2006 IEEE International Symposium on Signal Processing and Information Technology

A Statistical Thresholding Method for Cell Tracking

Nezamoddin N. Kachouie[1], Paul Fieguth[1]
John Ramunas[2] and Eric Jervis[2]
[1]Department of Systems Design Engineering
[2]Department of Chemical Engineeribg
University of Waterloo
Waterloo, Canada

Abstract- *Tracking the motion of cells in culture is a task, which often still is undertaken manually, and for which automated methods are strongly desirable. Researchers visually perform cell motion analysis, observe cell movements and cell shape changes for hours to discover when, where and how fast it moves, splits or dies. Hematopoietic Stem Cells (HSCs) proliferate and differentiate to different blood cell types continuously during their lifetime, and are of substantial interest in gene therapy, cancer, and stem-cell research. In this paper a statistical method is introduced to track HSCs over time. A statistical thresholding method is combined with joint probabilistic data association in the proposed HSC tracker.*

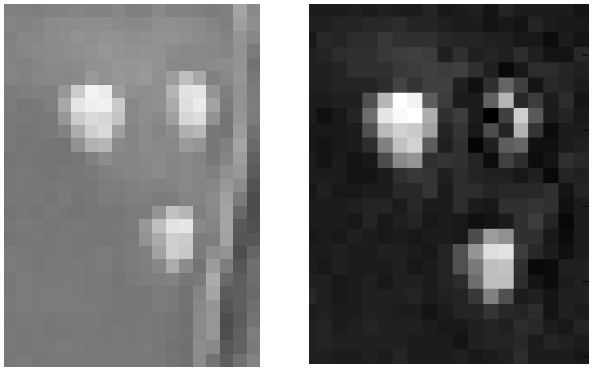
Keywords- *Tracking, Segmentation, Background Estimation, Statistical Model, Thresholding, JPDA, Stem Cell.*

1. INTRODUCTION

Drug and disease research strongly rely on cell behavior studies. Cell size, shape, and motility may play a key role in stem-cell specialization or cancer development. Traditional manual methods have been used to infer these values from image sequences. Given the increasing amount of cell data being collected, the manual methods are onerous and automated cell tracking and segmentation methods are in high demand.

Automatic cell tracking is essentially an object tracking problem. This problem is yet an attractive and challenging task which is closely related to the computer vision and image processing research communities [1, 2, 3, 4, 5]. Applying advanced techniques in digital image processing and pattern recognition to a huge number of bio-cellular images, can improve our understanding of cellular and inter-cellular events so that significant progress and new discoveries in biological and medical research may be achieved.

The challenging task in microscopic cell image segmentation is to adapt and extend available image segmentation approaches to the applications of cell imaging. A variety of semi-automatic or automatic methods have been proposed to segment cell boundaries [6, 7]. These methods include thresholding, watershed, nearest neighborhood graphs, mean shift procedure and deformable models.



(a) (b)

Fig. 1. (a) Close-up of an original HSC phase contrast microscopic image $phenotype_2$. (b) HSC $phenotype_2$ phase contrast image after background subtraction.

Mean shift procedure method was proposed by Comaniciu et al [8] for cell image segmentation for diagnostic pathology. Watershed have been used by Markiewicz et al [9] for segmentation of the bone marrow cells. Geusebroek et al [10] introduced a method based on Nearest Neighbor Graphs to segment the cell clusters. Meas-Yedid et al [11] proposed a method to quantify the deformation of cells using snakes. Kittler [12], Otsu [13] and Wu [14] have used thresholding methods.

This paper proposes a statistical cell tracking method by combining a thresholding approach and probabilistic data association.

2. CELL IMAGING AND ANALYSIS

To produce the data for this study, HSC samples are first extracted from mouse bone marrow and cultured in custom

arrays having up to forty wells. HSCs are then imaged using manual focusing through a 5X phase contrast objective using a digital camera (Sony XCD-900). Images were sampled every three minutes over the course of several days. A small fraction of a typical HSC microscopic image is depicted in Fig. 1.

