CELL DETECTION METHODS FOR THE IMAGES FROM HOLOGRAPHIC MICROSCOPE

Tomáš Vičar
Doctoral Degree Programme (1), FEEC BUT
E-mail: xvicar03@stud.feec.vutbr.cz

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

Abstract: Microscopical cell image analysis is widely used for cell behavior and morphology study. In dense cell cultures precise detection (separation) of a single cell is challenging task and it is important step for automatic cell analysis methods. There are a variety of methods, but most of them are less accurate for non-circular cells. This paper describes the common approaches for cell detection applied on images from holographic microscope. Linear discriminant analysis is used for combining results of these methods to obtain new more precise and robust approach.

Keywords: Holographic microscope, cell detection, LoG filter, distance transform

1 INTRODUCTION

Despite a distinctive effort in a cancer research, this disease is still considered a major health problem in developed countries. Although initial stages of solid tumors are relatively easy-to-manage, advanced stages (which are usually not localized to tissue of origin and frequently form metastases) are developing resistance to treatment. Therefore, precise understanding of individual subpopulations of cells in the tumor is thus highly needed. Nevertheless, in accordance with developments in technology, novel microscopy techniques were introduced in last decade (high-resolution, two-photon, holographic microscopy, etc.).

MHM (Multimodal Holographic Microscopy) is a relatively new modality suitable for high throughput long-term label-free live cell imaging. This approach opens new promising applications, but manual processing of such amounts of data is impracticable. Automatic segmentation of single cells in images from MHM is a complex and not trivial task, because of various shapes and sizes, creation of cell clusters and strong cell interactions.

A critical prerequisite of automatic segmentation of these cell image is then precise cell detection - identification of single cells. There are a huge number of methods for cell detection. The most popular approaches are based on distance transform [7], morphology operations, LoG filter [1, 3], maximally stable extremal region [5, 4], Hough transform, radial symmetry-based voting, etc. and now-day attention often focuses on supervised learning methods (see review [2]). One of the aims of this study is to compare some cell detection techniques on given data and determine their suitability on images from MHM. Because of the unsatisfactory results of these methods and their relatively small computational costs, combing of these methods for higher accuracy is another purpose of this study.

PC-3 cells acquired with MHM (example in Figure 1) were selected for methods evaluation . This data was chosen because its part of active research [6] and its automatic analysis is highly needed. 18 images with 638 cells were manually labeled with centroids and divided to train data set (4 images - 160 cells) a test data set (14 images - 478 cells). Train data set is used for parameters setting of individual methods and for creating methods combination. The results are evaluated on the test data.

2 METHODS

2.1 FOREGROUND SEGMENTATION

In images from MHM, image background can be easily segmented by thresholding. Threshold can be manually set to the same value for all images. All detection techniques can then use this image and restrict cell search only to foreground pixels. The accuracy can be increased if morphological erosion is applied to the binary foreground image and small binary objects are removed.

2.2 MULTI-SCALE LOG FILTER DETECTOR

The LoG filter is one of the most popular methods for blob object detection in medical image analysis. There is many modifications of this detector - simple LoG, multi-scale LoG, generalized LoG, hessian LoG. The simple LoG filter can be used only for detection of blobs with the same known size - filter scale σ. LoG filter at a scale σ is defined by equation

$$LoG(\mathbf{x},\sigma) = \nabla^2 G(\mathbf{x},\sigma) = \frac{\sigma^2 - ||\mathbf{x}||^2}{2\pi\sigma^6} e^{\frac{-||\mathbf{x}||^2}{2\sigma^2}}, \quad (1)$$

where G is 2D Gaussian function, $\mathbf{x} = (x,y)$ and $||\cdot||$ is euclidean norm [1]. In principle it is matched filter for detection of blobs.

In cell detection using only single scale filter is not suitable, because there is a large variation in cells size. Multi-scale LoG filtration uses bag of LoG filters with different sigma and for each position pick most appropriate scale. For picking the highest filter response the filter scale must be normalized $LoG(\mathbf{x},\sigma)_{norm} = \sigma^{\gamma} LoG(\mathbf{x},\sigma)$. Parameter $\gamma = 2$ for scale invariance, but this parameter can be tuned for preferring larger or smaller objects. Maximum Intensity Projection of the series of LoG filtered images $MIP(\mathbf{x}) = \max_{\sigma}(LoG_{norm}(\mathbf{x},\sigma))$, produced parametric image. Centers of cells can be obtained by thresholding or local maxima finding. Local maxima finding shown to be appropriate for used images.

2.3 HESSIAN ANALYSIS OF LOG DETECTOR

HLoG (Hessian analysis of LoG) uses the same bag of LoG filtered images, only optimal scale identification and cell center detection is different. It is known, that local Hessian Matrix for blob-like structure is positive definite. The Hessian H (computed from LoG filtered image) at position (x,y) can be approximated with differences in 2x2 neighborhood. H is considered as positive definite matrix if both, H_{11} and determinant H are positive (positive leading principal minors). The single cell is identified as every connected region with positive definite Hessian.

