MULTI-CELL 3D TRACKING WITH ADAPTIVE ACCEPTANCE GATES

Michael Landau[1], Ekaterina Koltsova[2], Klaus Ley[2], and Scott T. Acton[1]

[1]Dept. of Electrical Engineering, University of Virginia, Charlottesville, VA 22908;
[2]Division of Inflammation Biology, La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037

ABSTRACT

Recent advances in multiphoton microscopy have allowed the capture of dendritic and T cells on video in a 3D volume. This paper reports on an approach for automatically detecting and tracking the cells and collecting statistics on their characteristics of motion and interaction durations. A novel method to extend the track longevity is presented, where an adaptive acceptance gate is computed based on the local target density. Results are provided that show that the total number of track segments was reduced by 22% and the percentage of tracks that lasted longer than half the video increased from 12% to 17% of the total tracks.

Index Terms— Kalman Filter, Target Tracking, Biological Image Analysis

1. INTRODUCTION

Characteristics observed during interactions between dendritic cells (DCs) and T cells are relevant to the field of inflammation biology. For example, long cell interaction times are thought to be productive for antigen presentation. The desired post-processed information includes statistics on the change in speed, direction of movement, tortuosity, confinement ratio, motility, and interaction duration. Novel techniques to capture the video depicting the physical interactions of these cells are being explored. One such method is to image the cells with multiphoton microscopy in the presence of antigen. By way of multiphoton micoscopy, the DCs and T cells are filtered into two separate video streams. The DCs are displayed as green masses, and the T cells are displayed as red masses traveling across a 3-dimensional cellular space (Figure 1).

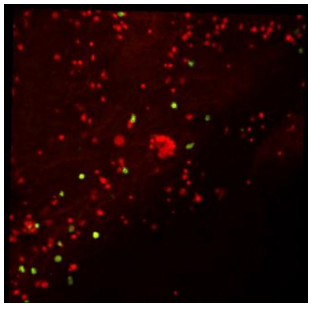

Figure 1: 3D cellular space containing DCs and T cells.

With this distinct separation of colors, the tasks of detecting and tracking the cells are achievable. This paper reports on a process that automatically tracks DCs and T cells and provides useful information about their characteristics of motion. Previously, our research group has examined the problem of tracking cells in sequences of 2D images using active contours [1] [2] and particle filters [3]. In this paper, we consider the problem of tracking a large quantity of cells captured in 3D movies.

Various methods are utilized to initiate and retain tracks for each of the cells that exist during a recorded video. This is not straightforward, however, because there are problems with cell congregation and arbitrary cell motion. Often, a large number of T cells tend to migrate around a single DC, making the detection of each individual T cell difficult to discern. Additionally, when T cells are not interacting with DCs, they tend to travel alone for much longer distances at fast and non-linear rates. Because of this, tracks are difficult to maintain and are often lost. Adapting the tracking function separately to each cell's local density condition is hypothesized to increase total track time performance and is a key focus of this paper.

The number of objects in a covariance volume has been shown to be the key parameter for predicting track data association performance [4]. This parameter depends on the product of covariance volume with target density. Target density, however, is not generally computed and exploited by multi-target tracking systems. The approach taken in this paper is to use a local track density estimate, computed separately for each track, to compute an adaptive measurement association gate. Performance results utilizing an adaptive gate are compared against the results obtained utilizing fixed measurement gates.

Bhattacharyya and Chakrabarti derived a formula for computing the mean nth nearest neighbor distance of objects in a uniform D-dimensional Euclidean space as a function of target density [5]. Their equation yields a robust estimator for density as a function of the nth nearest neighbor distance. The method proposed in this paper is to average the density estimates corresponding to N of the nth nearest neighbor distances.

2. METHOD AND THEORY

The system that implements the multi-cell tracking (MCT) algorithm contains two major components: target detection

and target tracking as shown in Figure 2. These are described in detail in the following sections.