Although cell staining techniques may be used to increase the contrast between cell and background, different parts of tissue are undesirably stained unevenly, causing inhomogeneity. Moreover to minimize photo-toxicity and to keep cells alive and healthy, light exposure must be controlled during their life cycle. Limiting the light exposure in each frame, and sampling the frames as far apart as possible lead to infrequent, poorly-contrasted images, directly at odds with the data desired for easy tracking: frequent, high-contrast images.

3. THE PROPOSED METHOD

In our previous work [15] we characterized a typical HSC in microscopic image as an approximately circular object with a darker interior and a bright boundary. Although the proposed cell model works well to localize a specific HSC phenotype, its performance drops with

- Significant illumination variations during phase contrast imaging.
- Different HSC phenotypes.

To improve the previous cell model and resolve its shortcomings, a more general and robust cell model must be designed. The goal here is modelling HSC so that with slight changes in the model parameters it can represent different HSC phenotypes and be robust against illumination variations as well.

The cell detection is essentially one of anomaly detection, the localization of groups of pixels inconsistent with the random behavior of the image background. Our proposed method consists of background estimation and subtraction in the first place, an anomaly detection model as the next step and data association at the end which will prove to be invariant to illumination variations and HSC phenotypes.

3.1. Background Estimation

Assume S and I are pure and corrupted sequence of K, $N \times M$ images. S is corrupted by spatial illumination variations v over time and temporal additive noise n in each frame. Moreover we assume that v and n are identically distributed, independent from each other and both independent from I

$$I = S + b \tag{1}$$

where $b = v + n$. Each pixel I_{ijk} represents a pixel in 3 spatio-temporal dimensions such that

$$I_{ijk} = S_{ijk} + b_{ijk} \tag{2}$$

where

$$b_{ijk} = v_{Spatial}^{\{[1:K]\}} + n_{Temporal}^{\{[1:N][1:M]\}} = v_{ij}^{\{[1:K]\}} + n_{k}^{\{[1:N][1:M]\}} \tag{3}$$

Furthermore we assume that the spatio-temporal background noise can be modelled as a three dimensional Gaussian density function. We consider a $2-D$ spatial Gaussian over time as

$$p(v_{ij}|\,\theta_{ij}),\;\; \theta_{ij} = (\mu_{ij}, \Sigma_{ij}) \tag{4}$$

where $v_{ij} \sim \mathtt{N}(\mu_{ij}, \Sigma_{ij})$, and a $1-D$ temporal Gaussian for each image frame k

$$p(n_k|\,\theta_k),\;\; \theta_k = (\mu_k, \Sigma_k) \tag{5}$$

where $n_k \sim \mathtt{N}(\mu_k, \Sigma_k)$.

Spatial background intensity variations for pixel I_{ij} are caused by the spatial illumination variations and can be estimated by spatial mean μ_{ij} of spatial Gaussian over time $(1:K)$ as $\hat{v}_{ij} = \mu_{ij}$. Temporal background intensity variations of pixel I_k are caused by the temporal additive noise and can be estimated by temporal mean μ_k of temporal Gaussian on each frame as $\hat{n}_k = \mu_k$. Therefore the pure signal S can be estimated by

$$\hat{S} = I - \hat{b} \tag{6}$$

where for each spatio-temporal pixel S_{ijk} we have $\hat{S}_{ijk} = I_{ijk} - \hat{b}_{ijk}$.

3.2. Cell Model

The next essential step is detecting and localizing HSCs in the uniform background. Figs. 2(b) and 3(b) show typical HSC images, $phenotype_2$ and $phenotype_1$ after background subtraction respectively. As we can see, HSC can be characterized as approximately circular object with high intensity variations against the uniform background. Except for $phenotype_1$, the other HSC phenotypes cannot be modelled as an object with dark interior and bright boundary, however intensity variations in cell locations can be used as an unique characterization for all HSC phenotypes. Thus a general model for HSC can be characterized by the following criteria

1. All HSC phenotypes have an approximate circular shape.
2. Cell pixels are significantly different from the uniform background pixels, hence HSCs can be localized by detecting the pixels with significant intensity variations in the image.

