ROBUST CELL NUCLEI TRACKING USING GAUSSIAN MIXTURE SHAPE MODEL

Tomáš Vičar
Doctoral Degree Programme (2), FEEC BUT
E-mail: vicar@feec.vutbr.cz

Supervised by: Radim Kolář
E-mail: kolarr@feec.vutbr.cz

Abstract: The life cell microscopic imaging is a standard approach for studying of cancer cell morphology and behaviour during some treatment. In the dense cell cultures, tracking each cell nucleus is challenging task due to cell overlap and interactions. Moreover, for time-lapse sequences (lasting typically 20-30 hours) the robust automatic cell tracking is needed. This paper describes new method for fluorescence nuclei tracking based on Gaussian mixture model (GMM), and additionally, GMM modification allowing application to the images is also introduced. Method is mainly designed for robustness - tracking the highest possible number of nuclei in the whole sequence. Proposed algorithm proved to by very reliable with 80% of correctly tracked nuclei.

Keywords: Fluorescence nuclei images, nuclei tracking, Gaussian mixture model

1 INTRODUCTION

The detection of the individual nuclei and, in particular, separation of the clustered nuclei is complicated and still unresolved problem [1]. Moreover, detected cells are further utilized in the cell tracking methods, where errors from the individual frames are cumulated, because each detection error leads to interruption of track and thereby loss of information about this cell for whole sequence.

Method presented in this paper was mainly designed for the fluorescence images from Multimodal Holographic Microscope (MHM) [2], which is able of simultaneous acquisition of fluorescence and quantitative phase images. The main motivation for this paper is to design a robust method for tracking of nuclei trough the whole image sequence. The result of nuclei tacking than can be simply used as seed-points for the segmentation of the individual cells in each frame. We are only interested in cells that we can track all the time without cell division. This is due to the design of the experiment, which is intended to study different cell parameters related to cell death.

Many tracking algorithms exist (e.g. [3] and [4]), but they all incorporate cell division and do not focus on reliable tracking of the initial set of cells. The proposed method for the tracking of the fluorescence nuclei can significantly improve beyond them due to active search of all cell from the last frame. However, the main problem is separation of clustered nuclei, which is solved by assumption of the same number of nuclei in each pair of neighboring frames. This separation is achieved with Gaussian Mixture Models (GMM), which very robustly divide each cluster to the right number of Gaussian shapes with relatively equal size of the resulting parts. For the direct application of GMM on images, GMM algorithm was also slightly modified.

2 MATERIALS AND METHODS

2.1 DATASET

Method has been tested and evaluated on 5 field of view of the PNT1A cells, which were treated with inductor of the apoptosis Staurosporin and nuclei were fluorescently labeled with Hoechst 33258. Images were acquired on MHM with Objective Nikon Plan $10x/0.3$ and CDD camera XIMEA MR4021MC. Each field of wiew have the size 600x600px, with the frame rate 0.2 frame/min and 467 frames.

2.2 FOREGROUND SEGMENTATION AND NUCLEUS CLUSTERS EXTRACTION

The first step of proposed method is segmentation of nucleus pixels, but simple thresholding cannot be used, due to large variability in image values between experiments, therefore Maximally Stable Extremal Region (MSER) [5] approach was used. MSER identifies connected components in thresholded binary image, which are not changing their area with respect to threshold perturbation. Efficient algorithms for MSER extraction exist, and it is robust to variability of images in intensity changes and with same parameters setting. The binary masks from MSER method contain a large number of nuclei, which are connected, thus each connected component (nuclei cluster) is subjected to the subsequent GMM analysis.

2.3 GAUSSIAN MIXTURE MODEL TRACKING

The shape of each cell nucleus can be well approximated with 2D Gaussian function, thus each nucleus cluster can be modeled using the GMM [6]; moreover, GMM parameters can by estimated using Expectation-Maximization (EM) method. GMM probability density is given by the sum of K weighed Gaussian functions [6]

$$P(\mathbf{x}|\theta) = \sum_{k=1}^{K} w_k p(\mathbf{x}|\theta_k) = \sum_{k=1}^{K} w_k (2\pi)^{-d/2} |\Sigma_k|^{-1/2} exp\left[-\frac{1}{2}(\mathbf{x}-\mu_k)^T \Sigma_k^{-1}(\mathbf{x}-\mu_k)\right] \quad (1)$$

