# Tracking of Cell Populations to Understand their Spatio-Temporal Behavior in Response to Physical Stimuli

David House[1], Matthew L. Walker[2], Zheng Wu[1], Joyce Y. Wong[2] and Margrit Betke[1]

Departments of Computer Science[1] and Biomedical Engineering[2]

Boston University

{house, mwalk79, wuzheng, jywong, betke }@bu.edu

## Abstract

*We have developed methods for segmentation and tracking of cells in time-lapse phase-contrast microscopy images. Our multi-object Bayesian algorithm detects and tracks large numbers of cells in presence of clutter and identifies cell division. To solve the data association problem, the assignment of current measurements to cell tracks, we tested various cost functions with both an optimal and a fast, suboptimal assignment algorithm. We also propose metrics to quantify cell migration properties, such as motility and directional persistence, and compared our findings of cell migration with the standard random walk model. We measured how cell populations respond to the physical stimuli presented in the environment, for example, the stiffness property of the substrate. Our analysis of hundreds of spatio-temporal cell trajectories revealed significant differences in the behavioral response of fibroblast cells to changes in hydrogel conditions.*

## 1. Introduction

Hundreds of thousands of living cells are recorded in time-lapse phase-contrast microscopy video for research studies in biomaterial engineering [5]. Interpreting these vast amounts of data via manual analysis is time-consuming, costly, and prone to human error. In this paper, we propose image understanding methods that automatically analyze the behavior of live cells as extracted from time-lapse phase-contrast microscopy video. For this analysis, we computed the shape, orientation, and movement characteristics of large groups of cells, as we tracked them automatically over the course of a day. We used two approaches to tracking and compared their performance. One approach solves the two-dimensional data assignment problem, i.e., the assignment of current measurements to cell tracks, optimally in a probabilistic, iterative, and online manner. The other approach is a two-phase batch algorithm that solves the data association problem in an efficient, but suboptimal way. We also introduce cost functions that evaluate the likelihood of the set of assignment.

Our segmentation and tracking algorithms yielded quantitative data about the characteristics of cells and their interactions and enabled us to develop new descriptions of cell behavior. Our system allowed us to reason about cell motility, i.e., the cell's ability to move spontaneously and actively. The motion of cells, i.e., cell migration, has been described in the biomaterials research community by a random walk model [6, 1]. We show that cell behavior is not represented well by this theoretical random-walk model and propose a data-driven approach to quantifying cell migration.

Since understanding the directed cell migration, i.e., the directed motion of cells in response to particular stimuli is of great interest to biomaterial research, we focused on analyzing the motion of fibroblast cells, i.e., cells that contribute to the formation of connective tissue fibers, on hydrogel substrates with different stiffness properties that mimic the physiological environment. We note that the few commercial tools for analysis of living cell images require that fluorescent cell tags are used (that can change cell behavior, which would be undesirable) or substrates with glass coverslips. To our knowledge, this paper is the first to propose an accurate, reproducible, and automated method that can measure both durokinetic and durotaxic effects [17, 22, 23] that are due to the stiffness properties of the substrate. In particular, our method quantifies statistically significant differences in cell behavior in response to the level of stiffness of uniformly rigid substrates, as well as cell movement along a rigidity gradient present in the substrate.

In summary, our original contributions are

- **Cell Tracking Algorithms:** We propose a novel cost function that encodes the response of cells to conditions of their environment.

- **Cell Behavior Analysis:** We introduce a comprehensive set of metrics for migration analysis.

- **Model Building:** We reveal the shortcomings of a widely-used theoretical model of cell migration and propose an alternative, data-driven way to predict cell behavior in response to environmental conditions.

Recently, there have been significant efforts by the computer vision community to develop methods for tracking cells in microscopy video, e.g., [10, 11, 12, 13, 14, 15, 16, 18, 21, 25]. We would like to highlight the seminal work by Li et al. [13, 14, 15, 16], which resulted in a multi-target tracking system that used level-set contour tracking, interactive multiple-model filters, and trajectory management modules to compile and link tracks of hundreds of osteosarcoma cells and thousands of amnion epithelial stem cells in time-lapse fluorescence microscopy. Notable is also the cell tracking system by Smith et al. [21], which employs a Markov chain Monte Carlo batch processing approach to track a hundred migrating neurons in two-photon excitation microscopy. Xie et al. [24] recently solved the problem of tracking *Escherichia coli* bacteria in microscopy video using a greedy assignment algorithm [20]. They introduced a matching criteria that compares intensity histograms of newly detected and previously tracked cells to address the issue that the bacteria were imaged with very-low contrast boundaries.

Unlike previous work, we have focused on developing measures that can evaluate the activity of cells and quantify their behaviors. To accomplish this, we developed accurate methods for segmenting and tracking living cells and then computed the shape characteristics and motion trajectories of thousands of sample cells.

## 2. Methods

Our approach consists of two phases. In the first phase, our system detects and tracks cells in microscopy video and outputs shape characteristics and trajectories of groups of cells. In the second phase, our system reasons about the behavior of individual and groups of cells (Fig. 1). The measured statistics of cell behavior can provide feedback to the tracking system, allowing adaptation of the techniques for detection, segmentation, state prediction, and data association to the expected migration properties of cells.

