Tracking of cell population from time lapse and end point confocal microscopy images with multiple hypothesis Kalman smoothing filters

Tracking of Cell Population from Time Lapse and End Point Confocal Microscopy Images with Multiple Hypothesis Kalman Smoothing Filters

Lee-Ling S. Ong
Singapore-MIT Alliance for Research and Technology,
3 Science Drive 2,
Singapore 117543.
sharon.ong@smart.mit.edu

Marcelo H. Ang Jr
Department of Mechanical Engineering,
National University of Singapore,
Singapore 117576.
mpeangh@nus.edu.sg

H. Harry Asada
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, MA, 02139
asada@mit.edu

Abstract

This paper describes an automated visual tracking system combining time-lapse and end-point confocal microscopy to aid the interpretations of cell behaviors and interactions, with the focus on understanding the sprouting mechanism during angiogenesis. These multiple cells exhibit stochastic motion and are subjected to photobleaching and the images acquired are of low signal to noise ratio. Hence, following time-lapse imaging, high resolution end-point images are acquired. Our approach applies a probabilistic motion filter (a backward Kalman filtering followed by track smoothing) which incorporates end-point and all available time-lapse information in a mathematically consistent manner to obtain trajectory and phenotype information of multiple individual cells simultaneously. An extension of this algorithm, track smoothing with a Multiple Hypothesis Testing (MHT) data association, is proposed to improve association of multiple close contact and proliferating cells across images acquired from different time points to existing track trajectories. Our methodology was applied to tracking endothelial cell sprouting in three-dimensional micro-fluidic devices.

1. Introduction

Cell proliferation, differentiation, and migration are essential for the conception, development and maintenance of any living organism. Automated tracking of multiple cell interactions in *in vitro* in time lapse microscopy would allow more accurate automated interpretations of cell behaviors from experimental data, by providing trajectory and phenotype information of each individual cell, particularly when robust probabilistic techniques are applied. Following time-lapse imaging, fixed end point staining protocols may be applied to provide more information on cell behavior [19].

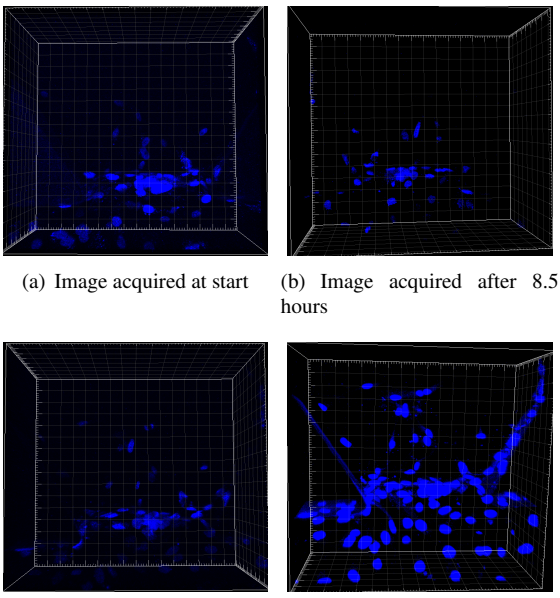
A significant amount of data can be produced from typical experiments captured by time-lapse microscopy images, particularly from 3D confocal images. Manual cell tracking of such data would be a time-consuming chore and subjected to human-to-human variance. Automated image analysis is therefore a more favourable alternative for a more efficient and accurate data assessment. Tracking in biological systems has been applied at all scales of microscopic observations, from molecules such as tracking intra-cellular particles[9, 18, 20], to cells [11, 13] and up to organisms [8]. The focus of our tracking system is to aid investigations into angiogenesis in *in vitro* systems in high-throughput devices which generate a large amount of data [4, 22].

Angiogenesis is the formation of new blood vessels from a monolayer of cells or by the reorganization of capillaries via morphogenesis. When exposed to growth factors, endothelial cells forming a monolayer, undergo stochastic phenotype transitions such as migrating, quiescent, proliferation or death. Explorations of the angiogenic sprouting mechanism to determine how a population of cells could sprout out creating a new vascular network structures requires efficient and accurate image analysis to provide estimates of the cell trajectories and phenotypes over time.

A number of issues are faced in automating cell tracking in addition to the extensive volume of the experimental data. These issues include the low signal to noise ratio of the data and varying cell densities with a denser group of cells forming the monolayer and a more sparely populated cells migrating from the monolayer. As a result complex cellular topologies are encountered due to close contact and partial overlap of cells and the ability of the cell to deform and alter its shape. Cell densities also differ due to the proliferation

978-1-4244-7030-3/10/$26.00 ©2010 IEEE

and death of cells. Another issue specific to fluorescent confocal microscopy is photo-bleaching. As more images are acquired over time, more photochemical destructions occurs in the stained fluorescent molecules in the cells. Excessive light exposure stimulating these fluorescent molecules may even cause photo-toxicity. As a results, the sampling time and the acquisition step size throughout the height of the cell migration regions is limited. This motivates the use of end-point imaging after the cells have been fixed. Given a large number of cells with stochastic behaviors, association between identified cells from one image frame to another image frame captured in the next time point and between identified cells from the images from the time-lapse data and end-point data, where examples are shown in Figure 1, becomes an issue.



