

Tracking Colliding Cells

Nhat Nguyen¹, Steve Keller², Toan Huynh², and Min Shin¹

¹ Department of Computer Science, University of North Carolina at Charlotte

² Department of General Surgery, Carolinas Medical Center

{nhnguye1,mcshin}@uncc.edu , {Toan.Huynh,Steven.Keller}@carolinashhealthcare.org

Abstract

The motion of leukocytes is significant in studying the inflammation response of the immune system. In inflammation conditions, some leukocytes slow down and eventually adhere to vessel walls. With many cells moving at a variety of speeds, collisions occur. These collisions result in abrupt changes in the motion and appearance of leukocytes. In this paper, we propose a novel method of tracking multiple cells undergoing collision by modeling the collision states of cells and testing multiple hypotheses of their motion and appearance. We demonstrate the results on in vivo intravital microscopy image sequences. The proposed method results in 20% more accurate cell position and 82% longer tracks than a previous approach.

1. Introduction

The study of leukocyte behaviors is an important area of physiology. For example, rolling velocity and the number per unit area of leukocytes are used to learn about the inflammatory process as well as to design and test anti-inflammatory drugs [5]. Traditionally, detection and tracking of leukocytes in microscopic images are performed manually. This manual process is tedious and time-consuming, thus limiting the amount of data for analysis.

Leukocytes usually flow along with blood. In inflammation conditions, leukocytes slow down to roll along blood vessels and eventually adhere to the vessel walls. Moving at such a wide range of speeds, these cells often collide with each other. The collisions cause abrupt changes in cells' motion and appearance.

Previous works has addressed the automatic detection and tracking of leukocytes from video microscopy. An active contour approach was proposed by Ray *et al.* [7]. Cui *et al.* [1] proposed a Monte Carlo approach to cell tracking that reduces the effect of vessel wall edges when cells are rolling along the vessel walls. These methods rely on

the intensity gradient of cell boundary. In the case of colliding cells, the gradient where the cells collide is minimal. In addition, these methods are only designed for tracking a single rolling leukocyte at a time. Eden *et al.* [2] proposed an automated method with smoothness constraints to extract leukocytes circulatory velocity. However, these smoothness constraints do not hold during collisions. Recently, Smith *et al.* [8] introduced constraints to link the movement of a target and its appearance. For example, one constraint linked elongated cells to high velocities. During collisions, the cells would appear merged and the elongation often reflects multiple cells rather than high velocity. Li *et al.* [6] provided a track linking module which attempted to link broken track segments to complete cell trajectory. This method improved the number of continuous tracks, however, it assumed two physical constraints: that a cell does not vanish unless it leaves the field of view and a cell does not appear unless it enters from outside. These constraints are not valid since our data sets are 2-D slices of complex 3-D vasculature, which means cells can enter and leave anywhere in an image. In this paper, we propose a method for robustly tracking cells undergoing collisions. The cells are tracked by assessing multiple hypotheses of their motion and appearance. These hypotheses are based on possible transitions from the cells' current collision state. The collision state of a cell is recognized based on their motion and appearance *change*. The performance of the method has been demonstrated with the analysis of four *in vivo* intravital microscopy image sequences of leukocytes undergoing collisions. Figure 1 shows how a collision is handled by a previously published method of Eden *et al.* [2] and the proposed method.

2. Methods

In inflammation conditions, more cells start rolling or adhering to the vessel walls, increasing the frequency of collisions. When cells collide, they come to a stop at the point of collision and appear merged, making it difficult to accu-