### 2.1. Detection and Segmentation of Cells

Accurate segmentation of fibroblast cells on hydrogel substrate is challenging, because heterogeneities in the substrate exist and the background of the cells cannot simply be subtracted. The undesirable inconsistencies in image intensities can be alleviated by a contrast adjustment procedure that re-distributes the intensity values the middle 1/3 of the dynamic range of each image. Our segmentation method then computes the intensity gradient of the adjusted image and applies adaptive thresholding of the gradient magnitude
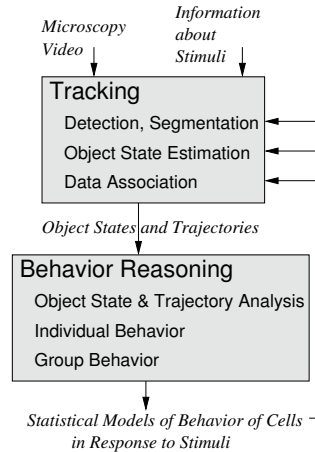


Figure 1. System overview.

image to produce a binary image in which cells and clutter are segmented from the background. In this binary image, small groups of pixels are classified as clutter ($< 300$ pixels) and disregarded in further processing. Our method classifies the larger components as cells and produces an accurate segmentation by closing gaps in their boundaries and holes within the binary components (Fig. 2). We use standard techniques [9] to describe characteristics of cells, such as size, shape, and orientation. Given the binary representation of a cell, our method computes its area $A$, centroid $\vec{x} = (x_x, x_y)$, and axes of least and most inertia, $a_{\min}$ and $a_{\max}$. Computing the sum $E_{\min}$ of squared distances of cell pixels to $a_{\min}$ and the sum $E_{\max}$ of squared distances of cell pixels to $\vec{a}_{\max}$ yields a measure for the circularity $c = E_{\min}/E_{\max}$ of the cell.

### 2.2. Tracking of Cells

In this section, we first formulate of our multi-cell tracking problem as a standard Bayesian filtering problem [3], then briefly describe the algorithms [12, 25] we used, and then focus on our contribution, the selection of a suitable cost function for our cell tracking problem.

Given that $n(t)$ cells are in image $I(x, y, t)$, state $x_j^{(t)}$ of the $j$th cell can be assumed to evolve in time according to the equations

$$x_j(t+1) = A_j x_j(t) + v_j(t), \text{ for } j = 1, ..., n_t, \text{ (1)}$$

as observed via $m_t$ measurements

$$z_i(t) = H_i\, x_j(t) + w_i(t), \text{ for } i = 1, ..., m_t, \quad (2)$$

where $v_j(t)$ and $w_i(t)$ are independent zero-mean Gaussian noise processes, $A_j$ is the state transition matrix, and $H$ the measurement matrix. For each cell, the true-positive detection rate is $P_{\det} \leq 1$ and the false-positive detection rate is $1/V$, where $V$ is proportional to the number of pixels in the image (e.g., 3%, which corresponds to 3 false detections per
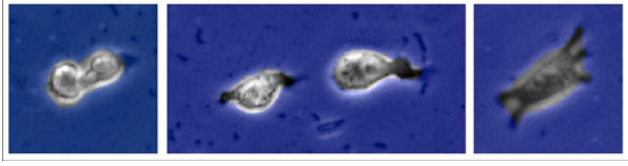
Figure 2. Segmented cells with discarded background region shown in blue. From left to right: a dividing cell, two cells in close proximity, a single spreading cell.

frame). We add a "dummy" measurement $z_0(t)$ to handle the case of missed detections. In particular, when cell $x(t)$ is not detected at time $t$, dummy measurement $z_0(t)$ is associated with the hidden state of cell $x(t)$. The likelihood that one of the measurements $z_i(t)$ describes object state $x_j(t)$ is given as $p(z_i(t)|x_j(t))$, and the corresponding assignment is denoted by the binary variable $a_{i,j}(t)$. A complete assignment of all measurements to states is the set $\mathbf{a}(t) = \{a_{i,j}(t)|i = 0, ..., m_t, j = 0, ..., n_t\}$. The optimal solution to the two-dimensional data assignment problem, i.e., the assignment of all current measurements to all cell tracks, is the complete assignment that maximizes the likelihood ratio:

$$\mathbf{a}_{\mathrm{opt}}(t) = \arg \max_{\mathbf{a}(t)} \prod_{i=0}^{m_t} \prod_{j=0}^{n_t} \left( \frac{p(z_i(t)|x_j(t))}{p(z_i(t)|x_0(t))} \right)^{a_{i,j}(t)}. \quad (3)$$

We use two algorithms to solve the assignment problem in Eq. 3: (1) an iterative probabilistic data association (PDA) algorithm [12] in combination with the auction algorithm [4], which produces an optimal solution, and (2) a two-phase batch algorithm [25], which produces a suboptimal assignment. The iterative PDA algorithm is a modification of the original PDA algorithm [3], which can only handle one-track-to-one-measurement assignments. In an iterative process, the algorithm "redistributes" association probabilities based on the likelihood of a match and the iteration number, allowing many-to-one and one-to-many associations. The batch algorithm is based on the nearest neighbor approach which, in the first phase of the algorithm, greedily assigns to each track its "favorite" current measurement, i.e., the measurement that is statistically closest to the predicted one. In the second phase, the batch algorithm matches tracks and measurements that have not been assigned in the first phase, including the dummy measurement.

