A multispectral data model for higher-order active contours and its application to tree crown extraction*

Péter Horváth^{1,2}

 ¹ University of Szeged, Institute of Informatics, P.O. Box 652, H-6701 Szeged, Hungary, Fax:+36 62 546 397, hp@inf.u-szeged.hu,
² Ariana (joint research group CNRS/INRIA/UNSA), INRIA, B.P. 93, 06902 Sophia Antipolis, France, Fax:+33 4 92 38 76 43

Abstract. Forestry management makes great use of statistics concerning the individual trees making up a forest, but the acquisition of this information is expensive. Image processing can potentially both reduce this cost and improve the statistics. The key problem is the delineation of tree crowns in aerial images. The automatic solution of this problem requires considerable prior information to be built into the image and region models. Our previous work has focused on including shape information in the region model; in this paper we examine the image model. The aerial images involved have three bands. We study the statistics of these bands, and construct both multispectral and single band image models. We combine these with a higher-order active contour model of a 'gas of circles' in order to include prior shape information about the region occupied by the tree crowns in the image domain. We compare the results produced by these models on real aerial images and conclude that multiple bands improves the quality of the segmentation. The model has many other potential applications, *e.g.* to nanotechnology, microbiology, physics, and medical imaging.

1 Introduction

Successful forestry management depends on knowledge of a number of statistics connected to forest structure. Among these are the number and density of trees in a forest, their average size, and changes in these quantities over time. High-resolution remote sensing images, and in particular colour infrared (CIR) aerial images, can facilitate the acquisition of these statistics by providing images from which tree crowns can be identified and counted, and their areas and shapes analysed. The task of manually extracting this information from aerial images, or worse, measuring the statistics in the field, is, however, labour intensive, which limits the extent to which it can be used. Typically, information can be extracted at tree stand resolution but not below, simply due to the time and cost involved. Image processing methods capable of extracting the same information automatically would therefore be of great use.

^{*} This work was partially supported by EU project MUSCLE (FP6-507752), Egide PAI Balaton, OTKA T-046805, and a HAS Janos Bolyai Research Fellowship. We thank the French National Forest Inventory (IFN) for the data.

The problem of inferring automatically the region in the image domain corresponding to tree crowns given the image data is, however, not simple to solve. Like all inference problems, it can be phrased probabilistically. The quantity of interest in P(R|I, K), the probability that region R in the image domain corresponds to tree crowns given the image data I and any prior knowledge K we may choose to include. This is proportional to P(I|R, K)P(R|K), and thus we must construct models of the image to be expected given knowledge of where the tree crowns are, and of the possible regions corresponding to tree crowns in the absence of the data. We discuss the latter first.

It might be thought that the region model P(R|K) could be rather generic, but this turns out not to be the case. Trees are not always easily distinguished from the background using the data alone, and so a prior model of R that incorporates enough knowledge to disambiguate these situations is required. Fortunately, we have a great deal of prior knowledge about the type of region to be expected. In this paper, we will focus on plantations, that is, collections of trees that do not often overlap and that are of the same species and roughly the same age. In this case, which is of great importance in practice, the region corresponding to tree crowns will consist of an unknown number of connected components corresponding to different trees, each connected component being a circular shape with a certain radius. Horváth *et al.* [1,2] addressed the extraction of tree crowns from CIR images. They constructed a model of such regions, called the 'gas of circles' model, using the higher-order active contour framework proposed in [3]. In this paper we use the same prior model.

Horváth *et al.* [1,2] also described a data model. This model described the behaviour of only one band of the three available bands in the CIR images. The model was Gaussian, with the values at different pixels independent, and with different means and variances for tree crowns and the background. While successful, this model, even with the strong region prior, was not capable of extracting accurately the borders of all trees. Some trees were simply too similar to the background. The purpose of this paper is to construct a new data model that makes use of all three bands in the CIR images. We study the improvement or otherwise of the extraction results produced by modelling the three bands as independent or as correlated. As we will see, even at the level of maximum likelihood, the inclusion of 'colour' information, and in particular, interband correlations, can improve the results, and in conjunction with the region prior, the full model is considerably better than that based on one band alone.

