33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Using Multimodal Information for the Segmentation of Fluorescent Micrographs with Application to Virology and Microbiology

Christian Held, Jens Wenzel, Rike Webel, Manfred Marschall, Roland Lang, Ralf Palmisano and Thomas Wittenberg

Abstract— **In order to improve reproducibility and objectivity of fluorescence microscopy based experiments and to enable the evaluation of large datasets, flexible segmentation methods are required which are able to adapt to different stainings and cell types. This adaption is usually achieved by the manual adjustment of the segmentation methods parameters, which is time consuming and challenging for biologists with no knowledge on image processing. To avoid this, parameters of the presented methods automatically adapt to user generated ground truth to determine the best method and the optimal parameter setup. These settings can then be used for segmentation of the remaining images. As robust segmentation methods form the core of such a system, the currently used watershed transform based segmentation routine is replaced by a fast marching level set based segmentation routine which incorporates knowledge on the cell nuclei. Our evaluations reveal that incorporation of multimodal information improves segmentation quality for the presented fluorescent datasets.**

I. INTRODUCTION

Fluorescence microscopy based experiments in microbiology or virology often require the evaluation of large datasets. Furthermore, manual examination of such complex data is prone to inter- and intra- observer errors. In order to improve both, reproducibility and objectivity of such experiments, (semi-) automatic image analysis software can be utilized. Thereby, the robust segmentation of the image content has shown to be the most challenging part.

Due to the variability of the experimental setup which includes the cell type, staining or microscope setup, segmentation methods must be able to adapt to the image content. This flexibility is maintained by using a parametric image processing chain. Thereby, the obtained segmentation performance depends on the combination of methods that is used for segmentation and on the adjustment of each method's individual parameters. Manual tuning of these parameters is time-consuming and requires expert knowledge on image processing. This knowledge has shown to be an obstacle for biomedical researchers in praxis.

C. Held, T.Wittenberg are with the Department of image Processing, Fraunhofer-Institute for Integrated Circuits, Erlangen, Germany

R. Palmisano is with the Department of Biology, University of Erlangen-Nuremberg, Erlangen, Germany

R. Webel and M. Marschall are with the Institute for Clinical and Molecular Virology, University of Erlangen-Nuremberg, Erlangen, Germany

J. Wenzel and R. Lang are with the Institute for Immunology and Hygine, University of Erlangen-Nuremberg, Erlangen, Germany

In order to perform an adjustment of the segmentation parameters without expert knowledge on image processing, Wittenberg et. al [1] have proposed an adaptive image analysis concept. Thereby, a small subset of reference images is selected from the complete dataset and annotated manually. This ground truth is then used for training of a so-called segmentation engine. During the training step, the most effective method is selected from a pool of available segmentation methods and its parameters are automatically adjusted to the ground truth. This knowledge can then be used for segmentation of the remaining images of the dataset. Consequently, a high quality segmentation result is obtained without manual selection of a specific segmentation method and without manual adjustment of its parameters.

For the improvement of this adaptive segmentation scheme, we carry out research on segmentation methods that enable the robust segmentation of different cell types and stainings by adjusting the workflow's parameters. Based on the observation that experienced users often use multiple modalities for the recognition and splitting of confluent cells, this procedure shall be imitated by using multimodal segmentation methods. Assuming that each cell contains exactly one nucleus, this knowledge about the cell nuclei can be utilized for segmentation of the cells. Due to the automatic parameter adaption, the time required for segmentation of the cells should be less than $2s$. We have therefore selected the fast marching level set method (FM) [2] and investigate its applicability for the segmentation of multimodal image data. Particularly, appropriate speed functions are discussed. Furthermore, the multimodal method is compared to a mono modal watershed transform (WT) [3] based scheme.

