EVALUATION OF INTRAVITAL TRACKING ALGORITHMS

Jinshan Tang, Gang Dong, Nilanjan Ray, Scott T. Acton

Department of Electrical and Computer Engineering,
University of Virginia,Charlottesville
Virginia 22904 USA
E-mail:{jt6w, gd7p,nr2b,acton}@virginia.edu

ABSTRACT

We are exploring automated tracking for cells observed *in vivo* by video microscopy. In the past, three tracking algorithms have been utilized for tracking cells: the normalized cross correlation tracking algorithm, the centroid tracking algorithm and the snake based tracking algorithm. This paper focuses on the analysis, evaluation and comparison of the above three tracking algorithms using some performance measures. These performance measures include: (1) the average number of consecutive frames tracked; (2) the localization error; (3) correlation with ground truth, and (4) tracking time. One hundred *in vivo* (intravital) videos were used in our experiments for analysis, evaluation and comparison. The results show that snake based tracking provides the best performance, and that automated tracking of leukocytes is indeed possible.

1. INTRODUCTION

The analysis of instantaneous and average rolling leukocyte velocity is crucial to the study of inflammatory disease. In order to record features associated with leukocyte rolling, the leukocyte position must be tracked [1]. Currently leukocyte tracking is carried out manually. Although fairly accurate, this method of manual data collection is extremely time-consuming. Therefore, we have designed and implemented an image processing system for automated tracking of rolling leukocytes *in vivo*. In automated tracking system, the design of tracking algorithms is a key step because it affects the performance of the system. In the past, three tracking algorithms have been proposed. They are the centroid tracker [2], the correlation tracker [3] and the snake tracker [4]. Centroid trackers use the intensity "center of mass" of an object to track the position. The centroid is a reliable feature in the absence of noise, occlusion, background movement and clutter. Correlation trackers use the correlation of a fixed template with the image to find instances of the target cell (or other particle of interest). Correlation methods are insufficient for tracking deformable targets and are also susceptible to false acquisitions in the presence of noise and clutter. Both centroid trackers and correlation trackers seem to encounter acquisition problems when the background is heavily cluttered or when the background is moving. In contrast, the snake tracker is tolerant of low SCR and high registration error. In this paper, we mainly focus on the analysis, evaluation and comparison of the above three tracking algorithms.

2. THE TRACKING ALGORITHMS

The three trackers discussed in this paper are correlation tracker, centroid tracker, and snake tracker. The working principles are summarized briefly in this section.

Correlation tracker: The correlation tracker works on the principle of template matching. First, a template $T(x,y)$ of the object is created. Within a search window, the correlation tracker then computes the normalized cross-correlation of the image, I, and the template, T. Then the tracker selects the position of maximal normalized cross-correlation. Let the search window of height h and width w be placed at location, (x_0, y_0) in the image I. Further let h_t, and w_t be the height and the width of the template T. The position of maximum normalized cross-correlation, (x_c,y_c), is the object center location given below:

$$(x_c,y_c)=\underset{x_0\le x\le x_0+w,\ y_0\le y\le y_0+h}{\operatorname{argmax}} \frac{\int_{-\frac{h_t}{2}}^{\frac{h_t}{2}}\int_{-\frac{w_t}{2}}^{\frac{w_t}{2}} I(x+u,y+v)T(u,v)\,du\,dv}{\sqrt{\left(\int_{-\frac{h_t}{2}}^{\frac{h_t}{2}}\int_{-\frac{w_t}{2}}^{\frac{w_t}{2}} I^2(x+u,y+v)\,du\,dv\right)\left(\int_{-\frac{h_t}{2}}^{\frac{h_t}{2}}\int_{-\frac{w_t}{2}}^{\frac{w_t}{2}} T^2(u,v)\,du\,dv\right)}} \quad (1)$$