(a) Image acquired at start (b) Image acquired after 8.5 hours

(c) Image acquired after 17 hours (d) End-point Imaging Results

Figure 1. Fluorescent intensities images of cell nuclei of time lapse and end-point image acquisition. Photobleaching occurs prolonged light exposure required to stimulate the fluorescent molecules in the cells, causes the photochemical destruction of these molecules. However, these molecules may be restained at end point to provide more information.

Our approach to multi-cell tracking addresses these challenges via probabilistic methodologies. A Kalman filtering combined with Multiple Hypothesis Testing (MHT) and smoothing/retrodiction is proposed to allow tracking of varying cell dynamics and account for clutter due to close contact cells. In addition to that, probabilistic techniques are used to incorporate fixed end-point imaging data with time-lapse information in a mathematically consistent manner.

2. Related Work

The simplest approach to cell association in tracking is known as the "nearest-neighbor association", where each cell in one frame is associated with the spatially nearest cell in the next frame within a pre-fixed range threshold [6]. This is similar to the mean shift process [5], where the threshold is defined by a uniform kernel. These techniques often fail when tracking multiple cells or rapid cell movements. Another approach to tracking of time-lapse images is a combination of active contours and optical flow [12, 13, 14]. A review of existing automatic cell tracking software available [10] showed that active contours and nearest neighbour feature matching were the popular approaches for time-lapse image analysis. The performance of these methods are poor when there are partial overlaps and close contact of cells. To account for these issues, probabilistic methods have been used particularly for tracking at the molecular level [9, 18]. At the cellular level, Kalman filtering [23], MHT [2] and Interacting Multiple Models (IMM) [11] have been applied for tracking. To our knowledge, our work is the only application of Kalman filtering *with* MHT *and* smoothing on cell tracking. MHT allows more accurate inferences of a cell location and identification of proliferation in each time frame. These inferences are made on all available information; observations of cells in upcoming image frames, previous image frames in time and end point imaging via smoothing.

3. Incorporation of End-Point and Future Time Lapse Information to Motion Filtering

The standard Kalman filter, used in most tracking operations forms the posterior estimate over time given only observations up to time k. Images for the whole time lapse as well as end point information are available. Incorporation of this information into the filter would provide a more accurate estimate of the trajectory and aid in the identification of cell phenotypes such as proliferation. In this application, our tracked state at time k, for an individual cell is the three-dimensional cell centroid position and velocity, ($\mathbf{x}_k = (x_k, \dot{x_k}, y_k, \dot{y_k}, z_k, \dot{z_k})$) and the measurement state is the cell centroid ($\mathbf{z}_k = (x_k, y_k, z_k)$).

3.1. Filtering with End-Point Observations

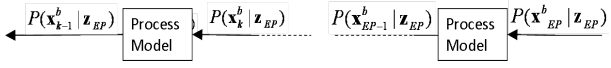


Figure 2. Incorporation of end point information to a time k (where $k < t_{EP}$) involves propagation backwards through the process model

As shown in Figure 2, the estimate based on the observations at end point is initialized at end point time ($k = t_{EP}$) with the end point observation. In order to align this estimate with the estimates of each cell using each discrete time lapse acquisition period, backward filtering is applied. As the state posterior is represented by a Gaussian function, $p(\mathbf{x}_k|\mathbf{z}_{EP}) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}(k|EP), \mathbf{P}(k|EP))$, backward Kalman filtering is applied by back-propagation of the state posterior mean and covariance via the process model as shown in Equation 1.

$$\hat{\mathbf{x}}(k|k+1) = \mathbf{F}_k^{-1}\hat{\mathbf{x}}(k+1|k+1) + \mathbf{B}_k^{-1}\mathbf{u}(k) \quad (1)$$

$$\mathbf{P}(k|k+1) = \mathbf{F}_k^{-1}\mathbf{P}(k+1|k+1)\mathbf{F}_k^{-T} + \mathbf{G}_k^{-1}\mathbf{Q}_k\mathbf{G}_k^{-T} \quad (2)$$

where $\mathbf{F}_k^{-1}$ is the inverse of the state transition model which is applied to the subsequent state $\hat{\mathbf{x}}(k+1|k+1)$, $\mathbf{u}(k)$ represents some independent state vector such as a control input, $\mathbf{B}_k^{-1}$ is the control-input model which is applied to $\mathbf{u}(k)$ and $\mathbf{G}_k^{-1}\mathbf{Q}_k\mathbf{G}_k^{-T}$ denotes the covariance matrix of the process noise. Cell movements consists of two components, a directed migration in response to the growth factors and a stochastic factor due to membrane adhesion and cytoskeleton contraction [16]. This movement is modeled as a constant velocity model [1]. Note that as observation measurements from end point information are obtained only at $k = t_{EP}$, the state posterior at each time k given observations up to time k equals to the state posterior given observations up to time $k+1$ ($\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k+1)$ and $\mathbf{P}(k|k) = \mathbf{P}(k|k+1)$).

3.2. Filtering with Future Time-Lapse Observations

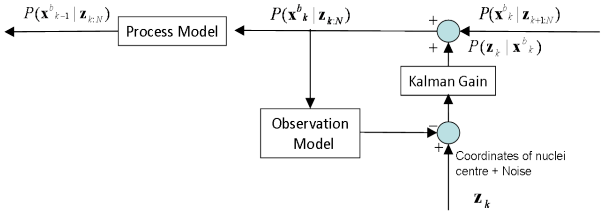


Figure 3. Incorporation of upcoming time lapse information to a current time involves back-propagation through the process model and a Kalman update.