II. MATERIALS AND METHODS

For evaluation of the described methods, two different datasets are used. Both datasets consist of two modalities whereof one modality depicts the cells and the other modality shows the corresponding cell nuclei. For segmentation of the nuclei, a WT [3] based segmentation routine is used. Then, cells are segmented by utilizing a FM method, where the initialization is based on the segmentation of the cell nuclei. For the automatic adjustment of each method's parameters, a genetic algorithm (GA) [4] is used including three fold cross validation. Performance of this multimodal method can then

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

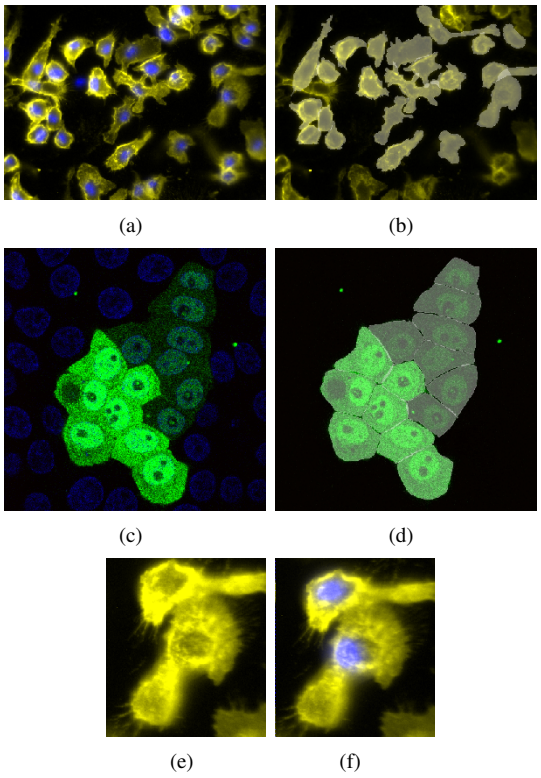


Fig. 1. Representative examples for the Macrophage (a) and the HeLa dataset (c) including manually annotated ground truth cells (b) and (d). Note that cells whose boundaries could not be determined unambiguously were excluded from the ground truth. Image (e) shows macrophages which are erroneously split by most human annotators. In this case nucleus information (f) is required to perform a correct annotation of the macrophages.

be compared to the results obtained by application of the WT based routine to the cell image.

A. Image data

The first dataset consists of 21 CD11b/APC stained images of murine bone marrow macrophages that were stimulated with LPS. Furthermore, a DAPI staining is used for visualization of the cell nuclei. In this dataset, 421 cells have been manually annotated by an experienced microbiologist. The second dataset consists of 6 micrographs containing GFP stained HeLa cells and the corresponding DAPI stained cell nuclei. In this dataset 85 manually annotated cells serve as ground truth. For representative examples of the datasets see Figure 1.

Especially for the macrophage dataset, an accurate segmentation of all cells is hardly possible without incorporation of nucleus information, even for experienced users. Such an ambiguous example is shown in Figure 1(e). Furthermore, some cells depicted in this dataset cannot be separated unambiguously by an experienced human user. Such cells are not included in the manually generated ground truth as can be seen in Figure 1.

B. Mono modal segmentation method

As the automatic parameter optimization incorporated in our segmentation scheme requires fast and robust segmentation methods, we have implemented a WT based segmentation method. For pre-processing, Gaussian smoothing and morphological opening by reconstruction [5] are applied, where the strength of the Gaussian smoothing step is adjustable by a parameter σ_n determining the standard deviation of the Gaussian function. The applied morphological opening by reconstruction method is influenced by the radius of the utilized structuring elements r_n. After pre-processing, fore- and background pixels are separated by a threshold operation. Therefore, a k-means clustering based method is used, performing an efficient histogram based clustering of the image's intensity values. Then, all clusters except the darkest cluster are regarded as foreground. This enables efficient adjustment of the applied threshold value by variation of the number of clusters k_n, while preserving a small parameter space (usually $2 \leq k_n \leq 7$).

After pre-processing and binarization of the image the WT [3] is applied for separation of touching cells. The WT is either applied to the distance transformed binary image, to the pre-processed image or to the gradient image. Our implementation enables adjustment of the used input images by adjusting a parameter m_n which is integrated into the optimization step.

C. Multi modal segmentation of cells

Because WT based methods suffer from over-segmentation, knowledge about the cell nuclei is used for segmentation of the cells to eliminate this problem. This knowledge can efficiently be integrated by the FM method [2]. The FM algorithm can be described as an initial wave front that starts moving at the boundary of each cell's nucleus. The front is expanding with a speed function $F(\mathbf{x}) > 0$, where $\mathbf{x}$ denotes the position of the current pixel. Pixels that are likely to represent the boundary of a cell (e.g. pixels with large gradients) can hardly be passed. This is implemented by a small value of $F(\mathbf{x})$. Cell boundaries can then be described as points where wave fronts with different starting labels touch each other. This enables splitting of non overlapping cells. The appropriate choice of $F(\mathbf{x})$ is consequently crucial to obtain a good splitting performance.

