Classifier-based Cell Segmentation from Confocal Microscopy Images

Monica Marcuzzo[1], Pedro Quelhas[1], Ana Campilho[3,4] Ana Maria Mendonça[1,2], Aurélio Campilho[1,2]
[1]INEB - Instituto de Engenharia Biomédica, Divisão de Sinal e Imagem, Campus FEUP
[2]Universidade do Porto, Faculdade de Engenharia
Departamento de Engenharia Electrotécnica e Computadores
[3]Department of Molecular Genetics, University of Utrecht

Abstract

In vivo observation and tracking of cell division in the Arabidopsis thaliana *root meristem, by time-lapse confocal microscopy, is central to biology research. This paper discusses an automatic cell segmentation method, which selects the best cell candidates from a starting watershed segmentation. The selection of individual cells is obtained using a Support Vector Machine (SVM) classifier, based on the shape and edge strength of the cells' contour. The result is an improved segmentation, which is largely pruned of badly segmented cells.*

1 Introduction

Cell division in plants is concentrated in specialized regions known as meristems [3]. In the *Arabidopsis*, the most important meristem is located at the tip of the root and perpetuates its pattern by cellular division. However, the mechanism by which cell division is controlled is not completely known. Development biologists studying roots find difficult to cope with the great amount of data, which requires the development of image analysis tools to automatically extract information, such as identifying cell division and growth.

The first step for automated cell division identification is cell segmentation in the root images. Segmentation is a difficult problem due to the image acquisition process, the data's variability and the noise. We present a first step to the segmentation by designing a method that selects well segmented cells from a set of possible segmented regions.

This paper is organized as follows: Section 2 describes the image data acquisition. Section 3 describes the proposed approach. Section 4 presents and discusses the obtained results. Finally, conclusion is presented in Section 5.

2 Image acquisition

The database used in this work was obtained using an automated confocal microscope image-acquisition process. Images are acquired every 10 or 15 minutes and the experiments last 10 hours to more than a day. Green Fluorescence Protein (GFP) is used to mark the cell walls. One problem with this type of image acquisition is the bleaching of the images caused by the degradation of the protein compound.

[4]Present address: Institute of Biotechnology, University of Helsinki

3 Methodology

Our system has three main stages:

- Pre-processing: image registration and filtering;
- Segmentation: watershed segmentation, contour extraction and description;
- Classification: Support Vector Machine (SVM) contour classification.

This system is novel in its structure and introduces novel parts such as collaborative filtering and SVM contour classification.

3.1 Pre-processing

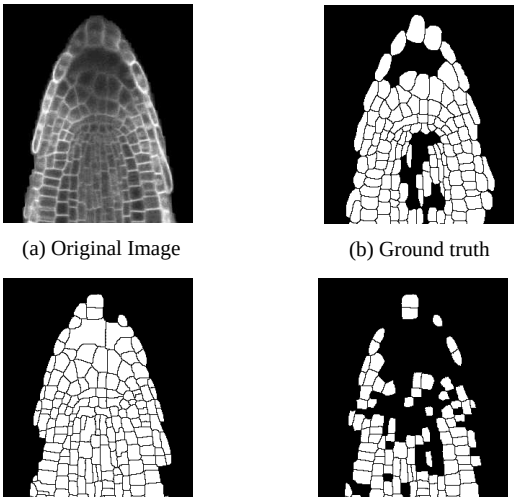
The direction of the root in the acquired images is not constant, due to the irregular growth of the root. To obtain a normalized image I_r from each input image I, we use an estimation of the central line as in the method described by Garcia *et al.* [5] to perform a rotation.

To improve the quality of the images, prior to segmentation we apply a denoising by sparse 3D transform-domain collaborative filtering (BM3D) [4] using an appropriate variance σ^2_{filt} for the estimated noise.

3.2 Watershed segmentation

In order to segment the cells, we apply a watershed transform to the filtered images. The resulting segmentation is the set of n regions R_i $(i = 1, ..., n)$ obtained from the watershed transform.

Usually the direct application of the watershed transform to an image leads to over-segmentation. In the *Arabidopsis* confocal images, it is difficult to establish a strategy to avoid over-segmentation which remains valid for a whole experiment. This is due to the variability of the root size in the image and to the bleaching effect. Our approach is to prune badly segmented regions after the segmentation step.