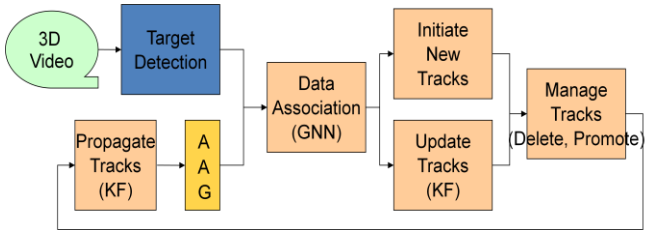

Figure 2: Multi-cell tracking system with detection and tracking functions.

2.1. Target Detection

For each incoming 3D video frame, the MCT's detection processing creates the new measurements for the DC and T cells used by the tracker. This is accomplished by way of binary clustering of cell intensities that separates the desired moving targets from the unwanted background pixels that contribute to the video space. Then spatial morphological techniques generate a list of binary connected components (BCCs) which further filters out any unwanted noise primarily leftover from residual tissue. 3D target centroid locations as well as 3D size and shape features are extracted for each detected BCC. After this routine is complete, these measurements are sent to the MCTs tracking function.

2.2. Propagation and Updating of Tracks

The Kalman filter (KF) is used to provide the target estimation portion of the video tracker. The KF sequentially propagates and updates the kinematic state x_k, which consists of position and velocity, of each moving target that travels across the 3D target space spanned by the fixed camera. The KF assumes the measurement and motion models, (1) and (2), respectively.

$$Z_k = Hx_k + \tilde{n}_k \quad (1)$$
$$x_k = \Phi x_{k-1} + \tilde{v}_k \quad (2)$$

Z_k is the measurement vector obtained from the target detection routine at time k. The density functions of the measurement noise $\tilde{n}_k$ and the target process noise $\tilde{v}_k$ are assumed to be Gaussian and zero mean. H is the observation matrix that sifts the positional part of the state vector. Φ is the one-step propagation matrix for position and velocity. Given a prior predicted state estimate, denoted by $\hat{x}_k^{(-)}$, the updated state estimate $\hat{x}_k^{(+)}$ can be calculated as a weighted average using a gain matrix K.

$$\hat{x}_k^{(+)} = (I - KH)\hat{x}_k^{(-)} + KZ_k \quad (3)$$

From the definition of the track error covariance matrix

$$P_k^{(+)} = E\left[\tilde{x}_k^{(+)}\tilde{x}_k^{(+)^T}\right] \quad (4)$$

the updated covariance can be described by

$$P_k^{(+)} = (I - KH)P_k^{(-)}(I - KH)^T + KRK^T \quad (5)$$

R is the covariance of the measurement noise $\tilde{n}_k$. The Kalman gain matrix K is then derived as a minimum mean-square error estimator:

$$K = P_k^{(-)}H^T\left(HP_k^{(-)}H^T + R\right)^{-1} \quad (6)$$

Finally, the one-step propagation of the discrete state estimate and error covariance matrix can be described through (7) and (8).

$$\hat{x}_k^{(-)} = \Phi\hat{x}_{k-1}^{(+)} \quad (7)$$
$$P_k^{(-)} = \Phi P_{k-1}^{(+)}\Phi^T + Q \quad (8)$$

The addition of the covariance Q to (8) accounts for uncertainty due to unpredictable process noise $\tilde{v}_k$ [6].

2.3. Data Association

After the tracks are propagated, they are assigned (in a new frame) to the measurements provided by the detection processor. This is accomplished via the global nearest neighbor (GNN) algorithm [7]. The likelihood of the association of the measurement Z_k with the track state $\hat{x}_k^{(-)}$ is given by (9).

$$f = \frac{1}{(2\pi)^{\frac{D}{2}}|S|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}\left(Z_k - H\hat{x}_k^{(-)}\right)^T S^{-1}\left(Z_k - H\hat{x}_k^{(-)}\right)\right\} \quad (9)$$
where $S = HP_k^{(-)}H^T + R$ (10)

