

3D Segmentation and Quantification of Mouse Embryonic Stem Cells in Fluorescence Microscopy Images

N. Harder¹, M. Bodnar², R. Eils¹, D. L. Spector², K. Rohr¹

¹Dept. Bioinformatics and Functional Genomics, Biomedical Computer Vision Group, University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg

²Spector Lab, Cold Spring Harbor Laboratory (CSHL), NY, USA

`n.harder@dkfz-heidelberg.de`

Abstract. We present an automatic approach for 3D segmentation of mouse embryonic stem cell nuclei based on level set active contours. Due to the specific properties of these cells, standard methods for cell nucleus segmentation and splitting of cell clusters cannot be applied. Our segmentation approach combines information from two different channels, which represent the nuclear region and the nuclear membrane, respectively. Moreover, we perform segmentation of gene loci within two other channels which enables single cell quantification of gene distances.

1 Introduction

Embryonic stem cells (ES cells) are pluripotent cells, which can differentiate into any cell type of the adult body. This striking property makes ES cells a highly interesting study target, likewise for basic research as well as for potential clinical and therapeutic applications. To better understand the principles of pluripotency and differentiation it is crucial to study the specific mechanisms of gene expression in differentiating ES cells.

In this work, multi-channel and multi-cell 3D images of fixed ES cell nuclei have been acquired to examine gene positioning in differentiating ES cells. To quantify and statistically analyze the characteristics of selected gene loci, the nuclear regions, the nuclear membranes, and the gene loci are imaged in different channels. First, individual nuclei have to be segmented to enable per-cell analysis of the multi-cell images. Next, gene loci have to be segmented and gene distances as well as nuclear volumes have to be determined. Manual segmentation and quantification of such multi-channel 3D image data is difficult and error prone since information from different channels and different spatial dimensions has to be considered at once. To assist and accelerate the analysis of the image data we developed an automatic image analysis approach comprising methods for segmentation and quantification. However, automatic segmentation of ES cell nuclei is very challenging because of their specific properties: First, ES cell nuclei have a high nuclear-to-cytoplasmic ratio, i.e. nuclei lie very close to each other, as there is almost no cytoplasm between nuclei. In addition, these cells grow in

colonies and thus form dense cell clusters. Second, ES cell nuclei have highly irregular shapes, including foldings and invaginations of the nuclear membrane (Fig. 1, lamin A and C). Consequently, a priori assumptions about the nuclear shapes cannot be used for splitting up clusters of cell nuclei.

A number of approaches for 3D segmentation of cell nuclei have been described, e.g., based on watershed transform (e.g., [1]) or active contours (e.g., [2]). However, these approaches deal with other types of cells which do not exhibit the challenging properties of mouse ES cells as described above. On the other hand, previous approaches developed for ES cell segmentation have only been used for 2D images and for data where cell nuclei are less densely clustered (e.g., [3, 4]).

2 Materials and Methods

To study gene positioning during embryonic stem (ES) cell differentiation based on 3D multi-channel multi-cell images, we have developed an automatic approach for segmenting cell nuclei as well as gene loci. To determine distances between gene loci and statistics about properties of the cell nucleus we developed a software tool which allows performing individual postprocessing and quantification.

