Title

Automated Nuclei Tracking in C. elegans based on Spherical Model Fitting with Multiple Target Tracking

Authors

Kang, S
Giurumescu, C A
Chisholm, A D
et al.

Publication Date

2012-04-01

Peer reviewed

AUTOMATED NUCLEI TRACKING IN C. ELEGANS BASED ON SPHERICAL MODEL FITTING WITH MULTIPLE TARGET TRACKING

Sukryool Kang[1], Claudiu A. Giurumescu[2], Andrew D. Chisholm[2], and Pamela Cosman[1],

[1] Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA, 92037-0407, USA
Email: srkang@ucsd.edu, pcosman@ucsd.edu
[2] Division of Biological Sciences, Section of Cell and Development Biology
Email: cgiurumescu@ucsd.edu, chisholm@ucsd.edu

ABSTRACT

We present a method for tracking nuclei of C. elegans cells in three dimensional time lapse (4D) data during embryogenesis. It is based on spherical model fitting with multiple target tracking. The algorithm tracks nuclei in a local search space, detects cell division, and searches for newborn sister cells locally. It tracks multiple candidates and postpones the decision to select the newborn sister cell until it becomes bright. This approach results in low error rates. We evaluate the accuracy of the algorithm by comparing to an existing approach on 4D data sets acquired from laser scanning confocal microscopes.

Index Terms— spherical model fitting, Caenorhabditis elegans, embryogenesis, cell tracking, mutiple target tracking

1. INTRODUCTION

The nematode Caenorhabditis elegans (C. elegans) is a free-living worm that inhabits soil and feeds on bacteria. C. elegans is widely used as a model organism in developmental and molecular biology. It develops from one cell to 588 cells during embryogenesis. The embryonic lineage has been determined and is highly invariant, describing the fates of cells to the adult stage [1][2]. This trait makes C. elegans suitable for studies on cell differentiation and related subjects. Rapid advances in microscopy technology enable high resolution 3D image sequences. Especially, fluorescence microscopy provides sufficient spatial and temporal resolution to facilitate analysis of embryonic cell lineages in living embryos [2].

In recent years, various segmentation and tracking methods have been developed for fluorescence microscopy [3][4][5][6]. In this paper, we propose a new algorithm to track nuclei in fluorescently labeled C. elegans embryos. It is based on spherical model fitting [2] and the algorithm consists of nuclei tracking and cell division detection. The nuclear position is tracked from the position at a previous time point in a narrow search space without considering cell division events, making the tracking algorithm simple. This approach shows robust tracking results even in low SNR. In the cell division detection algorithm, we apply a multiple target tracking framework. The detection of newborn cells is made not at the frame when cell division happens, but at a later frame when the cell is more recognizable. Multiple target tracking and postponed detection of newborn cells reduces incorrect detection of nuclei.

2. MATERIALS AND METHODS

2.1. Image acquisition

Fluorescently-labeled C. elegans embryos were recorded with Zeiss LSM510 or LSM700 laser scanning confocal microscopes. We followed previously published approaches to acquire 4D data sets [7]. 3D stacks of 512×275×35 voxels with a resolution of 0.125×0.125×0.90μm³ were acquired every minute. The imaging starts from the 4 to 8 cell stage which indicates the number of cells in the frame and continues until 480 minutes of development. Each stack is rendered isometric by replicating missing z-slices due to the low resolution along the z-axis. After replicating z-slices, each stack has 512×275×280 voxels.

2.2. Nuclei tracking algorithm

Tracking of nuclei consists of several steps. The spatial positions of nuclei are identified manually in the first frame with a visualization tool. The number of nuclei could be from 4 to 8 in the first frame depending on the time when the image is recorded. The radii of the cells are pre-defined. In our images, the radius of the AB cell is 26 pixels and those of EMS and P2 cells are 24 pixels. Whenever cell division happens, the radius is assumed to decrease by 2 pixels. This pre-defined radius is off by 33% at the 28 cell stage in the worst case.