Centroid Tracker: The centroid tracker assumes that the object under tracking is bright and is on a dark background. Within a search window on the image, the tracker computes the center of mass of the image. The center of mass is considered to represent the object location in this case. Let $I(x,y)$ be the image. Further, let the search window be a rectangle of height h and width w. If the search window is placed at some location, (x_0,y_0), in the image, then the centroid, $(\bar{x},\bar{y})$, of the search window is given below:

$$\bar{x} = x_0 + \frac{\int_0^h\int_0^w xI(x,y)\,dx\,dy}{\int_0^h\int_0^w I(x,y)\,dx\,dy} \quad (2) \qquad \bar{y} = y_0 + \frac{\int_0^h\int_0^w yI(x,y)\,dx\,dy}{\int_0^h\int_0^w I(x,y)\,dx\,dy} \quad (3)$$

The centroid $(\bar{x},\bar{y})$ is the location of the object reported by the tracker.

Snake Tracker: A snake or active contour is an elastic, flexible curve. When placed on an image, it can deform by some force (typically computed from the image data) to lock itself onto some desired image features *e.g.*, edges. Let $(p(s),q(s))$ denote a point on the active contour parameterized through $s \in [0,1]$. If we desire the active contour to adhere to a leukocyte (white blood

0-7803-7523-8/02/$17.00 ©2002 IEEE

cell) edge, then we compute the best active contour position that minimizes the energy functional [4]:

$$E_s = \lambda_1 E_{shape} + \lambda_2 E_{size} + E_{ext} \tag{4}$$

where the $\lambda's$ weight the shape and size constraints and the external energy. The shape and size constraints encourage the snake to take a circular shape of certain radius, K [4]:

$$E_{shape} = \frac{1}{2}\int_0^1 \left(p(s) - \int_0^1 p(r)dr - \overline{R}\cos(2\pi s) \right)^2 ds + \frac{1}{2}\int_0^1 \left(q(s) - \int_0^1 q(r)dr - \overline{R}\sin(2\pi s) \right)^2 ds,$$

$$E_{size} = \frac{1}{2}(\overline{R} - K)^2.$$

$\overline{R}$ is the average radius and is defined as

$$\overline{R} = \int_0^1 \sqrt{\left(p(s) - \int_0^1 p(r)dr\right)^2 + \left(q(s) - \int_0^1 q(r)dr\right)^2}\, ds.$$

The external energy is computed from image data and it encourages the active contour to cling to edges [4]:

$$E_{ext} = \int_0^1 |\nabla I(p(s), q(s))| ds.$$

To get the active contour that minimizes (4), variational calculus is applied to obtain the following gradient descent equations for snake horizontal position p and vertical position q:

$$\frac{\partial p}{\partial t} = \lambda_1 \left(p(s) - \int_0^1 p(r)dr - \overline{R}\cos(2\pi s)\right) + \lambda_2 \left(\frac{(\overline{R} - K)\left(p(s) - \int_0^1 p(r)dr\right)}{\sqrt{\left(p(s) - \int_0^1 p(r)dr\right)^2 + \left(q(s) - \int_0^1 q(r)dr\right)^2}} \right) + \frac{\partial}{\partial x}|\nabla I(p,q)|, \tag{5}$$

$$\frac{\partial q}{\partial t} = \lambda_1 \left(q(s) - \int_0^1 q(r)dr - \overline{R}\sin(2\pi s)\right) + \lambda_2 \left(\frac{(\overline{R} - K)\left(q(s) - \int_0^1 q(r)dr\right)}{\sqrt{\left(p(s) - \int_0^1 p(r)dr\right)^2 + \left(q(s) - \int_0^1 q(r)dr\right)^2}} \right) + \frac{\partial}{\partial y}|\nabla I(p,q)|, \tag{6}$$

Starting from an initial active contour (*i.e.*, an initial values of p and q), (5) and (6) are solved to find out the stationary state of p and q. This stationary state solution is the desired active contour that minimizes (4). The process of reaching to the stationary state from an initial contour is referred to as snake evolution.