3. Neighboring pixels on the cell boundary have symmetric intensities.

Circular Mean Square Model. HSC is modelled as a circular anomaly which is represented by a set of pixels with significant intensity variations against the uniform background. Assuming (x_c, y_c) and r as center coordinates and radius of the cell, the continuous circular cell is spatially discretized as

$$|(x_i - x_c)^2 + (y_i - y_c)^2| \leq r^2, \tag{7}$$

where (x_i, y_i) are coordinates of cell pixels. We construct the function $S(x_c, y_c, r, I)$, which is a vector returning the set of inside cell pixels as

$$S(x_c, y_c, r, I) = \{I_{ij} |\ (x_c - i)^2 + (y_c - j)^2 \leq (r)^2\ \}, \tag{8}$$

from which we extract sample mean of square-intensities of cell pixels

$$\bar{S} = \frac{\sum_i {S_i}^2}{|S|} \tag{9}$$

Two Class Classification: Cell and Background. To recognize cells from the uniform background, first (9) is applied to the cell image and circular mean square is computed. Assuming $|S| = L$, (9) can be rewritten as $\bar{S} = (\sum_i {S_i}^2)/L$. Moreover the variance is

$$\sigma^2 = \frac{\sum_i (S_i - \mu)^2}{L} = \frac{\sum_i {S_i}^2 - \sum_i \mu^2}{L} \tag{10}$$

after simplification we have

$$\bar{S} = \frac{\sum_i {S_i}^2}{L} = \sigma^2 + \mu^2 \tag{11}$$

For uniform background $\sigma^2 = \sigma^2_{bkg}$, $\mu^2 = \mu^2_{bkg} = 0$ and we have

$$\bar{S}_{bkg} = \frac{\sum_i {S_i}^2}{L} = \sigma^2_{bkg} \tag{12}$$

where $S_i \in \{background \mid\ i = [1, L]\}$. In cell locations, $\sigma^2 = \sigma^2_{cell}$, $\mu^2 = \mu^2_{cell}$ and we have

$$\bar{S}_{cell} = \frac{\sum_i {S_i}^2}{L} = \sigma^2_{cell} + \mu^2_{cell} \tag{13}$$

where $S_i \in \{cell \mid\ i = [1, L]\}$. For different cell phenotypes one of the σ^2_{cell}, μ^2_{cell} or both are significantly higher than those of background, therefore σ^2_{bkg} is very small in comparison with $\sigma^2_{cell} + \mu^2_{cell}$ and as a result circular mean square, $\bar{S}$, can be used to detect HSC in uniform background. To discriminate cells from background, the circular mean square image is classified to two classes, cell and background, by minimizing the inter-class variance

$$\sigma^2(T) = l_{cell}(T) \cdot \sigma^2_{cell}(T) + l_{bkg}(T) \cdot \sigma^2_{bkg}(T) \tag{14}$$

where $l_{bkg}(T)$ and $l_{cell}(T)$ are the number of pixels in the background and the cell classes, $\sigma^2_{cell}(T), \sigma^2_{bkg}(T)$ and $\sigma^2(T)$ are variance of background, variance of cell class and inter-class variance considering the threshold (T).

Anomalous Background Pixels. Considering the HSC as a circular anomaly in the proposed method it can be concluded that the cell center has the maximum distance to the cell boundary in comparison with any other pixel in the cell area. Thus to fit a circular shape to the classified anomaly, we compute the Euclidean distance of anomaly pixels from the background

$$D_{cell} = \sqrt{(x_{cell} - x_{bkg})^2 + (y_{cell} - y_{bkg})^2} \tag{15}$$

where (x_{cell}, y_{cell}) and $(x_{bkg}$,$y_{bkg})$ are cell and background pixel coordinates respectively and D_{cell} is the Euclidean distance of cell pixels from the nearest background pixel. As we can observe in Figs. 2(c) and 3(e) after computing the Euclidean distance for the thresholded image the resultant image has the highest values in the cell center locations.