As with the end point observations, backward prediction is performed for each track in order to incorporate upcoming time-lapse observations. The filters is initialized with observations from the last time lapse information and back prediction via the process model is also performed. Observations of the state of the cell centroids would be available every time frame and are updated to the filter as shown in Figure 3 to create a new state estimate $\hat{\mathbf{x}}_b(k|k)$ of the state at time k given all future information up to and included the observation at time k with the corresponding covariance in the estimate via the following equations:

$$\hat{\mathbf{x}}_b(k|k) = \hat{\mathbf{x}}_b(k|k+1) + \mathbf{W}_k[\mathbf{z}(k) - \mathbf{H}_k\hat{\mathbf{x}}_b(k|k+1)] \quad (3)$$

$$\mathbf{P}_b(k|k) = \mathbf{P}_b(k|k+1) - \mathbf{W}_k\mathbf{S}_k\mathbf{W}_k^T \quad (4)$$

where $\mathbf{z}(k) - \mathbf{H}_k\hat{\mathbf{x}}_b(k|k+1)$ is the error between the back-predicted and actual observation and is referred to as the innovation, and the weighting or gain matrix $\mathbf{W}_k$ is selected to minimise the mean squared error in the estimate which is $\mathbf{W}_k = \mathbf{P}_b(k|k+1)\mathbf{H}_k^T\mathbf{S}_k^{-1}$ where $\mathbf{S}_k$ is the innovation covariance given as $\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_b(k|k+1)\mathbf{H}_k^T + \mathbf{R}_k$ and $\mathbf{R}_k$ denotes the covariance matrix of the observation noise.

3.3. Smoothing

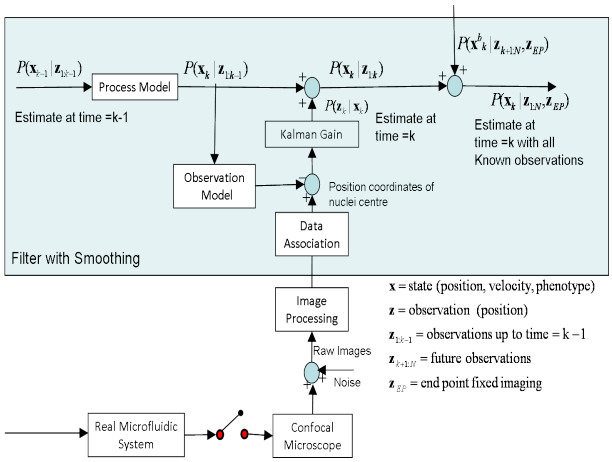


Figure 4. The system overview at each time frame is shown. The noisy raw images obtained at each time frame are processed. The position coordinates of the nuclei centre is updated to the filter to obtain the state posterior estimate. The smoothed estimate is then obtained through combination of this 'forward' estimate with the 'backward' estimate and end point information

The objective of smoothing is to obtain the best estimate of state at time k given all observations from time 1 to N and the end point observations at time EP, where $N < EP$ and $k < N$ which is defined by:

$$\hat{\mathbf{x}}(k|1:N, EP) = \arg\max_{\mathbf{x}(k)} p(\mathbf{x}_k|\mathbf{z}_{1:N,EP}) \quad (5)$$

Using Bayes Theorem, the smoothed estimate may be computed as a combination of the standard forward Kalman filter and an estimate computed from a backward running Kalman filter [7] as illustrated in Figure 4. The smoothed estimate conditioned on all observations at time k is given by:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{z}_{k+1:N}) = \frac{1}{c}p(\mathbf{x}_k|\mathbf{z}_{1:k})\,p(\mathbf{x}_k|\mathbf{z}_{k+1:N}) \quad (6)$$

where the constant c, is the normalizing factor, such that the volume of the probability distribution is one $\int_{\infty}^{\infty} p\left(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{z}_{k+1:N}\right) \mathrm{d}\mathbf{x} = 1$ and is independent of the experimental data. Using the Gaussian assumption, the results is a multiplication of two Gaussian functions, where $\hat{\mathbf{x}}_f\left(k|k\right)$ and $\mathbf{P}_f\left(k|k\right)$ are the mean and covariance of the forward Kalman filter at time k updated with $\mathbf{z}_{1:k}$ (observations up to time k) and $\hat{\mathbf{x}}_b\left(k|k+1\right)$ and $\mathbf{P}_b\left(k|k+1\right)$ are the mean and covariance of the backward Kalman filter at time k, conditioned only with observations from $k+1$ to N.

$$p\left(\mathbf{x}_k|\mathbf{z}_{1:k}\right) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_f\left(k|k\right), \mathbf{P}_f\left(k|k\right)) \quad (7)$$

$$p\left(\mathbf{x}_k|\mathbf{z}_{k+1:N}\right) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_b\left(k|k+1\right), \mathbf{P}_b\left(k|k+1\right)) \quad (8)$$

The smoothed estimate for each image acquisition time is then:

$$\mathbf{P}_s\left(k|1:N\right) = \left[\mathbf{P}_f^{-1}\left(k|k\right) + \mathbf{P}_b^{-1}\left(k|k+1\right)\right]^{-1} \quad (9)$$

$$\begin{aligned}\hat{\mathbf{x}}_s\left(k|1:N\right) =& \mathbf{P}\left(k|1:N\right)\left[\mathbf{P}_f^{-1}\left(k|k\right)\hat{\mathbf{x}}_f\left(k|k\right) + \right. \quad (10)\\ & \left.\mathbf{P}_b^{-1}\left(k|k+1\right)\hat{\mathbf{x}}_b\left(k|k+1\right)\right]\end{aligned}$$