S is the covariance of the residual $Z_k - H\hat{x}_k^{(-)}$, and D is the dimension, which in this experiment is set to 3. For each track detection pair, a score Ψ is computed that is twice the negative log association likelihood f, which is related to the Mahalanobis distance D_M via (11).

$$\Psi = -2\log(f) = D_M + \log|S| + D\log(2\pi) \quad (11)$$
where $D_M = \left(Z_k - H\hat{x}_k^{(-)}\right)^T S^{-1}\left(Z_k - H\hat{x}_k^{(-)}\right)$ (12)

Starting with promoted tracks, the tracks are scanned and the measurement with the smallest (best) Ψ is assigned and marked as unavailable for the other tracks. Promoted tracks are defined as tracks that have had more than a user specified number of updates. This differentiates mature tracks from those that have just been started. However, the assignment is rejected if the D_M of the track assignment falls

outside of a fixed chi-squared measurement gate-threshold, T_{Fixed} (displayed in Figure 3). Only if the track assignment is accepted is the track updated. After all tracks are processed, unassigned measurements are used to initiate new tracks, while the unassigned tracks have their miss counters incremented.

The manage track function promotes tracks that have the required number of updates but also deletes tracks that have not had updates for a significant period of time. The track deletion count is an adjustable parameter, and was set to 2 frames for the T cell and DC tracks. Tracks on cells that maneuver outside of the measurement gate are generally lost but could have been preserved if the chi-squared threshold for the Mahalanobis distance were larger. Larger gates can be deployed for targets that are in locally less dense environments. In more dense environments, cell tracks can be susceptible to swapping with other cell tracks. In this inflammatory biology application, track swapping would result in unreliable cell interaction characteristics and therefore must be minimized.

2.3.1. Local Density Estimation and the Adaptive Acceptance Gate Extension to MCT

The likelihood function in (9) can also be thought of as a probability density function for the new measurement Z_k given that it belongs to the same target that has the state estimate $\hat{x}_k^{(-)}$. An analogous null hypothesis H_0 would provide a probability density for the new measurement given that it belongs to a different target. The assumption that targets can exist anywhere in the 3D space motivates the desire for using a uniform density model for this H_0. The height of the uniform density function (β) would be the inverse of a 3D volume element corresponding to the typical separation of objects. So, the volume element is the inverse of the local target density. As shown in [8], we can use this density in a ratio hypothesis test. In the context of [8], in which the hypothesis test is derived, we also make the assumption that all cells are detected (no missed detections), and all detections come from real cells (no false detections).

The goal now is to compute an adaptive acceptance gate (AAG) threshold T_{AAG} for D_M that depends on local target density. In other words, we will determine a threshold for D_M that would guarantee that the alternate hypothesis H_a likelihood is greater than the H_0 likelihood. For the ratio of these likelihoods to be greater than 1,

$$f > \beta \tag{13}$$

Applying (11) to the ratio in (13), we obtain

$$\Psi < -2\log(\beta) \tag{14}$$

Solving for D_M, (14) becomes

$$D_M < -2\log\left(\beta|S|^{1/2}\right) + D\log(2\pi) = T_{AAG} \tag{15}$$

Mori *et al.* [4] derived a related expression for the probability of correctly assigning the correct measurements to tracks, P_{ca}, when an acceptance gate is not deployed, i.e. when it is infinite. This is described in (16) and (17) as follows:

$$P_{ca} = \exp\left(-C_D\beta|S|^{1/2}\right) \tag{16}$$

$$C_D = \left(2\sqrt{\pi}\right)^{(D-1)} * \frac{\Gamma\left(\frac{D+1}{2}\right)}{\Gamma\left(\frac{D}{2}+1\right)} \tag{17}$$

Here, C_D is a constant that depends on the dimension D and Γ is defined as the Gamma function. From (15) and (16) we see that a key parameter is the product of density and covariance volume $\beta|S|^{1/2}$. The derivation for P_{ca} assumed that targets are located uniformly on a D-dimensional sphere and that the number of targets follow a discrete Poisson density.

