Alexander Ihlow and Udo Seiffert

Haustoria segmentation in microscope colour images of barley cells

***URN*:** urn:nbn:de:gbv:ilm1-2014300021

Haustoria Segmentation in Microscope Colour Images of Barley Cells

Alexander Ihlow and Udo Seiffert

Leibniz Institute of Plant Genetics and
Crop Plant Research (IPK) Gatersleben
Corrensstr. 3, 06466 Gatersleben, Germany
Pattern Recognition Group
{ihlow, seiffert}@ipk-gatersleben.de
http://mue.bic-gh.de

Abstract. Dyed barley cells in microscope colour images of biological experiments are analysed for the occurrence of haustoria of the powdery mildew fungus by a fully automated screening system. The region of interest in the images is found by applying Canny's edge detector to the hue channel of the HSV colour space. For the segmentation of potential haustoria within the dyed cells, two different methods are considered: A clustering in RGB colour space using the Expectation Maximisation (EM) algorithm, and morphological contrast enhancement of the colour image with subsequent hysteresis thresholding in the saturation channel of the enhanced images. The second approach seems to be more viable because of its robustness and more promising results.

1 Introduction

Automating the screening and the analysis of biological experiments is a challenging research area in the field of bioinformatics and engineering. This paper is related to a project studying resistance mechanisms of crop plants against the powdery mildew fungus from the genetical point of view. In the experiments, young barley leaves are bombarded with DNA-coated particles to "switch on or off" desired genes in cells. For analysis purposes, an additional reporter gene[1] is expressed in cells that were hit by a particle. This dyes the affected genetically transformed cells greenish blue and allows their identification by bright field microscopy [7]. The task is to evaluate the susceptibility of the genetically transformed cells to the powdery mildew fungus under the impact of different test genes. A successful penetration of the fungus into the cell is indicated by the development of a haustorium – a dark object consisting of a "waist" with "fingers" that is located between the cell wall and the cell membrane and feeds the fungus by leaching the cell (see Figure 2). These objects have to be counted in an automatic analysis procedure.

Since there are many genes to be considered for a potential resistance of the plant against pathogens, a big number of experiments has to be performed to obtain a sufficient statistical confidence. Therefore, an automated image acquisition system and an

[1] β-glucuronidase (GUS) reporter gene

automatic analysis procedure is needed. Manual screening is a tedious, subjective and time-consuming task that cannot be handled by laboratory assistants due to that huge amount of data. For an automatic image acquisition, the microscope slides are mounted on an x-y table which scans a number of preparations fully automatically under the control of a computer, e.g., overnight. Now, finding genetically transformed cells and therein assessing the development status of the haustoria without human interaction is the task and the challenge of the analysis procedure.

This paper deals with the problem of segmenting the scarcely outstanding haustoria from the remaining cell tissue. It is organised as follows: Section 2 introduces the properties of the image material and explains how the regions of interest, i.e., genetically transformed cells, are found in the images. Afterwards, Section 3 reviews the Expectation Maximisation (EM) clustering technique and presents its results on two examples. Section 4 describes a segmentation technique based on thresholding in the saturation channel after morphological contrast enhancement of the colour image and shows its results on four examples, before Section 5 concludes the paper.

2 Preprocessing of the image material

Figure 2 shows two typical cutouts of microscope images, both containing one dyed genetically transformed cell with two haustoria of the powdery mildew fungus inside. By default, the microscope camera produces images of 2600×2060 pixel in 24-bit colour.

In [3] we have shown that these dyed cells can be reliably detected by applying Canny's edge detector [1] to the hue channel of the HSV colour space. It is advantageous to cut out the cells exactly at its edges first before processing the images by morphological contrast enhancement. This prevents potential distortions from regions outside the cell during operations with large neighborhood-masks. Since Canny's edge detector produces *either* smooth and closed edges at the expense of a significant spatial uncertainty *or* rather disrupted but certainly positioned edge elements at fine scales, an edge linking procedure across multiple scales is considered in tradeoff to get both a *closed* as well as *correctly located* cell boundary. For the EM approach, this step is omitted and the rectangular region of interest, containing a dyed cell, is processed directly.