Probability Density Function of Cell Centers. An exponential density function is proposed to integrate circular mean square and cell-background Euclidean distance functions as

$$P_{cell}(x_c, y_c, r, I) = 1 - exp\{-\lambda \cdot D_{cell} \cdot \bar{S}\} \tag{16}$$

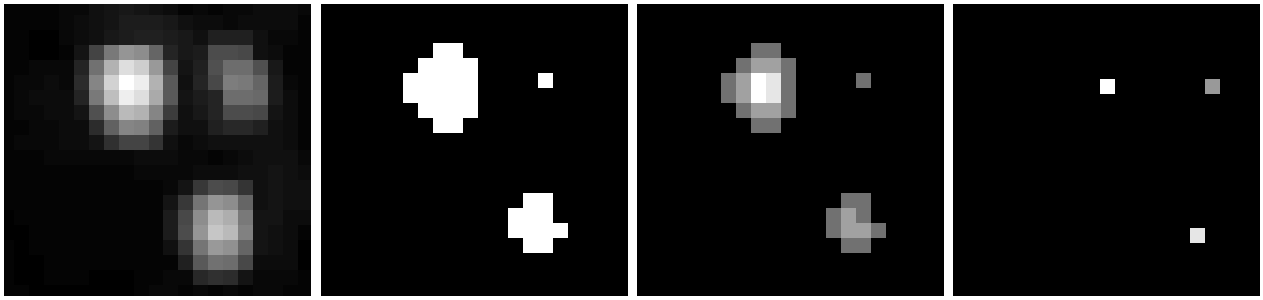
where λ is a constant. A probability map is generated by computing P_{cell} for image I. Each value p in the generated probability map shows the probability of a cell located at (x_c, y_c) with radius r.

3.3. Cell Tracking

Cell center candidates are obtained by applying the proposed cell model to the HSC image sequence, generating a probability map, finding the local maxima which are at least E_d apart in the probability map and thresholding the local maxima map to obtain a set S_τ of local maxima which are E_d apart and are greater than a threshold τ. The threshold τ has been obtained empirically.

To associate the cell center candidates we apply a gated joint probabilistic data association approach [16, 17, 18] in which to reduce the number of association hypothesis, a validation gate is computed to obtain a distance vector D_e for each identified cell center in time step $t - 1$ based on its Euclidian distance from the cell center candidates in time step t. Assuming a random walk for cell motion, cell dynamics is modelled as

$$P_V \sim N(0, \sigma) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{D_e^2}{\sigma^2}\right) \tag{17}$$



(a) (b) (c) (d)

Fig. 2. (a) Circular mean square of HSC phenotype 2. (b) Classifying circular mean square to cell and background classes by minimizing the inter-class variance. (c) Euclidean distance of cell pixels from the background. (d) Cell center locations after thresholding the maxima map.

The proposed gated joint probabilistic data association is illustrated as

$$P(x_t|z_t) = \prod_{C_{div}} P_{State} \cdot \prod_{j\in[1,M_{t-1}]} P_V \cdot \prod_{j\in[1,M_{t-1}]} max(P_D \cdot P_{De}^{j,i}, P_F) \cdot \prod_{New} max(P_D \cdot P_{De}^{j,i}, P_F) \quad (18)$$

where $P_{De}^{j,i}$ is probability of associating the cell center candidate i in time step t to the cell j in time step $t-1$ based on their Euclidean distance. P_D and P_F are probability of cell detection and false alarm respectively and are assumed to be constant. P_{State} is cell state probability and is used to penalize cell division based on size and age of the cell. C_{div} is the set of dividing cells, New is the set of new detected cell candidates, and M_{t-1} is the number of cells in the previous time step $t-1$.