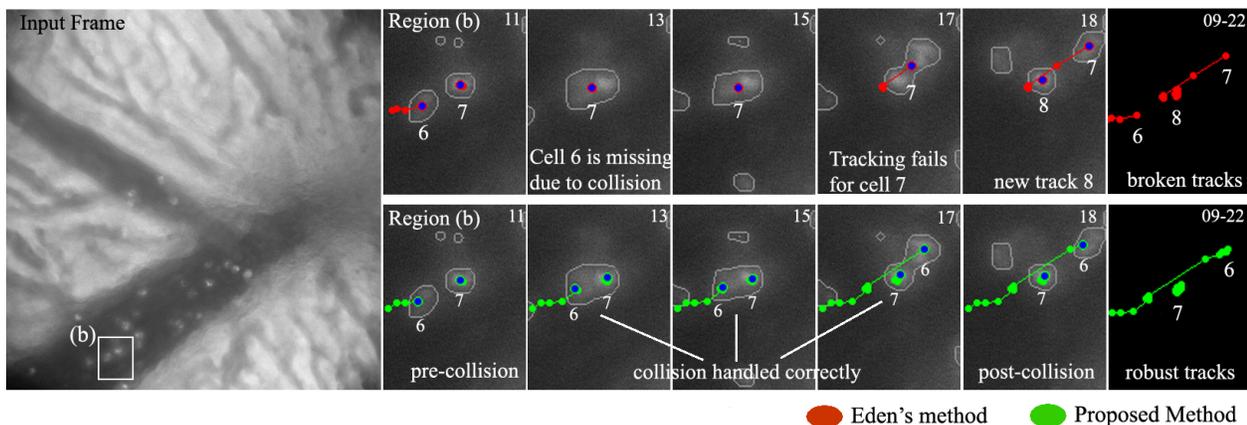


Figure 1. The left image contains a frame of rolling leukocytes in a blood vessel from an intravital microscopy video. Region (b), where a collision takes place, has been enlarged. Cells have been outlined based on the detection results. The results from smoothness constraints method proposed by Eden *et al.* [2] are shown in the first row. The tracks get broken during abrupt change in cell appearance and motion due to collision (top-right image). Our method, is shown in the bottom row, was able to achieve reliable tracking.

rately recognize them as individual cells. In addition, fast-moving cells decelerate quickly to the speed that is almost zero. After collision, the cells often resume their original speed, and they no longer appear merged. Collisions cause two challenges in automatic tracking. First, the motion and appearance during collision could be different from before and after collision. Second, the *change* in cell motion and appearance into as well as out of collision is abrupt. Therefore, the smooth transition between different motion patterns used in traditional motion filtering techniques cannot be assumed.

In this paper, we propose to improve the tracking accuracy by (1) having separate collision states to describe cells during collision as well as before and after collision, and (2) creating and testing multiple hypotheses of cell motion and appearance as transitions between abrupt motion patterns. Note that the recognition of collision in itself is not important. The goal is to improve tracking of cells undergoing collisions by creating and testing multiple hypotheses based on the recognized collision states.

The next sections describe the proposed method in detail. First, the cell detection procedure is illustrated in section 2.1. The collision states are explained in section 2.2. Then, multiple hypotheses of cell motion and appearance as transitions between the collision states is elaborated in section 2.3. Next, section 2.4 illustrates how our method corresponds cells from frame to frame. Finally, the track update process is described in section 2.5. The overview of the proposed method is given in Figure 2.

2.1. Cell Detection

The detection of leukocytes is challenging due to a high noise level in *in vivo* images and a wide ranging appear-

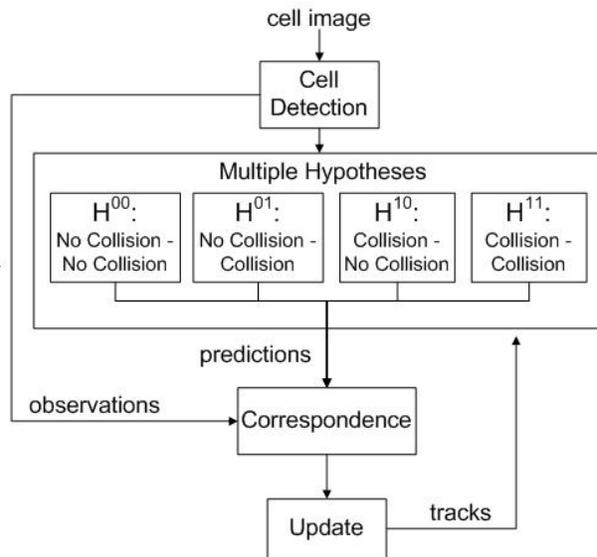


Figure 2. The overview of the proposed method.

ance of leukocytes in intensity and shape. To reliably detect leukocytes, we use a supervised learning approach to learn the appearance and classify each pixel p at location (x, y) as a cell or background. A set of features is computed from a window of $l_w \times l_h$ centered around p . Then, AdaBoost is used to train a classifier [3]. Unlike previous methods which use boosting for biomedical image classification and rely on generic image features [6], we select a set of features based on the perceived visual differences between pixel regions corresponding to cells and those of the background (e.g. vessels, other tissue). For an $l_w \times l_h$