Second, we use a local search to track the identified nuclei. Nuclei are spherical except during mitosis. This allows us to use a simple spherical mask to track nuclei [2]. Given an identified nucleus position at time t-1, we define a search space which is a cube centered at the nucleus location at time t-1 with edges of length 2.5R, where R is the pre-

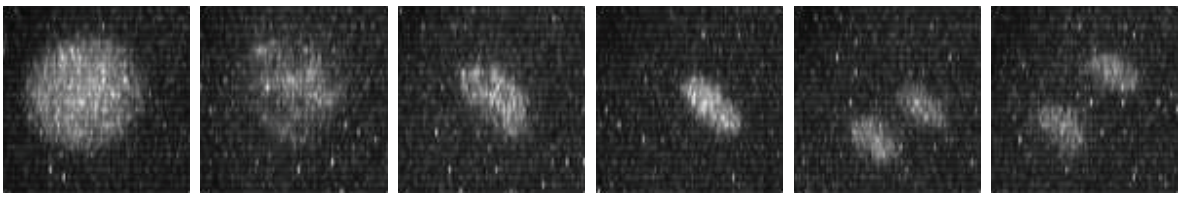
(a) time t (b) time t+1 (c) time t+2 (d) time t+3 (e) time t+4 (f) time t+5

Figure. 1. Projection image during mitosis (a) Nucleus is spherical, (b), (c), (d) Nucleus becomes progressively more ellipsoidal, (e), (f) Cell has divided into two sister cells.

defined nucleus radius. We use a uniform spherical mask of radius 0.75R to capture the core of the nucleus that is less affected by image noise. We convolve the search space with the spherical mask and find the new position where the convolved signal is maximum. Convmax is defined as the maximum value of the convolved signal. If there is more than one position at a time t which achieves the same global maximum value, we choose the one closest to the position at t-1. All identified nuclei at time t-1 are linked to nuclei at time t.

2.3. Cell division detection algorithm

In the tracking, we track only one nucleus in each local search space. After tracking all identified nuclei from time t-1 to time t, we apply the cell division algorithm.

First, we look for a mitotic nucleus and estimate the cell division time. From observation of 20 embryo data sets, the minimum cell division cycle is about 16 minutes, and mitosis typically takes 4 minutes. During mitosis, the nucleus shape changes from a sphere to an ellipsoid, and we can observe condensed mitotic chromosomes [2] (Fig. 1). We use three conditions to detect cell division. First, we set the minimum cell division cycle as 10 minutes. If nuclei last more than 10 minutes, they are candidates for cell division. Second, bright pixels are aggregated in a small area during cell division, and Convmax decreases. We measure the variation of Convmax using Convmax Rate, defined as:

$$Convmax\ Rate_t = \frac{\max_{i \in T} Convmax_i - Convmax_t}{\max_{i \in T} Convmax_i}$$

where the range of T is from time t-4 to time t. Because mitosis typically takes 4 minutes, we use Convmax values from time t-4 to time t to measure the decrease in Convmax values. A threshold of 0.15 is applied to detect the decrease in Convmax. Third, we use the change of nucleus shape during mitosis. We apply 4^3 ellipsoidal matched filters, and the standard deviation of filter outputs is used to detect the change of cell shape. We generate the ellipsoidal filters by decreasing the minor axis length by half of the cell diameter and the filter is rotated by 45 degrees along the X, Y and Z axes. The tracked nucleus location is used as the center of the matched filters. The output values are normalized by the maximum value of the filter output and the standard deviation is calculated. When the cell becomes ellipsoidal, the variance of the matched filters output increases. The threshold value of 0.035 is applied. If the identified nucleus satisfies the three conditions of nucleus age, Convmax rate and standard deviation of ellipsoidal filters, the algorithm declares a possible cell division frame.

After such declaration, we search for candidate sister cells of the nucleus. First, we set two cube search spaces, which are S_S (small space) and S_L (large space). The edge length of S_S is 4R and the edge length of S_L is 8R. S_L is used to find a newborn sister cell that moves far from the center nucleus and S_S is used to select candidates near the identified cell. After setting search spaces, all previously tracked and identified nuclei are removed in the search spaces. Since the pre-defined cell radius could be smaller than the actual cell radius, we use a sphere with 2.5R radius to remove bright pixels around the edge of the identified nuclei. For the center nucleus, the pre-defined radius R is used to keep pixel information around the center cell (Fig. 2 – (a), (b), and (c)).