In the next subsection 1.1, we briefly review previous work on tree crown extraction and on region modelling. Then in section 2, we recall notions of higher-order active contours and describe the 'gas of circles' model, with the emphasis on a method to fix all but one of its parameters. In section 3, we study four possible 'colour' data models and compare them. In section 4, we describe the gradient descent algorithm used to minimize the full active contour energy. In section 5, we show experimental results on CIR images. In section 6, we sum up.

1.1 Previous work

The problem of delineating, locating or counting individual trees in high resolution aerial images has been studied in several papers. Different approaches have been proposed based on template matching [4], collections of rules [5], contours [6,7], math-

ematical morphology [8] and stochastic geometry [9]. Although the input to many of these approaches consists of multispectral images, usually only the infrared band is used for tree crown extraction. One way to use the multispectral information is through spectral signatures of various types, thoroughly reviewed and compared in [10]. For example, 'tree colour lines' approximate the cigar-shaped distribution of tree crown pixels in 'colour space' by a line. In this paper we study probabilistic models, based on multispectral histograms and Gaussian distributions.

Previous work on region modelling, for example [11–14], is in general not suitable for the tree crown extraction problem. This is because it focuses on small variations of a region around a small number (usually one) of template regions. This means that regions with high probability lie in a bounded subset of the space of regions close to the template(s). The regions corresponding to tree crowns have an unknown number of connected components, and hence cannot easily be described by such approaches.

'Higher-order active contours' (HOACs) [3] provide a complementary approach. In active contour models, higher-order or not, a region R is represented by its boundary ∂R . P(R|I, K) is constructed implicitly, via an energy functional $E(\partial R) = E_g(\partial R) +$ $E_i(\partial R, I)$, where E_g and E_i correspond to prior and likelihood. In classical active contours[15], $E_{\rm g}$ is constructed from single integrals over the boundary or region. Euclidean invariance then requires that E_{g} be a linear combination of the length of the boundary and the area of the region, since these are the only Euclidean invariant energies that can be constructed using single integrals over the boundary if the curvature is not used. Thus these energies incorporate only local differential-geometric information. HOACs generalize these classical energies to include multiple integrals over ∂R . Thus HOAC energies explicitly include long-range interactions between tuples of boundary points without the use of a template, which allows the inclusion of sophisticated prior knowledge while permitting the region to have an arbitrary number of connected components, which furthermore may interact amongst themselves. Euclidean invariance is intrinsic, with no pose estimation required. The approach is very general: classical energies are linear functionals on the space of regions; HOAC energies include all polynomial functionals. Rochery *et al.* [3] applied HOACs to road network extraction, while Horváth et al. [1, 2] extended the model to describe a 'gas of circles', and applied it to tree crown extraction. We describe this model in the next section.

2 The 'gas of circles' model

HOAC energies generalize classical active contour energies by including multiple integrals over the contour. The simplest such generalizations are quadratic energies, which contains double integrals. There are several forms that such multiple integrals can take, depending on whether or not they take into account contour direction at the interacting points. The Euclidean invariant version of one of these forms is [3]

$$E_{g}(\gamma) = \lambda L(\gamma) + \alpha A(\gamma) - \frac{\beta}{2} \iint dp \, dp' \, \mathbf{t}(p) \cdot \mathbf{t}(p') \, \Psi(r(p, p')) \,, \tag{1}$$

where γ is the contour, a representative map in the equivalence class of embeddings representing ∂R , and thereby R; p and p' are parameters for γ ; L is the length of ∂R ; A

	ょ	Ĭ	*	I	.\	I	• •	I	••	•
*	そ	×	、	::	•	••	:•	••	:•	••

Fig. 1. Gradient descent evolution using E_g alone, from initial (left) to final, stable (right) state $(r_0 = 5, \alpha = 2, \beta = 1.69 \text{ and } d = 5)$.

is the area of R; $r(p, p') = |\gamma(p) - \gamma(p')|$; $\mathbf{t} = \dot{\gamma}$ is the (unnormalized) tangent vector to the contour; and Ψ is an interaction function that determines the geometric content of the model. With an appropriate choice of interaction function Ψ , the quadratic term creates a repulsion between antiparallel tangent vectors. This has two effects. First, for particular ranges of α , β , and d ($\lambda = 1$ wlog), circular structures, with a radius r_0 dependent on the parameter values, are stable to perturbations of their boundary. Second, such circles repel one another if they approach closer than 2d. Regions consisting of collections of circles of radius r_0 separated by distances greater than 2d are thus local energy minima. The model with parameters in the appropriate ranges is called the 'gas of circles' model [1].