3.4. Multi Cell Tracking

Multiple centroids of cells are extracted from each image at each time frame. In order to simultaneously obtain multiple cell trajectories, multi-target tracking algorithms [1] are applied. In this algorithm, a filter is allocated to each cell. Each track is initialized by the first or first few measurements of the cell position. At each time frame, validation gating is applied each observation of the cell to determine whether to update that observation to an existing track or to create a new track. The gate is a chi-square random variable in n degrees and is centered on the predicted measurement of the cells. The single best matched observation that falls into the gate is selected to be associated to the filter.

Validation gating of the observation to a track is derived from the Normalised Innovation Covariance between track j and observation i.

$$d_{ij}^2 = \mathbf{v}_{ij}^T\left(k\right)\mathbf{S}_{ij}^{-1}\left(k\right)\mathbf{v}_{ij}\left(k\right) \quad (11)$$

where $\mathbf{v}_{ij}\left(k\right)$ is the innovation between the observation i and prediction observation j is $\mathbf{v}_{ij}\left(k\right) = \mathbf{z}_i\left(k\right) - \mathbf{H}_k\left(\hat{\mathbf{x}}_j\left(k|k-1\right)\right)$ and the innovation covariance is $\mathbf{S}_{ij}\left(k\right) = \mathbf{H}_k\mathbf{P}_i\left(k-1|k-1\right)\mathbf{H}_k^T + \mathbf{R}\left(i,k\right)$.

If there are no observations with a normalised innovation less than chi-square thresh-old, no observation is associated. Filter maintenance is performed to eliminate any filters initialized from spurious measurements. An indication of spurious measurements is when no observations are associated with the track for an extended period of time resulting in an increase of track uncertainty. When the track uncertainty reaches a set threshold, the filter is pruned.

To perform smoothing over multiple cell trajectories simultaneously, multi-target tracking is performed both in the forward and backward direction. Association between tracks of forward and backward filters for smoothing is performed using the Mahalanobis distance measure which is:

$$D_M\left(k\right) = \sqrt{\left(\mathbf{x}_1 - \mathbf{x}_2\right)^T\mathbf{P}^{-1}\left(\mathbf{x}_1 - \mathbf{x}_2\right)} \quad (12)$$

where $\mathbf{x}_1 = \hat{\mathbf{x}}_f\left(k|k\right)$,$\mathbf{x}_2 = \hat{\mathbf{x}}_b\left(k|k\right)$ and $\mathbf{P} = \mathbf{P}_f\left(k|k\right) + \mathbf{P}_b\left(k|k\right)$. The pair of filters which the minimum distance under a chi-square threshold is selected for smoothing.

4. Multiple-Hypothesis Cell Tracking

The standard multi-target tracking where only the single best matched observation is updated assumes the rate of detection of the cell is high and the "clutter" is low. When tracking multiple cells, issues arise when cells come in close contact or when cell proliferation occurs. This is addressed with a multi-hypothesis testing approach with delayed decision. Multiple possible tracks are maintained and hypotheses with low probability are gradually merged or pruned as more information becomes available. Given multiple matches (hypotheses), the probability density function of the target state vector is hence represented by Gaussian mixture shown in Equation 13.

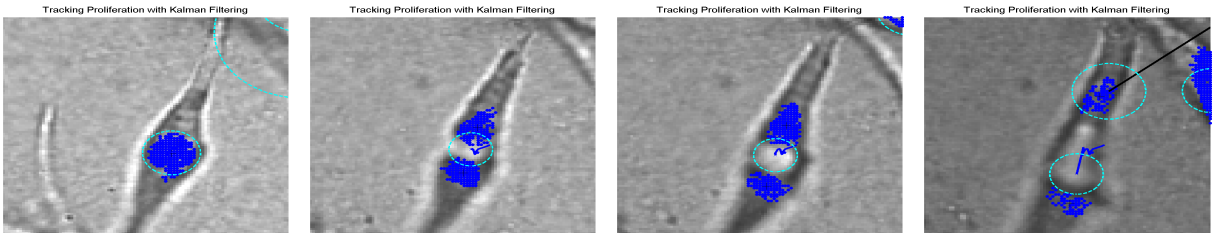
$$\begin{aligned}p\left(\mathbf{x}_k|\mathbf{Z}^k\right) &= \sum_{i=0}^{N_\theta} p\left(\mathbf{x}_k|\Theta_i^k, \mathbf{Z}^k\right) p\left(\Theta_i^k|\mathbf{Z}^k\right)\\ &= \sum_{i=0}^{N_\theta} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_i\left(k|k\right), \mathbf{P}_i\left(k|k\right)) p\left(\Theta_i^k|\mathbf{Z}^k\right)\end{aligned} \quad (13)$$

where Θ_i^k is a particular hypothesis and is based on the current hypothesis $\theta_i\left(k\right)$ and its previous hypothesis Θ_i^{k-1} and N_θ is the number of hypothesis at time k. The first term in Equation 13, $p\left(\mathbf{x}_k|\Theta_i^k, \mathbf{Z}^k\right)$ is a Gaussian component of the distribution in which Kalman filtering equations are applied and the conditional probability of the hypothesis is given as:

$$\begin{aligned}p\left(\Theta_i^k|\mathbf{Z}^k\right) = \frac{1}{c}\lambda_F\lambda_N &\prod_{i=1}^{m_k}\left[\mathcal{N}_{t_i}\left(z_i\left(k\right)\right)\right]^{\tau_i}\\ &\prod_t \left(P_D^t\right)^{\delta_t}\left(1 - P_D^t\right)^{1-\delta_t}\\ &\left(P_\chi^t\right)^{\chi_t}\left(1 - P_\chi^t\right)^{1-\chi_t} p\left(\Theta_i^{k-1}|\mathbf{Z}^{k-1}\right)\end{aligned} \quad (14)$$