where model parameters θ consist of set of parameters for each Gaussian $\theta_k = \{w_k, \mu_k, \Sigma_k\}$, which are Gaussian weight, mean vector and covariance matrix, respectively, and d denotes dimensionality. EM is statistical iterative method for the estimation of maximum a posteriori (MAP) estimate of model parameters, which consist of iterations between two steps: (1) In the E-step, expected MAP for all samples is updated and (2) in the M-step, model parameters that maximize MAP are estimated, where both steps have closed form solution for GMM [6]; however, it can leads to the local optima, which highly depend on the initialization. Furthermore, GMM can be used for following classification, simply by assignment of the sample to the Gaussian with highest probability.

GMM is designed for the estimation of the probability density function, $p(\mathbf{x}|\theta)$, based on density of observed samples. It is possible to convert image $I(\mathbf{x})$ into samples by creation of $I(\mathbf{x})$-times replicated sample at each $\mathbf{x}$-position (i.e. for every pixel is created as many samples as is value of the pixel) as in [7], but it leads to increase of computational demands and is impractical. Therefore, new modification of GMM, weighted GMM (wGMM), is proposed it this paper, where each sample have weight λ, which affects importance of this sample in EM algorithm. E-step of EM algorithm is not affected, thus posterior probability for i-th sample and j-th Gaussian can be computed as

$$p_{ij} = p(\theta_k = \theta_j | \mathbf{x} = \mathbf{x}_i) = \frac{w_j p(\mathbf{x} = \mathbf{x_i} | \theta_k = \theta_j)}{\sum_{l=1}^{K} w_l p(\mathbf{x} = \mathbf{x_i} | \theta_k = \theta_l)} \quad (2)$$

Modification of the algorithm is in M-step, where we add λ into computation of j-th Gaussian parameters

$$w_j = \frac{\sum_i p_{ij}\lambda_i}{K\sum_i \lambda_i}, \quad \mu_j = \frac{\sum_i p_{ij}\lambda_i \mathbf{x}_i}{\sum_i \lambda_i p_{ij}}, \quad \Sigma_j = \frac{\sum_i p_{ij}\lambda_i (\mathbf{x}_i - \mu_j)(\mathbf{x}_i - \mu_j)^T}{\sum_i \lambda_i p_{ij}} \tag{3}$$

where i are indexes of the samples. It leads to only few simple modifications in standard implementations of GMM algorithm. This wGMM can be applied on the image simply by use of position on the pixel grid $\mathbf{x}$ as samples $\mathbf{x}_i$ and pixel values $\mathrm{I}(\mathbf{x})$ as weights λ_i. However, application on the non-image data can be also considered.