Next, we pick the top 3 candidate sister nuclei in each search space. After convolving search spaces with a spherical mask of radius 0.75R, 3 candidate sister nuclei are selected in S_S and 3 more candidate nuclei are selected in S_L. In each space, we select convolution maximum points that have distance more than 2R from other maximum points to avoid overlap. If candidate sister nuclei overlap in the two search spaces, we keep the brighter candidate and remove the other candidate cell (Fig. 2 – (d), (e), and (f)). This set of candidate sister cells which are found at time t is denoted F_t.

All candidate sister nuclei are tracked from time t-1 to time t with the nuclei tracking algorithm. Let C_t denote the set of candidate sister cells in existence at time t. The construction of C_t is determined recursively from C_{t-1} and F_t. First, we run the nucleus tracking algorithm on all the candidate sister cells in C_{t-1} to determine a set T_t of tracked candidate sister cells at time t. Next, we check overlap between T_t and all identified nuclei to remove candidates that were the perimeters of identified nuclei. Next, F_t and T_t are combined to construct C_t. If candidates in F_t overlap with candidates in T_t, the candidates that have higher Convmax are included in C_t. When more than 2 candidates at time t-1 are merged into one candidate at time t, the candidate that has less movement is included in C_t. If there are candidates in F_t not corresponding to candidates in T_t, they are also included in C_t for further tracking.

After tracking candidate sister nuclei continuously, the

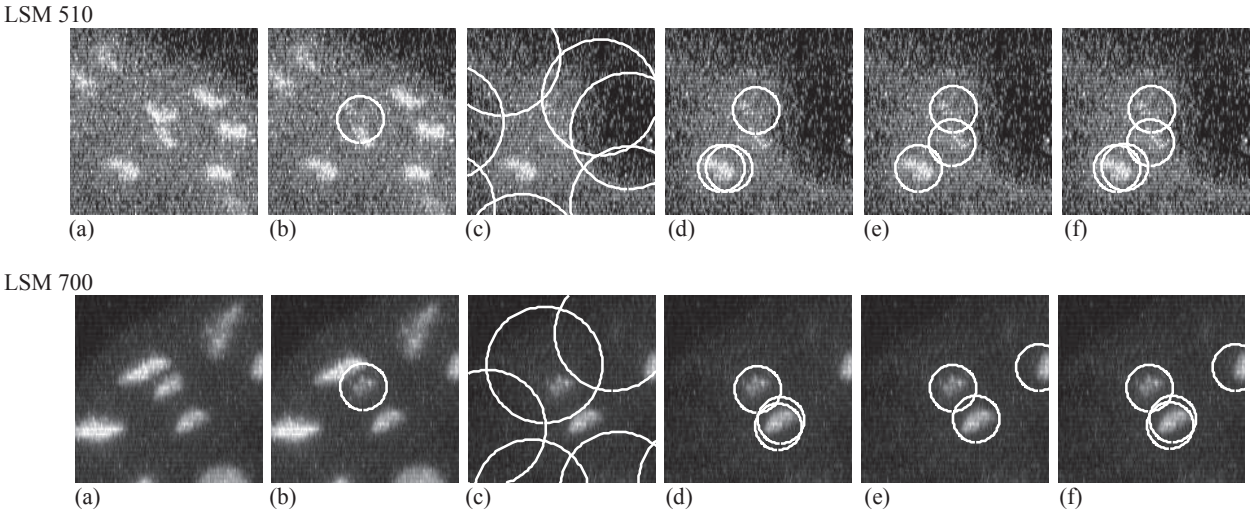

Figure. 2. Search for candidate sister cells (a) projection image of cube search space with edge of length 8R - center cell divides into two sister cells, (b) projection image after identified center cell is removed by using radius R, (c) projection image after identified neighbor cells are removed by using radius 2.5R, (d) three candidates in S_S, (e) three candidates in S_L, (f) candidates in S_S and S_L; even if some candidates look overlapped on projection images, candidates do not overlap in 3D search space.

algorithm looks for the true sister nucleus among the candidates. The decision is based on the fact that two sister nuclei are located closely when the cell division happens. The detection of the sister cell is made when any candidates reach age 10 minutes. In this case we wlll declare a true sister cell among the candidates whose ages are more than 5 minutes and whose Convmax values are greater than 30% of the Convmax of the identified sister nucleus at time t. If the estimation of the possible cell division time is correct, 5 minutes is long enough to find a true sister nucleus. However, we postpone the decision point until one of candidates reaches age 10 minutes to include more possible candidates. After selecting possible candidates that satisfy the age and brightness conditions, we look for the candidate sister nucleus that has the minimum distance from the identified sister nucleus when the candidate sister nuclei are first found. After deciding the true sister cell at time t, the selected sister cell is added to the identified nuclei from the time when it started to the current time t. The radius and lineage name of the cell are also updated. We repeat the nuclei tracking algorithm and cell division algorithm at every time point.

