

Tracking of *Arabidopsis thaliana* root cells in time-lapse microscopy

Monica Marcuzzo¹, Pedro Quelhas¹
¹INEB - Instituto de Engenharia Biomédica
Divisão de Sinal e Imagem, Campus FEUP

Ana Maria Mendonça^{1,2}, Aurélio Campilho^{1,2}
²Faculdade de Eng., Universidade do Porto
Dep. de Eng. Electrotécnica e Computadores

Abstract

In vivo observation of cells in the *Arabidopsis thaliana* root, by time-lapse confocal microscopy, is central to biology research. The research herein described is based on large amount of image data, which must be analyzed to determine the location and state of individual cells. Automating the process of cell tracking is an important step to create tools which will facilitate the analysis of cells' evolution through time. Here we introduce a confocal tracking system designed in two stages. At the image acquisition stage, we track the area under analysis based on point-to-point correspondences and motion estimation. After image acquisition, we compute cell-to-cell correspondences through time. The final result is a temporal structured information about each cell.

1. Introduction

Biologists study plant cells *in vivo* to attempt to understand the factors which control the cellular growth and division mechanisms. Many studies focus on the analysis of the *Arabidopsis thaliana* development. Cell growth is analyzed using different approaches, such as mathematical models [7] and motion estimation methods [12]. The relation between cell division and elongation in the regulation of organ growth rate is also investigated [1]. This research shows that *in vivo* imaging of the root is a valuable tool. However, development biologists studying roots find difficult to cope with the lack of suitable technology to analyze root meristem growth *in vivo* [2]. The great amount of data produced requires the development of image analysis tools to automatically extract useful information, such as identifying cell division and growth. Furthermore, cell features should be extracted to be processed statistically, and an easy way to track individual cells should be pursued.

In some studies it is possible to use the root's tip shape to base the tracking of the time-lapse imaging system [4]. However, this is limiting since it requires

the whole root tip to be in frame, inhibiting the studies along time of a particular cell region.

We present an approach to both track the root area under study and track individual cells through time. Tracking of the root area is designed to be as generic as possible, so that it can supply accurate results under different imaging conditions, like protein markers, zoom and time frame. Cells are detected by image segmentation, which is prone to errors due to noisy images. Individual cells' identity is tracked based on shape. We improve cell detection by imposing a temporal cell stability criterion, based on biology assumptions. The tracking process is fully unsupervised and can be adapted to other types of image sequences.

This paper is organized as follows: Section 2 describes the *in vivo* image data acquisition. Section 3 explains the process of individual cell tracking. In section 4 we present and discuss results. Finally, section 5 concludes this paper.

2. Real time root tracking

The database under analysis was obtained using a confocal microscope with a motorized stage, which is controlled to compensate the relative translation introduced by the root's growth. The time duration of the experiments ranged from 10 hours to more than a day, with images being acquired every 10/15 minutes. To visualize the cell walls, Green Fluorescence Protein (GFP) markers were used. When excited with a laser wavelength of its excitation spectrum, GFP emits light in the emission spectrum.

If no compensation is performed, the area under study will grow out of the field of view. Depending of the zoom this can happen in as little as 15 minutes. In Figure 1, we can see an example of images acquired without motion compensation, in this case the area under study leaves the frame in less than two hours. Higher zoom causes the area under study to leave the frame even faster. This makes it necessary to obtain automated tracking to relieve the biology researcher from

spending up to a day in front of the microscope performing adjustments every 10 minutes.

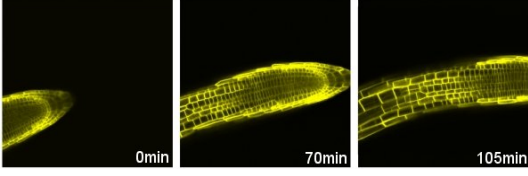


Figure 1. Uncontrolled time-lapse in vivo imaging.

In order to automatically track cells, we first follow the area under study over time performing motion compensation. This process is now presented.