The performance of both algorithms greatly depends on the selection of a cost function. We define the cost $c(t)$ of an assignment $\mathbf{a}$ at time $t$ to be the negative log-likelihood ratio:

$$c(t) = \sum_{i=0}^{m_t} \sum_{j=0}^{n_t} \left( a_{i,j}(t)(-\ln \frac{p(z_i(t)|x_j(t))}{p(z_i(t)|x_0(t))} \right), \quad (4)$$

which simplifies to three cases: If $j = 0$, i.e., no cells have been tracked, the cost is zero. If $m = 0$, the cost of a missed detection of a cell is $-\ln(1 - \mathrm{P}_{\mathrm{det}})$. If $n > 0$ and $m > 0$, the cost depends on the negative log-likelihood that previously tracked cells are again detected, which is the product of the true-positive detection rate $\mathrm{P}_{\mathrm{det}}$, the inverse false-positive detection rate $V$, and the tracker-computed likelihood $c_{i,j}(t)$ of the assignment $a_{i,j}$.

We propose three cost functions $c_{i,j}(t)$ for our cell tracking problem:

$$c_{i,j}^{(1)}(t) = \frac{S_{i,j,\mathrm{cos}}(t)}{C_{i,j,\mathrm{dist}}(t)}, \quad (5)$$

$$c_{i,j}^{(2)}(t) = \lambda \, C_{i,j,\mathrm{dist}}(t) + (1 - \lambda) \, C_{i,j,\mathrm{dir}}(t), \quad (6)$$

$$c_{i,j}^{(3)}(t) = \lambda(\xi) \, C_{i,j,\mathrm{dist}}(t) + (1 - \lambda(\xi)) \, C_{i,j,\mathrm{dir}}(t), \quad (7)$$

where $S_{i,j,\mathrm{cos}}(t)$ is the cosine similarity, i.e., the dot product, of previously-estimated and currently-measured centroids of a cell [25], $C_{i,j,\mathrm{dist}}(t)$ is the Euclidean distance between the centroids, and $C_{i,j,\mathrm{dir}}(t)$ is the movement direction computed from two consecutive frames. We designed cost function $c^{(2)}$ so that it depends on a fixed regularization parameter $\lambda$, which we set to 0.6 in our experiments. Similarly, we designed cost function $c^{(3)}$ so that it depends on a regularization parameter that is a function of the level $\xi$ of physical stimuli presented in the environment. For example, we used $\lambda(\xi) = 0.1$ for a substrate with a high level $\xi=150$ kPa of stiffness.

## 2.3. Characterization of Cell Behavior

To analyze cell migration behavior of both individual cells and groups of cells, we propose to use the set of measures of cell morphology and spatio-temporal trajectories, and group behavior defined in Table 1. We use the measures of cell orientation and area overlap to determine if a cell is motile or immotile. We compare measures of cell movements, in particular, displacement, path length, speed, and changes and persistence in movement direction, to describe cell migration. We are particularly interested in analyzing migration behavior on uniform substrates with varying stiffness and gradients in stiffness to quantify durokinesis and durotaxis [22, 8]. Durokinesis is defined as the dependence of individual movement on a scalar stimulus (in this case, substrate stiffness), and durotaxis is defined as the dependence of individual movement on a directional stimulus or signal related to the movement direction. There are a number of theories that have been proposed to describe cell migration. By being able to obtain actual measurements of cell movement, we can evaluate how realistic the random cell walk model [6] is and develop a data-driven approach to describe cell migration.

The expected squared displacement of a motile cell for a time period $t$ has been defined by the following random

walk model:

$$E[d_{0 \to t}^2] = 2\, S_t^2\, T_p\, (t - T_p(1 - e^{-t/T_p})), \qquad (8)$$

where persistence time $T_p$ is the period during which the cell moved consistently in one direction and $S_t$ is the average speed of the cell on the path $\mathcal{P}_{0 \to t}$. DiMilla et al. [5] used the model to estimate $T_P$ and $S_t$ of a cell by fitting Eq. 8 to experimental data, i.e., average measurements of cell displacement. For sake of comparison, we use our definitions for $T_p$ and $S_t$ from Table 1 to evaluate the validity of Eq. 8 for our data. Note that we define directional persistence as the longest period $T_p$ during which a migrating cell does not change its direction by more than $\beta$ degrees (we set $\beta = 15^o$ in our experiments).