4. RESULTS

In our experiments, first the proposed background estimation - subtraction has been applied to different HSC phenotypes to eliminate the noise and illumination variations and to obtain images with uniform background. The proposed probabilistic cell model (16) is then applied to the subtracted resultant images and the probability map of the cell centers is computed for each frame in which HSC is characterized as a circular shape with radius $r \in [r_1, r_2]$. To further identify the cell centers, the probability map is thresholded and local maxima are located. A $3-D$ local maxima map based on $r \in [r_1, r_2]$ is generated as measurement hypotheses which will be used in the tracking algorithm.

Figs. 1(a) and 3(a) show the original HSC images $phenotype_2$ and $phenotype_1$ respectively. As we can observe HSCs of $phenotype_1$ have dark inside, bright and uniform boundary while HSCs of $phenotype_2$ are approximately evenly bright circular cells.

The results obtained for $phenotype_2$ and $phenotype_1$ after background subtraction using the proposed method are depicted in Figs. 1(b) and 3(b) respectively. Figs. 2(a) and 3(c) show circular mean square images and classified images by minimizing the inter-class variance are depicted in Figs. 2(b) and 3(d) for $phenotype_2$ and $phenotype_1$ respectively. Circle fitting on detected anomalies is illustrated on the classified images by computing the Euclidean distance of detected anomalous pixels from the background as depicted in Fig. 2(c) and Fig. 3(e) for $phenotypes_2$ and $phenotypes_1$ respectively. A probability map is computed using the product of the circular mean square and Euclidean distance images in which HSC centers are identified by thresholding and locating the local maxima. HSC center candidates are depicted in Fig. 2(d) and Fig. 3(f) for the respected phenotypes. Fig. 4(a) shows the original HSC $phenotype_2$ sequence. Fig. 4(b) and (c) show local maxima maps and tracking sequence which are obtained by applying the cell model and gated joint probabilistic data association respectively. Cell centers are correctly localized and associated for mature, dividing and new-born cells by applying the proposed method. Different colors show the cell center association over time.

5. CONCLUSIONS AND DISCUSSIONS

Measure and extract cell properties from large volumes of microscopic cell images demands for practical approaches. As an important application of biomedical research, stem cell biology and digital microscopic image processing, this paper presents a statistical thresholding method for blood stem cell tracking to locate and track different HSC phenotypes in culture in phase contrast microscopic images. The proposed method, which is constructed by observing HSCs