2.1. Pre-processing

Images acquired by confocal-microscopy I_{raw}^t have a high level of noise. To improve the quality of the images, prior to processing, we filter the noisy image, resulting the image I_{filt}^t ,

$$I_{filt}^t = \mathcal{F}(I_{raw}^t, \sigma_{filt}^2), \quad (1)$$

where \mathcal{F} is the noise reducing filter, σ_{filt}^2 is the assumed noise variance and f is the frame number. Larger σ_{filt}^2 values will result in a smoother image. Based on the comparison between several filtering techniques [9] we choose to use denoising by sparse 3D transform-domain collaborative filtering (BM3D) [3].

2.2. Motion estimation/prediction

Assuming that we have both the previous I_{filt}^{t-1} and current I_{filt}^t filtered images. We want to measure the relative movement between images to predict and compensate root growth by moving the microscope stage. Motion of the area under study is defined as the difference between its center coordinate \mathbf{x} in consecutive frames.

Motion between frames is limited to rotation and translation. Equation 2 defines the coordinates transform for this system [6].

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} * \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} tx \\ ty \end{bmatrix} \quad (2)$$

where (x, y) are the coordinates for frame t and (x', y') are the projected coordinates for frame $t + 1$.

The estimation/prediction of motion is performed as an iterative process:

Estimation Relative image motion is estimated by obtaining correspondences between images I_{filt}^{t-1} and I_{filt}^t based on pixel information. For each image we obtain a set of corner points:

$$\mathcal{L}(I) \mapsto P^t = \{p_i, i = 1, \dots, n\}, \quad (3)$$

where \mathcal{L} is the Harris corner detector [5], P^t is the set of resulting n corner points $p_i = (x_i, y_i)$. Coordinates are represented in the microscope stage coordinate system.

Based on cross-correlation distance we obtain putative correspondences between point sets from both images (P^{t-1} and P^t) [6]:

$$C^{t-1}(p_i, p_j) = \begin{cases} 1 & D_{cc}(p_i, p_j) < D_{cc}(p_i, p_{j'}) \forall j' \neq j \\ 0 & otherwise \end{cases} \quad (4)$$

where D_{cc} is the cross-correlation distance [6], i and j are indexes for point sets P^{t-1} and P^t respectively. Cross-correlation is performed over a 9×9 area, centered on each corner point. Search for possible putative matches is limited to a 50 pixel window around the predicted point's position.

Using a RANSAC estimator [6], based on C^{t-1} , we estimate the motion parameters (tx, ty, θ) . The RANSAC algorithm is based on an exact solution to equation 2, using random sample correspondences from C^{t-1} . This results in a tentative solution for the motion parameters: $\tilde{tx}, \tilde{ty}, \tilde{\theta}$. Using this solution, we project points in the previous image P^{t-1} in the coordinates of the new image (P^t). We then compute the distance between the projected points P^t and the corresponding points P^t , according to C^{t-1} . If the coordinate distance between projected and putative matched points is within a threshold ($th_{ransac} = 15$), we classify the correspondence as an inlier of the tentative solution. This process is repeated and the solution with the largest number of inliers is kept as the estimate for the motion parameters: $\hat{tx}, \hat{ty}, \hat{\theta}$. The estimate for the center of the area under study \mathbf{x}^t can now be calculated.

Prediction We use a FIR 2nd order filter \mathcal{F} to estimate the next position of the area under study:

$$\hat{\mathbf{x}}^{t+1} = \mathcal{F}(\mathbf{x}^t, \mathbf{x}^{t-1}, \mathbf{x}^{t-2}), \quad (5)$$

where t is the current frame.

The filter coefficients are adapted using a Kalman filter [11]. The Kalman adaptive process is based on the previous positions of the area under study.

Using the knowledge from $\hat{\mathbf{x}}^{t+1}$ and \mathbf{x}^t we can predict tx and ty , i.e., the parameters used to move the microscope stage before the next image acquisition and parameter estimation step.

For the initial 3 frames no prediction is available, and the stage is moved after the image acquisition, using the motion measure (\hat{t}_x, \hat{t}_y) .

After the end of the time-lapse experiment, having acquired all images, we perform a final stabilization of rotation between images. Due to some instability in the estimation of the rotation parameter θ , we perform a 1D Gaussian filtering step on θ before stabilization. The final images I_{st}^t , are assumed to represent the same area (the area under study).