To analyze cell migration behavior of groups of cells that are placed in close proximity to each other on the substrate, we propose to use the measure of mean dispersion distance $D_t$, which is defined by the mean distance between each member of the group and the centroid of the group at time $t$.

Table 1. **Quantitative Measures of Cell Behavior**

| Measure | Definition |
|---|---|
| Instantaneous Measures of Cell Behavior at Time $t$ | |
| Location | $\vec{x}(t) = (x_x(t), x_y(t))$ |
| Orientation | $\alpha(t) = \mathrm{atan}(a_{\min,y}(t)/a_{\min,x}(t))$ |
| Circularity | $c(t) = E_{\min}(t)/E_{\max}(t)$ |
| Velocity | $\vec{v}(t) = \frac{d\vec{x}(t)}{dt} \approx \vec{x}(t) - \vec{x}(t-1)$ |
| Speed | $S(t) = |\vec{v}(t)| = \sqrt{v_x^2 + v_y^2}$ |
| Movement Direction | $\theta(t) = \arctan(v_y(t)/v_x(t))$ |
| Change in Direction | $\dot{\theta} = d\theta(t)/dt \approx \theta(t) - \theta(t-1)$ |
| Measures Computed for Cell Path $\mathcal{P}_{0 \to T}$ | |
| Centroid Displacement | $d_{0 \to T} = |\vec{x}(T) - \vec{x}(0)|$ |
| Area Overlap | $A_{t \to t'} = A_t \cap A_t'$ |
| Path Length | $l_T = \sum_{t=0}^{T-1} d_{t \to t+1}$ |
| Average Speed | $S_T = 1/t \sum_{t=1}^{T} s(t)$ |
| Mean Direction of Movement | $\bar{\theta}_T = \mathrm{acos}(X/r) = \mathrm{asin}(Y/r)$, where $X = 1/T \sum_{t=1}^{T} \cos\theta(t)$, $Y = 1/T \sum_{t=1}^{T} \sin\theta(t)$, and $r = \sqrt{X^2 + Y^2}$ |
| Angular Deviation of Movement Direction | $\sigma_\theta = \sqrt{2(1-r)}$ |
| Period of Time with Directional Persistence | $T_p = \arg\max\{T' \mid |\dot{\theta}(t'+i)| < \beta,$ for some $t' \in \{0,...,t{-}1\}$, and all $i = 1,...,T'$, where $T' \leq T\}$ |
| Measures Computed for Group of Cells $x_1, ..., x_N$ | |
| Group Position at $t$ | $\vec{g}(t) = 1/N \sum_{i=1}^{N} \vec{x}_i(t)$ |
| Mean Dispersion Distance | $D_t = 1/N \sum_{i=1}^{N} \|\vec{g}(t) - \vec{g}(0)\|$ |

# 3. Experiments and Results

Our data includes 75 image sequences of approximately 1,000 Balb/c fibroblast cells on polyacrylamide hydrogel substrates with varying mechanical properties. The living cells were seeded on the substrates at a density of 1,000 cells/cm$^2$ [7, 19]. The image sequences were acquired with a Princeton Instruments D1299421 camera mounted on a Zeiss Axiovert S100 microscope at 15-minute intervals over the course of 24 hours. Each sequence consists of 82 frames with an image dimension of $1300 \times 1030$ pixels. Each pixel has a width of 1.52 $\mu$m. There are approximately 15–30 cells per image.

## 3.1. Detection and Segmentation Accuracy

We verified by inspection of 75 processed image sequences of about 1,000 cells that our detection method recognized every cell and did not misidentify any clutter as a cell. This means our true-positive detection rate was $P_{\det}$=1 and our false-positive detection rate approached zero.

To evaluate the accuracy of our segmentation method, we compared the computed centroids of migrating cells to the centers of these cells as hand-marked by an expert observer. The differences between computed and hand-marked cell positions were at most 15 pixels for 12 cells in 10 sequences ($12{\times}82 = 984$ samples). The cell sizes range from 400 to 700 pixels.

We also compared the estimates of the centers of 7 cells in 5 sequences ($7{\times}82$=574 samples) that were hand-marked by three independent observers: one expert, one skilled and one novice cell biologist. We found that inter-observer differences in marking cell centers was large, i.e., 15–25 pixels. Because the differences between observer estimates were larger than the upper bound on the difference between the position estimate derived from the segmentation algorithm and the expert observer estimate, we conclude that an estimate of cell position by the segmentation algorithm was at least as reliable as an estimate by an expert observer.

## 3.2. Tracking Accuracy

The rapidly changing state of migrating cells (alive/mitotic/dead) makes the data association task, i.e., the task of matching measurements to tracks, challenging. The success of a tracking approach depends on its ability to accurately update and link tracks as migrating cells divide, die, or move into and out of the field of view of the camera. We show examples of tracked cell paths in Figs. 3 and 6.