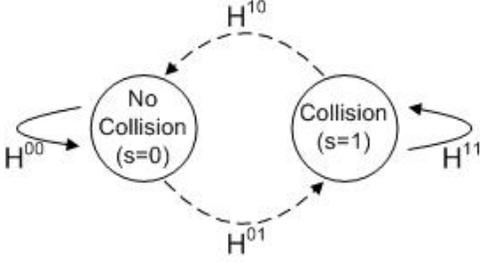


Figure 3. The translation of two collision states. The continuous arrow denotes the smooth change in motion while the dashed arrow denotes the abrupt change in motion. H^ϕ is the appropriate hypothesis for each state transition.

window, our features are: (1) mean intensity, (2) standard deviation of intensity, and (3) normalized radial mean. The normalized radial mean is defined as $\frac{\mu_1}{\mu_1 + \mu_2 + \mu_3}$ where μ_1 , μ_2 , and μ_3 are mean intensities of pixels within radius range of $[0, \frac{\omega}{3}]$, $[\frac{\omega}{3} + 1, \frac{2\omega}{3}]$, and $[\frac{2\omega}{3} + 1, \omega + 3]$, respectively, and ω is the diameter (in pixels) of a typical leukocyte.

We select a set of 100 cell and 100 background samples from various locations of the training sequence. These samples are evenly divided into training and validation sets. Based on the returned decision rule, each pixel is classified as cell or background. Groups of cell pixels are clustered into cell observations using connected component labeling. For cell observation k , the cell location (x_k, y_k) and a_k area are calculated as the centroid and the number of pixels of k , respectively. In frame t , we define the observation vector as $\mathbf{z}_k(t) = [x_k, y_k, a_k]^T$. In the next sections, we present the proposed tracking process, from the definition of collision states to the development of multiple hypotheses of cell motion and appearance.

2.2. Collision States

To handle different cell motion and appearance during collision as well as before and after collision, we use separate collision states: no collision ($s = 0$) or collision ($s = 1$). In the first frame, all cells are initialized to $s = 0$. Let $\phi = s_t s_{t+1}$ denote the transition between two collision states in time t and $t + 1$. We define two types of transitions: staying in the same state ($\phi = \{00, 11\}$) or moving to another state ($\phi = \{01, 10\}$). When $\phi = \{00, 11\}$, the motion pattern is mostly smooth and can be predicted by a traditional Kalman filter [4]. However, when $\phi = \{01, 10\}$, the motion pattern is no longer smooth since the collision results in abrupt changes in motion. The collision state transitions are illustrated in Figure 3. We incorporate multiple hypotheses of motion and appearance to handle transitions between collision states in the next section.

2.3. Multiple Hypotheses

Let H^ϕ be the hypothesis for the state transition ϕ . Since there are four possible state transitions, the multiple hypotheses approach consists of four possible hypotheses: H^{00} , H^{01} , H^{11} , and H^{10} . Each H^ϕ predicts a motion and appearance model. The motion and appearance models are described as

$$\mathbf{x}(t+1) = \mathbf{A}^\phi \mathbf{x}(t) + \mathbf{B}^\phi \mathbf{u}(t) + \mathbf{q}^s(t) \quad (1)$$

where $\mathbf{x}(t) = [x, y, u, v, a]^T$ is the state vector describing cell motion and appearance at time t , (x, y) is the cell position, (u, v) is the cell velocity, and a is the cell area. \mathbf{A}^ϕ and \mathbf{B}^ϕ are the state transition and control input matrices for the hypothesis H^ϕ , $\mathbf{u}(t)$ is a control input vector which is used to accommodate sudden changes in cell motion and appearance, $\mathbf{q}^s(t)$ is the process noise vector, which is a Gaussian process with zero mean and covariance \mathbf{Q}^s . The motion and appearance models share a measurement model:

$$\mathbf{z}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{r}(t) \quad (2)$$

where $\mathbf{z}(t)$ is the measurement vector containing the position (x, y) and the area a , \mathbf{C} is the measurement transition matrix, and $\mathbf{r}(t)$ is the measurement noise vector, which is a Gaussian process with zero mean and covariance \mathbf{R} .

