Robust Estimation of Stem Cell Lineages Using Local Graph Matching

Min Liu, Amit K. Roy-Chowdhury ∗
Department of Electrical Engineering
University of California, Riverside

G. Venugopala Reddy †
Department of Botany and Plant Sciences
University of California, Riverside

Abstract

In this paper, we present a local graph matching based method for tracking cells and cell divisions. This will allow us to estimate the lineages of the cells in a 4D spatio-temporal image stack obtained using fluorescence imaging techniques. We work with plant cells, where the cells are tightly clustered in space and computing correspondences in space and time can be very challenging. The local graph matching method is able to compute the lineages even when significant portions of the images are corrupted due to sensor noise in the imaging process or segmentation errors. The geometric structure and topology of the cells’ relative positions are efficiently exploited to solve the tracking problem using the local graph matching technique. The process not only computes the correspondences of cells across spatial and temporal image slices, but is also able to find out where and when cells divide, identify new cells and detect missing ones. Using this method we show experimental results to track the properly segmented cells, and compute cell lineages from images captured over 72 hours, even when some of those images are highly noisy (e.g., missing cells).

1. Introduction

Local spatio-temporal co-ordination of cell growth and cell division is important for proper development of organs in both plant and animal systems. The subject of this study, the shoot apical meristems (SAMs) also referred to as the stem-cell niche, is the most important part of the plant body plan because it supplies cells for all the above ground plant parts such as leaves, branches and stem, and at the same time maintains its stable size. Therefore, a tight spatio-temporal co-ordination between cell division and differentiation of progeny cells into organs is critical to maintain the stability of SAMs. However, the causal link between cell growth and cell division and how they, in turn, affect organ formation is not well understood. This is because of the lack of quantitative measurement of cell growth patterns. Therefore, the development of computational platforms capable of automatically tracking cells and cell division patterns from fluorescent 4D (3D spatial slices + time) images acquired by using laser scanning confocal microscopy is very important. It will lead to computational models to explain the causal relationships between cell deformation dynamics, cell growth and cell division patterns. This is a computational challenge that has universal application to all developmental fields, including both animals and plants.

In this work, we present a method to automatically compute cell lineages (i.e., cell tracks through cell divisions) from the 4D image stacks obtained over days of continu-
ous observations of the growth of the plant stem cell niche. The SAMs of model plant *Arabidopsis thaliana* consist of approximately 500 cells located in multiple layers. We observed the same set of SAMs repeatedly for 3 days at 3 hour intervals. A few examples of images used in this study are given in Figure 1. The images obtained from live-imaging study can be noisy (due to inherent limitations of imaging physics); thus, the tracking method has to be robust enough to account for this noise. Also, the cells are highly clustered together and it is very challenging to maintain their individual correspondences over extended periods of time.

1.1. Proposed Approach and Relation to Existing Work

Given sets of segmented cells at different time instants, cell tracking is essentially a kind of point matching problem, which has been widely studied. One of the most popular solutions is the Softassign Procrustes algorithm [2, 7], which has been applied to compute cell lineages, and has been improved to take care of cell divisions [5]. The Softassign method uses the information of point location, and solves the global correspondence problem as well as the affine transformation problem between two point sets iteratively at the same time. It however does not take into account the information about the topological inter-relationship between the points, especially the cells’ local neighborhood structures [11], which are very important for tracking similarly shaped cells in close proximity and in noisy images, like those in Figure 1. By considering this inter-relationship, we are able to handle high level of noise in the imaging and segmentation process.

In this paper, we propose a graph matching method, where instead of finding the correspondences of cells globally, we would like to solve the matching problem *locally* by cell, through local graph matching techniques [3, 4]. We will exploit the local geometrical and topological structure of the relative position of the cells. By matching the relative position information of cells with respect to their nearest neighbors, such as the length and orientation of the edges in the local graph, we find the most similar cell pair. Starting from this seed pair, we grow the number of matched cells by computing the similarities between local regions in the graph [8]. The proposed method has greater robustness over the global matching methods [5], especially when there are missing cells due to noise in imaging, as shown in Figure 1, or poor segmentation, as shown in Figure 8. The local graph matching method can be very flexible also. In situations, where it is hard to find the first corresponding cell pair (seed) automatically, we can identify the first corresponding pair, and then continue the rest of the tracking process. (In our application, the seeds were chosen automatically by the algorithm.) In fact, the local graph matching technique is motivated by object recognition approaches that rely on local feature matching to overcome the problems of noise and partial occlusion [3, 10].