The working principle with snake tracker is as follows. The initial contour on the current frame is obtained by taking the evolved snake from the previous frame. On the current frame one lets this snake evolve to find out the leukocyte under tracking. In the next frame, this evolved snake position is used as the next initial snake.

3. EXPERIMENTAL RESULTS

3.1. Test sequences

In order to evaluate and compare the three trackers, 100 *in-vivo* video clips were chosen from different *in-vivo* video sequences, which were obtained by digitizing the *in vivo* tapes from the Department of Biomedical Engineering at U.Va. Each video clip contains 91 frames, and the frame rate of all the 100 clips is 30 frames/second. Given the cell position, the frame rate and scale of the image, the computation of cell velocity and acceleration is possible [1]. Figure 1 shows one sampled from an *in-vivo* sequence.

Figure 1. A sample frame from an in vivo clip

3.2. Performance measures used in the experiments

There are five different quantiative measures used in our experiments. They are:

i) Average number of consecutive frames tracked: For each clip, let the radius of the rolling leukocyte be r, the center of the rolling leukocyte in the n-th frame be (c_x^n, c_y^n), and the center obtained by tracking algorithm in the n-th frame be (d_x^n, d_y^n). Then the number of the consecutive frames tracked in a clip is determined as follows:

Step1 Compute the distance between the real center and the tracked center of the rolling leukocyte in the n-th frame.

$$D_n = \sqrt{(c_x^n - d_x^n)^2 + (c_y^n - d_y^n)^2} \tag{7}$$

Step2 For some k, if $D_n \le r$ $(n = 1,2,..,k)$ and $D_{k+1} > r$, then we say that the number of the consecutive frames tracked in a clip is equal to k.

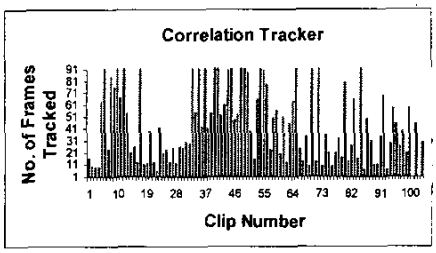
ii) Localization error: To measure the accuracy in position provided by the three trackers, we computed the average localization error for each clip. We used the root mean square error (RMSE) to compute the average localization error.

iii) Correlation: The Pearson correlation coefficient is used in our experiments to evaluate the feasibility of automated data collection by using the automated tracking algorithms. Given an ensemble of feature data $\{A_i\}$ $(i = 0,1,..., N-1)$ obtained from automated tracking and a collection of feature data $\{M_i\}(i=0,1,...,N-1)$ obtained from manual analysis, then the Pearson correlation coefficient can be computed as follows:

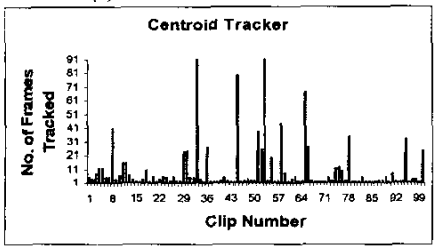
$$r = \frac{\sum_{i=0}^{N-1}(A_i - \overline{A})(M_i - \overline{M})}{\sqrt{\sum_{i=0}^{N-1}(A_i - \overline{A})^2 \sum_{i=0}^{N-1}(M_i - \overline{M})^2}} \tag{8}$$

where $\overline{A}$ is the mean of $\{A_i\}$, $\overline{M}$ is the mean of $\{M_i\}$, and N is the number of observations.

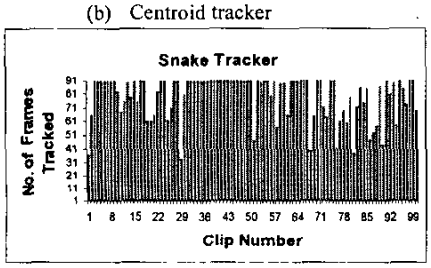
iv) Tracking Time: Time is an important factor in terms of the laboratory efficiency. The time required for the analysis of *in-vivo* video clips determines whether the automated tracking can be used to replace manual tracking system.