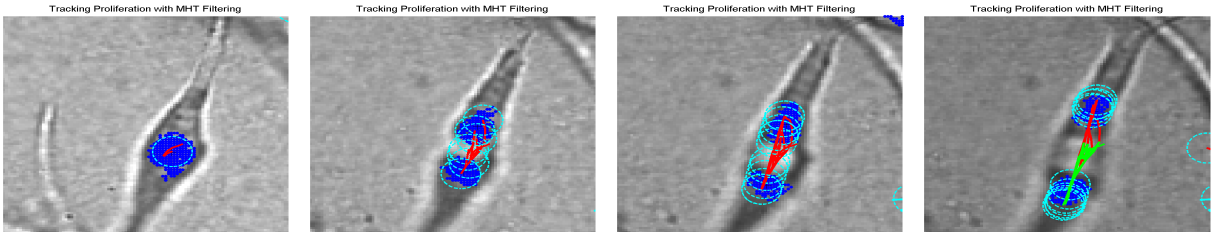
where λ_F and λ_N are the number of false alarms and new features, P_D^t and P_χ^t are the probabilities of detection and determination of the track and $\mathcal{N}_{t_i}\left(z_i\left(k\right)\right)$ is the likelihood function given by:

$$\begin{aligned}\mathcal{N}_{t_i}\left(z_i\left(k\right)\right) = &\left|2\pi\mathbf{S}^{t_i}\left(k\right)\right|^{-\frac{1}{2}}\\ &\exp^{-\frac{1}{2}\left(\mathbf{v}_{ij}^T(k)\left[\mathbf{S}^{t_i}(k)\right]^{-1}\mathbf{v}_{ij}(k)\right)}\end{aligned} \quad (15)$$



(a) Image acquired after 5 hours (b) Image acquired after 6.5 hours (c) Image acquired after 7.5 hours (d) Image acquired after 10 hours

Figure 5. Tracking of cell proliferation with association of the single best match observation.



(a) Image acquired after 5 hours (b) Image acquired after 6.5 hours (c) Image acquired after 7.5 hours (d) Image acquired after 9 hours

Figure 6. Tracking of cell proliferation with multiple hypothesis association.

where $\mathbf{v}_{ij}(k)$ is the innovation and $\mathbf{S}^{t_i}(k)$ is the innovation covariance of the up-date of the hypothesis. At each time state the predicted observations of each track is used to establish a validation gate. If there are n measurements found in the gate, a single track may be split into $n+1$ tracks, one track for each measurement and one track (denoted as 0) for a non-association hypothesis. Each of the new tracks is then treated independently and used to generate new predictions for the next time step, resulting in an exponentially increasing number of tracks. Hypothesis maintenance may be performed through track pruning or merging of hypothesis [15, 21] with low likelihood weights.

4.1. Tracking Cell Proliferation Results

Figure 5 shows the results of the application of selection of the best matched observation to update to a filter. The filter ends up tracking one of the proliferated cells while the observation measurement from the other cell is updated to another filter. With the application of multiple hypothesis data association, separate hypotheses are created to track both cells as shown in Figure 6. To identify that proliferation has occurred, the Mahalanobis distance measure is applied to cluster the hypothesis. Should the distance measure between clusters exceed the chi-square threshold, track splitting occurs as shown in Figure 6(d).

4.2. Multiple Hypothesis Smoothing

In order to incorporate future time-lapse and end-point information to a filter that tracks proliferation, a multiple hypothesis smoothing operation is proposed. Equation 6 shows that smoothing is a multiplication operation between the forward and backward posteriors.

$$\begin{aligned} p\left(\mathbf{x}_k|\mathbf{z}_{1:N}\right) &= \frac{1}{c} p\left(\mathbf{x}_k|\mathbf{z}_{1:k}\right) p\left(\mathbf{x}_k|\mathbf{z}_{k+1:N}\right) \\ &= \frac{1}{c} \sum_{i=0}^{N_{\theta,f}} p\left(\mathbf{x}_k|\Theta_i^{1:k}, \mathbf{z}_{1:k}\right) p\left(\Theta_i^{1:k}|\mathbf{z}_{1:k}\right) \\ &\sum_{i=0}^{N_{\theta,b}} p\left(\mathbf{x}_k|\Theta_i^{k+1:N}, \mathbf{z}_{k+1:N}\right) p\left(\Theta_i^{k+1:N}|\mathbf{z}_{k+1:N}\right) \end{aligned} \tag{16}$$

As the probability distribution function of the target state vector for the posteriors are Gaussian mixtures, the smoothing equation then become a multiplication of Gaussian mixture functions at each time point.