In the context of tracking through cell divisions, level set approach is another possible method. However, it is difficult to apply when the cells are in close contact and it is hard to distinguish a cell from its neighboring cells because they have similar shapes and almost the same features, and when large parts of the image may be missing in one instance due to sensor noise. This is why method like [6] can not be adapted to our problem. Although the level set method is not suitable for cell tracking in our case, we can adopt it for cell segmentation where the objective is not to find correspondences [1].

Another unique feature of our approach is that we can obtain the tracks by integrating the tracking results over the entire 4D image stack, unlike methods which do single-layer tracking or surface-layer analysis only [5] [8]. This allows us to get better tracking results since each cell is typically imaged in more than one slice at each time point.

1.2. Broad Overview of Solution Strategy

The level set segmentation will provide information like the centroids of the cells, their area, orientation and so on to the tracker. The local graph matching based tracking for any two time-consecutive cell image stacks can now be described as follows, a diagrammatic representation of which is given in Figure 2.

1. Find the seed cell pair by local graph matching.
2. Starting from this seed pair, find the correspondences of the neighboring cells by computing the similarity of local regions in the graph.
3. Detect cell divisions by detecting changes in the topology of the graph.
4. Fuse correspondence information across both time and slices to obtain cells’ tracks in the 4D image stack.

Figure 2. Diagrammatic representation of local graph-matching based cell tracking for two time-consecutive image stacks.
2. Detailed Methodology

2.1. From Cells to Graphs

We describe here how to create the graphical abstraction given a collection of cells in an image. Every cell is represented by a vertex in the graph and neighboring vertices are connected by an edge. We regard those cells that are within a certain distance around cell c as its neighboring cells, denoted by the neighborhood cell set N(c), which is ordered in the counter-clockwise direction. An example of the graph superimposed on the cells is shown in Figure 3.

2.2. Local Graph Matching Method

The graph structure automatically includes the relative position information of the cells, such as the relative distance between two neighboring cells (the edge length) and the edge orientation. Usually, if there is no cell division and the images are not corrupted, the local graph topology should not change. When this condition is satisfied, our tracking method can find the corresponding cell pairs by matching their related local graphs. The case where the graph structure changes due to cell divisions is described in the later section.

Given two local graphs G1 and G2, corresponding to cell ci at time t and cj at time t + 1 respectively, the distance measure DL(ci, cj) for these two graphs is composed of three parts: the normalized difference of the edge lengths at consecutive time instants t and t + 1, the difference of the orientation angles between the edges, and the cells’ normalized location difference information. This distance measure can then be mathematically expressed as

\[
DL(ci, cj) = \lambda \sum_{ck, cj \in N(ci), cj \in N(cj)} \left( \frac{1}{l_{ck, c}(t) l_{ck, c}(t+1)} \right)^2 \theta_{ck, c}(t) \theta_{ck, c}(t+1) + \frac{\|P_{ci}(t) - P_{cj}(t+1)\|^2}{\Delta^2}
\]

where cki is a neighboring cell of ci, and ckj is a neighboring cell of cj, lcki(t) and lckj(t + 1) are the edge lengths, \(\theta_{cki, cj}(t)\) and \(\theta_{ckj, cj}(t + 1)\) are the orientation angles in radians of the edges measured relative to a horizontal axis, \(P_{ci}(t)\) and \(P_{cj}(t + 1)\) are the cell position vectors, and \(\Delta\) is the average distance between two neighboring cells. If two local graphs match (i.e., the distance measure is small), we can say that the central cells (i.e., ci and cj) in those two local graphs are a corresponding cell pair. This process is depicted diagrammatically in Figure 3.

2.3. Finding the seed (first) cell pair

Although the local graph matching method is effective in finding some corresponding cell pairs, it is based on the assumption that the local graph topology does not change along time, and the geometry of the local graph changes little. However, this is not always true, since the cells are growing and every cell may not be imaged properly, so our tracking strategy is to first find the most similar cell pair, and then starting from it, to grow the matching process spatially.

We now describe how to find the first cell pair.

Let us assume that the first image has N cells, and the second image has M cells. According to the distance function in Equation (1), we can get the distance value \(DL(ci, cj)\) of every cell pair (ci, cj), and pick the most similar cell pair \((c_a, c_b)\) that satisfies

\[
DL(c_a, c_b) = \min_{ci, cj} \{DL(ci, cj), i = 1, \ldots, N, j = 1, \ldots, M\}
\]

Alternatively, we could find the top few most similar cell pairs instead. Then the tracking process will be faster, because we can start from more corresponding cell pairs (more seeds). This procedure is used in this paper. However, in some situations, it may be better to find the seed manually, which can be easily integrated into our approach.