3 Clustering in RGB using the EM algorithm

The Expectation Maximisation (EM) algorithm [2, 6] solves the general problem of classifying a number of N d-dimensional data vectors $\mathbf{x}_n \in \mathbb{R}^{d \times 1}$ from the entire data set $\mathbf{X} \in \mathbb{R}^{d \times N}$ into K classes. Therefore, a parametric model for the class distributions is chosen whose initial parameters are adapted in an iterative procedure. A popular, powerful and simple model is the multivariate Gaussian distribution, whose equal-probability surfaces describe (hyper)ellipsoids in the d-dimensional space. The center of an ellipsoid is given by the mean vector $\boldsymbol{\mu}$, whereas its shape and orientation

is described by the covariance matrix $\boldsymbol{\Sigma}$.

$$p(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{\sqrt{\det \boldsymbol{\Sigma}_k (2\pi)^d}} \mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}-\boldsymbol{\mu}_k)} \tag{1}$$

Using this model, the EM algorithm is run in the following manner: The probability (at iteration step t) of each data vector $\mathbf{x}_n$ to belong to class k is calculated (expectation step) by

$$P^t(k|\mathbf{x}_n) = \frac{P^t(k)\, p(\mathbf{x}_n|\boldsymbol{\mu}_k^t, \boldsymbol{\Sigma}_k^t)}{\sum_{j=1}^{K} P^t(j)\, p(\mathbf{x}_n|\boldsymbol{\mu}_j^t, \boldsymbol{\Sigma}_j^t)}. \tag{2}$$

A new parameter set for the iteration step $t+1$ containing the prior probabilities, mean vectors and covariance matrices for each class is calculated according to (maximisation step)

$$P^{t+1}(k) = \frac{1}{N} \sum_{n=1}^{N} P^t(k|\mathbf{x}_n) \tag{3}$$

$$\boldsymbol{\mu}_k^{t+1} = \frac{1}{N P^{t+1}(k)} \sum_{n=1}^{N} P^t(k|\mathbf{x}_n)\, \mathbf{x}_n \tag{4}$$

$$\boldsymbol{\Sigma}_k^{t+1} = \frac{1}{N P^{t+1}(k)} \sum_{n=1}^{N} P^t(k|\mathbf{x}_n)\, (\mathbf{x}_n - \boldsymbol{\mu}_k^{t+1})\, (\mathbf{x}_n - \boldsymbol{\mu}_k^{t+1})^T. \tag{5}$$

The algorithm is terminated at some stopping criterion, e.g., when the resulting labelling of the data vectors does not change anymore.

We examined this method in [4] and argued that a direct approach of this technique is not feasible even in case of a very good initialisation parameters. Therefore, a modified approach was introduced: Only data samples that were assigned *reliably* to a class are iterated on. This prevents the EM algorithm from deviating too much from its initial parameter set and segments potential haustoria regions quite well. Figure 3 shows the segmentation results using this constrained version of the EM algorithm. The pixel-labels are depicted in a soft-output manner, i.e., the vector of the posterior probabilities $[P(k=3|\mathbf{x}),\, P(k=2|\mathbf{x}),\, P(k=1|\mathbf{x})]^T$ is assigned to the RGB value of each pixel, making the saturation of the colour follow the reliability of the estimate.

Nevertheless, we have been looking for less fragile methods providing better segmentation results, since there are some interferences in the outputs so far.

4 Morphological contrast enhancement

In addition to other applications like texture description, gradient calculation, etc., the powerful techniques of mathematical morphology [8] can be used to gain a significant contrast enhancement of images. This is basically done by *adding enhanced bright features* to the original image I while *subtracting enhanced dark features*

$$\kappa(I) = I + \underbrace{WTH_B(I)}_{\substack{\text{enhanced bright} \\ \text{features}}} - \underbrace{BTH_B(I)}_{\substack{\text{enhanced dark} \\ \text{features}}}. \tag{6}$$

Both, the enhanced bright and dark features are extracted by the morphological top-hat operations

$$\text{white top-hat (opening top-hat)} \qquad WTH_B(I) = I - \gamma_B(I) \tag{7}$$

$$\text{black top-hat (closing top-hat)} \qquad BTH_B(I) = \phi_B(I) - I\,, \tag{8}$$

respectively, using the morphological opening γ_B and closing ϕ_B. In turn, the opening $\gamma_B(I) = \delta_B[\epsilon_B(I)]$ and closing $\phi_B(I) = \epsilon_B[\delta_B(I)]$ are based on the fundamental morphological operations dilation δ_B and erosion ϵ_B, using the structuring element B. Its shape and size has to be chosen according to the available knowledge about the structures to be filtered. Advantageously, morphological top-hats operate contrary – they extract structures that *cannot* contain the structuring element. Therefore, a simple rectangular or quadratic structuring element, being a little bit larger than the "waist"of a haustorium, is a good choice to enhance the haustoria as the wanted structures.