Let $\mathbf{x}_i(t) = [x_i, y_i, u_i, v_i, a_i]^T$ be the state vector of cell i . The proposed method initializes the state vector $\mathbf{x}_i(0) = [x_i^0, y_i^0, u_i^0, v_i^0, a_i^0]^T$ of cell i without explicit tracking. While the position (x_i^0, y_i^0) and area a_i^0 can be initialized directly from the observation vector $\mathbf{z}_k(0)$, the velocity (u_i^0, v_i^0) is initialized based on the adherence probability ρ_i . Let $\{\forall \mathbf{z}_k(t) = [x_k, y_k, a_k]^T, t = 1..T\}$ be the set of observations (obtained from Section 2.1) where T is the length of the sequence. Adherence probability ρ_i is defined as the ratio $\frac{n_f}{T}$. We compute n_f as the number of consecutive frames t where $\mathbf{z}_k(t)$ contains a location (x_k, y_k) such that the Euclidean distance between (x_k, y_k) and (x_i^0, y_i^0) is less than the leukocyte diameter ω . Note that when a cell has high ρ_i , it tends to move slowly or to remain adhered. Thus, the velocity is initialized to be a small value. When a cell has low ρ_i , it tends to move along with the blood flow direction θ . The velocity then is initialized as ψ , which is the average speed of leukocytes. We initialize the velocity as

$$[u_i^0, v_i^0] = (1 - \rho_i)\psi[\cos \theta, \sin \theta] \quad (3)$$

The blood flow direction θ can be automatically estimated using vessel segmentation. However, the vessel segmentation method is out of the scope of this paper. We estimate the blood flow direction of a vessel manually by clicking two positions along the vessel.

The predicted state vector $\hat{\mathbf{x}}_i^\phi(t+1)$ and predicted covariance $\hat{\mathbf{P}}_i^s$ of hypothesis H^ϕ are computed using Kalman filter's prediction step [4]:

$$\hat{\mathbf{x}}_i^\phi(t+1) = \mathbf{A}^\phi \mathbf{x}_i(t) + \mathbf{B}^\phi \mathbf{u}_i(t) \quad (4)$$

$$\hat{\mathbf{P}}_i^s = \mathbf{A}^\phi \mathbf{P}_i^s \mathbf{A}^{\phi T} + \mathbf{Q}^s \quad (5)$$

where \mathbf{P}_i^s is the covariance of cell i . In the next subsections, we derive \mathbf{A}^ϕ and \mathbf{B}^ϕ of Equation 4 to reflect the changes in each hypothesis H^ϕ .

2.3.1 H^{00} : No Collision \rightarrow No collision.

When cell i is not colliding, its motion is rather smooth. We derived \mathbf{A}^{00} so that Equation 4 uses a constant velocity motion model.

$$\mathbf{A}^{00} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Since \mathbf{B}^{00} and $\mathbf{u}_i(t)$ are not required in describing smooth motion, we set them to zero matrix and vector, respectively.

2.3.2 H^{01} : No Collision \rightarrow Collision.

When cell i collides with cell j , both cells usually stop at the point of collision (\hat{x}, \hat{y}) . We first estimate $\hat{\mathbf{x}}_i^{00}(t+1)$ of cell i and $\hat{\mathbf{x}}_j^{00}(t+1)$ of cell j using hypothesis H^{00} . Let L_i be the trajectory line formed by the location component (x_i, y_i) of the state vector $\mathbf{x}_i(t)$ and (\hat{x}_i, \hat{y}_i) of the predicted state vector $\hat{\mathbf{x}}_i^{00}(t+1)$. Similarly, L_j is the trajectory line of cell j . Figure 4 illustrates a collision location of two cells. Since a typical cell is ω pixels in diameter, the collision location is not the intersection of L_i and L_j ; it could happen even if L_i and L_j did not intersect. Thus, we define $d_i(x, y)$ and $d_j(x, y)$ as the distance transforms at pixel location (x, y) of line L_i and L_j , respectively. The point of collision can be estimated as

$$(\hat{x}, \hat{y}) = \arg \max_{x, y} (\Delta(x, y)) \quad (6)$$

where $\Delta(x, y) = e^{-d_i(x, y)^2/\omega} * e^{-d_j(x, y)^2/\omega}$.