2.4. Growing cell correspondences from a seed pair

Given a matching cell pair, we now describe the process to compute matches for the neighboring cells. This distance function is composed of two parts: the normalized difference of the edge lengths and the difference of the orientation angles between the edges. Let ci and cj be two correctly matched cells respectively at time instants t and t + 1, and cki and ckj be two neighboring cells around them respectively. Then the distance measure function is

\[
DN(ci, cj; cki, ckj) = \frac{l_{cki, ci}(t) l_{ckj, cj}(t + 1)}{l_{cki, ci}(t) l_{ckj, cj}(t + 1)} \theta_{cki, ci}(t) \theta_{ckj, cj}(t + 1)
\]

Figure 3. The matched local graphs \(G_1\) at t and \(G_2\) at \(t + 1\), the correspondence of the seed cell pair \((1, 1')\), as well as the correspondences of the neighboring cells, such as \((2, 2')\) and \((4, 4')\).
If the distance is small, $c_{k_i}$ and $c_{k_j}$ are identified as matching cells, and the process continues. This leads to a recursive procedure for matching cells from a seed to the neighbors. The process is illustrated in Figure 4 where the seeds identified in the top row are used to grow the tracking spatially using local graph matching.

2.5. Cell division detection

The topology of a certain local graph will change if a cell divides. We now explain how cell divisions are detected in the local graph matching approach. Take the local graphs illustrated in Figure 5 for example. Here a cell $M$ at time $t$ is divided into two sibling cells $A$ and $B$ at time $t + 1$.

We can first find the cell pair $(M, A)$ through the correspondence growing procedure starting from cell pair $(T, Q)$ and/or $(S, P)$. Then we check the area difference of cells $M$ and $A$. If the difference is over about half of the area of an adult cell, then we could say that there maybe a cell division event, and cell $M$ may be the mother cell and cell $A$ may be one of the daughter cells. If there is really a cell division, we should be able to find the other sibling cell in the neighboring cells of $A$. Specifically, we can search among the neighborhood around the original location $M'$ of the mother cell at time $t + 1$, and the cell $B$ with the shortest distance from $M'$ should be the candidate for the other sibling cell. Furthermore, this distance $|BM'|$ should be within about half of the average distance between two neighboring adult cells.

Formally, the procedure for detecting the cell division is as follows:

1. Find the corresponding cell pair $(M, A)$ using the local graph matching technique described in Section 2.4;
2. Check whether the area difference of $(M, A)$ exceeds half the area size of an adult cell; if true, then $(M, A)$ is a candidate pair of mother cell and sibling cell;
3. Search in the neighborhood of the daughter candidate $A$ to find the other daughter candidate (sibling) cell $B$;
4. Compare the sibling cell candidates $A$ and $B$; if the sum of their areas is about the same as the mother cell $M$, then we can say that $M$ divides into two daughter cells $A$ and $B$. We can also check the distance between $A$ and $B$, which should be within about half the average distance between two neighboring adult cells.

Tracking results with cell divisions are shown in Figure 7.

2.6. Correspondence across image slices

Using the local graph matching methods described above, we can compute correspondences across slices $S_k$.
and \( S_{k+1} \) at the same time instants, i.e., correspondences within the same stack. The distance measure function is

\[
D_S(c_i, c_j) = \sum_{c_r \in N(c_i), c_r \in N(c_j)} \frac{(l_{c_r, c_i}(S_k) - l_{c_r, c_j}(S_{k+1}))^2}{l_{c_r, c_i}(S_k)^2} + \lambda \sum_{c_r \in N(c_i), c_r \in N(c_j)} \left( \theta_{c_r, c_i}(S_k) - \theta_{c_r, c_j}(S_{k+1}) \right)^2 + \frac{||P_{c_i}(S_k) - P_{c_j}(S_{k+1})||^2}{\Delta^2} + \left| \frac{A_{c_i}(S_k) - A_{c_j}(S_{k+1})}{A_{c_i}(S_k)} \right|^2,
\]

where \( A_{c_i}(S_k) \) and \( A_{c_j}(S_{k+1}) \) are the area size of cells \( c_i \) and \( c_j \) respectively. Because in this case there should be no cell divisions, we add the cell area difference into the distance function to make it more accurate in finding the correspondences across slices. Tracking results across slices, with high sensor noise are shown in Figure 8.

