

SPATIO-TEMPORAL CELL SEGMENTATION AND TRACKING FOR AUTOMATED SCREENING

Dirk Padfield^{1,2}, Jens Rittscher¹, Badrinath Roysam²

¹GE Global Research, One Research Circle, Niskayuna, NY, 12309.

²Rensselaer Polytechnic Institute, 110 8th St., Troy, NY 12180.

ABSTRACT

A growing number of screening applications require the automated monitoring of cell populations including cell segmentation, tracking, and measurement. We present general methods for cell segmentation and tracking that exploit the spatio-temporal nature of the task to constrain segmentation. The images are de-noised and segmented by combining wavelet coefficients at various levels, thus enabling extraction of cells in images with low contrast-to-noise ratios. Each track of clustered cells resulting from association of nearby cells in the spatio-temporal volume is then split into individual cells by evolving sets of contours from other slices. The hypothesis whether to split or merge objects making up the cluster is tested using learned features trained from single track cells. Due to the difficult nature of generating ground truth, we also present a framework for edit-based validation whereby the user corrects the edits made by the automatic system rather than generating the truth from scratch. The results show the promise of the approach and demonstrate the ability of the algorithms to provide meaningful measurements of cell response to drug treatment in low-dose Hoechst-stained cells.

Index Terms— Spatio-temporal analysis, cell segmentation, cell tracking, edit-based validation.

1. INTRODUCTION

High-throughput automated screening of in-vitro systems is of great importance to biological research. Two core components of the algorithms required for this task are the segmentation and visual tracking of cells. Cell motility and division are important aspects of fluorescence microscopy applications in areas such as cancer research, immunology, and developmental biology. While such applications have specific requirements, they share several core challenges. Large-scale experimentation makes it difficult to optimize the imaging conditions for individual experiments. In addition, for time-lapse datasets, the contrast-to-noise ratio is often low as a result of the low concentrations of fluorescent dyes that can be used since all fluorescent DNA binding dyes inhibit DNA replication to a greater or lesser extent with inevitable toxicity. Model systems also typically contain a large number of

cells thus requiring the monitoring of far more targets than common surveillance applications. Specific biological events such as mitosis also need to be handled.

Existing algorithms for cell tracking can be roughly divided into two main approaches: independent detection with subsequent data association [1, 2] and model based tracking [3, 4]. Li *et al.* [5] combines these tasks by both segmenting each frame separately and using a multi-target tracking system using model propagation with level sets and a stochastic motion filter. Padfield *et al.* [6] approach the tracking task as a spatio-temporal segmentation task.

By segmenting the spatio-temporal volume directly, we solve the segmentation and tracking problems simultaneously. We utilize a sparse representation of the data using wavelets. The data is de-noised in the wavelet coefficient space, and the segmentation is found by extracting connected regions from a combined subset of wavelet levels. As opposed to segmenting the data in a single pass, we present an iterative approach for extracting the individual cell tracks. Since cells often touch, multiple cells can generate one single 3D spatio-temporal surface, and we apply model-based constraints to extract the individual cell tracks. The advantages of our approach are that it automatically trains the cell-like distribution based on measures from the single-tracks of each dataset, it uses the training data to test both whether objects are cells and whether cells at different time points match, and it separates objects that are merged using the temporal segmentation information from other slices. To address the problem of high-throughput validation in the absence of ground truth, we also present a framework for edit-based validation. Our overall approach is outlined in Figure 1.

2. SPATIO-TEMPORAL CELL SEGMENTATION AND TRACKING

2.1. Cell Segmentation

To denoise the images and segment the cells, we use an algorithm based on the shift-invariant wavelet frames transformation of the image as well as the filtering of non-salient wavelet coefficients. Prior research on biomedical data [7, 8] demonstrates that the à trous (with holes) wavelet transform is robust

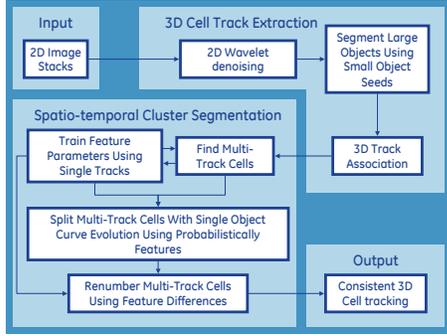


Fig. 1. Cell segmentation and tracking flowchart. The details are described in Section 2.

to local noise variations and discards low frequency objects in the background. The decomposition is represented as

$$I_i(x, y) = \sum_{m, n} h(m, n) I_{i-1}(x - 2^{i-1}m, y - 2^{i-1}n) \quad (1)$$

$$W_i(x, y) = I_i(x, y) - I_{i+1}(x, y) \quad (2)$$

where I_i and W_i represent the approximation and detail images, respectively, at each scale, i , and $h(m, n)$ denotes the scaling function.