The theory of mathematical morphology is well established for grayscale images, though the question arises how to deal with colour images. Developing a theory of mathematical morphology for colour images, the step from a scalar-valued function to a vector-valued function needs the definition of a supremum and infimum for vectors instead of scalars, which in turn requires the definition of an order relationship of multivariate samples, making the development of a general theory on "colour-morphology" difficult [5]. Nevertheless, as a straightforward extension to grayscale morphology, a component-wise approach could be deployed which processes each vector component independently. This approach is chosen in the sequel, where the RGB components are treated as independent grayscale images.

The considerable effect of this contrast enhancement method is depicted in Figure 4. In these examples, a (flat) quadratic structuring element of size 31×31 pixel was used. While the hue is almost unaffected (this is not shown here), the effect is most expressed in the saturation channel (see the miscoloured images). The saturation of the greenish blue cell tissue stays almost unchanged whereas the saturation of the haustoria is boosted in such a manner that it absolutely reaches its maximum value in some regions of each haustorium.

4.1 Thresholding

This effect is exploited by a hysteresis thresholding technique: Since each haustorium reaches maximum saturation somewhere, this maximum value defines an upper threshold, while a lower threshold as a second parameter needs to be tuned on the data. Everything above the lower threshold is considered to be a haustorium if it is *connected* to a region segmented by the upper threshold.

Of course, the segmentation results strongly depend on the selection of the lower threshold. As we mentioned above and as can be clearly seen in the miscoloured images in Figure 4, the saturation of the cell tissue is almost not affected by the contrast enhancement – but the haustoria regions are. Now, the lower threshold is gained directly from the data: Consider the set of pixels belonging to maximum saturation level in the *enhanced* image. Using this set of pixels, we go back into the saturation channel of the *original* image and calculate the mean value of the pixels belonging to the

set. The result gives the lower threshold. Figure 1 depicts the corresponding probability density functions of the examples from Figure 4. Note their variability and the different resulting lower thresholds.

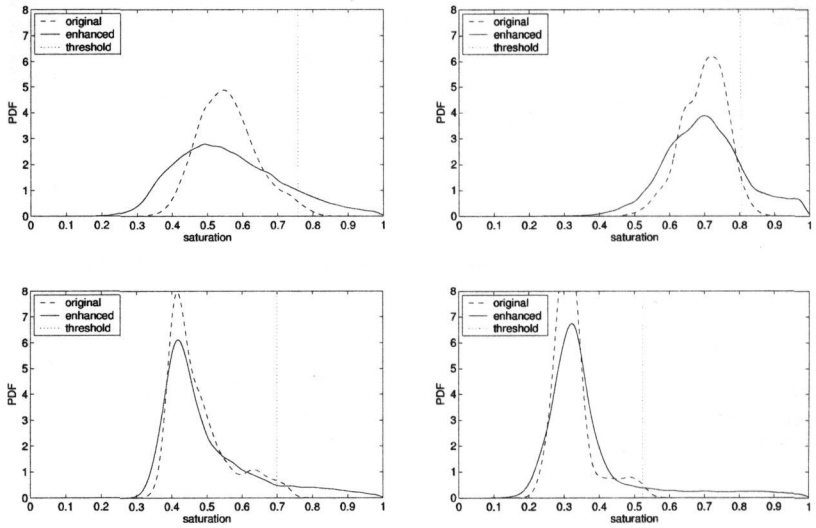

Fig. 1. Smoothed probability density functions of the saturation channel of the four exemplary cells of Figure 4. Maximally saturated pixels were excluded for convenience.
– dashed curve: PDF of saturation channel of original image
– solid curve: PDF of saturation channel of contrast enhanced image
– dotted vertical line: lower threshold used in hysteresis thresholding.

5 Conclusions

We applied two different segmentation techniques on microscope colour images of barley cells for the identification of small objects – so called haustoria – which stand out scarcely from the cell tissue. As the favorite solution, we propose to enhance the contrast of the cell image by morphological processing first and to perform a hysteresis thresholding in the saturation channel of the enhanced image afterwards. The upper and lower threshold is gained adaptively directly from the image data.

On the other hand, we presented in comparison a pixel-colour clustering technique based on the EM algorithm. Despite a good initialisation, this method needs a constraining mechanism to prevent a defection from the desired segmentation.

