COMPUTER VISION TRACKING OF STEMNESS

Kang Li[1], Eric D. Miller[1], Mei Chen[2], Takeo Kanade[1], Lee E. Weiss[1], Phil G. Campbell[1]

[1]Carnegie Mellon University
[2]Intel Research Pittsburgh

ABSTRACT

Clinical translation of stem cell research promises to revolutionize medicine. Challenges remain toward better understanding of stem cell biology and cost-effective strategies for stem cell manufacturing. These challenges call for novel engineering toolsets to study stem cell behaviors and the associated *stemness*. Towards this goal, we are developing a computer vision based system to automatically and reliably follow the behaviors of individual stem cells in expanding populations. This paper reports on significant progress in our development. In particular, we present a machine-learning approach for detecting spatiotemporal mitosis events without image segmentation. This approach not only improves tracking performance, but can also independently quantify mitoses and cellular divisions. We also employ bilateral filtering to improve cell detection performance. We demonstrate the effectiveness of this system on tracking C2C12 mouse myoblast stem cells.

***Index Terms*—** Computer vision, tracking, time-lapsed microscopy, stem cells, stemness

1. INTRODUCTION

Clinical translation of stem cell research holds the promise to revolutionize medicine. Beyond blood stem cell transplantation, which has been an established clinical therapy for decades, other stem cell sources, including fetal and adult stem cells, are currently under investigation for use in a broad array of applications. These applications include the treatment of heart disease, Parkinson's disease, stroke, and extensive tissue regeneration from trauma and disease. However, significant challenges remain to be addressed to make fetal or adult stem cell based therapies effective, safe, and predictable, as well as to improve the established stem cell therapies. Two universal challenges are the need for a better understanding of basic stem cell biology, and the need for improved and cost-effective strategies to produce sufficient stem cell numbers *in vitro*, to meet both current and expected clinical demands.

Basic biological discovery experimentation requires new engineering toolsets to reliably and automatically and reliably follow the complete spatiotemporal histories of individual stem cells in population environments, as they divide into daughter cells. With this information, which includes cell *family trees* (or *lineage maps*), division times, motion trajectories, quiescence and death, investigators will be able to more efficiently study cell behaviors in response to varying culture conditions. Similarly, stem cell production or *manufacturing* processes, which are used to expand relatively small numbers of stem cells into the millions required for therapeutic delivery, also demand innovative toolsets to automatically and reliably monitor cell numbers *in vitro* in real-time. Predictive models can be created with this information, which can be used to both optimize cell culture expansion methodologies and to provide quality control during production.

To realize these much-needed toolsets, we are developing a computer vision based cell tracking system that can track each and every cell in a dynamic and expanding population imaged with phase-contrast time-lapsed microscopy. The goal is to be able to automatically measure and report the behavior of each cell in a population in real-time. These behaviors include cell movement, division, quiescence, death, and eventually differentiation. Collectively these behaviors reflect *stemness*, i.e., the self-renewal capability of stem cells. Achieving this goal requires overcoming challenges such as increasing cell densities during population expansions, cells leaving/entering the field of view, and discriminating between crossing/overlapping cells.

Continuing the footprints of our previous developments [1], [2], we present significant improvements to our modular tracking system (Fig. 1). These improvements address the unique challenges for tracking stem cells and quantifying the associated *stemness* metrics.
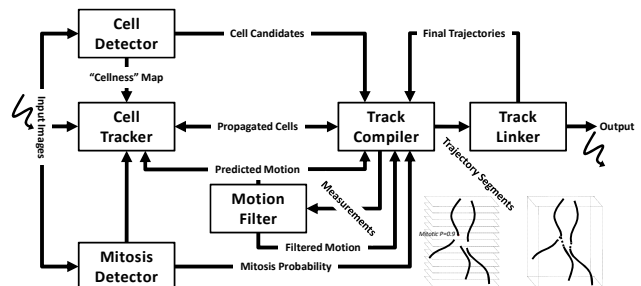

Fig. 1. Tracking system overview

One major improvement is the incorporation of a machine learning based mitosis event detector, which computes the probability of the occurrence of mitosis events at any spatiotemporal location in the image sequence. This information is useful for two purposes: 1) to compute an *energy* term in the level-set based cell tracker to improve the tracking performance of cell divisions; and 2) to inform the track compiler and linker of the locations where mitoses occur, thus assisting the correct establishment of mother-daughter relations among cells. This module improves the robustness of the tracking of cell divisions, which is critical for the reliable construction of cell lineages.

Another improvement goes to the cell detector, where we utilize bilateral filtering to reduce noise, as well as to smooth the filopodia of C2C12 cells that may interfere with detection. We also employ an edge-based approach to detect mitotic cells near interphase or prophase.