(a) Original Image (b) Ground truth

(c) Watershed segmentation (d) SVM pruned segmentation

Figure 1. Original image and resulting segmentation of the proposed method.

3.3 Classification

We obtain the contour c_i of each region R_i by the extraction of its points, starting from the leftmost one in a clockwise order. We describe each region's contour by its shape and underlying image pixel's edge strength. The shape is given by the Discrete Cosine Transform (DCT) of the distance between each contour point and the contour's centroid. The edge strength at the contour's pixels is characterized by the phase symmetry measure $PhSym(c_i)$ [6].

In order to prune the watershed segmentation, we classify each region R_i into cell C_j or non-cell, based on its vector descriptor, using a Gaussian kernel SVM [2]. The classifier training and testing is performed as follows:

Training: For each image, we applied the segmentation using different filtering sigmas σ_{filt}. For each image we labeled segmentation regions that correspond to cells and those which are clearly wrong (non-cell). We do not perform full annotation since some cases are ambiguous.

Testing: using the SVM model, given a new segmentation region's descriptor D_i, we can automatically classify that region as cell C_j or non-cell. Performing this operation for all regions, we obtain an SVM pruned segmentation image with all the regions classified as cells (Fig.1(d)).

4 Results and Discussion

We selected images from 16 experiments, from which 9 were used for training and 7 for test. In total, we used 68 images for training, containing 5125 manually selected cells. For test, 12 images were used, with 1421 manual segmented cells, used as ground truth for evaluation.

Applying the methodology described, we obtain an image with an SVM pruned cell segmentation (Fig.1(d)). Now we compare the SVM pruned result with the direct result of the watershed segmentation.

Table 1. Performance results of the SVM classification and the watershed segmentation.

method	watershed			SVM		
σ_{filt}	20	30	40	20	30	40
F-measure	0.817	0.821	0.825	0.831	0.836	0.837
FP	156.7	123.5	107.5	61.7	57.3	51.1
FN	25.8	24.7	26.9	64.6	52.9	51.8
performance	0.38	0.43	0.46	0.51	0.55	0.56

For evaluation of the segmentation results, we map each region classified as cell C_j to the best fitting ground truth region GT_k (Fig.1(b)). To obtain this mapping we use the F-measure F [1]. Cell region C_j is mapped to GT_k if the F-measure between them $F(C_j, GT_k)$ is above a threshold $th = 0.6$. If there is one and only one cell region C_j mapped to a ground truth GT_k, we consider that cell region as well classified. All measures were calculated for each image and averaged over all images. The results are presented in Table 1, where performance is the ratio of correctly classified cells regions according to the ground truth.

We can conclude that, using our approach, we are able to reduce false positives in at least 50% and create a segmentation which has at least 10% more correctly segmented regions. We obtain an increase of the F-measure with the SVM pruned segmentation (1.5% approx.), even if it does not modify the segmentation of the cells.

5 Conclusion

In this work we introduced an approach to automatically select the segmentation of cells in plant confocal microscopy using an SVM classifier. Using this approach we are able to prune most of the wrongly segmented cells improving the performance of the resulting segmentation.

ACKNOWLEDGEMENTS

The authors acknowledge the funding of Fundação para a Ciência e Tecnologia, under contract ERA-PG/0007/2006.

References

[1] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. ACM Press, 1999.

[2] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[3] A. Campilho, B. Garcia, H. Toorn, H. Wijk, A. Campilho, and B. Scheres. Time-lapse analysis of stem-cell divisions in the arabidopsis thaliana root meristem. *The Plant Journal*, 48:619–627, 2006.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3d transform-domain collaborative filtering. *IEEE Trans. on Image Processing*, 16(8):2080–2095, Aug. 2007.

[5] B. Garcia, A. Campilho, B. Scheres, and A. Campilho. Automatic tracking of arabidopsis thaliana root meristem in confocal microscopy. In *Proc. of the Int. Conf. on Image Analysis and Recognition, Springer LCNS 3212*, pages 166–174, 2004.

[6] P. Kovesi. Symmetry and asymmetry from local phase. In *Proc. of the 10th Australian Joint Conf. on Artificial Intelligence*, Dec. 1997.