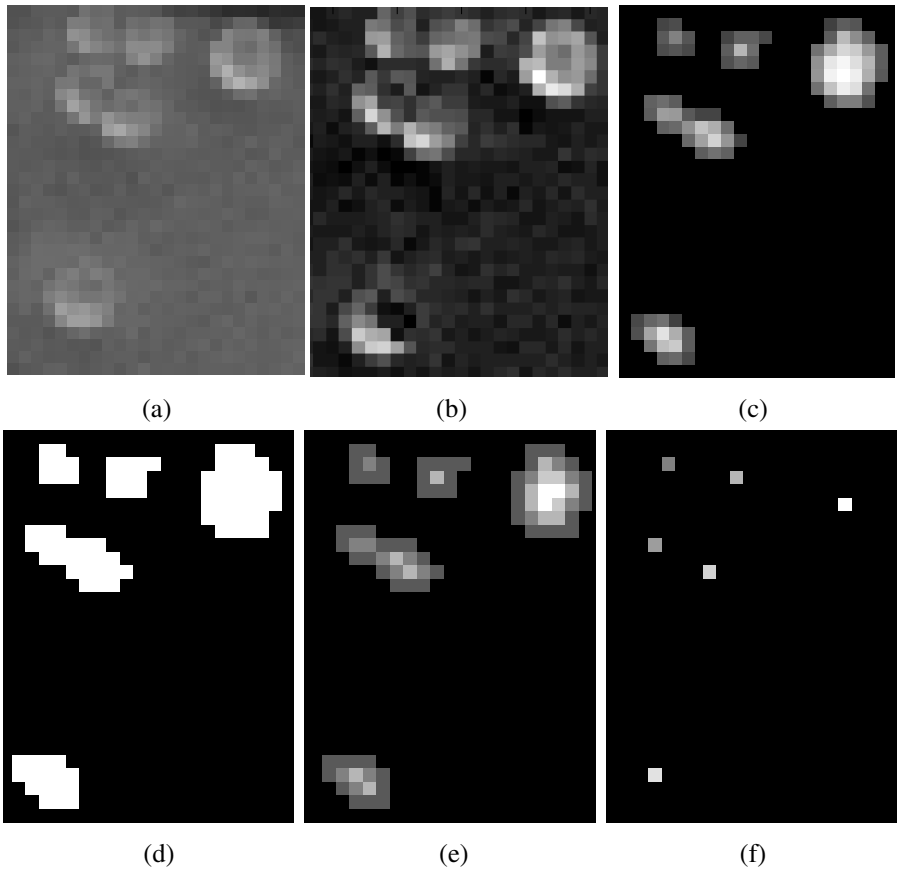


Fig. 3. (a) Original HSC image. (b) HSC image after background subtraction. (c) Circular mean square model. (d) Classification of circular mean square to cell and background classes by minimizing the inter-class variance. (e) Euclidean distance of cell pixels from the background. (f) Cell center locations after thresholding the maxima map.

in typical image sequences, detects and associates HSCs by extracting their key properties over time. After background estimation-subtraction a probability map of cell centers is generated by matching the image data with the cell model for each frame. Cell centers are located by finding and thresholding the local maxima in the probability map. Cell center candidates are further associated using a gated probabilistic data association. There are some cases where associating the right cell centers over time is ambiguous, for example having several close by dividing cells. Our future work includes designing a parametric cell shape with more degrees of freedom to adapt the proposed model to other non spheroidal cell types and improving the tracking method so that more challenging association cases can be resolved correctly.

6. ACKNOWLEDGEMENTS

This research has been funded by Natural Science and Engineering Research Council of Canada (NSERC).

7. REFERENCES

[1] A. Blake and A. Yuille, *Active Vision*, MIT Press, 1992.

[2] D. A. Forsyth and J. Ponce, *Computer Vision - A Modern Approach*, Prentice Hall, 2003.

[3] A. Blake and M. Isard, *Active Contours*, Springer-Verlag, 1997.

[4] B. K. P. Horn, *Robot vision*, MIT Press, 1986.

[5] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Freeman and company, 1982.

[6] D. Comaniciu, D. Foran, and P. Meer, "Shape-based image indexing and retrieval for diagnostic pathology," in *Int'l Conf. on Pattern Recognition*, pp. 902–904, 1998.