Via a stability analysis, Horváth *et al.* [1] found the ranges of parameter values rendering circles of a given radius stable as functions of the desired radius. Stability, however, created its own problems, as circles sometimes formed in places where there was no supportive data. To overcome this problem, in [2], the criterion that circles of a given radius be local energy minima was replaced by the criterion that they be points of inflexion. As well as curing the problem of 'phantom' circles, this revised criterion allowed the fixing of the parameters α , β , and d as functions of the desired circle radius, leaving only the overall strength of the prior term, λ , unknown. For energy-based models, parameter adjustment is a problem, so this is a welcome advance.

To illustrate the behaviour of the prior model, figure 1 shows the result of gradient descent starting from the region on the left. Note that there is no data term. The parameter values in these experiments render the circles involved stable. With the parameter values calculated in [2], they would disappear.

3 Aerial images and image models

The previous section described the prior model of regions that we will use. In this section, we examine the data, and study data models that use all three bands of the CIR image data. The wavelength of the three bands are between 520nm and 900nm approximately, with shifted colour bands used to add false colour to the images. Notice that the blue band in the images corresponds to green in reality, green to red, and red to photographic or very near infrared (700–900nm).

Figure 2(a) shows a typical CIR aerial image falsely coloured for display purposes. It is of a poplar stand. Figure 2(b) shows the infrared band of the image. To see how colour can help, note that the bright pixels in the spaces between the trees are light grey



Fig. 2. (a): typical CIR aerial image of a poplar plantation; (b): greyscale version of the image; (c): ground truth used for statistics; (d): another CIR image; (e): corresponding ground truth. Images ©French National Forest Inventory (IFN).

in the colour image, while the trees are red. In the greyscale image, they have roughly the same intensity, making separation of trees and background difficult. Although the prior model helps to disambiguate these situations, it is not always successful, and it makes sense to consider a data model that uses the available information to the full.

We want to construct a data model for the observed CIR image, given that region R corresponds to tree crowns. We can divide the image I (a three-component quantity) into two pieces, I_R and $I_{\bar{R}}$ corresponding to the tree crowns and the background. Then we have that $P(I|R, K) = P(I_R, I_{\bar{R}}|R, K)$. Without further information, I_R and $I_{\bar{R}}$ are not independent given R: illumination for example will link them together. However, we may introduce parameters for the two regions, θ_R and $\theta_{\bar{R}}$, so that the two pieces of the image become independent given these parameters. We note that the size of the tree crowns (~10 pixels), coupled with the resolution of the image, does not allow the definition of meaningful texture features. Thus we will assume, without real justification, that the image values at different pixels are independent. Refinements to this assumption, for example, tree crown profiles, will be considered in future work. The data model then takes the form

$$\mathbf{P}(I|R,\theta_R,\theta_{\bar{R}},K) = \prod_{x \in R} \mathbf{P}(I_R(x)|\theta_R,K) \prod_{x \in \bar{R}} \mathbf{P}(I_{\bar{R}}(x)|\theta_{\bar{R}},K) \ .$$

To help us design the model for individual pixels, we examine the statistics of the pixel values in the different bands. Figure 3 shows histograms of the pixel values in figure 2(a) for all three bands, separated into tree crown and background based on a manual labelling shown on figure 2(c). As expected, the infrared band shows the largest separation. Can adding the other two bands help?

To test this idea, we performed four different types of maximum likelihood classification, based on four different estimates of the probability distributions for individual pixels of each class. Two of these estimates use raw histograms with different bin sizes. Of these, one is constructed as a product of the individual histograms for each band (independent bands), called HI for short, while the other uses the colour histogram (HC). The other two estimates use Gaussian models, either with covariances diagonal in colour space (independent bands), called GI, or with full covariances (G3D). The models parameters were learned from figure 2(a) and figure 2(d) based on the manual labelling. The resulting models were then used to classify the image in figure 2(a).