After all, the segmentation via morphological contrast enhancement provides rather good results; Haustoria are segmented properly. Now, a further shape analysis of the

segmented objects is needed to distinguish haustoria from other segmented objects. This is currently in progress.

Acknowledgements. We thank Patrick Schweizer and the laboratory assistants for their support concerning the biological background. Thanks also to Christian Schulze and Tobias Czauderna for fruitful discussions. This work was supported by the German Ministry of Education and Research (BMBF) under grant 0312706A.

References

1. John F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 8(6):679–698, November 1986.
2. Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.
3. Alexander Ihlow and Udo Seiffert. Microscope color image segmentation for resistance analysis of barley cells against powdery mildew. In *9. Workshop "Farbbildverarbeitung"*, ZBS Zentrum für Bild- und Signalverarbeitung e.V. Ilmenau, Report Nr. 3/2003, pages 59–66, Ostfildern-Nellingen, Germany, October 2003.
4. Alexander Ihlow and Udo Seiffert. Automating microscope colour image analysis using the Expectation Maximisation algorithm. In *DAGM'04: 26th Pattern Recognition Symposium*, pages 536–543, Tübingen, 30th Aug – 1st Sept 2004, September 2004. Springer.
5. Mario Köppen, Christoph Nowack, and Gert Rösel. Pareto-morphology for color image processing. In *11th Scandinavian Conf. on Image Analysis (SCIA99)*, pages 195–202, Kangerlussuaq, Greenland, June 1999.
6. Richard A. Redner and Homer F. Walker. Mixture densities, maximum likelihood, and the EM algorithm. *SIAM Review*, 26:195–239, 1984.
7. Patrick Schweizer, Jana Pokorny, Olaf Abderhalden, and Robert Dudler. A transient assay system for the functional assessment of defense-related genes in wheat. *Molecular Plant-Microbe Interactions*, 12(8):647–654, 1999.
8. Pierre Soille. *Morphological Image Analysis*. Springer, 2002.

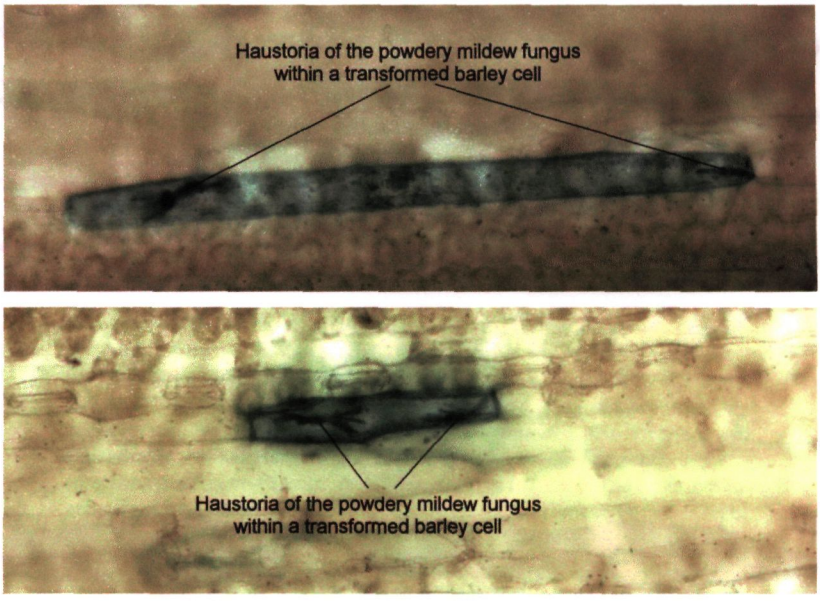

Fig. 2. Two cutouts of typical microscope colour images: Both transformed (greenish blue coloured) barley cells contain two haustoria of the powdery mildew fungus. These objects have to be detected automatically by a high throughput screening system.