$$\begin{aligned} p\left(\mathbf{x}_k|\mathbf{z}_{1:N}\right) &= \sum_{i=0}^{N_{\theta,f}} \sum_{i=0}^{N_{\theta,b}} p\left(\Theta_i^{1:k}|\mathbf{z}_{1:k}\right) p\left(\Theta_i^{k+1:N}|\mathbf{z}_{k+1:N}\right) \\ &\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{i,f}(k|k), \mathbf{P}_{i,f}(k|k)) \\ &\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{i,b}(k|k), \mathbf{P}_{i,b}(k|k)) \\ &= \sum_{ij=0}^{N_{\theta,f} N_{\theta,b}} p\left(\Theta_{ij}^{1:N}|\mathbf{z}_{1:N}\right) \\ &\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{ij,s}(k|k), \mathbf{P}_{ij,s}(k|k)) \end{aligned} \tag{17}$$

where the resultant mean and covariance from the multiplication of the ith component of the forward estimate and jth component of the backward estimate is the same as in

Equation 9 and

$$P\left(\Theta_{ij}^{1:N}|\mathbf{z}_{1:N}\right) = P\left(\Theta_{i}^{1:k}|\mathbf{z}_{1:k}\right) P\left(\Theta_{j}^{k+1:N}|\mathbf{z}_{k+1:N}\right) \tag{18}$$

To associate forward and backward MHT filters for smoothing, the Kullback-Liebler divergence or relative entropy measure [3] is used. The divergence between two estimates $p_f(\mathbf{x})$ and $p_b(\mathbf{x})$ is defined as

$$D_{KL}\left(p_f(\mathbf{x})\,||p_b(\mathbf{x})\right) = \sum_{\mathbf{x}\in X} p_f(\mathbf{x}) \log \frac{p_f(\mathbf{x})}{p_b(\mathbf{x})} \tag{19}$$

The result is always non-negative and associations with more overlap would result in a lower divergence value and selected for smoothing.

5. Experimental Procedures

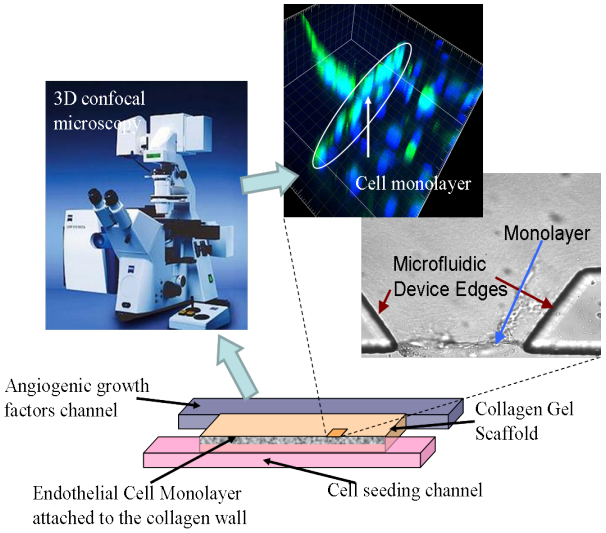


Figure 7. Experiments are carried out in a microfluidic flow chamber. Each region with a monolayer was captured by a confocal microscope. The captured image on the top right shows the acquisition results of two fluorescent channels, one of the nuclei and one of the cytoplasm.

Angiogenesis experiments in *in vitro* environments are typically performed using traditional on-the-gel dish experiments, which significantly different from the actual *in vivo* environment. An illustration of experimental setup is shown in Figure 7. Our experiments are carried out in a microfluidic flow chamber. In these devices, collagen gel is placed between the two flow channels. Through one channel, the cells are seeded and attached to the collagen forming a monolayer. Angiogenic growth factors, supplied in the other channel, diffuse through the collagen gel and bind to the receptors in the cells. This drives the cells to proliferate and migrate into the collagen gel.

5.1. Cell Staining and Imaging

Prior to seeding, the cells were stained with a Hoechst nuclear stain and a CMFDA cytosolic stain. Confocal fluorescent images of two channels (one which detects the nuclear stain and another which detects the cytosolic stain) and intensity images of each migration region in the gel (marked by two posts) were acquired at approximately 30 minute intervals. 4 hours following the time-lapse microscopy phase, the cells were fixed and stained with a nuclear stain (DAPI) and a stain to detect cell adhesion (VE-cadherin stain). End Point Imaging was then applied. The images captured are processed to obtain observation measurements of the cell centroid location. This in formation is updated to the filter at each time interval to capture the spatiotemporal state of the cell.

5.2. Image Processing

Image processing was performed via thresholding as the focus of this work is on the application of probabilistic motion filtering rather than image processing techniques. K-means clustering was applied to classify the image pixels into cell or background regions. These pixels are clustered based on connected components clustering. The 3-dimensional centroid of each cluster is obtained. The focus of this work is on studying cell sprouting and therefore our interest lies in tracking cell trajectories above the monolayer. Segmentation of the cell monolayer was performed on the intensity images through a combination of edge detection and Hough transforms. Another issue that arises is the alignment of the time-lapse image frames and the end point image frames due to the removal of the microfluidic device from the microscope for fixing and restaining. As our imaging region along the device were performed between two trapezoidal posts, the alignment offset was determine from the location of these posts from the intensity images.