Fig. 3. Histograms of pixel values in the three bands of figure 2(a), based on the manual labelling shown in figure 2(c). Green is background; blue is tree crown.

The results of maximum likelihood classification on the same image are shown in figure 4. The images have four different colours: black and dark grey correspond to correct classification of background and tree crowns respectively, while light grey and white correspond to incorrect classifications in which tree crowns were classified as background and vice-versa respectively. Table 1 left shows the resulting classification error rates.



Fig. 4. Maximum likelihood classifications of figure 2(a) using the different models trained on the same image.

Naturally, the results using HC are almost perfect. The number of bins is very large, and this means that there are unlikely to be more than one or two pixels in each bin. Consequently, any given pixel is very likely to have zero probability to be in the incorrect class. Equally clearly, the results using HI are poor: the different bands are not independent. This is confirmed by the result for GI. G3D, however, produces a reasonable performance, second only to the HC results. Bearing in mind that G3D has 3 + 6 = 9 parameters, while HC has the same number of parameters as bins, this is encouraging.



Fig. 5. The same classification trained on figure 2(d).

These conclusions are confirmed by the label images, which clearly show the inferior classifications produced by the models with independent bands.

To test the generalization ability of the models, we used a different image to learn the model parameters, and used them to classify figure 2(a). The new training image is figure 2(d), along with a manual labelling. Figure 5 shows the results, while table 1 right shows the error rates.

Method	$B \rightarrow F$	$F \rightarrow B$	error (%)	Method	$B \rightarrow F$	$F \rightarrow B$	error (%)
HI (64 bins)	446	404	9.64	HI (64 bins)	748	242	11.22
HI (128 bins)	446	399	9.58	HI (128 bins)	752	253	11.39
HC (64^3 bins)	121	214	3.8	HC (64^3 bins)	1028	490	17.21
HC (128^3 bins)	19	97	1.32	HC (128 ³ bins)	1747	1277	34.52
GI	470	426	10.16	GI	1106	123	13.93
G3D	256	328	6.62	G3D	841	85	10.5

Table 1. Error rates for the maximum likelihood classification of figure 2(a), using models trained on the same image (left) and on figure 2(d) (right).

It is not a surprise that the error rates are larger. The histogram-based methods do not generalize well, and produce more errors than both Gaussian models. The Gaussian results are naturally not as good as in the previous test, but are adequate in the absence of a prior energy. The model with dependent bands performs considerably better than the independent band model in both cases. In particular, the independent band models, whether histogram-based or Gaussian, consistently confuse certain types of inter-tree background with the tree crown foreground.

4 Data model and energy minimization

Our full energy functional for tree crown extraction is a combination of the energy associated to the likelihood, $E_i(\gamma, I) = -\ln P(I|R, \theta_R, \theta_{\bar{R}}, K)$, and the HOAC 'gas of circles' prior geometric term E_g given in equation (1): $E(\gamma, I) = E_g(\gamma) + E_i(\gamma, I)$. In the last section, we established that the Gaussian model with full covariance provides the best compromise between precision and generalization. In this section, we describe this data term and how we minimize the full energy E.

The parameters of E_i are learnt from samples of each class using maximum likelihood, and then fixed. We denote the mean vectors inside and outside as M_{in} and M_{out} and the covariance matrices Σ_{in} and Σ_{out} . We define the energy as we wrote above

$$E(\gamma) = E_{\rm g}(\gamma) - \int_{R} dp \, \ln \left[\det^{-1/2} (\Sigma_{\rm in}/2\pi) e^{-\frac{1}{2} (I(p) - M_{\rm in})^{T} \Sigma_{\rm in}^{-1} (I(p) - M_{\rm in})} \right] - \int_{\bar{R}} dp \, \ln \left[\det^{-1/2} (\Sigma_{\rm out}/2\pi) e^{-\frac{1}{2} (I(p) - M_{\rm out})^{T} \Sigma_{\rm out}^{-1} (I(p) - M_{\rm out})} \right].$$