3. Cell tracking

Assuming that we have the filtered and stabilized images I_{st}^t , we now want to follow each cell within each image in other images of the same time-lapse experiment. In this section we explain how to obtain cell segmentation, cell-to-cell correspondence and how to correct some inaccuracies that may exist in cell tracking/segmentation.

3.1. Segmentation

In order to segment image I_{st}^t into a set of possible cells, we apply a watershed transform [8]. The resulting segmentation is the set of regions $R_j^t = \{r_j^t, \dots, j = (1, \dots, n)\}$ obtained from the watershed transform. These regions are possible segmented cells, however may also be the result of noise or bad initialization of the watershed process.

The direct application of the watershed transform to an image usually leads to over-segmentation. This happens because there are more image minima than objects, that is, not all minima really represent true objects, due to noise [10]. We have the advantage of knowing that cells evolve in a stable way, changing their shape slowly. This will allow us to use temporal information to prune some badly segmented regions.

3.2. Cell correspondence

We want to obtain the correspondence of each region r_j^{t-1} to the best fitting region r_j^t , for all $t > 1$. We first obtain a tentative mapping based on pixel overlap between regions using the F-measure, also known as coverage measure, defined by:

$$F(r_j^t, r_j^{t-1}) = \frac{2\nu(r_j^t, r_j^{t-1})\rho(r_j^t, r_j^{t-1})}{\nu(r_j^t, r_j^{t-1}) + \rho(r_j^t, r_j^{t-1})}, \quad (6)$$

where ν is precision and ρ is recall defined from $t - 1$ to t by:

$$\nu(r_j^t, r_j^{t-1}) = \frac{r_j^t \cap r_j^{t-1}}{r_j^t}, \rho(r_j^t, r_j^{t-1}) = \frac{r_j^t \cap r_j^{t-1}}{r_j^{t-1}} \quad (7)$$

Region r_j^{t-1} is mapped to r_j^t if the F-measure between them $F(r_j^t, r_j^{t-1})$ is above a certain threshold $th_{overlap}$. For the evaluation presented here, we consider $th_{overlap} = 0.6$.

Since overlap is not a reliable similarity measure, we use the region's shape to prune the region mapping. For each region r_j , we obtain the contour c_j by extracting the regions border points, starting from the leftmost one in a clockwise order. To characterize the contour's shape we use the Discrete Cosine Transform (DCT) of the distance between each contour point and the contour's centroid. The contour's descriptor vector is:

$$f_j = DCT(\|c_j - centroid(c_j)\|), \quad (8)$$

where $\|c_j - centroid(c_j)\|$ is the vector of distances between each contour point and the contour's centroid. In order to have the same descriptor dimension, we resample the contour to 40 points. This was found to describe the cell's shape with enough detail. Using this new descriptor and Euclidean distance we prune the coverage mapping. If the distance between the two regions' shape $d(f_j^{t-1}, f_j^t)$ is below a threshold ($th_{shape} = 50$) they are considered to correspond to the same cell.

Due to instability of the segmentation, regions can change randomly in some isolated frames, which leads to tracking failure since no correspondence is found. A basic biologic assumption is that cells change slowly through time, this means that rapid changes are most likely due to segmentation errors. To correct for this case, we search for a possible matching cell in the frame at $t + 1$, if that exists we consider the cell to be tracked across frame t . The region at frame t is approximated by the previous region r_j^{t-1} .

We also assume that since cells vary slowly a valid cell should be tracked during a minimum number of frames, and ignore small tracks. Tracks with less than 4 frames in length were removed from the final result.

The result is the definition of cell tracks $T_{cell} = i^l$, $l = 1, \dots, w_j$, where w_j is the number of frames for which a specific cell r_j was successfully tracked.

4. Results

As the root tracking result, we obtained images from the same root region, see Figure 2, correctly compensating the situation showed in Figure 1. As the cell tracking result, we can automatically follow the same cell under study over a sequence of frames, as shown in Figure 2(top). Furthermore we filter the segmentation result using temporal stability criteria, selecting regions more likely to be cells, as show in Figure 1(bottom).