(a) Correlation tracker



(b) Centroid tracker



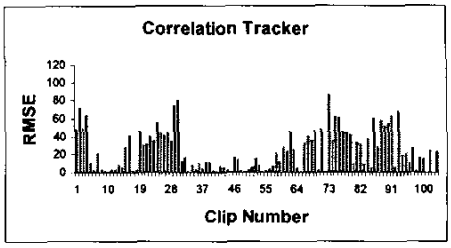
(c) Snake tracker

Figure 2. *The average number of the consecutive frames tracked per clip*

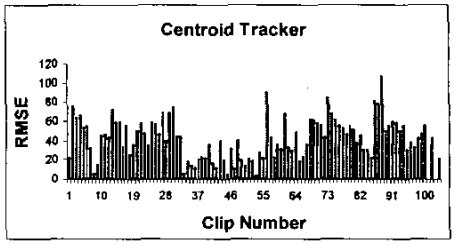
3.3. Experimental results

3.3.1. Tracking accuracy comparison of the three Trackers

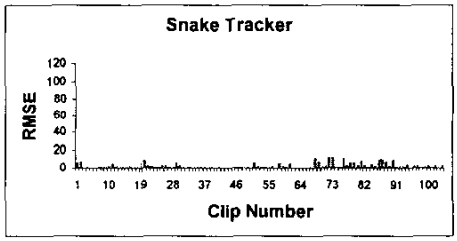
We first compared the performance of the three trackers by computing the number of consecutive frames tracked for each clip for a given tracker. Figure 2 displays the number of the consecutive frames tracked per clip for the 100 clips using the three different trackers. The average number of the frames tracked for the entire 100-clip dataset is 44.1 frames for correlation tracker, 9.7 frames for centroid tracker, 80.0 frames for snake tracker. The standard deviation of the number of the frames tracked using three trackers is 30.0 frames for correlation tracker, 17.7 frames for centroid tracker and 15.8 frames for snake tracker. According to related research [1], the intravital tracker is deemed successful if the average number of the consecutive tracked frame is at least 30 so that the cell may be observed for a least one second. Thus the centroid tracker is not successful, because the average number of the consecutive tracked frame is less than 30 frames. The snake tracker and the correlation tracker satisfy the 30-frame requirement. But the snake tracker is much better than the correlation tracker.



(a) Correlation tracker



(b) Centroid tracker



(c) Snake tracker

Figure 3. *The root mean squared localization error per clip*

Second, we compared the three trackers by the average localization error for each clip and for the whole 100-clip data-set. We used the root mean square error (RMSE) in microns to compute the average localization error [1]. Figure 3 displays the localization error per clip obtained using the three trackers. The average RMSE for the entire dataset is 25.0 μm for correlation tracker, 43.3 μm for centroid tracker and 2.8 μm for SNAKE tracker. From Figure 2 and the average RMSE values, the experiments how that the snake tracker significantly decreases the localization error.

3.3.2. Tracking efficiency

In this subsection, we compare the efficiency of the automated tracker as compared to manual tracking of leukocytes in terms of

the time required for the analysis of 20 video clips by trained technicians.

We timed manual tracking as conducted by an electrical engineer and by a biomedical specialist. The technicians used a computer mouse to track the cells manually. The computer recorded the time (in seconds) required to track each of the 20 cells. The time needed to identify the initial cell position was not included in the overall time, and the time needed for breaks between clips was not included. Next, we used the automated tracker to track the same 20 video clips and also recorded the time for each clip. The average tracking time for one clip was 153.1 s for the electrical engineering expert, 90.1 s for the biomedical expert, and only 8.7 s for automated tracker. The automated tracker is more than ten times faster than manual tracking. Thus, if an automated tracking is used to replace manual tracking, it will improve the tracking efficiency greatly.