Optimal scale is in [3] selected as scale where the mean intensity of the LoG filtered image on positive definite locations is maximal. This is inappropriate for analyzing cell images due to variability in cell size. The method here is slightly modified, by identification of optimal scale for every cell cluster. Cell clusters can be obtained by foreground segmentation and connected components are considered as cell clusters. This method is not optimal, but size of cells is more probably being more similar in cell clusters than in the whole image.

2.4 DISTANCE TRANSFORM DETECTOR

Very popular and simple method for cell detection using DT (distance transform) of the foreground image. DT of foreground image is defined as distance to nearest background pixel (Euclidean distance is chosen as metric). Local maxima of the generated distance map are considered as cells. This method often detects many false cells. For this purpose h-maxima transform is used [7]. H-maxima

transform using grayscale morphology for elimination of small local maxima, where parameter h set depth of local maxima which are eliminated and it is used as a parameter of this method.

2.5 MAXIMALLY STABLE EXTREMAL REGION DETECTOR

The MSER (Maximally Stable Extremal Region) detector is also often used for cell detection. Extremal regions of gray-value image is defined as connected components of thresholded image $I_t = I > t$ for some t. MSER detector described in [4] produced stable extremal regions of image which are stable in sense of area variation w.r.t. changing threshold t. Minimal stability of extracted region can be set with parameters threshold step δ (% of intensity range) and maximal relative area change with this step. This method generate many regions that overlap and only smallest region generated with highest threshold are picked as in [5], where these region are found by dynamic programing on graph constructed from overlapped regions. Same results is here achieved by detection of local maxima in image, which is created as sum of all extremal regions binary images.

2.6 CELL CENTER REFINEMENT

For combining cell centers detected by multiple methods is appropriate to move centers to the same position. This can be done by moving the center to the cell intensity maxima, but for cell maxima identification complete cell segmentation (pre-segmentation) is needed. For method training even ground true centers are refined to the cell intensity maxima.

Binary foreground segmentation and position binary image (ones only at cell centers) is used for center refinement. Cell refinement consist of these steps:

- DT - with distances from cell centers (DT of inverse position binary image)
- watershed segmentation of DT image
- foreground division with watershed results (spiting foreground with watershed dams)
- maxima finding in each region (of divided foreground) and assign maxima as new cell centers

When Euclidean distance is used, then it may cause problems when center point of the neighbor cell is closer to some pixels of actual cell, than center of actual cell (but cells are separated with background). The better distance metric seems to be a geodesic distance [9] - distance to the center pixels, when only foreground is used for moving.

2.7 METHODS PARAMETERS SELECTION

Each of four methods has only one relevant parameter to be set, which are summarized in Table 1. These parameters were optimized w.r.t F1-score on the training data set, with complete search in typical range (20 steps were used). For LoG methods is the filter scale σ set in range of minimal an maximal cell diameter (assumed that cell size range is known). Maximal relative area variation of MSER was set to 0.25.

2.8 METHODS COMBINATION

Results of all detection methods are position binary images. Because of the shifting to the cell centers to maxima, when cell is detected, the center pixels are in the same position. The next step is to combine results from individual methods to achieve better accuracy than individual methods.

This problem can be defined as search for a best linear combination of cell center images or as a classification problem with position binary image as features. This linear combination for classification

Table 1: Methods parameters summary.

Method	Parameter	Parameter meaning	Typical range	Optimal value
Multi-scale LoG	γ	Filter scaling penalization	1 - 3	2.8
HLoG	γ	Filter scaling penalization	1 - 3	2.1
DT	h	Minimal depth of maxima	based on cell size	5
MSER	δ	Threshold step	0.5 - 5	3.6

Table 2: Methods combination weights from LDA, detection accuracy and computation time.

Method	w_n	Precision	Recall	F1 - score	computation time
Multi-scale LoG	5.260	0.8644	0.7823	0.8213	0.68s
HLoG	6.371	0.8520	0.7914	0.8206	12.56s
MSER	9.099	0.8047	0.7927	0.7987	1.55s
DT	11.883	0.8771	0.7862	0.8291	0.18s
Methods combination	t=-13.870	0.9002	0.8256	0.8613	14.97s + 0.62s

can be found with LDA (linear discriminant analysis) - this classifier is picked for its simplicity and fast evaluation of classification. LDA is able to found the ideal linear combination of features (projection to 1D space) to achieve the best discrimination and corresponding threshold for classification [8]. The LDA can be trained on the training data set - from position binary images of methods and true labels it produces a vector of weights $\vec{w}$ and threshold t. Resulting image can than be simply computed as tresholded linear combination of position binary images.

3 RESULTS AND DISCUSSION

Due to refinement of cell centers only centers on the same position as ground truth is evaluated as correctly detected - TP (true positive). Cell is consider as FN (false negative) if it is not detected and as FP (false positive), if cell center is detected and does not match the ground truth. The results are reported in terms of $Precision = TP/(TP+FP)$, $Recall = TP/(TP+FN)$ and $F1-score = 2TP/(2TP+FP+FN)$. *Precision* refers about falsely detected centers, *Recall* refers about not detected centroids and $F1-score$ refer about a combination of both.