Now, also making the uniform and Poisson statistics assumption, Chakrabarti *et al.* [5] derived a formula for mean nth nearest neighbor distance $\langle d_{nn}(n)\rangle$ as a function of local density β in a D-dimensional space given by:

$$\langle d_{nn}(n)\rangle = \frac{\Gamma^{1/D}\left(\frac{D}{2}+1\right)}{\sqrt{\pi}} * \frac{\Gamma\left(\mathrm{n}+\frac{1}{D}\right)}{\Gamma(\mathrm{n})} * \frac{1}{\beta^{1/D}} \tag{18}$$

Solving (18) for β provides an equation that depends on the nearest neighbor, n. Here we exploit Bhattacharyya and Chakrabarti's formula by averaging the density estimates for each of N nth nearest neighbor distances which results in the following robust estimator for density:

$$\hat{\beta}_i = \frac{\Gamma\left(\frac{D}{2}+1\right)}{N\sqrt{\pi}^D} * \sum_{n=1}^{N} \frac{\Gamma^D\left(\mathrm{n}+\frac{1}{D}\right)}{\Gamma^D(\mathrm{n})} * \frac{1}{(d_{nn}(n)_i)^D} \tag{19}$$

The AAG extension to MCT computes the acceptance gate size for each track separately. Here, $d_{nn}(n)_i$ is the nth nearest neighbor distance from track i. For each track, N $d_{nn}(n)_i$ are computed to find $\hat{\beta}_i$, and (15) and (19) are used to solve for the track's D_M threshold. Figure 3 plots the correct association probability P_{ca} and the track miss probability P_{miss} as a function of normalized density $\beta|S|^{1/2}$. P_{miss}, in blue, represents the probability that the correct measurement will fall outside the adaptive acceptance gate leading to a potential track drop. P_{ca} reflects the probability of associating the correct measurement when no gating is employed. Miss-associations can result in track swaps which is especially undesirable in this application. Therefore, the gating feature of MCT serves to prefer track drops to track swaps. The figure also shows that the baseline MCT without AAG employs a fixed D_M threshold set to correspond to a P_{miss} of 0.2. This region displays where the AAG extension will reduce track drops and where it will reduce track swaps by virtue of its adaptive nature.

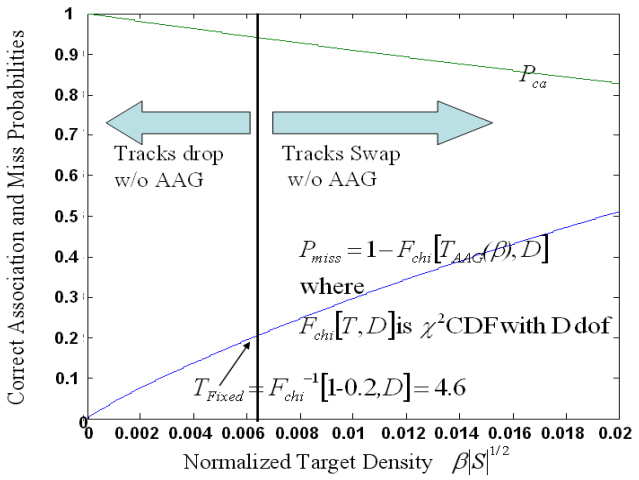

Figure 3: The correct association and miss probability is plotted against target density.

3. RESULTS AND DISCUSION

The input video used to test the tracking program is recorded DCs and T cells traveling through a 3D space inside the aortic wall of a mouse. The size of the target space is 512 pixels by 512 pixels by 15 pixels, which corresponds to a 777µm x 777µm x 225µm box, and the video exists for 60 frames, which corresponds to 62.05 minutes of recorded time. It is estimated that about 400 T cells and about 40 DCs exist in the target space for the duration of the video.