The following sections elaborate on these improvements, and demonstrate the effectiveness on tracking and analyzing the lineage of C2C12 mouse myoblast stem cells. We refer the readers to [1], [2] for an overview of our system and details on the other modules.

2. TRACKING SYSTEM

2.1. Spatiotemporal Mitosis Event Detection

While the appearances of stem cells could be complex and time-varying, cells that undergo mitosis typically exhibit a series of distinctive and highly-regulated changes, which are observable using phase-contrast microscopy. These changes include an initial increase of brightness, an increase of circularity, and a decrease of size. These are followed by a reverse process and eventually cytokinesis occurs, whereby a cell completes its division into two daughter cells.

These phenomena could be captured by classifying several features extracted from the segmented cell regions [1], [2], [7]. However, the robustness of such approaches is both dependent on and limited by the accuracy of segmentation, which may vary among different cell types. To improve robustness, we propose a segmentation-free approach for detecting mitosis using machine learning. We consider mitoses as spatiotemporal event patterns, and aim to detect them in any spatiotemporal volumes of appropriate scales in the image sequence. For this purpose, we trained an efficient cascade classifier using collected examples.

2.1.1. Training Examples

To collect positive examples, we manually identified 295 mitosis events in one image sequence. We extracted spatiotemporal sub-volumes of 24×24×5 pixels centered near *telophase*, where the mitotic cell forms a distinctive

Fig. 2. Volumetric Haar-like filters used to extract features for mitosis detection. The features are computed by subtracting the sum of pixel values in the black boxes from the sum of pixel values in the white boxes.

figure-eight structure in most cases. Volumes of this size are sufficient to capture salient features of mitosis events. We create negative examples automatically, both *offline* and *online*, from the training sequences excluding the identified sub-volumes containing mitoses. We acquire an initial set of 1,000,000 negative examples offline. Among these, 1,380 examples were extracted at the cell centroids detected using the cell detector. We rotate these examples at 90°, 180° and 270° along the time axis to produce 4,140 additional examples. The remaining 994,480 examples were generated by randomly sampling the spatiotemporal sub-volumes from the training sequences. During training iterations, we can bootstrap supplementary negative examples online by running the already-trained partial classifier on the training sequence and extracting all false detections.

During mitosis, a cell may divide at any orientation. To avoid the curse of dimensionality, it is desirable to align the positive examples to reduce data complexity. However, it is both undesirable and unrealistic to align the examples perfectly, as it would require us to train classifiers for all specific orientations. To avoid artifacts induced by image registration, we utilize the following approach. First, we binarize the middle frame of each positive example using Otsu thresholding. Then, we classify each example into one of two categories, [0, 90°) or [90°, 180°), according to the orientation of the major axis of the connected component at the center. We flip the examples in the second category horizontally, such that all examples have orientations in the range of [0, 90°). We double the size of the positive training set with 180°-rotated copies of the examples. This procedure avoids interpolation and preserves the original pixel intensities. The resulting examples cover 50% of the space of possible orientations. Due to the rotational symmetry of the Haar-like filters, once a classifier is trained on the aligned examples, the classifier that covers the other orientations can be derived automatically.

2.1.2. Features

Inspired by the work of Ke *et al* [3], we draw on the concept of *integral volumes*. An integral volume V at pixel location (x, y) and time t is evaluated as the sum of pixels at coordinates less than or equal to (x, y, t):

$$V(x, y, t) = \sum_{x' \leq x, y' \leq y, t' \leq t} I(x', y', t'),$$
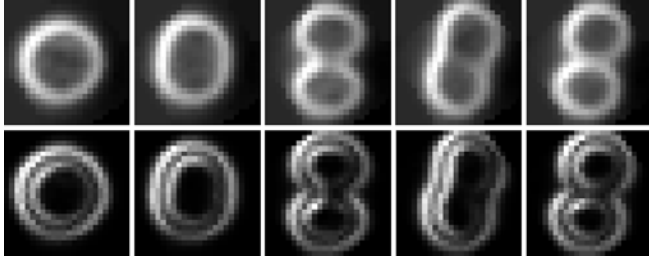
Fig. 3. A training example for mitosis event detection at the canonical orientation. The top row shows the original image frames; the bottom row shows the local variance maps.

where $I(x, y, t)$ is a pixel in the original image at time t. The integral volume representation allows a set of volumetric Haar-like filters (Fig. 2) to be evaluated efficiently using array lookups and subtractions. By applying these filters to a training template at various scales and aspect ratios, we obtain an over-complete set of 80,364 candidate features.