The optimal tracking algorithm performed better than the suboptimal favorite-matching algorithm (Table 2). Averaging the results for the optimal algorithm across population densities and environmental conditions, we obtained a rate of correct measurement-to-track assignments of 92.8% when implemented with cost function $c^{(2)}$ and 94.1% when implemented with cost function $c^{(3)}$. On the other hand, averaging the results for the suboptimal algorithm with cost function $c^{(1)}$ across population densities and environmental
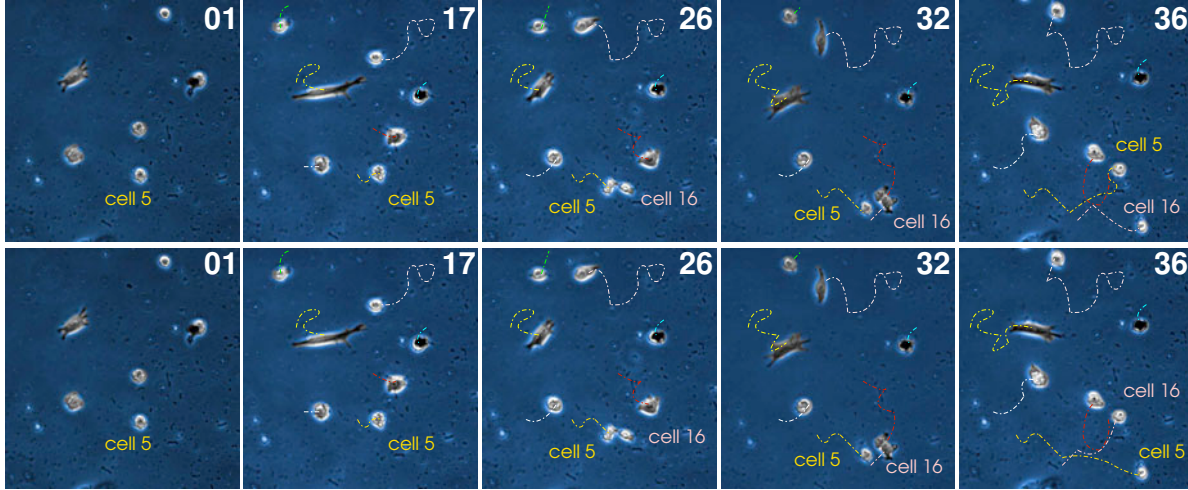
Figure 3. Tracking results for the suboptimal (top) and optimal algorithm (bottom). The optimal algorithm correctly resolves the track-to-measurement assignments for cell 5 and 16 in frame 36. The suboptimal algorithm causes a track switch.

conditions yields a correct assignment rate of only 73.6%. The suboptimal algorithm works well for low-density sequences with cells that do not occlude each other, or divide or die in close proximity to one another (100 pixels), achieving a 95.2% rate of accurate data associations in this case.

Table 2. **Accuracy of Data Association.** The rates of correct measurement-to-track assignments for 180 cells in 12 sequences (i.e., 14,760 samples) are given for low- and high-density cell populations and environmental conditions that included two levels of substrate stiffness.

| Method and | 500 cells/cm$^2$ | | 1,000 cells/cm$^2$ | |
| --- | --- | --- | --- | --- |
| Cost Function | 10 kPa | 150 kPa | 10 kPa | 150 kPa |
| Subopt. Alg. $c^{(1)}$ | 94.2 % | 96.1 % | 47.7 % | 56.2 % |
| Optimal Alg. $c^{(2)}$ | 100 % | 98.8 % | 85.7 % | 86.6 % |
| Optimal Alg. $c^{(3)}$ | 100 % | 99.4 % | 86.8 % | 90.3% |

For the optimal algorithm implemented with cost function $c^{(3)}$, the incorrect associations were due to (1) "track switching," where the measurements of positions of cells $a$ and $b$ were associated to the tracks of cells $b$ and $a$ (10.2% of incorrect associations), (2) "false track initiations/terminations," where the measurement of a previously tracked cell was matched to a new track, ending the previous track (35.9% of incorrect associations), and (3) "false lineage identification" (53.9% of incorrect associations), where cell division was misinterpreted, i.e., daughter cells were not matched up correctly with their parent cell.

For both tracking algorithms, track switching becomes more frequent as cell density increases. For sequences with a seeding density of 1,000 cells/cm$^2$, the ability to perform complex associations drops to 51.8% for the suboptimal approach with cost function $c^{(1)}$ and 88.0% for the optimal approach with cost function $c^{(3)}$.

## 3.3. Results for Cell Behavior

We analyzed the behavior of cells migrating over a 24 hour period on both stiff (150 kPa) and soft (10 kPa) substrates (Table 3). The purpose of our analysis of these cell paths was to identify possible migration trends relating cell motility and directional persistence with substrate stiffness. Our experiments revealed that cells that migrated across stiff substrates exhibited a longer persistence time $T_p$ on average than cells which migrated across non-stiff substrates. Cells on stiff substrates were more likely to exhibit high directional persistence as opposed to a random "wandering" behavior. Our experiments revealed statistically significant different responses to substrate conditions using the following measures of cell behavior: mean displacement $d$, mean speed $S$, deviation of movement direction $\sigma_\theta$, and duration of directional persistence $T_p$.