To predict the state vector using Equation 4, \mathbf{A}^{01} and \mathbf{B}^{01} are set as below:

$$\mathbf{A}^{01} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{B}^{01} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Since \mathbf{A}^{01} contains all zeros in the first four rows, the first four components of $\hat{\mathbf{x}}_i(t+1)$ do not depend on $\mathbf{x}_i(t)$. Rather, they depend on the vector $\mathbf{u}_i(t) = [\hat{x}, \hat{y}, u_i, v_i, a_j]^T$ where (\hat{x}, \hat{y}) is the predicted collision location calculated in Equation 6, (u_i, v_i) is the current velocity of cell i and are

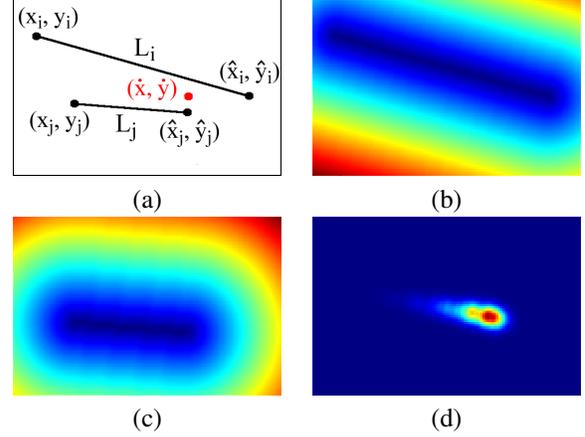


Figure 4. Finding collision location of two cells. (a) Trajectory lines L_i and L_j using H^{00} . The collision location (\hat{x}, \hat{y}) denoted by the red dot. In (b) (c) and (d), we visualize in color scale using jet color-map (blue presents lower value and red presents higher value) (b) Distance transform of L_i . (c) Distance transform of L_j (d) Values of Δ at all (x, y) locations. Note that the reddish blob contains the maximum value of Δ , and the location of that value can be estimated as the collision location.

estimated from its state vector, a_j is the area of cell j and is estimated from its state vector. Thus, Equation 4 becomes $\hat{\mathbf{x}}_i^{01}(t+1) = [\hat{x}, \hat{y}, 0, 0, a_i + a_j]^T$. Here, the cell velocity is adjusted to zero to compensate for sudden deceleration. Since colliding cells appear merged, the hypothesized area is estimated as the sum of areas of colliding cells.

2.3.3 H^{11} : Collision \rightarrow Collision.

During collision, cells move at much slower speed. However, their motion is still smooth. Therefore, the state transition and optional control matrices are the same as described in Section 2.3.1. However, the collision state s has been set to 1. In order to preserve the cell motion though out the collision, a separate covariance matrix \mathbf{P}^s is used.

2.3.4 H^{10} : Collision \rightarrow No Collision.

When colliding cells split away from one another, they often accelerate to their original velocity and appear as individual cell. The state transition and optional control matrices are designed so that cells' original velocity and area are used as prediction:

$$\mathbf{A}^{10} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{B}^{10} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Thus, Equation 4 becomes $\hat{\mathbf{x}}_i^{10}(t+1) = [x_i + u_i, y_i + v_i, u_i, v_i, a_i - a_j]^T$. The original velocity (u_i, v_i) , which

is preserved in \mathbf{u}_i , is used as the predicted velocity. The original area is also predicted as the different between the merged cell area and the area of the other colliding cell j , which is also preserved in \mathbf{u}_i . In the next section, the predictions $\hat{\mathbf{x}}_i^\phi(t+1)$ from H^ϕ is compared with the observations $\mathbf{z}_k(t+1)$ from section 2.1 to correspond cells.

2.4. Correspondence

To track multiple colliding cells, we establish correspondences between cells in frame $t-1$ and the observations in frame t . In frame t , multiple hypotheses are generated for each cell i . If cell i 's collision state $s=0$, hypotheses H^{00} and H^{01} are generated. While there's only one H^{00} , multiple H^{01} hypotheses can be generated for all pairs of i and j such that the Euclidean distance between locations of i and j is less than 2ψ where ψ the average speed of leukocytes. If cell i 's collision state $s=1$, hypothesis H^{11} and hypothesis H^{10} are generated. Given $\hat{\mathbf{x}}_i^\phi(t)$ from each hypotheses, the measurement model is described as

$$\hat{\mathbf{z}}_i^\phi(t) = \mathbf{C}\hat{\mathbf{x}}_i^\phi(t) \quad (7)$$

where $\hat{\mathbf{z}}_i^\phi(t) = [x_i, y_i, a_i]^T$ is the hypothesized measurement vector, \mathbf{C} is the measurement transition matrix.

