide shape estimates even in occluded regions. We are eager to apply similar models in more powerful inference frameworks and to investigate unsupervised methods for learning multi-region probabilistic shape models.

References

- [1] Y Boykov and V Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(9):1124–1137, 2004.
- [2] Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
- [3] Soumya Ghosh and Erik B Sudderth. Nonparametric learning for layered segmentation of natural images. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2272–2279, 2012.
- [4] Walter R Gilks, Sylvia Richardson, and David Spiegelhalter. *Markov Chain Monte Carlo in practice: interdisciplinary statistics*, volume 2. Chapman & Hall/CRC, 1995.
- [5] Laure Tougne Antoine Vacavant Guillaume Cerutti and Didier Coquin. A Parametric Active Polygon for Leaf Segmentation and Shape Estimation. pages 1–12, August 2011.
- [6] John F Hughes, Andries Van Dam, James D Foley, and Steven K Feiner. *Computer graphics: principles and practice*. Pearson Education, 2013.
- [7] V Kolmogorov and R Zabih. What energy functions can be minimized via graph cuts? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(2):147–159, 2004.
- [8] Hamid Laga, Sebastian Kurtek, Anuj Srivastava, and Stanley J Miklavcic. Landmark-free statistical analysis of the shape of plant leaves. *Journal of Theoretical Biology*, 363:41–52, 2014.
- [9] M Minervini, H Schar, and S A Tsafaris. Image Analysis: The New Bottleneck in Plant Phenotyping. *IEEE Signal Processing Magazine*, 32(4):126–131, July 2015.
- [10] Massimo Minervini, Mohammed M Abdelsamea, and Sotirios A Tsafaris. Image-based plant phenotyping with incremental learning and active contours. *Ecological Informatics*, 23:35–48, 2014.
- [11] Jean-Michel Pape and Christian Klukas. 3-D Histogram-Based Segmentation and Leaf Detection for Rosette Plants. In Lourdes Agapito, Michael M Bronstein, and Carsten Rother, editors, *Computer Vision - ECCV 2014 Workshops*, pages 61–74. Springer International Publishing, 2015.
- [12] Bo Peng, Lei Zhang, and David Zhang. A survey of graph theoretical approaches to image segmentation. *Pattern Recognition*, 46(3):1020–1038, 2013.
- [13] S Reece, R Garnett, M A Osborne, and S J Roberts. Anomaly detection and removal using nonstationary Gaussian processes. *University of Oxford, Oxford, UK, Tech. Rep.*, 2009.
- [14] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "GrabCut": interactive foreground extraction using iterated graph cuts. *SIGGRAPH '04: SIGGRAPH 2004 Papers*, August 2004.
- [15] Hanno Schar, Massimo Minervini, Andreas Fischbach, and Sotirios A Tsafaris. Annotated image datasets of rosette plants. Technical report, 2014.
- [16] Erik B Sudderth and Michael I Jordan. Shared segmentation of natural scenes using dependent Pitman-Yor processes. In *Advances in Neural Information Processing Systems*, pages 1585–1592, 2009.
- [17] Chin-Hung Teng, Yi-Ting Kuo, and Yung-Sheng Chen. Leaf segmentation, classification, and three-dimensional recovery from a few images with close viewpoints. *Optical Engineering*, 50(3), 2011.

- [18] Zhuowen Tu and Song-Chun Zhu. Image segmentation by data-driven Markov chain Monte Carlo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):657–673, 2002.
- [19] Caroline Weight, Daniel Parnham, and Richard Waites. LeafAnalyser: a computational method for rapid and large-scale analyses of leaf shape variation. *The Plant Journal*, 53(3):578–586, 2008.
- [20] CKI Williams and CE Rasmussen. *Gaussian processes for machine learning*. MIT Press, 2006.
- [21] Chenxi Zhang, Mao Ye, Bo Fu, and Ruigang Yang. Data-driven Flower Petal Modeling with Botany Priors. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 636–643, Washington, DC, USA, 2014. IEEE Computer Society.