Some of the depicted HeLa cells show a high noise level. This impedes computation of robust features that can be used for detection of the cell boundaries and for splitting of the cells. In order to reduce the noise level, a Gaussian filter with standard deviation σ_c is applied to the fluorescent image.

To prevent the fast marching level set method from running into background tissue, foreground- and background pixels are separated by a threshold method. In order to combine flexibility with efficiency we use the already described k-Means clustering based threshold scheme. The resulting binary image is denoted as $I_B(\mathbf{x})$ with:

$$I_B(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \text{ foreground} \\ 0, & \text{else} \end{cases}. \quad (1)$$

After separation of cells from the background, individual cells have to be segmented and separated from each other.

In order to improve the performance of the splitting process, the previously segmented cell nuclei are utilized as an initialization for the FM algorithm.

Due to the known drawback that the fast marching level set method tends to leaking effects, the curvature term $\kappa(\mathbf{x})$ proposed by Nilsson [6] is modified and included in our speed function. Modification is required because our image contains multiple fronts starting at different cell nuclei. Therefore, the motion of each front is tracked in an additional label image $I_{\mathrm{L}}(\mathbf{x})$. For estimation of the curvature $\kappa(\mathbf{x})$, n equally spaced points $\Delta\mathbf{x}_i, i = 1,...,n$ that are located on a discrete circle with radius r are analyzed. Curvature is then estimated by counting the number of points that share the label of the current pixel:

$$\kappa(\mathbf{x}) = \frac{1}{n}\#\{i : I_{\mathrm{L}}(\mathbf{x}) = I_{\mathrm{L}}(\mathbf{x}+\Delta\mathbf{x}_i)\} - \frac{1}{2}. \tag{2}$$

Combining the estimated curvature with gradient information leads to the speed function F_1:

$$F_1(\mathbf{x}) = \frac{1+\alpha\kappa(\mathbf{x})}{I_{\mathrm{G}}(\mathbf{x})} I_{\mathrm{B}}(\mathbf{x}), \tag{3}$$

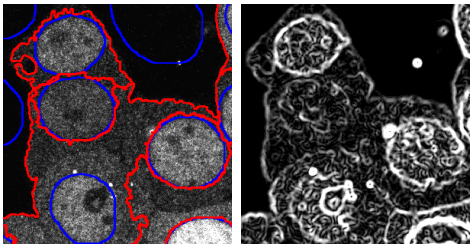
where $I_{\mathrm{G}}(\mathbf{x})$ denotes the gradient magnitude in the cell image at position $\mathbf{x}$ and α determines the influence of the curvature term.

Depending on the staining, cell nuclei most often show different intensities than the cytoplasm of a cell. In the Macrophage dataset, most cell nuclei appear darker than the surrounding tissue, whereas for the HeLa cell dataset most cell nuclei show an increased GFP intensity. This intensity deviation leads to additional edges in the image, especially for the HeLa cell dataset. Due to the variable magnitude of these additional gradients, some cell nuclei cannot be traversed by the wave front (see Figure 2). To decrease liability to edges that are aroused by cross talk, gradient magnitude should be decreased for pixels that are very close to a cell nucleus. This effect is achieved by incorporation of the distance transformed nucleus image $I_{\mathrm{D}}(\mathbf{x})$. The obtained speed function F_2 is given by:

$$F_2(\mathbf{x}) = \frac{1+\alpha\kappa(\mathbf{x})}{I_{\mathrm{G}}(\mathbf{x}) * I_{\mathrm{D}}(\mathbf{x})^d} I_{\mathrm{B}}(\mathbf{x}), \tag{4}$$

where d denotes a weighting factor that determines the strength of the applied a priori distance term.