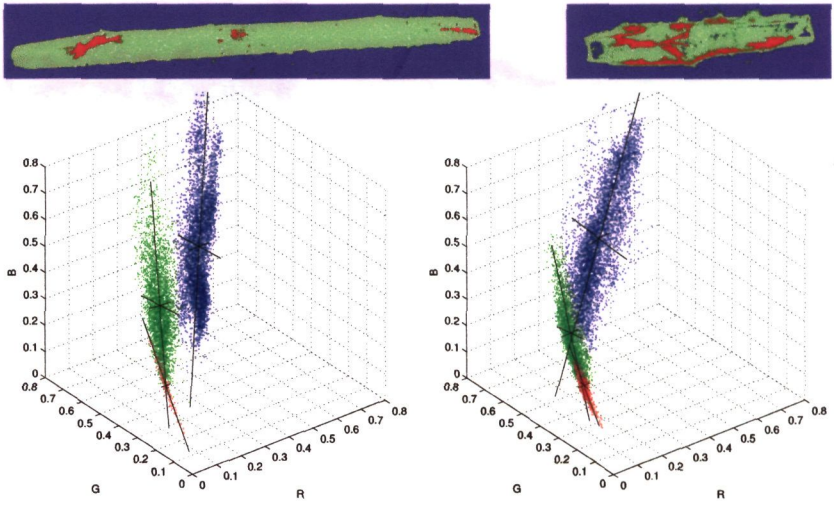

Fig. 3. Segmentation results of the two cells from Figure 2 via EM-clustering in the RGB colour space into *"background"* (blue), *"cell"* (green), and *"haustorium"* (red).

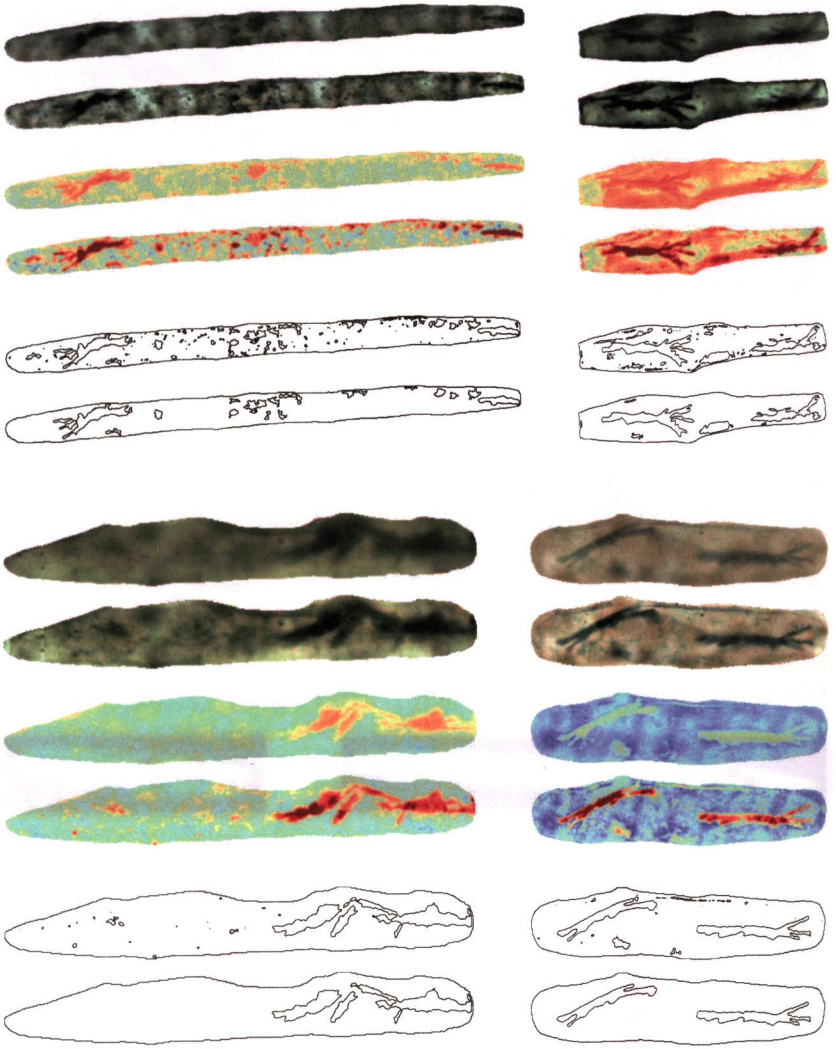

Fig. 4. Four examples for haustoria segmentation via morphological contrast enhancement of the colour image and subsequent thresholding in the saturation channel
(for each of the four examples from from top to bottom):
– cell image as acquired by the microscope camera
– cell image after contrast enhancement
– saturation channel of original cell image
– saturation channel of contrast enhanced image
– object boundaries (from sat. channel of enhanced image) by *simple* thresholding
– object boundaries (from sat. channel of enhanced image) by *hysteresis* thresholding.