2.7. Fusion of spatial and temporal tracks

We now show how to fuse the spatial and temporal correspondences in order to get the final cell tracks. As explained in Figure 1, every cell usually occupies 3 slices, so we need to integrate the tracking results in every 3 consecutive slices. The fusion process is described in Figure 6 and below.

In the tracking process, we build a correspondence matrix \( C \) to denote whether two cells \( c_i, c_j \) in two images (may be at different times or different slices at the same time) are the same cell or not. If the two cells are the same, we set the correspondence value \( C(i, j) \) to 1, otherwise we set it to 0. If cell \( c_i \) splits into cells \( c_j \) and \( c_{j2} \), we will set \( C(i, j_1) \) and \( C(i, j_2) \) to 2.

Let us denote the correspondence across time instants \( t \) and \( t + 1 \) at slice \( S_k \) as \( C_k(t, t + 1) \), and the correspondence across slices \( S_k \) and \( S_{k+1} \) at time \( t \) as \( C_k(k, k + 1) \). The fusion process is composed of the upward fusion matrix \( C_{k-1,k}(t, t + 1) \) and downward fusion matrix \( C_{k,k+1}(t, t + 1) \), which can be easily proved to be products of the individual correspondence matrices as follows:

\[
\begin{align*}
C_{k-1,k}(t, t + 1) &= C_k(k, k - 1)C_{k-1}(t, t + 1)C_{t+1}(k - 1, k); \\
C_{k,k+1}(t, t + 1) &= C_k(k, k + 1)C_{k+1}(t, t + 1)C_{t+1}(k + 1, k).
\end{align*}
\]

The final set of the correspondences at slice \( S_k \), denoted by \( C^*_k(t, t + 1) \), is obtained as follows:

1. \( C^*_k(t, t + 1) = C_k(t, t + 1); \)
2. For any new cell pair identified by \( C_{k-1,k}(t, t + 1) \) or \( C_{k,k+1}(t, t + 1) \), but not in \( C_k(t, t + 1) \), add this pair to \( C^*_k(t, t + 1) \);
3. For any cell pairs identified in \( C_{k-1,k}(t, t + 1) \) and \( C_{k,k+1}(t, t + 1) \), but with conflicting results (i.e., the matching cell pairs have one cell in common but the other cell is different in the upward fusion and downward fusion matrices), do not add to \( C^*_k(t, t + 1) \).

With this fusion process, we can find new cell pairs, which are not identified by single slice tracking, but can be found through the adjacent slices. With this process, we can track most of the properly segmented cells. The effect of the fusion in identifying additional cells is shown in Figure 9.

3. Experimental Results

We have tested our method on three data sets of SAMs. For limitation of space, we show detailed results on only one of them. The experimental results are shown on 4D cell image stacks obtained from plant cells and observed along 24 consecutive time instances, with the time interval of 3 hours between two consecutive instances. The cells were segmented by existing level set segmentation methods [1]. There are about 500 cells located in different layers. Registration is done by the existing method such as the alignment method of Maximization of Mutual Information [9]. The code for the tracker will be released publicly with this paper.

**Choice of the parameters in the tracking system**: The average distance between two neighboring cells is about 50 pixels, so if the relative distance of two cells is within 50 pixels, they are regarded as neighbors in the implementation. On the average, if two cells are the same cell in different images, the normalized length difference of the same edge is less than 0.1, the orientation difference of the same edge is within 0.17 radians (10°), the cells’ normalized location difference is within 0.2, and the normalized cell area difference is within 0.2. Based on these and the fact that one cell usually has 5 to 8 neighbors, we set \( \lambda = 0.33 \) to normalize the distance measure for different physical parameters.
Figure 7. Tracking results with cell division, where the mother cells are denoted by purple dots in the first frame, while the sibling cells are in white circles in the second frame.

Figure 8. Tracking results across slices in noisy cell images, where some cells in the central part are totally missing.

If two cells are considered to be the same one, the distance measure value $D_L$ should be less than 0.2 in Equation (1), $D_N$ less than 0.02 in Equation (3), and $D_S$ less than 0.24 in Equation (4). The average cell area is 400 pixel$^2$, so the area of the divided daughter cell is about 200 pixel$^2$. If two matched cell pair’s area difference is beyond 200 pixel$^2$, we consider that there may be a cell division. And if there really is a cell division, the distance between the two daughter cells should be about half of the average distance between two neighboring adult cells, i.e., 25 pixels.