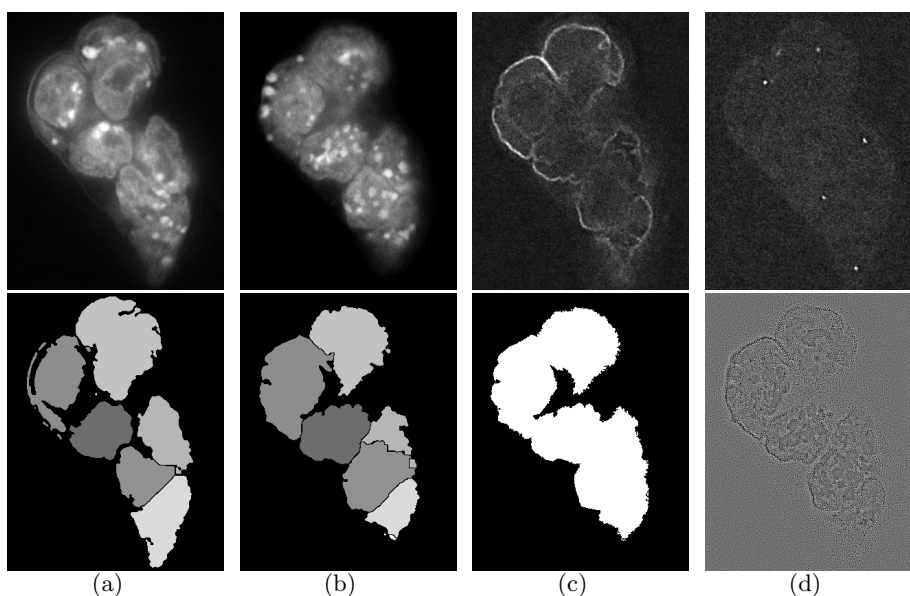


Fig. 1. Examples of single z-slices. (a) top: DAPI channel ($z=30$), bottom: corresponding final segmentation result, (b) top: DAPI channel ($z=41$), bottom: corresponding final segmentation result, (c) top: lamin B channel ($z=41$), bottom: result of region-adaptive thresholding on (b, top), (d) top: one of the two FISH channels ($z=41$), bottom: combined Laplace image of (b, top) and (c, top).

2.1 Image Data

High-resolution four-channel 3D images of clusters of fixed mouse ES cells have been acquired using widefield fluorescence microscopy, and gene loci have been labeled using fluorescence in situ hybridization (FISH). The first channel represents the nuclear regions (DAPI staining), while the second channel shows the nuclear periphery (lamin B channel) which has been labeled by immunofluorescence microscopy. The third and fourth channel provide FISH signals of two different gene loci. For each channel, 3D images have been acquired with a resolution of 512×512 pixels in the x-y plane and 73 to 87 z-slices (voxel size $0.1 \times 0.1 \times 0.2 \mu\text{m}$, 16 bit). Each image includes about five to eight cells (Fig. 1).

2.2 Automatic 3D Segmentation

Images were preprocessed and then the nuclear regions as well as the FISH signals were segmented and labeled in 3D to allow quantification of single cells. Prior to the segmentation of the nuclear regions we removed bright regions of condensed chromatin in the DAPI channel. To this end, we performed a slice-wise segmentation of these regions based on a tophat transform followed by automatic thresholding using the Renyi entropy [5] and median filtering. Next, the segmented regions were masked with a locally determined mean gray value and the edges of the masked regions were smoothed with a mean filter. As a result we yield relatively homogenous nuclear regions as well as a segmentation of the condensed chromatin regions.

To segment the nuclear regions we developed the following two-step approach. In the first step, initial contours are determined based on region-adaptive thresholding, and in the second step, the initial contours are refined using Laplacian-based active contours [6]. First, 3D region-adaptive thresholding was applied on the preprocessed DAPI images, providing a segmentation of the complete cell cluster (Fig. 1c, bottom). Local gray value thresholds were determined for overlapping image regions using Otsu's method. Afterwards, a 3D Euclidean distance transform followed by 3D watershed transform was applied to roughly subdivide the cellular region, providing initial contours.

In the second step, the initial contours were evolved to match the true nuclear regions using a 3D level set active contour method operating on the Laplacian image of the preprocessed DAPI channel. Since level set-based active contours are topologically flexible, the initial contours can merge or split during contour evolution. To further improve the performance of this method we additionally included information from the lamin B channel (Fig. 1d, bottom). This was done by combining the normalized Laplacian image of the DAPI channel with the normalized Laplacian image of the lamin B channel by a slice-wise summation. Since computing the Laplacian image is relatively sensitive to noise we smoothed all images using an anisotropic diffusion filter [7] prior to Laplace filtering. Finally, the resulting refined regions were split using a 3D watershed transform based on the 3D Euclidean distance map (Fig. 1a-b, bottom).

For segmentation of the FISH signals we first applied a slice-by-slice tophat transform after Gauss filtering to enhance the small bright spots, and second, we performed thresholding. The threshold for a 3D image was automatically determined based on the histogram of the brightest z-slice of the image using the Renyi entropy [5]. To split clustered FISH signals we performed a 3D watershed transform after Euclidean distance transform.