The energy is minimized using gradient descent. The descent equation is

$$\hat{\mathbf{n}} \cdot \frac{\partial \gamma}{\partial t}(p) = -\lambda \kappa(p) - \left[\alpha + \frac{1}{2} \ln \left(\frac{\det(\Sigma_{\text{in}})}{\det(\Sigma_{\text{out}})} \right) \right] + \beta \int dp' \, \hat{\mathbf{r}}(p, p') \cdot \mathbf{n}(p') \, \dot{\Psi}(r(p, p')) \\ - \frac{1}{2} \left\{ (I(p) - M_{\text{in}})^T \, \Sigma_{\text{in}}^{-1}(I(p) - M_{\text{in}}) - (I(p) - M_{\text{out}})^T \, \Sigma_{\text{out}}^{-1}(I(p) - M_{\text{out}}) \right\},$$

where κ is the curvature of the contour, a dot indicates derivative, $\mathbf{r}(p, p') = \gamma(p) - \gamma(p')$, and $\hat{\mathbf{r}} = \mathbf{r}/r$. To evolve the contour we use the level set framework [16] extended to cope with the nonlocal forces arising from higher-order active contours [3].

5 Experimental results

We tested the new approach on CIR aerial images of poplar stands located in the Saône et Loire region in France, provided by the French National Forest Inventory (IFN). We compare three models: the new model, which uses the multispectral data term with the 'gas of circles' prior; the model in [2], which uses only the infrared band of the CIR image with the 'gas of circles' prior; and a classical active contour model, which uses the multispectral data model, but only the length and area terms of E_g , *i.e.* $\beta = 0$. There is thus no prior shape information in this third model. In all experiments, the contour was initialized to a rounded rectangle slightly larger than the image domain.

Figure 6(a), (b), and (c) show the results obtained on the image shown in figure 2(a), using the new model, the model in [2], and the classical active contour model respectively. The new model is the most successful, separating trees that are not separated by the other models.

Figure 6(d), (e) and (f) show the results obtained on the image shown in figure 2(d). None of the results is perfect, all the models failing to separate some trees, but the new model detects several trees that are not detected by the model in [2]. The classical active



Fig. 6. Top row: results obtained on the image shown in figure 2(a), using the new model (a), the single band model of [2] (b), and the classical active contour model combined with the multi-spectral data term (c). Bottom row: results obtained on the image shown in figure 2(d), using the new model (d), the single band model of [2] (e), and the classical active contour model (f). In (a), (b), (d), and (e), the stable radius was set to $r_0 = 2.5$. ©French National Forest Inventory (IFN)

contour model was not be able to separate all the crowns, and found a large connected area at the bottom right, due to the missing prior shape information.

Figure 7(a) shows a difficult image with a field at the top, and strong shadowing. The result with the new model, shown in figure 7(b), is good, detecting all the trees and ignoring the field and shadows. The result with the model of [2], shown in figure 7(c), is not as good. Some trees are missed, but more importantly, the fact that the field has a similar IR response to the tree crowns means that a large incorrect region is produced. The result with the classical active contour model, shown in figure 7(d), avoids this error thanks to the multispectral information, but the lack of prior shape information means that some trees are merged.

Figure 7(e) shows a different type of image, of isolated trees in fields. The result with the new model, shown in figure 7(f), is correct, ignoring the field, for example. The result with the model of [2] is not as good, with one large false positive, and smaller errors on each of the detected trees, due to confusion between the field and parts of the road and the tree crowns(figure 7(g)). Figure 7(h) shows the result obtained using the multispectral data term combined a classical active contour model. The result is almost as good as the new model, except that the contours are slightly less smooth, and there is a small false positive area in the upper right corner, which was not detected by the new model, presumably because it is not circular.