3. RESULTS AND CONDITIONS

We tested our algorithm on 8 embryos (6 of LSM 510 and 2 of LSM 700) covering until the 180 cell stage. Using a 2.8 GHz Intel Xeon desktop computer, processing time takes less than 1 minute for each frame until 100 minutes and less than 3 minutes for each frame from 100 minutes to 200 minutes due to the increase in the number of cells.

We compare the results of our algorithm to manually identified nuclei of 8 embryos until the 180 cell stage and evaluate the errors by counting the false positives (incorrectly detected nuclei) and false negatives (incorrectly undetected nuclei). We measure FNR (false negative rate) and FPR (false positive rate) in three stages which are 4-51, 51-102 and 102-180 cell stages. We compute FNR = FN/(TP+FN) and FPR = FP/(TP+FN) where FN, TP, and FP are the total numbers of false negatives, true positives, and false positives in each stage [6]. TPR (True positive rate) = 1-FNR. The result is shown in Fig. 3. At the 4-51 stage the error rates are under 2% and the error rates increase to 4%~10% at the 102-180 cell stage. The majority of errors found in our experiment are from incorrectly estimated cell division frames. When the algorithm misses a cell division frame, there is no search for a newborn sister cell and it produces a false negative. Also, false detection of a cell division frame produces a false positive. Additionally, when multiple nuclei are located closely, multiple nuclei can be tracked to a single bright nucleus in our tracking algorithm. If the distances between two tracked nuclei are less than the sum of nuclear radii, we define this event as a cell collision even if nuclei in the image stack don't collide.

We compared our algorithm to the Starrynite algorithm [2] on LSM 510, LSM 700, and the Bao et al. data set in [2]. The Bao et al. data set is the one used to report results for Starrynite in [2]. The comparison result is shown in Fig. 4. Our algorithm shows that TPR exceeds 97% and FPR is less than 2.54% for three data sets (Fig. 4-A). Starrynite produces 99.64% TPR and 0.03% FPR on the Bao et al. data set using the parameters provided with the software (Fig. 4-B). For the other data sets, we optimized the parameters to minimize Error over the first 50 minutes. Error is defined as:

$$Error = 1 - \frac{TP}{TP+FN} \times \frac{TP}{TP+FP}$$

TPR = TP/(TP+FN) represents the fraction of actual

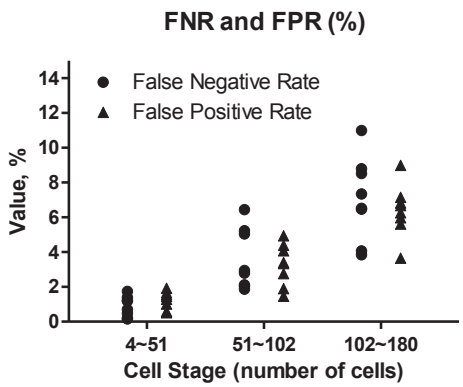

Figure .3. FNR and FPR for 8 data sets.