Assuming that the image noise is additive, the corresponding wavelet transformation results in coefficients generated by the underlying signal W^I and those that correspond to image noise W^N . To approximate the signal term, we threshold the image stack with an Amplitude-scale-invariant Bayes Estimator (ABE) using Jeffreys' non-informative prior

$$W_i^I(x, y) \approx \delta^{ABE}(W_i(x, y)) = \frac{(W_i(x, y)^2 - 3\sigma_i^2)_+}{W_i(x, y)} \quad (3)$$

where σ_i^2 is the estimated noise variance at a given scale. In order to further reduce noise and enhance objects that extend across multiple resolutions, we compute a correlation stack $C_s(x, y)$, which is the multiplication of a subset of the denoised wavelet coefficients corresponding to the selected scales

$$C_s(x, y) = \prod_{i=j_l}^{j_s} W_i^I(x, y)_+ \quad (4)$$

The segmentation is then obtained as the extraction of connected regions in the correlation stack. Figure 2 shows an example segmentation of a low-dose Hoechst-stained section demonstrating the accuracy of the segmentation approach even with a very low contrast-to-noise ratio. The segmented spatio-temporal volume is then labeled in 3D. This results in a 3D set of segmented “tubes” corresponding to cells moving through time.

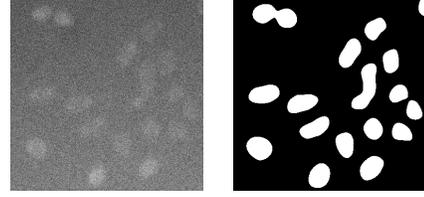


Fig. 2. Datasets with low contrast-to-noise. The figure on the left shows a cropped view of a typical low-dose Hoechst-stained image, and the figure on the right shows the segmentation using the wavelet approach. Contrast-to-noise (CNR) is defined as $\frac{\mu_S - \mu_N}{\sigma_N}$, where μ_S is the signal mean, μ_N is the noise mean, and σ_N is the noise standard deviation. For these images, $CNR = 0.5$, indicating that the average intensity difference between the signal and the noise is only half of the standard deviation of the noise. The intensity range is only 20 gray-scale levels.

2.2. Spatio-Temporal Cluster Segmentation

The 3D labeled “tubes” comprise two types of tracks: those corresponding to single cells (single-tracks) and those corresponding to multiple cells clustered together (multi-tracks). The goal of the spatio-temporal cluster segmentation step is to segment the multi-tracks into individual tracks by utilizing learned object descriptors. The distribution p of learned features that capture the appearance of cells f_i^t (where i is the cell number and t is the time slice) can be readily trained from the single-tracks. Alternatively, if few single-tracks are present in the dataset, these features can be trained using other datasets, assuming that the experimental conditions are kept constant.

Once the multi-tracks have been identified, the cell clusters are segmented by taking advantage of temporal context for determining the best way to split cells. The temporal information is utilized by propagating higher likelihood cell segmentations to slices with lower likelihood. We incorporate a statistical test to determine which segmented objects should be propagated across slices, and we use a curve evolution algorithm that finds the best segmentation such that it retains the relative area of the cell regions across slices.

For a given multi-track, the algorithm propagates the segmented objects from slices with greater number of objects to those with fewer. At the completion of the curve evolution, the number of objects on the adjacent frame will be the same as the reference frame, and at least one object will have been segmented. The features f_i^t of each of the split segments are then used to calculate the likelihood of the segment being cell-like using the trained distribution parameters μ and Σ

$$G(f_i^t) \propto \exp \left[-\frac{1}{2} (f_i^t - \mu)^T \Sigma^{-1} (f_i^t - \mu) \right] \quad (5)$$

Given the curve evolution, the likelihoods, and the fact that the reference frame has more objects than the adjacent

frame, we can determine whether to merge the labels on the reference frame or split the labels on the adjacent frame by comparing the likelihood of the full object to the average of its parts

$$L_{merge} = G(f_{i \in I}^t) - \frac{1}{|I|} \sum_{i \in I} G(f_i^t) \quad (6)$$

$$L_{split} = \frac{1}{|I|} \sum_{i \in I} G(f_i^{t+o}) - G(f_{i \in I}^{t+o}) \quad (7)$$

where o is an offset from the reference frame, and I is the set of labels under consideration for merging/splitting. If $L_{merge} > L_{split}$, the labels are merged on the reference frame; otherwise they are split on the adjacent frame. L_{merge} is high when the combined labels on the reference frame have a higher likelihood than the separate labels, and L_{split} is high when the separate labels on the adjacent frame have higher likelihood than when they are combined.