Figure 8(a) shows another CIR image with fields and some sparse trees. It is a difficult image, because some of the fields have a similar colour to the trees. The result with the new model, shown in figure 8(b), is good, detecting all the trees, and only merging two of them. The result with the model of [2], shown in figure 8(c), is not as



Fig. 7. Top row: (a), a CIR image; (b), result with the new model; (c), result with the model of [2] (stable radius $r_0 = 2.5$); (d), result with classical active contour model combined with the multispectral data term. Bottom row: (e), a CIR image; (f), result with new model; (g), result with model of [2] (stable radius $r_0 = 7.0$); (h), result with classical active contour model ($\beta = 0$) combined with the multispectral data term. ©French National Forest Inventory (IFN)

good. The greyscale level between some of the trees is too similar to the tree crowns to be separated, despite the prior shape information, meaning that several trees are merged. In addition, some non-tree objects were detected as tree crowns, again due to similarity of grey scale. The result obtained with the classical active contour and multispectral data model is slightly better, but due to the missing prior shape information several tree crowns are merged and a small non-tree area was detected.

6 Conclusion

We have presented a new higher-order active contour (HOAC) model for tree crown extraction from colour infrared (CIR) aerial images. The new data term takes into account the multispectral nature of the data, in contrast to almost all previous work. The interband correlations are modelled using a full covariance Gaussian distribution. The prior term is a HOAC model of a 'gas of circles', modelling regions consisting of a number of circles of approximately a given radius. Experimental results show that the new model outperforms both a model with the same prior shape information, but which uses only the IR band of the data (the model of [1, 2]), and models with the same multispectral data term, but including less prior shape information, to wit, a classical active contour model and maximum likelihood.



Fig. 8. (a), a CIR image; (b), result with the new model; (c), result with the model of [2] (stable radius $r_0 = 4.0$); (d), result with classical active contour model ($\beta = 0$) combined with the multispectral data term. ©French National Forest Inventory (IFN)

References

- Horváth, P., Jermyn, I.H., Kato, Z., Zerubia, J.: A higher-order active contour model for tree detection. In: Proc. International Conference on Pattern Recognition (ICPR), Hong Kong, China (2006)
- Horváth, P., Jermyn, I.H., Kato, Z., Zerubia, J.: An improved 'gas of circles' higher-order active contour model and its application to tree crown extraction. In: Proc. Indian Conference on Vision, Graphics and Image Processing (ICVGIP), Madurai, India (2006)
- Rochery, M., Jermyn, I.H., Zerubia, J.: Higher-order active contours. International Journal of Computer Vision 69 (2006) 27–42
- Larsen, M.: Finding an optimal match window for Spruce top detection based on an optical tree model. In Hill, D., Leckie, D., eds.: Proc. International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, British Columbia, Canada (1998) 55–66
- Gougeon, F.: A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. Canadian Journal of Remote Sensing, 21(3) (1995) 274–284
- Brandtberg, T., Walter, F.: Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. Machine Vision and Applications (1998) 64–73
- Gougeon, F.A.: Automatic individual tree crown delineation using a valley-following algorithm and rule-based system. In Hill, D., Leckie, D., eds.: Proc. International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, British Columbia, Canada (1998) 11–23
- Andersen, H., Reutebuch, S., Schreuder, G.: Automated individual tree measurement through morphological analysis of a LIDAR-based canopy surface model. In: Proc. International Precision Forestry Symposium, Seattle, WA, USA (2001) 11–21
- 9. Perrin, G., Descombes, X., Zerubia, J.: Tree crown extraction using marked point processes. In: Proc. European Signal Processing Conference (EUSIPCO), Vienna, Austria (2004)

- Gougeon, F.: Comparison of possible multispectral classification schemes for tree crown individually delineated on high spatial resolution MEIS images. Canadian Journal of Remote Sensing, 21(1) (1995) 1–9
- Cremers, D., Kohlberger, T., Schnörr, C.: Shape statistics in kernel space for variational image segmentation. Pattern Recognition 36 (2003) 1929–1943
- Leventon, M., Grimson, W., Faugeras, O.: Statistical shape influence in geodesic active contours. In: Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Hilton Head Island, SC, USA (2000) 316–322
- Paragios, N., Rousson, M.: Shape priors for level set representations. In: Proc. European Conference on Computer Vision (ECCV), Copenhagen, Denmark (2002) 78–92
- Riklin-Raviv, T., Kiryati, N., Sochen, N.: Prior-based segmentation by projective registration and level sets. In: Proc. IEEE International Conference on Computer Vision (ICCV). (2005) 204–211
- Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1 (1988) 321–331
- 16. Osher, S., Sethian, J.A.: Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics **79** (1988) 12–49