6. Results

Our proposed MHT filtering with smoothing was applied to track the nuclei of endothelial cells in an image sequence of cell proliferation and migration. The end result of this application is shown in Figure 8(c). The performance of this algorithm was compared with the results of Kalman filtering and Kalman filtering with smoothing. The results of the latter is shown in Figure 8(b). A cell trajectory is valid only if it follows the same cell throughout all the frames that the cell is visible. When the cells are spatially spaced out, such as the Track 11, the performance of both algorithms are the same. However, when there are many cells, and when proliferation occcurs, the performance of the MHT with smoothing provides more superior results as shown in Track 12, where proliferation occurred and the tracking re-

sulted in an invalid cell trajectory. Three-dimensional migratory trajectory information shown in Figure 9.

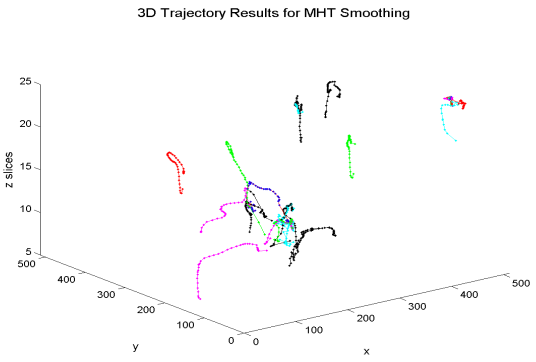


Figure 9. Three-dimensional migration trajectories from MHT smoothing.

One of the performance measures of the filters is entropy [17], which is defined as:

$$\begin{aligned} H_p(\mathbf{x}) &= -\sum_{\mathbf{x}\in X} p(\mathbf{x}) \log p(\mathbf{x}) \\ &= -\frac{1}{2}\log[(2\pi e)^n |P|] \end{aligned} \tag{20}$$

Entropy measures the information content of a distribution, where a lower entropy value indicates a distribution with more certainty. Figure 10 shows the entropy results for Track 11 of a cell migrating over time for the forward, backward and smoothing Kalman filters. Results indicate that the smoothing solutions are more informative due to the fusion of forward estimates $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ and backward estimates $p(\mathbf{x}_k|\mathbf{z}_{k+1:N})$ at each time k. Note that the entropy of the backward estimates are larger as it is based only on the upcoming estimates. The top figure shows the results without end-point information in which the filter uncertainty increases rapidly for both the forward and smoothed estimates. With the incorporation of end point information a more informative track is obtained particularly at the ends of the trajectories where the entropy is lesser that estimation without end-point information as shown in the bottom figure.

7. Conclusion

This paper presented a multiple cell tracking methodology based on Kalman filtering with multiple hypothesis testing and smoothing. It allows multiple cells to be tracked more accurately with the incorporation of all available observations included end-point information into each estimate and each time. Proliferation phenotypes were also detected through the use of MHT.

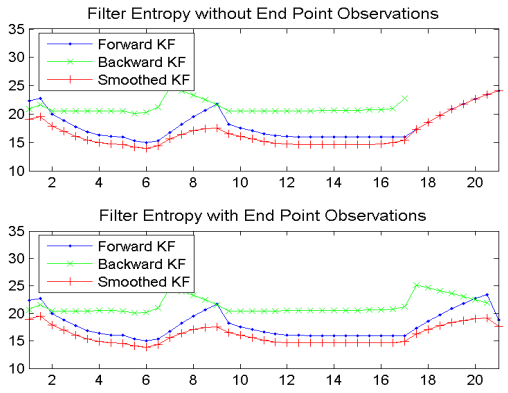


Figure 10. The entropy results for Track 11 of Figure 8(b) is shown, where smoothed solutions resulted in less entropy compared to the forward and backward solutions. Due to the low signal to noise ratio, feature segementation failed between the 6th and 9th hours, resulting in no updates and an increase in entropy. Results also show that the influence of the end-point is limited to a few steps.

Acknowledgements

The authors would like to thank Levi B. Wood for assistance in performing cell culturing and timelapse imaging. This work is supported by the Singapore-MIT Alliance for Research and Technology, Bio-Systems and Micromechanics Interdisciplinary Research Group (IRG).

References

[1] Y. Bar-Shalom. *Multitarget-multisensor tracking: advanced appplications*. Artech House, 1990. 3, 4

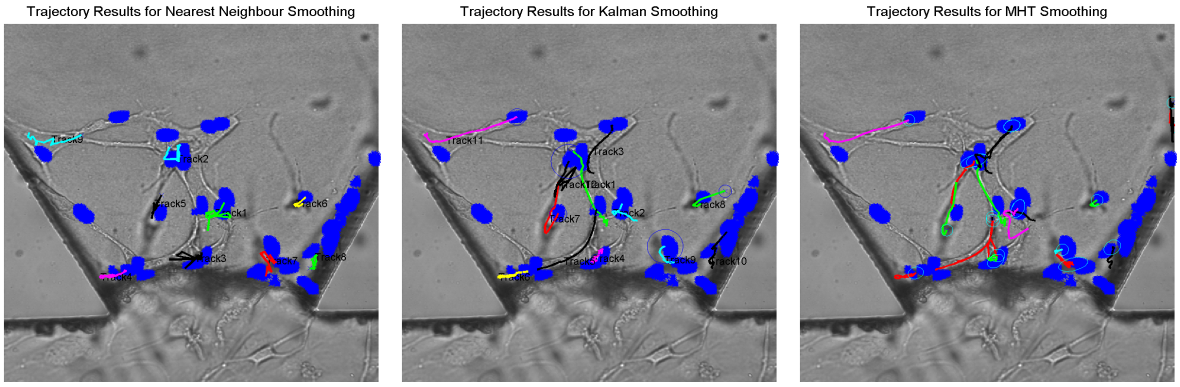
[2] F. Bunyak, K. Palaniappan, S. K. Nath, T. I. Baskin, and G. Dong. Quantitative cell motility for in vitro wound healing using level set-based active contour tracking. In *In: Proceedings of the Third IEEE International Symposium Biomedical Imaging (ISBI)*, pages 1040–1043, 2006. 2