Ke *et al* showed that computing features using the original pixel intensities led to poor performance, possibly due to shifts of intensity levels across frames. Instead, they relied on the optical flow of the training images to extract features. However, optical flow can be noisy and unreliable on highly-deformable objects such as cells. Therefore, we choose to use the local variance in the original image, which is invariant against shifts in intensity:

$$Var(x, y) = \frac{1}{|N|} \sum_{(x,y) \in N} [I(x, y) - \mu]^2,$$

$$\mu = \frac{1}{|N|} \sum_{(x,y) \in N} I(x, y),$$

where N represents a local neighborhood of size $|N|$. The local variance measures the contrast of local image textures, which captures the appearance and morphological changes of a cell during mitosis (Fig. 3).

2.1.3. Training the Classifier

Classifiers such as neural networks and the support vector machine are computationally expensive. Viola and Jones proposed the strategy of coarse-to-fine classification using a cascade classifier [5], which was proven to be highly efficient and effective. The cascade classifier consists of a sequence of AdaBoost *ensemble* classifiers with increasing complexity. This allows the majority of negative samples to be rejected quickly at early cascade nodes, thus achieving efficient classification.

However, the AdaBoost classifier suffers from a lengthy training time due to its tight coupling of two processes: feature selection and ensemble classifier formation. To overcome this drawback, Wu *et al* [6] proposed a fast

cascade learning framework, which decouples them into the forward feature selection (FFS) algorithm, and the linear asymmetric classifier (LAC). This approach reduces the training time by two orders of magnitude, enabling us to train the classifier in a reasonable amount of time despite the large number of examples and candidate features.

2.1.4. Mitosis Detection

We conduct mitosis detection during tracking at the same scale as the training examples. This is feasible because the image acquisition was performed with identical protocol, and because the integral volume can be computed sequentially (frame-by-frame) using simple recursions. Specifically, we perform detection by classifying all $24 \times 24 \times 5$ sub-volumes as either mitosis or non-mitosis in a sliding temporal window of 5 frames. We convolve the classification result with a 3-D Gaussian kernel with diagonal covariance, and scale the result to the range of $[0, 1]$ to represent the probability of mitosis events. As cell tracking proceeds, we load the next frame while dropping the oldest frame in the temporal queue, and update the integral volume accordingly. The numerical precision of the integral volume degrades gradually with sequential updates. Therefore, we recompute the integral volume from scratch periodically to maintain its precision.

2.2. Bilateral Filtering and Cell Detection

Now we detail the improved cell detector with bilateral filtering. The bilateral filter is a nonlinear filter that smoothes a signal while preserving strong edges. It is a *noniterative* alternative to anisotropic diffusion. The filter output $J(\mathbf{x})$ at each pixel is a weighted average of its neighbors. The weight of each neighbor decreases with its distance to the center pixel in space and intensity:

$$J(\mathbf{x}) = \frac{1}{S(\mathbf{x})} \sum_{\mathbf{x}' \in N} G_s(||\mathbf{x} - \mathbf{x}'||) G_r(|I(\mathbf{x}) - I(\mathbf{x}')|) I(\mathbf{x}'),$$

$$S(\mathbf{x}) = \sum_{\mathbf{x}' \in N} G_s(||\mathbf{x} - \mathbf{x}'||) G_r(|I(\mathbf{x}) - I(\mathbf{x}')|),$$

where $\mathbf{x} \equiv (x, y)$; G_s, G_r are Gaussian functions with standard deviations s and r, respectively. They control how much a neighbor is weighted due to spatial distance and intensity difference. Brute-force computation of the bilateral filter is expensive. We adopt the fast algorithm proposed by Paris and Durand [4], which achieves acceleration by downsampling and utilizing linear convolution in the joint space of space and intensity.

After bilateral-filtering, we apply the *rolling-ball* filter [2] to extract non-mitotic cells, which appear as darker regions surrounded by brighter halos. The rolling-ball filter is a grayscale morphological filter that is related to the top-hat filter by $I - \text{rollball}_h(I) \equiv \text{tophat}(I, ball_h)$, where h is the

radius of the rolling ball, and $ball_h$ is a non-flat (half-)ball-shaped *structuring element* with radius h. The parameter h is roughly equal to the average radius of cells to be detected.

Mitotic cells near interphase or prophase appear as bright, round objects. We employ an edge-based procedure to detect them, in three steps: 1) apply Canny edge detector to extract edges; 2) find and fill regions that are surrounded by closed edges; and 3) select the regions with solidity greater than 0.9, eccentricity smaller than 0.9, mean intensity greater than $\mu + 2\sigma$, and size within a specified range as mitotic cell regions. Here, μ and σ are the mean and standard deviation of the image, respectively.

3. CELL CULTURE AND IMAGING

C2C12 myogenic precursor cells (*American Type Culture Collection, Manassas, VA*) were observed with a Zeiss Axiovert 135TV inverted microscope using a 5X, 0.15 N.A. objective with phase optics (*Carl Zeiss, Inc., Thornwood, NY*). Images were acquired every 5 minutes for 92.5 hours using a 12-bit Qimaging Retiga EXi Fast 1394 CCD camera, at $500ms$ exposure with a gain of 1.01. Each image consists of 1392×1040 pixels with a resolution of 19 μm/pixel.