**Performance Evaluation:** Figure 7 shows the tracking results across time instants with cell division. The same cells are denoted by the same number while for the cell division, the mother cell and daughter cells are denoted by circles. Table 1 shows the number of cell divisions being detected manually and automatically by the proposed method.
Figure 9. The tracking results after the spatial and temporal track fusion process, where 9 newly tracked cells are denoted by white circles.

Figure 10. The cell lineages along 4 time instants are denoted by lines of different colors. Here the same number denotes the same cell at different time instants, and the divided daughter cells are denoted by the same number at the same time instant.

<table>
<thead>
<tr>
<th>Slice</th>
<th>Time (hours)</th>
<th>Manual</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>39</td>
<td>21 (20)</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>57</td>
<td>48 (49)</td>
<td>49</td>
</tr>
<tr>
<td>7</td>
<td>69</td>
<td>55 (53)</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 1. The comparison of cell divisions tracked manually (both in the unsegmented data and the segmented data, the latter is in the brackets), and those detected automatically with the proposed local graph matching method in different slices.

in different slices. Here the manually detected results act as the ground-truth. We can see that this local-graph matching method can detect 100% of cell divisions in the segmented data, and about 96% of the unsegmented data. Sometimes, imperfect segmentation will cause pseudo-divisions, as in the second row in Table 1.

Figure 8 is the case of tracking across image slices, with noise in the central part of the cell images, where local graph matching method demonstrates its robustness to find the corresponding cell pairs. From the tracking results, we can notice that when the cells in the central part are almost totally corrupted, the correctly segmented cells are effectively tracked. The robustness of this local graph matching method comes from the ability to grow the correspondences to the neighborhood from a seed pair. For example, if we assume the 10th cell pair as the seed pair, we can find the 34th cell pair by different correspondence growing paths, such as (10, 18, 22, 29, 34), (10, 17, 20, 32, 31, 34), (10, 9, 14, 23, 32, 31, 34) and so on. As long as we use right seed pairs are chosen, the tracking algorithm will automatically pick up a right path from the seed pair to any other cell pairs. Even if some of these paths are corrupted by noise, cell pairs can be found as long as there is one feasible path.

Figure 9 depicts the tracking results before the fusion of the spatial and temporal tracks and after the fusion process, respectively. We can notice some newly tracked cells (de-
Table 2. Cells being tracked manually (in the unsegmented data and in the segmented data, the latter is in the brackets) and automatically (before the fusion process and after the fusion process). We have chosen some representative examples from the data set.

<table>
<thead>
<tr>
<th>Slice</th>
<th>T (hours)</th>
<th>Manual</th>
<th>No fusion</th>
<th>With fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3-6</td>
<td>45 (37)</td>
<td>37</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>6-9</td>
<td>72 (64)</td>
<td>64</td>
<td>69</td>
</tr>
<tr>
<td>5</td>
<td>3-6</td>
<td>84 (74)</td>
<td>74</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>18-21</td>
<td>85 (72)</td>
<td>72</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>21-24</td>
<td>52 (41)</td>
<td>41</td>
<td>48</td>
</tr>
<tr>
<td>6</td>
<td>6-9</td>
<td>99 (77)</td>
<td>77</td>
<td>90</td>
</tr>
</tbody>
</table>

noted by white circles) identified by the fusion process. Table 2 gives a comparison on how many cells are tracked before the fusion process in two consecutive time instants, and how many cells are tracked after the fusion process. Also, the number of the cells can be tracked manually both in the unsegmented data and the segmented data are treated as ground truth. We can track all those cells that can be tracked manually in the segmented data in one single slice. By fusion of the cells correspondences across time and slices, we can identify more cell pairs, improving the tracking accuracy up to 90% of the manually tracking cells in the unsegmented data.

Finally, we show the tracking results on computing the cell lineages by repeating the two-frame tracking process over the entire data collection of 72 hours. Based on the correspondence results of cell tracking and cell division tracking, we can get the cell lineages, as illustrated in Figure 11. Figure 10 shows the example of cell lineages superimposed on the images.

4. Conclusions

In this paper, we presented a local graph matching based method to track cells and cell divisions from 4D (3D space + time) confocal microscopy images with significant sensor noise. The main challenge comes from the special structure of the SAMs where the cells adhere to each other. We address this problem by exploiting the geometric structure and topology of the cells’ relative positions. By matching the cells’ related local graphs we get the cells’ correspondences both spatially and temporally, and locate the cell divisions. Finally, we fuse the temporal and spatial tracking results together to get the cell lineage. In the experiment, we can track up to 90% of all the cells and can identify 96% of cell divisions.

References