Comparison of the cell and the nucleus modality for the macrophages shows that some of the cells do not express any GFP signal. Therefore, the cell is not visible in the GFP modality. Such cells are initialized with the segmentation result obtained by segmentation of the cell nuclei. Then, growing is impeded by the binary image based term $I_{\mathrm{b}}(\mathbf{x})$ except for some noise pixels. As these cells are not relevant for the evaluation, they are removed by a heuristic post processing step. Therefore, size of each segmented nucleus n_{nucleus} is compared to the size of the segmented cell n_{cell} and removed from the dataset if $n_{\mathrm{nucleus}} > \frac{n_{\mathrm{cell}}}{1.1}$.



(a) (b)

Fig. 2. Image (a) shows the segmentation result of a FM method with a gradient magnitude based speed function (red). Furthermore, cell nuclei are displayed in blue. This image illustrates that some cell nuclei cannot be passed by the wave front due to the presence of large gradients (b) and imperfect initialization.

D. Parameter optimization and cross validation

In order to enable a fair comparison of the described speed functions, the free parameters of each method are discretized and automatically optimized by using a genetic algorithm [4]. Therefore, the GALib [7] has been used. In order to avoid over-fitting and to separate training from testing data, a three fold cross validation is applied.

E. Performance Measurement

For assessment of the segmentation quality, each segmented cell region S_i, $i = 1,...,n$ is compared to the best matching manually annotated ground truth cell T_j, $j = 1,...,m$, where n denotes the number of segmented cells and m the number of ground truth cells using the Jaccard similarity measure [8]. The Jaccard similarity J_{ij} compares the intersection between S_i and T_j with the union of S_i and T_j by:

$$J_{ij} = \frac{|S_i \bigcap T_j|}{|S_i \bigcup T_j|}. \tag{5}$$

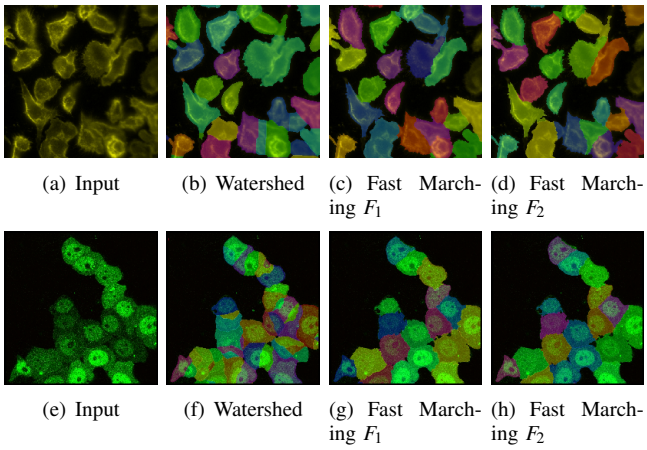
We have chosen this similarity measurement because corresponding research projects evaluate the size and intensity profiles of the cells [9]. Therefore, overlap based similarity is more important than shape based details that would be covered for example by the Hausdorff distance.

For optimization of the genetic algorithm, the average Jaccard similarity is used as objective function:

$$O = \sum_{i=1}^{n} J_{ij}. \tag{6}$$

III. RESULTS

After parameter optimization, the Jaccard similarity is evaluated for each cell. Analyzing the similarity for the HeLa cell dataset shows that usage of the 'mono modal' watershed transform based segmentation method results in a median similarity of 0.757. Incorporating knowledge on the cell nuclei by using a Fast Marching Level set method and the described speed function F_1, the median accuracy increases to 0.926. Segmentation performance of the speed function F_2 is comparable to performance of F_1 and results in median similarity of 0.919 for the HeLa cell dataset. Comparison



(a) Input (b) Watershed (c) Fast Marching F_1 (d) Fast Marching F_2

(e) Input (f) Watershed (g) Fast Marching F_1 (h) Fast Marching F_2

Fig. 3. Segmentation results for the Macrophage (a, b, c, d) and the HeLa dataset (e, f, g, h).

of the described segmentation methods on the Macrophage dataset achieve a median Jaccard coefficient of 0.719 for the mono modal watershed transform based method, 0.864 when using the Fast Marching Level Set method with F_1 and 0.851, for F_2. Figure 4 shows a box plot including minimum, maximum, 75%, 50% and 25% quantil values for both datasets. These results are consistent with the results obtained by manual inspection of representative images (see Figure 3).