For cell i and observation k , the error of each hypothesis $\varepsilon_{i,k}^\phi(t)$ is computed as the weighted sum of squared of different between its hypothesized measurement $\hat{\mathbf{z}}_i^\phi(t)$ and the observations \mathbf{z}_k . This error can be computed as

$$\varepsilon_{i,k}^\phi(t) = \mathbf{w}^\phi(\mathbf{z}_k(t) - \hat{\mathbf{z}}_i^\phi(t)) \quad (8)$$

where \mathbf{w}^ϕ is the weight used to regulate the importance of each component in the state vector. \mathbf{w}^ϕ depends on the hypothesis. For example, \mathbf{w}^{01} , emphasizes the area component since the increase in area is the main indication of a collision. Cell i is corresponded to observation k in a greedy manner by establishing correspondence which returns the lowest error first. When a cell is corresponded, all of its hypotheses are removed from further consideration.

The greedy method continues corresponding cells until $\varepsilon_{i,k}^\phi(t) > \delta$ where δ is an error threshold of unlikely (i, k) pairs. Any remaining observation $\mathbf{z}_h(t)$ which has $a_h > \pi \frac{\omega^2}{4}$, where ω is the diameter of typical leukocyte, is considered for tracking as a *new* cell. Any remaining cell g which does not have a correspondence is continued to be predicted in the next frame using its predicted state vector $\hat{\mathbf{x}}_g^\phi(t)$ without the update step (section 2.5). If cell g is not corresponded for three frames, it is considered to have disappeared from the video and is no longer tracked.

2.5. Update

After obtaining the predicted state vector $\hat{\mathbf{x}}_i^\phi(t)$, predicted covariance $\hat{\mathbf{P}}_i^s$, and the correspondence observation

$\mathbf{z}_k(t)$, we used Kalman filter's update step [4] to update the state vector $\mathbf{x}_i^\phi(t)$ and covariance \mathbf{P}_i^s :

$$\mathbf{x}_i^\phi(t) = \hat{\mathbf{x}}_i^\phi(t) + \mathbf{K}_i(t)(\mathbf{z}_k(t) - \mathbf{C}\hat{\mathbf{x}}_i^\phi(t)) \quad (9)$$

$$\mathbf{P}_i^s = \hat{\mathbf{P}}_i^s - \mathbf{K}_i(t)\mathbf{C}\hat{\mathbf{P}}_i^s \quad (10)$$

where $\mathbf{K}_i(t) = \hat{\mathbf{P}}_i^s \mathbf{C}'(\mathbf{C}\hat{\mathbf{P}}_i^s \mathbf{C}' + \mathbf{R})^{-1}$ is the Kalman gain, $\mathbf{x}_i^\phi(t)$ is the updated state vector, \mathbf{P}_i^s is the covariance of cell i in state s . Note that each cell i maintains two separate covariances $\mathbf{P}_i^s (s=0,1)$ and each gets updated depending on the cell current state s . Thus, the abrupt change in collision ($s=1$) does not effect the motion and appearance of a cell when it returns to no collision state ($s=0$).

The updated state vector $\mathbf{x}_i^\phi(t)$ is added to a track segment which is a set of state vectors that are corresponded to cell i . In the next time step, $\mathbf{x}_i^\phi(t)$ is used to compute the predictions.

3. Experiments

3.1. Data

Our data set consists of four *in vivo* image sequences of leukocytes in liver capillaries. These image sequences were acquired using an intravital microscope. Each sequence has an average of 200 frames captured at 1000 x 1000 pixel resolution and at 30 frames/second. We obtained a ground truth from a biology technician who manually tracked the cells' centroid in the main vessel region for a total of 124 cells (1936 cell positions).

An independent sequence along with its ground truth was also acquired using the same imaging protocol to use as training data. The training samples in the detection procedure are obtained from this sequence. In addition, using this training sequence, we estimate and fine-tune several parameters including the process noise covariance \mathbf{Q}^s , measurement noise covariances \mathbf{R} , set of weights \mathbf{w}^ϕ , and error threshold δ . The diameter of leukocyte ω and leukocyte's average speed ψ are estimated by the biology technician.

3.2. Evaluation

We evaluated and compared our method against two existing methodologies. First is the work by Eden *et al.* [2] is a method for tracking multiple cells in *in vivo* microscopic video similar to our data. Second is the single hypothesis Kalman-based method which is equivalent to the proposed method using only the constant velocity motion model in subsection 2.3.1. Two methods were chosen to access the effect of proposed multiple hypotheses approach on tracking colliding cells.