3.3.3. Analysis of feasibility of Automated Data Collection

Twenty video clips of 31 frames each (1 s of video) were used in our experiments to test the feasibility of the automated data collection. We used the snake tracker (automated tracking) to perform automated data collection. Four features were measured: instantaneous velocities, instantaneous accelerations, average velocities and average accelerations of the rolling leukocytes. Then, the velocities and accelerations obtained by the snake automated tracker were compared with the velocities and accelerations obtained by manual tracking. The Pearson correlation coefficient was used to compare results.

In terms of average velocity and average acceleration, we computed both features for each of the 20 clips using the cell positions obtained by manual analysis and by the automated tracker. The correlation between the manual and automated computation of average velocity is 0.92. For average acceleration, the correlation is 0.84.

For *instantaneous* velocity and acceleration, we first smoothed the position information obtained manually and by the automated tracker using an average filter that replaces the position in each frame with the average position over three frames. Then, we computed the instantaneous velocities and accelerations using the displacements over 6 frames (= 1/5 s). An interval of 6 frames was chosen so that slow moving cells (velocity of ~ 2 μm/s) would move at least one pixel width in the time interval. We obtained two correlation values for each clip: 1) the correlation between the instantaneous velocities obtained by automated tracking and the instantaneous velocities obtained by manual tracking; 2) the correlation between the instantaneous accelerations obtained by automated tracking and the instantaneous accelerations obtained by manual tracking. Tables 1 and 2 provide the correlation values. The overall correlation value for the instantaneous velocity measurements on the twenty 31-frame sequences is 0.93. The overall correlation value for the instantaneous acceleration is 0.81. Thus, the experiments show that the automated tracker can be a valuable tool for data collection and that future work is needed with regard to data acceptance.

In summary, the experiments show that automated tracking can replace manual analysis for data collection regarding cells observed *in vivo*. Of the three trackers tested, the snake tracker holds the greatest promise for successful tracking in this noisy, cluttered imaging environment.

Table 1. Correlation of instantaneous velocity for the 31-frame clips

Clip number	1	2	3	4	5
Correlation	0.98	0.85	0.96	0.97	0.99
Clip number	6	7	8	9	10
Correlation	0.97	0.98	0.98	0.99	1.00
Clip number	11	12	13	14	15
Correlation	1.00	1.00	0.97	0.57	0.94
Clip number	16	17	18	19	20
Correlation	0.78	0.82	0.97	0.97	0.90

Table 2. Correlation of instantaneous acceleration for the 31-frame clips

Clip number	1	2	3	4	5
Correlation	1.00	0.97	1.00	0.98	0.95
Clip number	6	7	8	9	10
Correlation	1.0	0.99	1.00	1.00	1.00
Clip number	11	12	13	14	15
Correlation	1.00	1.00	0.99	0.96	1.00
Clip number	16	17	18	19	20
Correlation	0.69	0.77	0.99	0.98	0.85

REFERENCES

[1] S.T. Acton, K. Wethmar, and K. Ley. (2002). Automatic tracking of leukocytes in vivo. *Microvascular Research*, 63:139-148.

[2] R.N. Ghosh and W.W. Webb. (1994). Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. *Biophysical Journal*, 66:1301-1318.

[3] A. Kusumi, Y. Sako and M. Yamamoto. (1993). Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured endothelial cells. *Biophysical Journal*, 65:2021-2040.

[4] N. Ray, S.T. Acton, and K. Ley. (2002). Tracking leukocytes in vivo with shape and size constrained active contours. Accepted for publication in the *IEEE Transactions on Medical Imaging* (for special issue on Drug Discovery and Clinical Trials).

[5] M. Kass, A. Witkin and D. Terzopoulos, (1987). Snakes: Active contour models. *International Journal of Computer Vision*, 1:321-331.