In an experiment with 61 frames, our approach found 691 tracks with more than 4 frames each. The max-

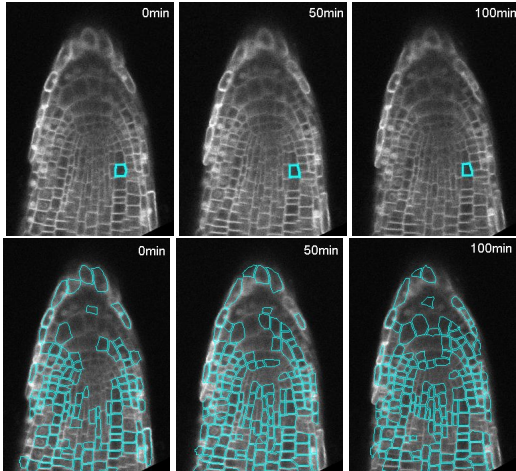


Figure 2. Images of the same cell tracked over time (top) and filtering of the segmentation result using temporal stability criteria (bottom).

imum tracking length is 52 frames, and the average length is 11.

We chose 40 of these tracks to analyze their validity. These were chosen randomly among those which contained correctly segmented cells (based on user observation). Of these, 26 (65%) were completely correct over time, 12 (30%) had some deformation in shape cell (mostly in the beginning or in the end of the tracking) and 2 (5%) had incorrect correspondences. Considering individual frame correspondences between cells, 93% of them were correctly tracked over time.

Regarding the criteria of discarding small tracks, we found that 68% of the tracks with 2 or 3 frames correspond to regions that do not represent true cells, due to noise or segmentation error.

The biologist user can choose any tracked cell to navigate through time and observe the cell's evolution. For each cell, characteristics like growth rate, elongation and relative movement can now be extracted in an automatic and unsupervised way.

5. Conclusion

In this work we introduced an unsupervised approach to automatically track cells in plant confocal microscopy. In a first step we track the region under study in real time during acquisition so that root growth can be compensated. This keeps the region under study within frame for the duration of the experiment. Secondly, after acquisition we process the cell images to obtain a cell-to-cell correspondence which can be used to automatically study the change of cell's characteristics through time.

One major difficulty in this work is that the water-

shed segmentation from which we start contains a high percentage of miss-segmented cells, which leads to cell tracking errors (from which we cannot recover). Work is underway to include temporal and neighborhood information at the segmentation stage, to improve stability of the image segmentation. We also face the ambiguity problem of 2D imaging of a 3D structure, which can eventually lead to loss of tracking of the specific cell. This is a result of the small depth of focus from confocal microscopy, which can only be solved by 3D image acquisition/tracking.

Acknowledgements

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

References

- [1] G. Beemster and T. Baskin. Analysis of cell division and elongation underlying the developmental acceleration of root growth in *arabidopsis thaliana*. *Plant Physiology*, 116:1515–1526, 1998.
- [2] 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.
- [3] 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, August 2007.
- [4] B. Garcia, A. Campilho, B. Scheres, and A. Campilho. Automatic tracking of *arabidopsis thaliana* root meristem in confocal microscopy. In *Proc. of the Inter. Conf. on Image Analysis and Recognition, Springer LCNS 3212*, pages 166–174, 2004.
- [5] C. Harris and M. Stephens. A combined corner and edge detector. *Proc. of the 4th Alvey Vision Conf.*, 1988.
- [6] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [7] A. Iwamoto, D. Satoh, M. Furutani, S. Maruyama, H. Ohba, and M. Sugiyama. Insight into the basis of root growth in *arabidopsis thaliana* provided by a simple mathematical model. *J. Plant Res.*, 119:85–93, 2006.
- [8] V. Luc and P. Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. on PAMI*, 13(6):583–598, 1991.
- [9] M. Marcuzzo, P. Quelhas, A. Campilho, A. M. Mendonça, and A. Campilho. A hybrid approach for cell image segmentation. *Proc. of the ICIAR, Springer LCNS 5112*, pages 739–749, 2008.
- [10] F. Meyer and S. Beucher. Morphological segmentation. *JVCIR*, 1(1):21–46, Sept. 1990.
- [11] S. Mohinder and P. Angus. *Kalman Filtering Theory and Practice*. Prentice Hall, 1993.
- [12] T. Roberts, S. Mckenna, N. Wuyts, T. Valentine, and A. Bengough. Performance of low-level motion estimation methods for confocal microscopy of plant cells in vivo. In *Motion07*, pages 13–13, 2007.