In this process, mitosis is a special case since a cell will be one object in one frame and two in the next. This case is handled implicitly by leaving the regions unchanged if both splitting the cell on the previous frame and merging the cells on the next frame lead to lower likelihoods. Here, the causal relationships are also taken into account since cells can split over time but not merge.

The step of renumbering the cells in the clusters to correctly associate cells across frames is accomplished using the parameters μ_d, Σ_d of a distribution trained from the single-tracks. The features f_j^{t+1} with $j = 1 : L$ of all cells on the next slice are compared to the features of each of the cells on the current slice f_i^t with $i = 1 : L$ using μ_d and Σ_d

$$G(d_{i,j}^t) \propto \exp \left[-\frac{1}{2} (d_{i,j}^t - \mu_d)^T \Sigma_d^{-1} (d_{i,j}^t - \mu_d) \right] \quad (8)$$

$$(9)$$

where $d_{i,j}^t = |f_i^t - f_j^{t+1}|$ is the absolute difference between the feature vectors. Using the assignment from the Hungarian algorithm that provides the optimal assignment from cells across frames, the cell labels are matched across the slices.

3. EDIT-BASED VALIDATION

Validation of high-throughput time-lapse cell segmentation and tracking algorithms is a challenging task since most cell assays have many cells and rapid acquisition times (leading to many time points) making manual validation time-consuming and tedious. We propose an edit-based validation framework that consists of the following elements: accepts a general input independent of the segmentation/tracking algorithm, draws the attention of the user to regions with probable errors, provides tools for correcting errors, and generates a

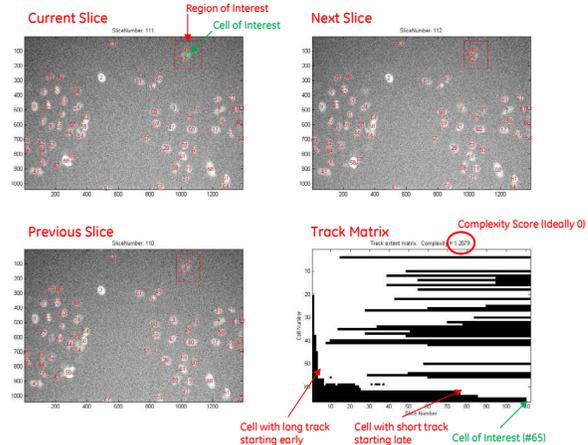


Fig. 3. Edit-based validation tool. The upper left image is the current image being edited, the lower left is the previous frame, and the upper right is the next frame. The lower right image is the track matrix that indicates the consistency of the tracks. The columns of the matrix correspond to different time points.

score for the algorithm based on the edits. Figure 3 shows a screenshot of the prototype system.

The system draws the attention of the user to probable errors in the automatic analysis in a multi-stage approach. In the first stage, it seeks to ensure continuity of cell tracks using a construct we call the “track matrix” (see lower right image in Figure 3). We characterize the consistency of the tracks as the average length of the run-length encoding of each row. As corrections are made to increase the consistency of the tracks, this score improves. In the second stage, the cell-likelihoods of each cell track are considered. The track having the lowest average cell-likelihood is calculated, and the lowest likelihood cell within this track is displayed. This targets tracks that consist of either clusters of cells or cells that were broken into smaller segments.

At each stage of the editing process, the system provides editing tools for the user to correct errors. These tools enable adding, removing, moving, and combining cells. The system also tracks all changes, which enables the tracking and undoing of changes as well as the ability to generate reports.

4. RESULTS

The cell tracking tool was applied to study the effect of the cell-cycle inhibitor compound Roscovitine. The biological expectation is that this inhibitor compound will arrest cells attempting to undergo mitosis. The algorithms were applied to four datasets, two with no treatment (control), and two treated with Roscovitine. Nuclei were stained with a relatively low concentration of the nuclear DNA stain HoechstTM33342 to

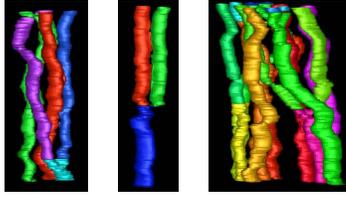


Fig. 4. Spatio-temporal trees. The figures show groups of touching cells, some of which undergo mitosis.