4. RESULTS

We tested the proposed mitosis event detector as a standalone tool on two image sequences independent of the training sequence. Fig. 4 shows the mitosis probability output of the detector overlaid on one of the sequences. The detector achieved good detection results in these sequences as per visual inspection. For quantitative assessment, we consider regions with detected mitosis probabilities greater than 0.5 as positive mitoses. Comparing to manual detection, the automated detector achieved precision and recall rates of 90.4% and 94.1%, respectively. We also evaluated the cell detection accuracy. We found that bilateral filtering (with parameters $s = 4.0$ and $r = 0.1$) reduces false positives by 11.3%, and also improves cell localization accuracy.

We applied the improved tracking system to track the C2C12 stem cells in the above two sequences, and constructed the corresponding lineage maps (Fig. 5). The
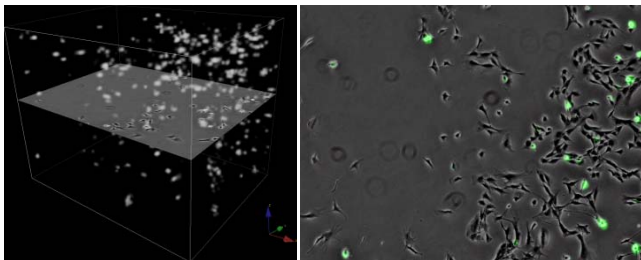

Fig. 4. Result of spatiotemporal mitosis event detection for an entire sequence (*left*), and in one frame (*right*).
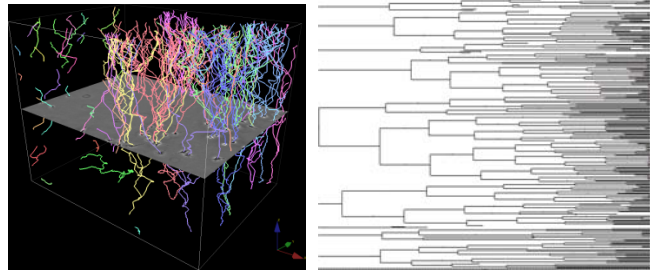

Fig. 5. Tracking result and the associated lineage map of C2C12 cells. *Left)* Spatiotemporal cell trajectories. Different colors represent different family trees. *Right)* Lineages of the cells that appear in the first frame and their descendants.

lineage map contains important metrics such as *symmetry* and *division time*. These metrics are useful for the predictive modeling of stem cell population expansions, as well as the design and optimization of adaptive subculturing strategies.

5. CONCLUSION

We described an improved cell tracking system for tracking a large number of cells in expanding populations imaged with phase-contrast microscopy. The improvements include a novel machine-learning approach to detect mitosis events without segmentation, and a more reliable cell detector with bilateral filtering. The system and its modules provide useful toolsets to study stem cell behaviors, as well as to other areas such as tissue engineering and drug discovery.

6. REFERENCES

[1] K. Li, M. Chen, and T. Kanade, "Cell population tracking and lineage construction with spatiotemporal context," In *Proc. 10th Int. Conf. Med. Image Computing and Comp.-Assisted Intervention.* Springer, Oct. 2007, pp. 295-302.

[2] K. Li, E. D. Miller, L. E. Weiss, P. G. Campbell, and T. Kanade, "Online tracking of migrating and proliferating cells imaged with phase contrast microscopy," In *Proc. IEEE Conf. Comp. Vision and Patt. Recog. Workshop*, Jun. 2006, p. 65.

[3] Y. Ke, R. Sukthankar and M. Hebert, "Efficient visual event detection using volumetric features," In Proc. *IEEE Int. Conf. Comp. Vision*, Oct. 2005, pp. 166-173.

[4] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," In *Proc. European Conf. Comp. Vision*, Jul. 2006, pp. 568-580.

[5] P. Viola and M. Jones. "Rapid object detection using a boosted cascade of simple features", In *Proc. IEEE Conf. Comp. Vision and Patt. Recog.*, Dec. 2001, pp. 511-518.

[6] J. Wu, S. C. Brubaker, M. D. Mullin, and J. M. Rehg, "Fast asymmetric learning for cascade face detection," *IEEE Trans. Pattern Anal. Machine Intell.*, May 2007, in press.

[7] F. Yang, M. A. Mackey, F. Ianzini, G. Gallardo, and M. Sonka, "Cell segmentation, tracking, and mitosis detection using temporal context," In *Proc. Med. Image Computing and Comp.-Assisted Intervention*, vol. I, Oct. 2005, pp. 302–309.