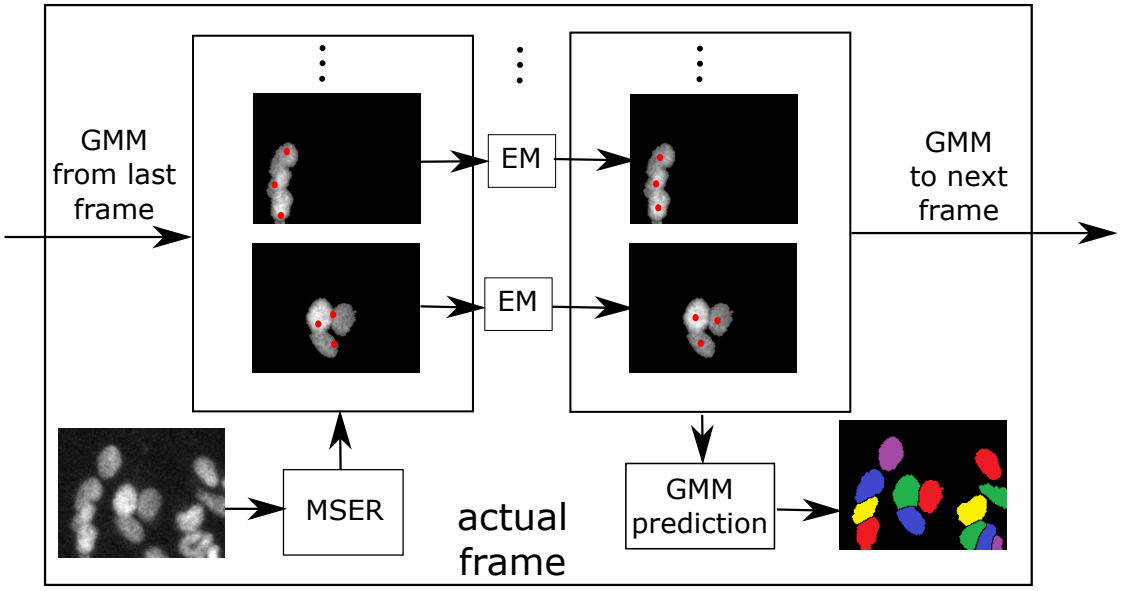

Figure 1: Block diagram of the process in one frame; red dots indicates Gaussians centers and number of EM runs depend on number of clusters identified in the image.

One of the problems of the GMM is that we need to estimate the number of Gaussians. In our application of nuclei tracking, the number of nuclei in each cluster is known from the previous frame. Furthermore, we can use all Gaussians parameters from previous frame for initialization of EM. EM algorithm than only adjust parameters of wGMM in each frame. Gaussians for each cluster are identified as Gaussian whose μ lies inside the mask cluster, and, in the second step, these Gaussians are adjusted by EM applied on the pixels under the cluster mask. Afterward, Gaussians from all the clusters are mixed together and move to the next frame. Processing of one frame is depicted on the Figure 1.

Additionally, GMM fits the normalized probability distribution, but not the image values, thus Gaussians weights w_i will not match between frames, if some Gaussian change cluster. Therefore, it is more appropriate to modify weights after EM in each cluster as

$$\mathbf{w} = \frac{\mathbf{w}\sum_{i=1} I(\mathbf{x}_i)}{\sum_k p_{x=x_i,\theta}} \tag{4}$$

where p_{ik} is evaluation of GMM at position of the pixels of the cluster x_i. In the other words, weights of the all Gaussians are divided by the sum of the image values predicted by the model and multiplied

Table 1: Results of proposed method and TrackMate as the number of correctly tracked cells; x/y represents number of the correctly tracked cells / total number of the cells.

Field of view	1	2	3	4	5	sum	%
proposed	52/67	21/26	19/25	61/73	32/40	185/231	80%
TrackMate [3]	22/67	6/26	7/25	22/73	15/40	72/231	31%

by sum of the values of original image. After this adjustment Gaussians fit the image values and can be validly reused in the next frame.

2.4 GMM INITIALIZATION

In the first frame, GMM is initialized from detected seed-point, which are obtained with LoG based cell detection algorithm described in [8], but the manual adjustments in first frame can be also considered way for higher precision. After detection, one Gaussian model is created for each initialization point i, where μ_i is set to point position, w_i is set to one and Σ_i is set to identity matrix. The newly appeared cells in next frames, are identified as binary mask without any Gaussian. These cells are also initialized and add to the GMM tracking.

3 RESULTS AND DISCUSSION

Due to strict requirement on the analysis of these types of experiment, only cells tracked through the whole sequence are considered as correct; moreover even temporarily false union of two nuclei is considered as error for the both nuclei. Because of complicated manner of such evaluation criterion and non-existence of a ground truth data, evaluation was done by the visual checking of each nuclei and classifying them as correct/incorrect. Results were compared to the state-of-the-art method TrackMate [3], which use LoG filter based detector and Linear Assignment Problem tracker for connection of detected points paths and for filtration of the false detections.

The results obtained by our approach and TrackMate are sumarized in the Table 1. Proposed method achieved significantly better results of 80% percent of correctly tracked cells. Above that, besides of totally insensitive parameters of MSER, proposed method have no parameters to be set. Besides of larger number of nuclei progress to the next analysis, wrongly tracked cells must be also filtered out in some post-processing step, thus this increase in the precision can significantly refine the results of the analysis. On the other hand, computational time on standard desktop machine (procesor Intel(R) Core(TM) i5-6500 CPI @ 3.20GHz) is 2.1s/frame and <0.1s/frame for proposed and TrackMate, respectively. The example of the results is depicted in the Figure 2(a), where is shown first frame with cell tracking paths through whole sequence.

Besides off very good detection and tracking, proposed method also provide segmentation of single nuclei (see example on Figure 2(b)). Tracked nuclei can be also used as cornerstone for the subsequent segmentation.