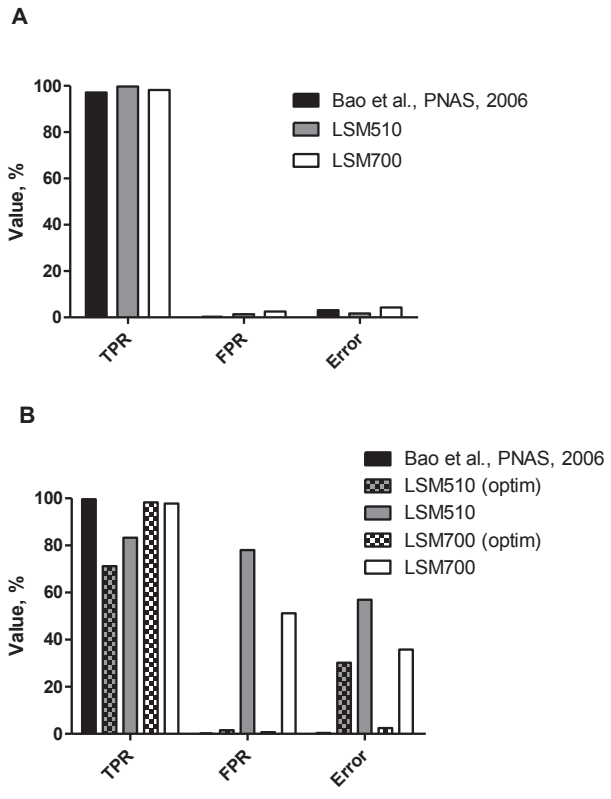

Figure. 4. Comparison to Starrynite [2]. We tracked nuclei in three data sets (original Bao et al. dataset, LSM510, and LSM700)
(A) The cumulative performance of our algorithm until 99 minutes. (B) Performance of Starrynite. We report the cumulative performance using original parameters from Bao et al. on the original Bao et al. dataset; for other data sets, we optimize parameters over the first 50 minutes (optim) and report TPR and FPR. Then, we report the performance until 99 minutes using these parameters.

positives which are declared by the algorithm to be positives. PPV (positive productive values) = TP/(TP+FP) represents the fraction of cases declared by the algorithm to be positives which actually are positive. Ideally, we want TPR to equal 1 (no false negatives) and we also want PPV to equal 1 (no false positives). So we define $1-TPR\times PPV$ as Error to be minimized. The LSM 510 data set has 71% TPR with 1.5% FPR over the first 50 minutes. As we use these parameters until 99 minutes, FPR increases to 78%. The FPR of the LSM 700 data set also increases to 51% until 99 minutes. The result shows that parameters of Starrynite are dependent on time. Even if the FPR of our algorithm is slightly higher than the FPR of STARRYNITE on the Bao et al. data set, our algorithm shows robust performance over multiple data sets and the parameters of our algorithm are less dependent on time.

4. DISCUSSION

We combine nuclei tracking and cell division based on spherical and ellipsoidal model fitting [2]. We use a local search based on the prior temporal information for each nucleus. This approach improves tracking performance in cluttered environments by reducing the influence of neighboring nuclei. In the cell division algorithm, we search for newborn sister cells locally when the algorithm estimates a cell division frame for an identified nucleus. We track multiple targets and look for one newborn sister nucleus when the nucleus is bright enough to be identified to avoid false positives.

5. REFERENCES

[1] J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, "The embryonic cell lineage of the nematode Caenorhabditis elegans," *Dev Biol*, vol. 100, no. 1, pp. 64–119, 1983

[2] Z. Bao, J. I. Murray, T. Boyle, S. L. Ooi, M. J. Sandel, and R. H.Waterston, "Automated cell lineage tracing in Caenorhabditis elegans," *Proc Natl Acad Sci USA*, vol. 103, no. 8, pp. 2707–2712, 2006.

[3] G. Lin, U. Adiga, K. Olson, J.F. Guzowski, C.A. Barnes, and B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," *Cytometry A*., vol. 56, no. 1, pp. 23-36, 2003.

[4] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, JC. Olivo-Marin, and C. Zimmer, "Segmentation and Tracking Fluorescent Cells in Dynamic 3-D Microscopy with Coupled Active Surfaces," *Image Processing, IEEE Transactions on*, vol. 14, no. 9, pp. 1396–1410, 2005.

[5] O. Dzyubachyk, R. Jelier, B. Lehner, W. Niessen, and E. Meijering, "Model-based approach for tracking embryogenesis in Caenorhabditis elegans fluorescence microscopy data," *Proc IEEE Eng Med Biol Soc*, pp. 5356-5359, 2009.

[6] N. Carranza, I. Smal, O. Dzyubachyk, W. Niessen, E. Meijering, "Automated lineage tree reconstruction from Caenorhabditis elegans image data using particle filtering based tracking," *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on,* pp.1921-1924, 2011.

[7] J. I. Murray, Z. Bao, T. J. Boyle, and R. H. Waterston, "The lineaging of fluorescently-labeled Caenorhabditis elegans embryos with StarryNite and AceTree," *Nat Protoc Proc Natl*, vol. 1, no. 3, pp. 1468-1476, 2006.