Table 3. **Analysis of Cell Behavior.** Migration properties are presented for 147 cells in 15 sequences of high-density cell populations (seeded at 1,000 cells/cm$^2$) and for two levels of substrate stiffness $\xi$ .

| Stiffness $\xi$ | 10 kPa | 150 kPa |
| --- | --- | --- |
| Mean displacement $d_{0 \to T}$ | 203 pixels | 463 pixels |
| Mean path length $l_T$ | 1,993 pixels | 2,129 pixels |
| Mean Speed $S_T$ | 21.6 $\mu$m/hr | 42.7 $\mu$m/hr |
| Angular deviation $\sigma_\theta$ | 78$^o$ | 57$^o$ |
| Mean persistence $T_p$ | 1:12 hr | 2:54 hr |

The difference in response to the physical stimuli presented in the environment, here substrate stiffness, applied not only to individual cells, but also to groups of cells (Fig. 4). We measured that cells disperse more rapidly on stiff substrates ($D_{315 \, \text{min}} = 225$ pixels with standard deviation $\sigma_D = 106$ pixels, for 3 cell groups) than on soft
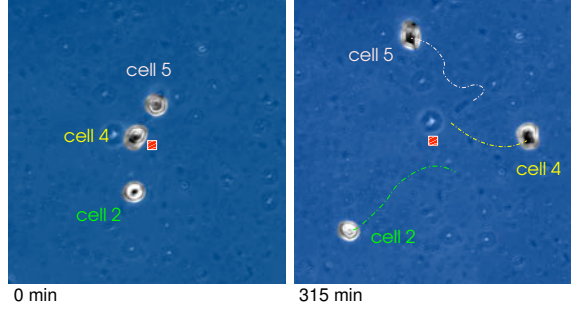
Figure 4. Dispersion of three cells with mean dispersion distances $D_0 = 65$ pixels and $D_T = 213$ pixels. The cell trajectories at time $T = 315$ min are shown. The initial and final group centroids are marked as red/white squares.

substrates ($D_{315\ \text{min}} = 72$ pixels with standard deviation $\sigma_D = 45$ pixels, for 2 cell groups).

We also conducted a preliminary test of the behavior of cells on a substrate with a stiffness gradient. We prepared the substrate so that it contained regions of increasing stiffness levels, starting from 10 kPa on one side of the substrate to 150 kPa on the other side in 10-kPa increments. Our test revealed that the motion of cells accelerated in the direction of the stiffness gradient. In future experiments, we will investigate if acceleration is sufficiently sensitive to detect and quantify a durotactic effect. Our preliminary gradient trial indicated that out of 10 cells tracked, 8 experienced a sudden change in acceleration when moving from a soft to a stiff substrate region.

### 3.4. The Random Walk Model and an Alternative Analysis

We verified that $\beta = 15^o$ is a reasonable threshold to define directional persistence $T_p$. By analyzing 42 migrating cells in 6 sequences for which we measured a persistence period of at least 1 hour, we found that, during their directionally persistent movement, cells change their directions by at most $\pm 14^o$ with a mean of $5.44^o$ and a standard deviation of $\pm 3.84^o$.

We then evaluated the validity of the random walk model, described in Eq. 8, for our data. We measure the root mean-squared (RMS) error

$$\sqrt{\frac{1}{k}\sum_{\text{cell}i=1}^{k}(2S_{T,i}^2 T_{p,i}(T-T_{p,i}(1-e^{-\frac{T}{T_{p,i}}}))-d_{0\to T,i}^2)^2}, \quad (9)$$

of the model given measurements of speed, persistence period and displacement of $k$ cells. We computed the RMS error to be 189,199 pixels$^2$ for $k = 14$ cells on 11 sequences. The square of the average maximum displacement of these 14 cells was $206^2$ pixels $= 42,630$ pixels$^2$. The RMS error was large for both soft and stiff substrates. Our experimental analysis shows that the random walk model, which has

been used extensively in the literature to describe cell migration, can produce large errors in predicting cell displacement. In fact the error in the displacement estimate is of the order of the displacement itself.

An alternative way to describe cell behavior in response to environmental conditions is to use the measures in Table 3, for which we measured statistically significant differences in behavior. These include four of the five measures, angular deviation, mean displacement, mean speed, and mean duration of the persistence period. Only the mean path length did not differ statistically significantly. In future work, we will investigate the relationship between behavior and level of the substrate stiffness and derive a functional connection.

The histograms of 550 measured instantaneous velocities of 10 cells, migrating on soft (10 kPa) and stiff (125 kPa) substrates, indicate that cells migrated faster on stiffer than on softer substrates (Fig. 5). The statistics of these distributions of instantaneous velocities (i.e., the mean and standard deviation) are significantly different from each other, which confirms that the velocity metric is sufficiently sensitive to detect and quantify a durokinetic effect.
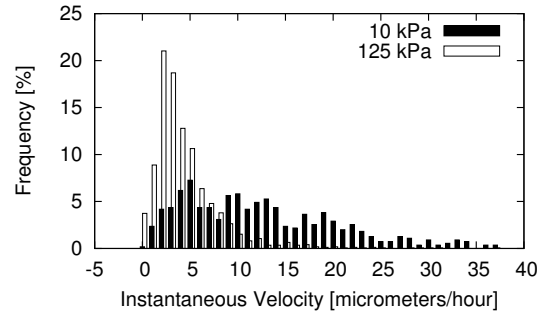


Figure 5. Histogram of instantaneous velocities measured for substrates with two stiffness conditions.