IV. DISCUSSION

The obtained performance values and the example images clearly show that the mono modal watershed transform based segmentation method is only able to provide a basic segmentation of the complex cells contained in the depicted datasets. By using multimodal segmentation methods which incorporate knowledge on the segmented cell nuclei, segmentation performance is improved for both datasets. This is achieved especially by the elimination of false positively and false negatively detected cells. Remaining over and under segmentation artifacts are caused by boundary cells whose nuclei are not contained in the image and erroneously segmented cell nuclei. Furthermore visual inspection of the image shows that most cells are split similar to the way a human user would split them. A high quality segmentation result which is comparable to a manually annotated dataset can therefore be obtained by removing erroneously split cells in a manual post processing step. This procedure is much less time consuming than manual annotation of the images.

V. CONCLUSIONS

We have demonstrated how multimodal information can efficiently be integrated into a fast marching level set based segmentation routine. Our evaluations clearly show that usage of this multi modal information improves the Jaccard similarity for the presented datasets compared to the watershed transform based mono modal scheme.

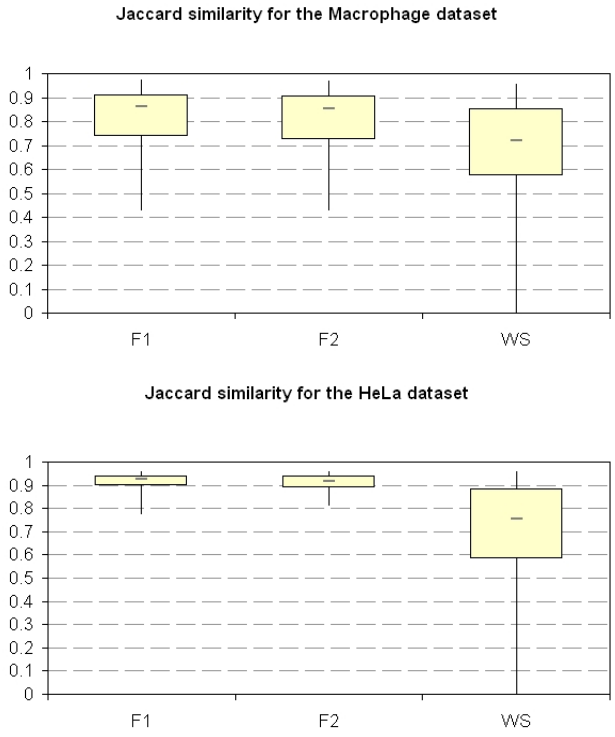


Fig. 4. Comparison of the segmentation performance for the HeLa and the Macrophage dataset. Performance is measured by using disjunctive training and testing data. The box plots show the performance median and 25% and 75% quantiles. Furthermore, the minimum and maximum outliers are displayed for each dataset.

VI. ACKNOWLEDGMENTS

We would like to thank the DFG for funding the SFB 796 in projects A4, C3, B6 and Z.

REFERENCES

[1] T. Wittenberg, F. Becher, M. Hensel, and D. G. Steckhan, "Image segmentation of cell nuclei based on classification in the color space," *IFMBE Proceedings*, vol. 22, no. 7, pp. 613–616, 2009.

[2] J. A. Sethian, "Level set methods and fast marching methods," *Cambridge University Press*, 1999.

[3] J. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," *Mathematical morphology*, p. 187, 2000.

[4] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, 1989.

[5] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *Image Processing, IEEE Transactions on*, vol. 2, no. 2, pp. 176–201, 1993.

[6] B. Nilsson and A. Heyden, "A fast algorithm for level set-like active contours," *Pattern Recognition Letters*, vol. 24, no. 9-10, p. 13311337, 2003.

[7] M. Wall, *GAlib: A C++ library of genetic algorithm components*, 1996.

[8] C. J. V. Rijsbergen, "Information retrieval," *Department of Computer Science, University of Glasgow*, 1979.

[9] R. Webel, J. Milbradt, S. Auerochs, V. Schregel, C. Held, K. Nbauer, E. Razzazi-Fazeli, C. Jardin, T. Wittenberg, H. Sticht, and M. Marschall, "Two isoforms of the protein kinase pUL97 of human cytomegalovirus are differentially regulated in their nuclear translocation," *The Journal of General Virology*, vol. 92, no. 3, pp. 638–649, 2011.