Each method is evaluated using the three metrics in which the first two were adapted from Cui *et al* [1]. First,

	colliding cells only			all cells		
	Eden [2]	Kalman	Proposed	Eden [2]	Kalman	Proposed
RMSE (pix)	7.1	7.5	5.7	7.4	6.6	5.9
PTP (%)	53.3	67.1	92.4	56.2	66.9	74.1
TL (frames)	7.1	14.0	19.2	9.1	15.1	16.7

Table 1. Comparisons of the proposed method to the Eden [2] and Kalman methods in accuracy (RMSE), number of positions tracked (PTP), and average track length (TL). The best performance value in each metric is bold faced.

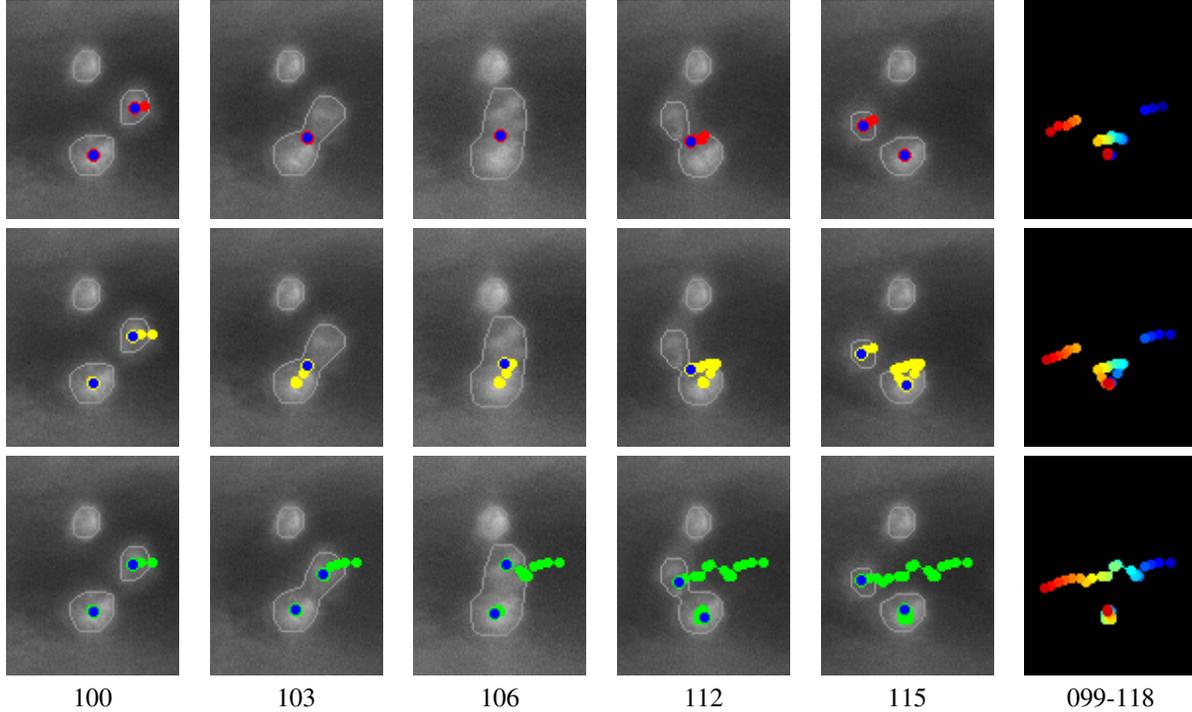


Figure 5. Example of the results on a cell colliding and rolling around an adhered cell. Cells have been outlined based on the detection results. Tracking results are provided in sequences in the 100th, 103rd, 106th, 112th and 115th frames. The last column illustrates the entire tracks in time order (blue denotes earlier frames and red denotes later frames). Result of Eden [2] (top row in red), Kalman (middle row in yellow) and our methods (bottom row in green) are shown. As cells collide (frame 103), both of the Eden and Kalman method failed to recognize individual cells. Our method was able to reliably track two cells. Also note the improvement in accuracy of position during collision (refer to right most images).

the root mean squared error (RMSE) of positions is computed. Second, the positions tracked percentage (PTP) of ground truth tracks is computed as the percentage of frames with automatic tracking within $\frac{\epsilon}{2}$ pixels from ground truth. Third, the average track length (TL) is computed for all automatic tracks corresponding to ground truth tracks. Note that as tracking gets broken up, the average track length will decrease.