## 4. Discussion and Conclusions

We have provided algorithms for detection, segmentation and tracking of large collections of cells and introduced a comprehensive set of metrics for analysis of cell migration. Our methods use the measured statistics of cell behavior to provide feedback to the tracking module, allowing adaptation of the cost function based on the conditions of the cell environment, in particular, substrate stiffness.

We plan to enhance the functionality of our current system so that it can detect a variety of cell types and also cell organelles. The latter will enable studies in cytoplasmic trafficking [2], i.e., studies that monitor the movement of cell organelles through a cell's cytoplasm.

Given the discrepancy between the theoretical random walk model and our experimental results, we suggest that research studies of cell behavior may benefit from a shift of

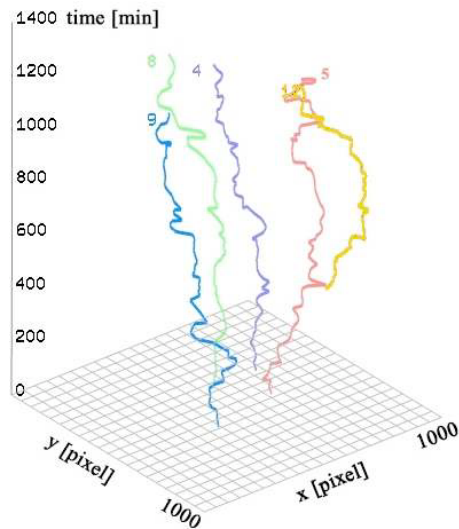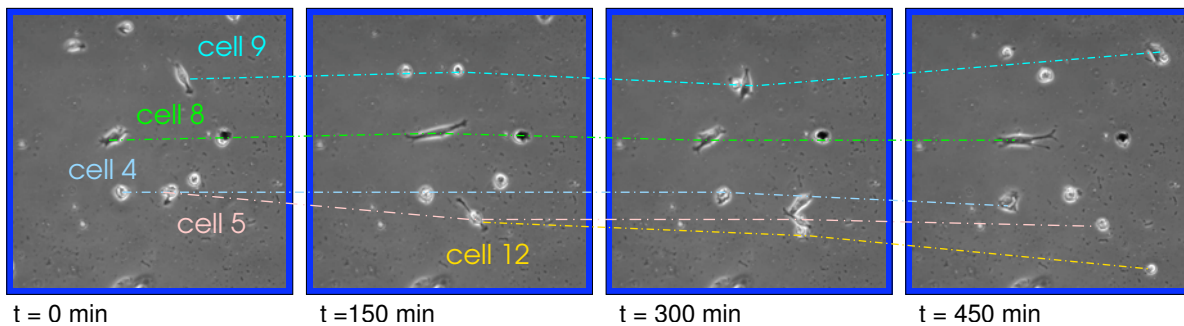t = 0 min    t =150 min    t = 300 min    t = 450 min



Figure 6. Tracking of multiple fibroblast cells on the surface of a hydrogel. The images (0.47 mm across) show the first 6 1/2 hours of the tracking process for a subset of 5 cells. The system detected the cell division and track-initiation event (yellow) and stores that cell 12 is the daughter of cell 5. The 3D plot visualizes the 5 tracks over the full imaging period of almost one day. Colors outline the correspondence between the 3D tracks and the 2D cell trajectories. Average migration speeds ranged from 5.4 $\mu$m/hr (cell 4) to 45.4 $\mu$m/hr (cell 9); average changes in direction of motion ranged from 2 to 12 degrees / 15 min.

focus: In the past, without automated image analysis, it was valuable to investigate the random walk equation, which relates the three key properties, i.e., cell displacement, persistence time, and average speed. Now, given the ability to analyze large collections of microscopy videos automatically, many more than these three properties, i.e., various instantaneous measurements and descriptions of the environmental conditions, should be used to characterize cell behavior. One of the goals of this paper was to make a first step into this data-driven direction by quantifying statistically significant differences in behavior in response to the level of substrate stiffness.

Understanding the movements of cells, their interactions, and their reaction to stimuli is essential for the effective management of the environments and materials in which these cells can thrive. Understanding the movement behavior of cell on specific substrates is essential for the development of biomaterial scaffolds, one of the central goals of tissue engineering.

Our work and similar efforts by other computer vision scientists are likely to have an important impact on biomaterial engineering. Because we provide image understanding tools that work with physiologically relevant substrates, our work is likely to reduce the barriers for data collection and significantly accelerate studies that interpret cell behavior on physiologically relevant substrates.

## References

[1] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *Journal of Mathematical Biology*, 9:147–177, 1980.