Figure 4 provides a reverse cumulative distribution plot for tracking longevity results when the data was run with the original MCT and when utilizing the AAG extension. The data demonstrate the desirable effect of fewer track drops via fewer total tracks, fewer short tracks and more long tracks. Highly maneuverable cells were tracked when they were in relatively low density neighborhoods. Moreover, the

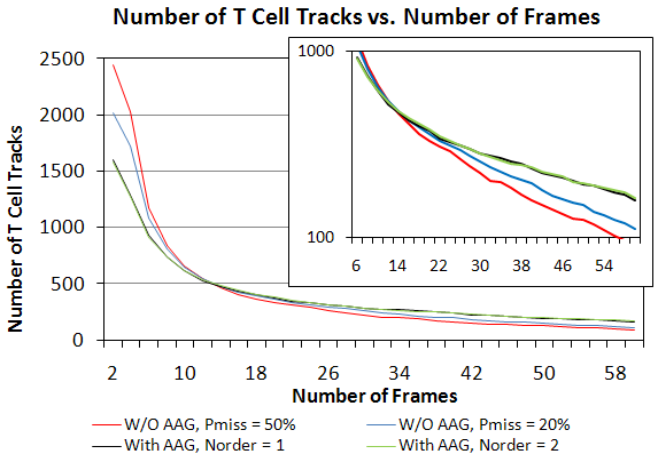

Figure 4: The number of tracks that existed for at least a number of frames is plotted for 4 cases.

AAG extension to MCT minimizes track swaps by adaptively shrinking the gate in higher density neighborhoods.

With the addition of the AAG, the total number of T cell tracks was reduced by 22% when comparing to the trial with a fixed P_{miss} of 20%. Moreover, tracks that lasted for more than 30 frames were increased from 12% to 17% of the total tracks.

4. CONCLUSION

Adapting the acceptance gate for new measurements based on averaging N of the nth nearest neighbor distances (i.e. target density) improves the longevity of tracks for highly maneuverable T cells, and reduces the chance of swapping tracks in highly dense areas. There are high density and high maneuvering instances in which the T cell track would inevitably be lost. These include the cases when T cells interact with DCs for a prolonged time where other T cells also congregate.

5. FUTURE WORK

The MCT software can be upgraded to include advanced data association techniques that will permit feature extraction from group tracks that are defined as closely spaced cells clustered together. Additionally, an adaptive filter such as an interacting multiple model (IMM) can be deployed to both smooth the tracks when the cells are quiescent and increase the filter agility when the cells are maneuvering.

6. REFERENCES

[1] N. Ray, S.T. Acton, and K. Ley, "Tracking leukocytes in vivo with shape and size constrained active contours," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 1222-1235, 2002.

[2] N. Ray and S.T. Acton, "Motion gradient vector flow: an external force for tracking rolling leukocytes with shape and size constrained active contours," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 1466-1478, 2004.

[3] J. Cui, S.T. Acton and Z. Lin, "Tracking rolling leukocytes in vivo by factored sampling," *Medical Image Analysis*, vol. 10, pp. 598-610, 2006.

[4] K.C. Chang, S. Mori, and C.Y. Chong, "Performance Evaluation of Track Initiation in Dense Target Environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 1, pp. 213-219, 1994.

[5] P. Bhattacharyya and B.K. Chakrabarti, "The mean distance to the nth neighbor in a uniform distribution of random points: an application of probability theory," *European Journal of Physics*, vol. 29, pp. 639-645, 2008.

[6] A. Gelb (editor), *Applied Optimal Estimation*, M.I.T. Press, 1974.

[7] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, British Library Cataloguing in Publication Data, London, England, 1999.

[8] Y. Bar-Shalom, S. Blackman, and R. Fitzgerald, "Dimensionless Score Function for Multiple Hypothesis Tracking," *Proceedings of SPIE Conference on Signal and Data Processing of Small Targets*, San Diego, CA, pp. 5913-50, 2005.