As shown in Table 2, DT with $F1-score$ 0.8291 is best from proposed individual methods - its large *precision* indicate a small number of falsely detected cells, but all methods shown very similar accuracy. Resulting F1-score of combination of methods 0.8613 is only slightly better than individual results. This is caused by a large part of the cells (8.26%), that were not detect by any method. LDA results $\vec{w}$ and t show methods predictive power. Computational time is evaluated for MATLAB implementations. HLoG is computational expensive due to scale selection for every cluster. Methods combination must compute all methods plus centers refinement and linear combination.

4 CONCLUSION

This paper describes four cell detection methods tested for MHM cell images data set. Methods are evaluated and an approach for its combination to one more robust detector is also presented. Accuracy improvement (0.0322 of $F1-score$) is relatively small, if computation complexity and need of learning set is considered. But this paper also shown that center refinement and LDA can properly combine more detection methods and with another methods, this might have great potential for cell detection improvement.

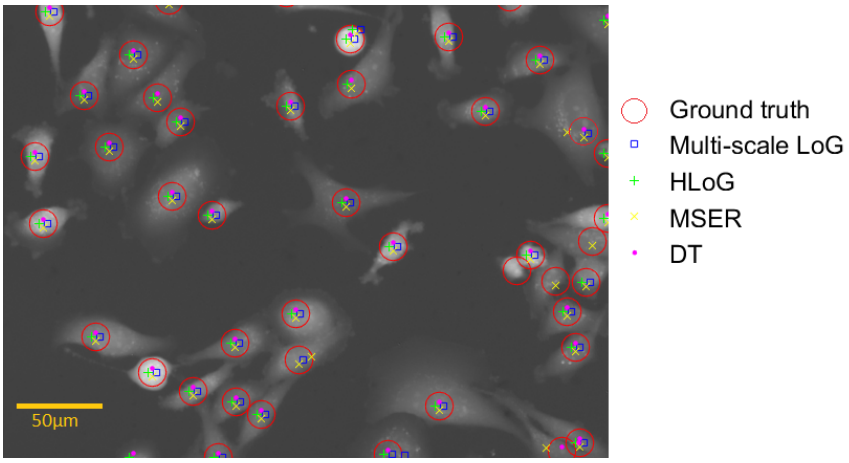

Figure 1: Cell detection methods results (cell image resolution 600x450px).

REFERENCES

[1] PENG, Huiming, Xiaobo ZHOU, Fuhai LI, Xiaofeng XIA a Stephen T.C. WONG. Integrating multi-scale blob/curvilinear detector techniques and multi-level sets for automated segmentation of stem cell images. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro [online]. IEEE, 2009, s. 1362-1365.

[2] XING, Fuyong a Lin YANG. Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review. IEEE Reviews in Biomedical Engineering [online]. 2016, 9, 234-263. DOI: 10.1109/RBME.2016.2515127. ISSN 19373333

[3] ZHANG, Min, Teresa WU a Kevin M. BENNETT. Small Blob Identification in Medical Images Using Regional Features From Optimum Scale. IEEE Transactions on Biomedical Engineering [online]. 2015, 62(4), 1051-1062. DOI: 10.1109/TBME.2014.2360154. ISSN 00189294.

[4] MATAS, J, O CHUM, M URBAN a T PAJDLA. Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing [online]. 2004, 22(10), 761-767. DOI: 10.1016/j.imavis.2004.02.006. ISSN 02628856.

[5] ARTETA, Carlos, Victor LEMPITSKY, J. Alison NOBLE and Andrew ZISSERMAN. Learning to Detect Cells Using Non-overlapping Extremal Regions. Miccai [online]. 2012,15(1), 348-56. DOI:10.1007/978-3-642-33415-3. ISSN 978-3-642-33414-6.

[6] BALVAN, Jan, Aneta KRIZOVA, Jaromir GUMULEC, et al. Multimodal Holographic Microscopy: Distinction between Apoptosis and Oncosis. PLOS ONE. 2015-3-24, 10(3). DOI: 10.1371/journal.pone.0121674. ISSN 1932-6203.

[7] THIRUSITTAMPALAM, Ketheesan, M. Julius HOSSAIN, Ovidiu GHITA a Paul F. WHELAN. A Novel Framework for Cellular Tracking and Mitosis Detection in Dense Phase Contrast Microscopy Images. IEEE Journal of Biomedical and Health Informatics [online]. 2013, 17(3), 642-653. DOI: 10.1109/TITB.2012.2228663. ISSN 21682194.

[8] BALAKRISHNAMA, Suresh; GANAPATHIRAJU, Aravind. Linear discriminant analysis-a brief tutorial. Institute for Signal and information Processing, 1998, 18.

[9] CÁRDENES, Rubén, Carlos ALBEROLA-LÓPEZ a Juan RUIZ-ALZOLA. Fast and accurate geodesic distance transform by ordered propagation. Image and Vision Computing [online]. 2010, 28(3), 307-316. DOI: 10.1016/j.imavis.2009.05.013. ISSN 02628856.