[2] W. Alt and M. Dembo. Cytoplasm dynamics and cell motion: two-phase flow models. *Mathematical Biosciences*, 156:207–228, Mar. 1999.

[3] Y. Bar-Shalom and X. R. Li. *Multitarget - Multisensor Tracking: Principles and Techniques*. YBS Publishing, 1995.

[4] D. P. Bertsekas. *Linear Network Optimization: Algorithms and Codes*. MIT Press, 1991.

[5] P. DiMilla, J. Quinn, S. M. Albelda, and D. A. Lauffenburger. Measurement of individual cell migration. *AIChE Journal*, 38(7):1092–1104, July 1992.

[6] G. A. Dunn. Characterising a kinesis response: Time averaged measures of cell speed and directional persistence. *Agents and Actions (Suppl.)*, 12:14–33, 1983.

[7] C. Gaudet, W. A. Marganski, S. Kim, C. T. Brown, V. Gunderia, M. Dembo, and J. Y. Wong. Influence of substratum type I collagen concentration on fibroblast spreading, motility, and contractility. *Biophysical Journal*, 85:3329–3335, 2003.

[8] D. Grünbaum. Advection-diffusion equations for generalized tactic searching behavior. *J. Math Biol.*, 38:169–194, 1999.

[9] B. K. P. Horn. *Robot Vision*. The MIT Press, Cambridge, MA, 1986.

[10] N. N. Kachouie and P. W. Fieguth. Extended-Hungarian-JPDA: Exact single-frame stem cell tracking. *IEEE Transactions on Biomedical Engineering*, 54(11):2011–2019, Nov. 2007.

[11] N. N. Kachouie, P. W. Fieguth, J. Ramunas, and E. Jervis. Probabilistic model-based cell tracking. *International Journal of Biomedical Imaging*, 2006:1–10, July 2006.

[12] T. Kirubarajan, Y. Bar-Shalom, and K. R. Pattipati. Multiassignment for tracking a large number of overlapping objects. In Y. Bar-Shalom, editor, *Multitarget-multisensor tracking – Vol. 3: Applications and advances*. Artech House, Norwood, MA, US, 2000.

[13] K. Li, M. Chen, and T. Kanade. Cell population tracking and lineage construction with spatiotemporal context. In N. Ayache, S. Ourselin, and A. Maeder, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007, 10th International Conference, Brisbane, Australia, October 29 - November 2, 2007, Proceedings, Part II, Lecture Notes in Computer Science, Volume 4792/2007*, pages 295–302. Springer Berlin / Heidelberg, 2007.

[14] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis*, 12(5):546–566, Oct. 2008.

[15] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell. Computer vision tracking of stemness. In *Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2008)*, pages 847–850, May 2008.

[16] 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 *2006 Conference on Computer Vision and Pattern Recognition Workshop on Mathematical Modeling in Biomedical Image Analysis (MMBIA 2006)*, pages 1–8, New York, NY, USA, June 2006.

[17] C. M. Lo, H. B. Wang, M. Dembo, and Y. L. Wang. Cell movement is guided by the rigidity of the substrate. *Biophys J*, 79(1):144–152, 2000.

[18] V. Rabaud and S. Belongie. Counting crowded moving objects. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, pages 705–711. IEEE Computer Society, 2006.

[19] P. Rajagopalan, W. A. Marganski, X. Q. Brown, and J. Y. Wong. Direct comparison of the spread area, contractility, and migration of balb/c 3T3 fibroblasts adhered to fibronectin- and RGD-modified substrata. *Biophysical Journal*, 87(4):2818–2827, 2004.

[20] K. Shafique and M. Shah. A noniterative greedy algorithm for multiframe point correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(1):51–65, 2005.

[21] K. Smith, A. Carleton, and V. Lepetit. General constraints for batch multiple-target tracking applied to large-scale videomicroscopy. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, pages 1–8, June 2008.

[22] R. T. Tranquillo and W. Alt. Glossary of terms concerning oriented movement. In W. Alt and G. Hoffmann, editors, *Biological Motion, Lecture Notes in Biomathematics. Vol. 89*, pages 584–565. Springer-Verlag, 1990.

[23] J. Y. Wong, A. Velasco, P. Rajagopalan, and Q. Pham. Directed movement of vascular smooth muscle cells on gradient compliant hydrogels. *Langmuir*, 19:1908, 2003.

[24] J. Xie, S. Khan, and M. Shah. Automatic tracking of escherichia coli bacteria. In D. Metaxas, L. Axel, G. Fichtinger, and G. Székely, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008, 11th International Conference, New York, NY, USA, September 6-10, 2008, Proceedings, Part I, Lecture Notes in Computer Science, Volume 5241/2008*, pages 824–832. Springer-Verlag Berlin Heidelberg, 2008.

[25] X. Zhou, J. Yang, M. Wang, and S. T. C. Wong. A novel cell tracking algorithm and continuous hidden Markov model for cell phase identification. In *IEEE/NLM Life Science Systems and Applications Workshop*, pages 1–2, July 2006.