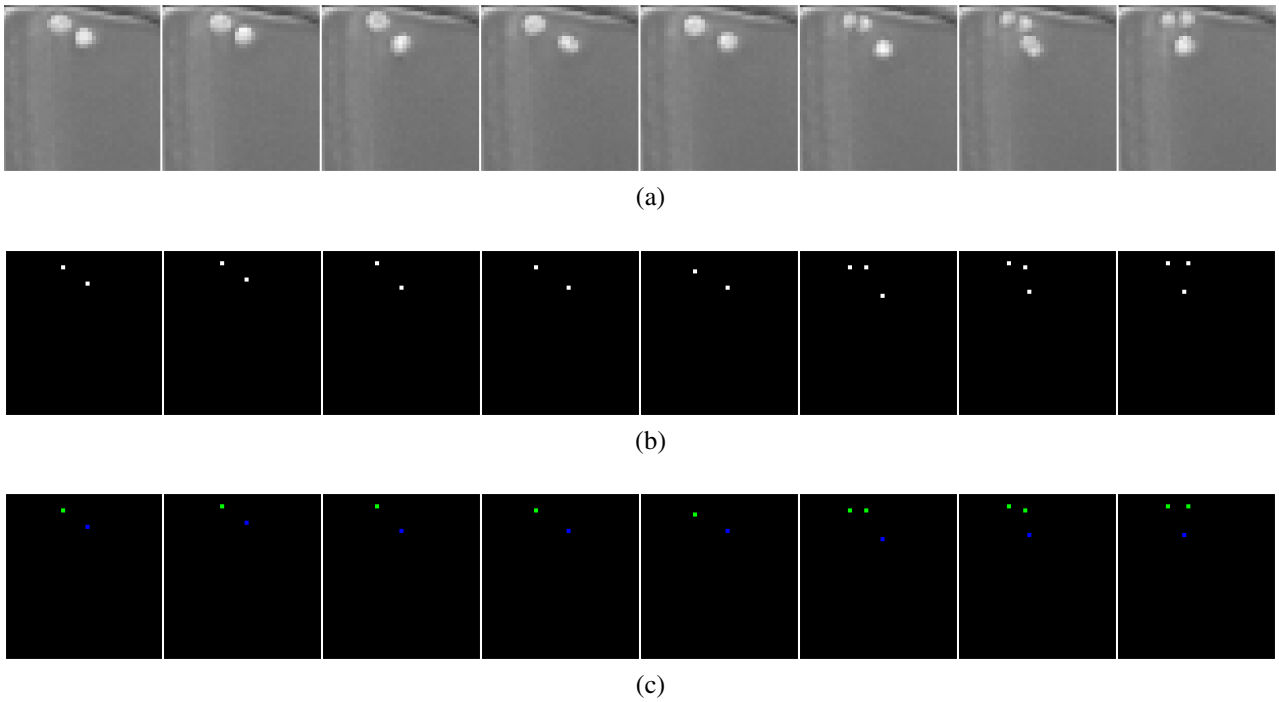


Fig. 4. HSC $phenotype_2$ (a) Original image sequence. (b) Cell center candidates which are obtained by finding and thresholding the local maxima in probability map. (c) Associated cell centers over time using gated joint probabilistic data association.

[7] E. Campo and E. J. E, "Mantle cell lymphoma," *Arch. Pathology Lab. Med.* **120**(1), pp. 12–14, 1996.

[8] D. Comaniciu and P. Meer, "Cell image segmentation for diagnostic pathology," *Advanced algorithmic approaches to medical image segmentation: State-of-the-art applications in cardiology, neurology, mammography and pathology* , pp. 541–558, 2001.

[9] T. Markiewicz, S. Osowski, L. Moszczyski, and R. Satat1, "Myelogenous leukemia cell image preprocessing for feature generation," in *5th International Workshop on Computational Methods in Electrical Engineering*, pp. 70–73, 2003.

[10] J. Geusebroek, A. Smeulders, and F. Cornelissen, "Segmentation of cell clusters by nearest neighbour graphs," in *Proceedings of the third annual conference of the Advanced School for Computing and Imaging*, pp. 248–252, 1997.

[11] V. Meas-Yedid, F. Cloppet, A. Roumier, A. Alcover, J.-C. Olivo-Marin, and G. Stamon, "Quantitative microscopic image analysis by active contours," in *VI 2001 Vision Interface Annual Conference - Medical Applications*, 2001.

[12] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition* **19**(1), pp. 41–47, 1986.

[13] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics* **9**(1), pp. 62–66, 1979.

[14] K. Wu, D. Gauthier, and M. Levine, "Live cell image segmentation," *IEEE Transactions on Biomedical Engineering* **42**(1), pp. 1–12, 1995.

[15] N. N. Kachouie, J. Li, and P. Fieguth, "A living cell segmentation model," in *IEEE Int'l Conf. on Image Processing*, pp. 901–904, 2005.

[16] Y. B. Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic-Press, 1988.

[17] D. Reid, "An algorithm for tracking multiple targets," *IEEE Transactions on Automation and Control* **24**(6), pp. 84–90, 1979.

[18] Y. Bar-Shalom and K. Birmiwal, "Consistency and robustness evolution of the pdaf for target trackingin a cluttered environment," *Automatica* **19**, pp. 431–437, 1983.