[3] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. Wiley, New York, 1991. 6

[4] A. Das, D. Lauffenburger, H. Asada, and R. Kamm. A hybrid contimuum-discrete modeling approach to predicting and controlling angiogenesis: Analysis of combinatorial growth factor and matrix effects on vessel sprouting morphology. *Accepted for publication in Phil. Trans. A.*, 2010. 1

[5] 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 Transactions on Medical Imaging*, 24:697–711, 2005. 2

[6] Z. Demou and L. McIntire. Fully automated three-dimensional tracking of cancer cells in collagen gels: Determination of motility phenotypes at the cellular level. *Cancer Research*, 62:5301–5307, 2002. 2



(a) Nearest Neighbour Smoothing (b) Kalman smoothing with the best matched observation measurement updated (c) Multiple Hypothesis Testing smoothing solution

Figure 8. End results for different smoothing solutions. The thresholding results from the end point imaging is shown by the blue clusters of points. Results show that tarcking with a nearest neighbour solution results in the most incorrect associations and trajectories in comparison with the Kalman filtering solution. The application of a multi-hypothesis solution resulted in the most number of trajectories and allowed cell-proliferation to be tracked.

[7] D. C. Fraser and J. E. Potter. The optimum smoother as a combination of two optimum linear filters. *IEEE Trans. Automatic Control*, AC-1 4:387–390, 1969. 3

[8] W. Geng, P. Cosman, C. Berry, Z. Feng, and W. Schafer. Automatic tracking, feature extraction and classification of *C. elegans* phenotypes. *IEEE Transactions on Biomedical Engineering*, 51:1811–1820, 2004. 1

[9] W. Godinez, M. Lampe, S. Wörz, B. Müller, R. Eils, and K. Rohr. Deterministic and probabilistic approaches for tracking virus particles in time-lapse fluorescence microscopy image sequences. *Medical Image Analysis*, 13:325–342, 2009. 1, 2

[10] A. Hand, T. Sun, D. Barber, D. Hose, and S. MacNeil. Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration. *Journal of Microscopy*, 234:62–79, 2009. 2

[11] K. Li, E. Miller, M. Chen, T. Kanade, L. Weiss, and C. P.G. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis*, 12:546–566, 2008. 1, 2

[12] B. Montrucchio, F. Lamberti, A. Gamba, and G. Serini. Tracking endothelial cells during blood vessel networks assembly using active contours. In *IEEE Workshop on Signal Processing Systems Design and Implementation*, 2005. 2

[13] D. Padfield, J. Rittscher, N. Thomas, and B. Roysam. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Medical Image Analysis*, 13:143–155, 2009. 1, 2

[14] A. Sacan, H. Ferhatosmanoglu, and H. Coskun. Celltrack: An open-source software for cell tracking and motility analysis. *Bioinformatics*, 24(14):1647–1649, 2008. 2

[15] D. Salmond, D. Atherton, and J. Bather. Mixture reduction algorithms for uncertain tracking. *IFAC Proceedings Series*, 2(8):775–780, 1989. 5

[16] W. Saltzman. *Tissue Engineering: Engineering Principles for the Design and Replacement Organs and Tissues*. Oxford University Press, July 2004. 3

[17] C. Shannon. A mathematical theory of communication. *Bell Systems Techical Journal*, 2,7:379–423,623–656, July and October 1948. 7

[18] I. Smal, K. Draegestein, N. Galjart, W. Niessen, and E. Meijering. Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: Application to microtubule growth analysis. *IEEE Transactions on Medical Imaging*, 27:789–804, 2008. 1, 2

[19] E. Storey, T. Spurck, J. Pickett-Heaps, K. Beyreuther, and C. Masters. The amyloid precursor protein of alzheimer's disease is found on the surface of static but not actively motile portions of neurites. *Brain Research*, 735(1):59–66, 1996. 1

[20] P. Vallotton, A. Ponti, C. Waterman-Storer, E. Salmon, and G. Danuser. Recovery, visualization, and analysis of actin and tubulin polymer flow in live cells: A fluorescent speckle microscopy study. *Biophysical Journal*, 85:1289–1306, 2003. 1

[21] J. L. Williams and P. S. Maybeck. Cost-function-based Gaussian mixture reduction for target tracking. In *Proc. Sixth Internatoinal Conference of Information Fusion*, pages 1047 – 1054, Australia, July 2003. 5

[22] L. Wood, A. Das, R. Kamm, and H. Asada. A stochastic broadcast feedback approach to regulating cell population morphology for microfluidic angiogenesis platforms. *IEEE Trans Biomed Eng.*, 6(9):2299–2303, 2009. 1

[23] X. Yang, H. Li, and X. Zhou. Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 53:2405–2414, 2006. 2