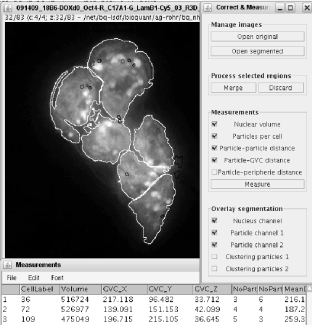
2.3 Semi-Automatic Postprocessing and Quantification

The final quantification of the FISH signals was performed semi-automatically for two main reasons. First, labeling of gene loci using FISH often produces nonspecific noise signals which cannot be distinguished from the target signal of the gene loci automatically. Consequently, to obtain accurate results biologists have to check the FISH channels and potentially select the relevant objects manually. Second, in some cases the automatic segmentation approach was not able to split closely clustered nuclei correctly. Thus, we developed a software tool to conveniently view the multi-dimensional data, to easily merge oversegmented nuclear regions or discard undersegmentations, or discard nonspecific FISH signals. Our software tool was implemented as a plugin for the public domain image processing software ImageJ [8] and creates on-the-fly custom overlays of segmentation results, provides synchronized views, and allows quick region inspection (Fig. 2, right). In addition, the tool computes per-cell measurements, such as the volumes of nuclei and FISH signals, the distances between single gene loci or clusters of loci, and distances to the nuclear center and periphery.

3 Experimental Results

To determine the accuracy of the 3D segmentation, we computed the Dice coefficient between automatic and manual segmentation in four 3D images. We merged oversegmented nuclei using our software tool for semi-automatic post-processing. For the total number of 23 nuclei, 5 nuclei were merged and one

Cell ID	Image 1		Image 2	
	Dice	Dist.	Dice	Dist.
1	0.94	0.30	0.92	1.19
2	0.93	0.78	0.93	1.37
3	0.91	1.28	0.92	2.59
4	0.94	0.55	0.96	5.91
5	0.88	3.81	0.90	3.42
6	0.84	3.54		
7	0.91	1.23		



The screenshot shows a 3D visualization of segmented nuclei in a dark environment. The software interface includes a 'Measurements' table at the bottom with columns for 'CellID', 'Volume', 'CNC-X', 'CNC-Y', 'CNC-Z', and 'NofPart/NofPart Meas'. The table contains three rows of data for cells 1, 2, and 3.

Fig. 2. Dice coefficients and distances of gravity centers between manual and automatic segmentation for 12 nuclei in two 3D images, and screenshot of the software tool.

nucleus was discarded in the postprocessing step. To provide ground truth for performance evaluation we manually segmented two 3D images including 12 nuclei. Top and bottom slices of the 3D images which only included background were discarded from the analysis. We yield a good agreement between the results of automatic and manual segmentation (Fig. 2); Dice coefficients range between 0.84 and 0.96, average is 0.92. The average distance of the nuclear gravity centers between manual and automatic segmentation is 2.9 voxels.

4 Discussion

We developed an automatic segmentation approach as well as a semi-automatic inspection and quantification software tool for the analysis of high-dimensional, high-content images of mouse embryonic stem (ES) cells. Automatic segmentation of ES cell nuclei is difficult because these cells are often densely clustered and have very irregular shapes. Our scheme can cope with these difficulties, and yields an average Dice coefficient of 0.92 for the segmentation of cell nuclei. Furthermore, the approach performs segmentation of gene loci in the FISH channels as well as a quantification on a single cell basis. In future work we aim to increase the level of automation and apply our approach to a larger number of 3D images.

Acknowledgement. Support of the BMBF NGFN+ project ENGINE is gratefully acknowledged. D.L.S. is supported by a grant from NIH/NIGMS 42694-21.

References

1. Lin G, Adiga U, Olson K, et al. A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. *Cytometry A*. 2003;56:23–36.
2. Dufour A, Shinin V, Tajbakhsh S, et al. Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Trans Image Process*. 2005;14(9):1396–1410.
3. Faustino GM, Gattass M, Rehen S, et al. Automatic embryonic stem cell detection and counting method in fluorescence microscopy images. *Proc IEEE ISBI*. 2009; p. 799–802.
4. Lowry N, Mangoubi R, Desai M, et al. Nonparametric segmentation and classification of small size irregularly shaped stem cell nuclei using adjustable windowing. *Proc IEEE ISBI*. 2010; p. 141–4.
5. Kapur JN, Sahoo PK, Wong AKC. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vis Graph Image Process*. 1985;29(3):273–85.
6. Sethian J. *Level Set Methods and Fast Marching Methods*. 2nd ed. Cambridge University Press; 1999.
7. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell*. 1990;12:629–39.
8. Rasband WS. *ImageJ*. Bethesda, MD, USA; 1997-2004. <http://rsb.info.nih.gov/ij/>.