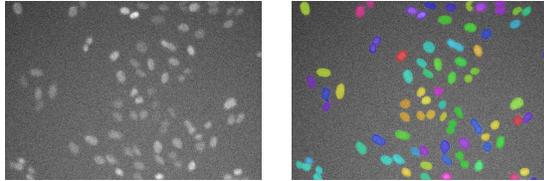


Fig. 5. Example segmentation results. An example cropped image slice is shown on the left, and the corresponding segmentation is shown on the right.

aid identification. Plates were imaged on a GE IN Cell Analyzer 1000 (10x objective) with environmental control (37°C, 5% CO₂). The images were taken every 15 minutes for 112 time points for a total duration of 28 hours.

Figure 4 shows several examples of spatio-temporal trees extracted by the algorithms. Although the cells move in clusters, touch one another, and undergo mitosis, the algorithms are able to successfully separate the clusters into individual tracks. These figures show that the algorithm can extract not only the tracks, but also the shape of the cells over time. Figure 5 shows an example segmentation result. Table 1 gives a numerical comparison of the number of mitotic events at the start and end of the four wells. We also calculated the distance from manually placed centroids to the centroids of the automatically segmented cells for every cell of one of the experiments. The mean distance was 5.5 μ m, and the standard deviation was 6.8 μ m; the average cell diameter in these experiments was approximately 35 μ m. The total number of manually identified cells was 6157, and 5758 were detected automatically, so that only 6% of the cells were under-segmented.

5. CONCLUSIONS AND FUTURE WORK

Automated analysis of high-throughput time-lapse data can provide statistically meaningful measures that are difficult to achieve by manual analysis. We have presented an automatic analysis approach that exploits the spatio-temporal nature of the data to constrain the segmentation and tracking problems. Our edit-based validation framework enables validation of large datasets with less effort than needed for manual ground truth generation.

Table 1. Mitotic cell count. The cell cycle inhibitor Roscovitine arrests cells as they attempt to undergo mitosis. Therefore, more cells will remain in the mitotic state at the end of the experiment. The number of cells in mitosis at the start and end of the 28-hour sequence for four wells is shown. While the number of mitotic cells is nearly the same for all wells at the start, more mitotic cells are present at the end for the Roscovitine inhibited wells.

	Control1	Control2	Inhibited1	Inhibited2
Mitotic start	0	0	0	1
Mitotic end	3	0	8	9

The topics for future work include measuring biologically relevant phenomena on more datasets and expanding the edit-based validation framework to incorporate a learning module that will make suggestions of changes based on past edits.

Acknowledgements: We would like to thank Elizabeth Roquemore and Angela Williams for working out the assay conditions and generating the image sets.

6. REFERENCES

- [1] O. Al-Kofahi, R.J. Radke, S.K. Goderie, Q. Shen, S. Temple, and B. Roysam, "Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells," *Cell Cycle*, vol. 5, pp. 327–35, 2006.
- [2] F. Yang, M. Mackey, F. Ianzini, G. Gallardo, and M. Sonka, "Cell segmentation, tracking, and mitosis detection using temporal context," in *MICCAI '05*, James Duncan and Guido Gerig, Eds., 2005, number 3749 in LNCS, pp. 302–309.
- [3] O. Debeir, P. Van Ham, R. Kiss, and C. Decaestecker, "Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes," *IEEE Trans. Med. Imaging*, vol. 24, no. 6, pp. 697–711, 2005.
- [4] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J.C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-d microscopy with coupled active surfaces," *IEEE Trans. Image Proc.*, vol. 14, no. 9, 2005.
- [5] K. Li, E. Miller, L. Weiss, P. Campbell, and T. Kanade, "Online tracking of migrating and proliferating cells imaged with phase-contrast microscopy," in *Proc. CVPRW*, 2006, pp. 65–72.
- [6] D. Padfield, J. Rittscher, T. Sebastian, N. Thomas, and B. Roysam, "Spatio-temporal cell cycle analysis using 3D level set segmentation of unstained nuclei in line scan confocal fluorescence images," in *IEEE ISBI*, 2006.
- [7] A. Genovesio, T. Liedl, V. Emiliani, W. J. Parak, M. Coppey-Moisan, and J.C. Olivo-Marin, "Multiple particle tracking in 3-d+t microscopy: method and application to the tracking of endocytosed quantum dots," *IEEE Trans. Image Proc.*, vol. 15, no. 5, pp. 1062–1070, 2006.
- [8] J.C. Olivo-Marin, "Automatic detection of spots in biological images by a wavelet-based selective filtering technique," in *ICIP*, 1996, pp. I: 311–314.