The results are shown in Table 1. The data support the proposed method's performance over both the Eden and Kalman methods in tracking colliding cells as well as tracking all cells. Considering only colliding cells, the location predictions in multiple hypotheses helped provide 20% more accurate in cell positions (in terms of RMSE) when

comparing with the Eden method and 25% more accurate when comparing with the Kalman method. Both the Eden and Kalman methods return less accurate results in RMSE because they predict locations based on smooth *change* in motion which does not hold in cell collisions. In addition, multiple hypotheses of cell motion and appearance enables the proposed method to track 92% of the total ground truth cell positions. This is 74% more cell positions than the Eden method and 38% more than the Kalman method. During collision when cells appear merged and are detected as one observation, the proposed method is able to recognize multiple colliding cells by generating and selecting the appropriate hypotheses. Average track length is also improved significantly from 7.1 frames in Eden and 14.0 frames in

Kalman to 19.2 frames in the proposed method. As multiple cells appear merged with each other during collision and then split away from each other after the collision, multiple broken tracks are created by both the Eden and Kalman methods. The proposed method recognizes the transition between different collision states and thus produces continuous and longer tracks. The visual comparisons are shown for a cell colliding and rolling around (Figure 5).

For all cells, the proposed method improves overall accuracy by reducing RMSE by 20% comparing with Eden and 10% with Kalman. Proposed method's TL remains relatively high (82% more than Eden and 11% more than Kalman). Although the improvements in PTP are still significant (32% and 11% to Eden and Kalman, respectively), the proposed method achieves only 74% of the total ground truth positions. This is due to two reasons. First, the detection fails when cells are going out of the imaging plane. Second, an incorrect selection of transition between collision states results in incorrect prediction of motion and appearance. The proposed method does not have a way to recover back to the previous state and the error simply accumulates. recover from incorrectly selected hypotheses.

Collisions of more than two cells are not explicitly addressed in this paper. They could be handled by generating additional hypotheses of already colliding cells to be collided with new ones. Note that with an increasing number of colliding cells, the correct tracking after splitting would be more difficult as the appearance and motion become less distinctive.

4. Conclusion

The leukocytes in inflammation conditions often begin to roll and eventually stop resulting in collisions. We proposed a method for reliably tracking multiple cells during collisions in *in vivo* microscopy image sequences. The method incorporated Kalman filter and multiple hypotheses for cell motion and appearance during as well as before and after collision. As part of future work, we plan to further improve tracking of leukocytes as follows. First, we are currently adding appropriate features to the supervised learning method to improve detection results. Second, we focus on incorporating a probabilistic approach to the transition between collision states to recover from incorrectly selected hypotheses. We plan to expand this frame work for tracking colliding cells with more complex and unpredictable motions.

Acknowledgments

This study was supported by grants from the National Science Foundation (NSF-DBI-0754748). We would like to thank Eric Norris for the ground truthing of data, and Richard Souvenir for implementation support.

References

- [1] J. Cui, S. T. Acton, and Z. Lin. A monte carlo approach to rolling leukocyte tracking in vivo. *Medical Image Analysis*, 10(4):598 – 610, 2006. Special Issue on Functional Imaging and Modelling of the Heart (FIMH 2005).
- [2] E. Eden, D. Waisman, M. Rudzsky, H. Bitterman, V. Brod, and E. Rivlin. An automated method for analysis of flow characteristics of circulating particles from in vivo video microscopy. *Medical Imaging, IEEE Transactions on*, 24(8):1011–1024, Aug. 2005.
- [3] Y. Freund and R. Schapire. A short introduction to boosting. *Japanese Society for Artificial Intelligence*, 14(5):771–780, 1999.
- [4] Kalman. A new approach to linear filtering and precision problems. *Trans. ASME: J. Basic Eng.*, pages 35–45, 1960.
- [5] K. Ley. Leukocyte recruitment as seen by intravital microscopy. *Physiology of Inflammation*, pages 303–337, 2001.
- [6] 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, 2008. Special issue on the 10th international conference on medical imaging and computer assisted intervention - MICCAI 2007.
- [7] N. Ray, S. Acton, and K. Ley. Tracking leukocytes in vivo with shape and size constrained active contours. *Medical Imaging, IEEE Transactions on*, 21(10):1222–1235, Oct. 2002.
- [8] K. Smith, A. Carleton, and V. Lepetit. General constraints for batch multiple-target tracking applied to large-scale videomicroscopy. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.