4 CONCLUSION

This paper describes the nuclei tracking and segmentation method tested for the fluorescence images acquired with MHM. The proposed method mainly focuses on tracking as many cells as possible throughout the whole image sequence, where it achieve 49% more correctly segmented than state-of-the-art method TrackMate. The proposed method use GMM for the fitting of the Gaussians like shape of the nuclei, where special modification of GMM model, wGMM, is introduced, therefore GMM can be applied on the images.

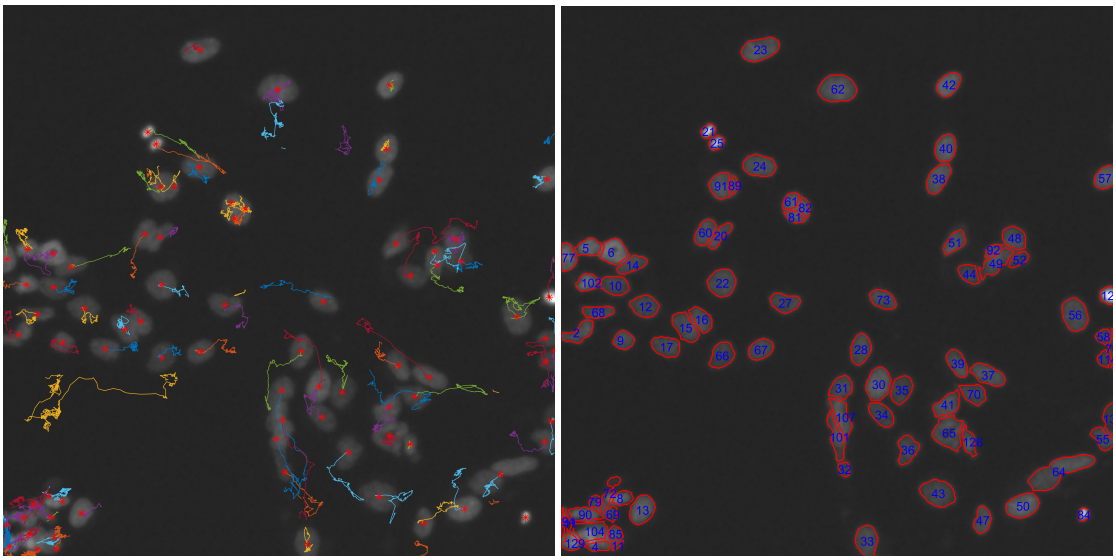

Figure 2: Example of the results; Image of fluorescence nuclei in the first frame (a) with the marked paths in whole sequence, (b) with the nuclei segmentation contours.

REFERENCES

[1] V. Ulman, M. Maška, K. E. Magnusson, O. Ronneberger, C. Haubold, N. Harder, P. Matula, P. Matula, D. Svoboda, M. Radojevic, *et al.*, "An objective comparison of cell-tracking algorithms," *Nature methods*, vol. 14, no. 12, p. 1141, 2017.

[2] P. Kolman and R. Chmelík, "Coherence-controlled holographic microscope," *Optics Express*, vol. 18, no. 21, pp. 21990–22004, 2010.

[3] J.-Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, S. Y. Bednarek, S. L. Shorte, and K. W. Eliceiri, "Trackmate: An open and extensible platform for single-particle tracking," *Methods*, vol. 115, pp. 80–90, 2017.

[4] J. Chalfoun, M. Majurski, A. Dima, M. Halter, K. Bhadriraju, and M. Brady, "Lineage mapper: A versatile cell and particle tracker," *Scientific reports*, vol. 6, p. 36984, 2016.

[5] D. Nistér and H. Stewénius, "Linear time maximally stable extremal regions," in *European Conference on Computer Vision*, pp. 183–196, Springer, 2008.

[6] G. McLachlan and D. Peel, *Finite mixture models*. John Wiley & Sons, 2004.

[7] C. Jung, C. Kim, S. W. Chae, and S. Oh, "Unsupervised segmentation of overlapped nuclei using bayesian classification," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2825–2832, 2010.

[8] H. Kong, H. C. Akakin, and S. E. Sarma, "A generalized laplacian of gaussian filter for blob detection and its applications," *IEEE transactions on cybernetics*, vol. 43, no. 6, pp. 1719–1733, 2013.