SEGMENTATION OF CELL NUCLEI USING INTENSITY-BASED MODEL FITTING AND SEQUENTIAL CONVEX PROGRAMMING

Leonid Kostrykin* Christoph Schnörr† Karl Rohr*

* Biomedical Computer Vision Group, Dept. Bioinformatics and Functional Genomics, University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg, Germany
† Image and Pattern Analysis Group, Heidelberg University, Germany

ABSTRACT

We introduce a convex model-based approach for the segmentation of cell nuclei, which exploits both shape and intensity information. The model is directly fitted to the image intensities. Previous shape-based approaches either are not globally optimal or require prior binarization of an image. Our approach relies on a fast second-order optimization scheme to solve a sequence of convex programs and estimate the globally optimal solution based on the image intensities. Model fitting is performed within image regions which are determined by exploiting the local image structure. We evaluated our approach using fluorescence microscopy images of two different cell types and performed a quantitative comparison with previous methods.

Index Terms— Fluorescence microscopy, cell segmentation, model fitting, convex optimization

1. INTRODUCTION

Automatic segmentation of individual cell nuclei in microscopy images is a central task in many biological studies. When using fluorescent markers, cell nuclei often appear as bright regions of roughly elliptic shape. Segmentation is challenging particularly in the absence of distinct borders and the presence of strong image noise, as depicted in Fig. 1.

In previous work, intensity thresholding was often used, however, this approach is prone to image noise and intensity inhomogeneities. When segmentation is formulated as a combinatorial problem of discrete energy minimization, the solution can often be computed close to global optimality (e.g., [1, 2]). Such approaches are robust because of the absence of local energy minima. Many cell segmentation methods are based on a continuous, variational framework (e.g., [3, 4, 5, 6, 7]), where object contours are represented as level sets of functions. Formulating the evolution of such functions as a convex program assures that a globally optimal solution is found reproducibly for any initialization (e.g., [5]).

To better cope with strong image noise and other distortions, cell segmentation methods were proposed which exploit shape information, like shape-regularized variational level sets (e.g., [4, 7]) or statistical shape models (e.g., [6]). Other approaches rely on elliptical models, which are fitted by marked point processes (e.g., [2, 10]) or snake energy minimization (e.g., [11]). None of these shape-based methods (except [2, 10]) yield globally optimal solutions. In [2], elliptical models are randomly sampled from uniform distributions. However, a large number of samples is needed to sample a globally optimal solution, which is computationally expensive. [10] uses only circular models, requires a binarization of the image, and does not use the image intensities.

In this contribution, we introduce a new approach to cell nuclei segmentation, which is based on convex optimization and exploits shape and intensity information. An elliptical model is directly fitted to the image intensities, thus binarization of the image is not required. In our approach, the globally optimal minimizer of a continuous energy is estimated using a sequence of convex programs. A fast second-order method is employed to numerically solve each convex program. We also present a robust method for automatic selection of image regions, where model fitting is performed. We have evaluated our approach using fluorescence microscopy images of two different cell types and performed a quantitative comparison with previous methods.

2. METHOD

In this section, we describe our convex shape-based approach for cell nuclei segmentation. Sect. 2.1 details the model and the optimization scheme based on sequential convex programming. Sect. 2.2 describes an approach for automatic selection of image regions for model fitting.

2.1. Model Fitting by Convex Programming

We formulate the shape model as the zero-level set $C_{s_p} = \{x | s_p(x) = 0\}$ of a $p$-parametrized function $s_p$, which maps...
an image point to a real value. With a symmetric matrix $A$, a vector $b$, and a scalar $c$, we choose the parameterization

$$s_p(x) = (x - b)^T A (x - b) + c,$$  

(1)

where $p = \{A, b, c\}$. Then, for a 2-D image, the zero-level set $C_{s_p}$ is the whole image plane $\mathbb{R}^2$, the empty set, a single dot, or takes either an elliptic, hyperbolic, or linear shape.

The model function (1) also induces the two distinct regions $I_{s_p}^+ = \{x | s_p(x) > 0\}$ and $I_{s_p}^- = \{x | s_p(x) < 0\}$, which identify the inside and the outside of the model if $C_{s_p}$ is elliptical. Given an image $g : \Omega \to \mathbb{R}$, which depicts an object and its background in an image region $\Omega \subset \mathbb{R}^2$, and an intensity threshold $\tau$, the intensity model

$$y(x) = g(x) - \tau,$$  

(2)

induces the two regions $I_{s_p}^+$ and $I_{s_p}^-$ analogously, corresponding to the imaged object and its background. To segment the object, we seek the model parameters $p$ such that $I_{s_p}^+$ matches $I_{s_p}^+$ and $I_{s_p}^-$ matches $I_{s_p}^-$. Formally, we minimize

$$\min_{x \in \Omega} L(y(x); s_p(x)),$$

(3)

which penalizes each sample $x$ with $\text{sgn} y(x) \neq \text{sgn} s_p(x)$. Since the energy (3) is non-smooth, we instead consider

$$\sum_{x \in \Omega} \phi_{s_p}(y(x); s_p(x)),$$

(4)

where $\phi_{s_p}(y(x); s_p(x)) = \log \left(1 + e^{-\frac{y(x)}{s_p(x)}} \right)$. The function $\frac{1}{\log 2} \phi_{s_p}$ is a minimal upper bound of $L$ from (1), which also is smooth, and moreover convex in the model $s_p$. The factor $\frac{1}{\log 2}$ is omitted in (4) because the minimizers of a function are invariant to positive constant factors. The value $\sigma > 0$ governs how strong samples $x$ with $\text{sgn} y(x) \neq \text{sgn} s_p(x)$ are penalized. Since (4) is non-convex in the model parameters $p$, we use an approximation scheme to estimate its global minimizer by solving a sequence of convex programs, as detailed below.

Using $p = p_0 + p_\delta$, where $p_0$ is a current estimate of the minimizer and $p_\delta$ is an increment, leads to the decomposition

$$s_p(x) = s_{p_0}(x) + d_{p_\delta}(x) + d'_{p_\delta}(x),$$

(5)

being linear in $p_\delta$, and a higher-order term

$$d'_{p_\delta}(x) = \langle A_0, x x^T - 2xb_0^T + b_0b_0^T \rangle + 2 \langle b_0, A_0(b_0 - x) \rangle + c_\delta.$$

Then, within the trust region $\|A_\delta\|_F \leq \epsilon_A$ and $\|b_\delta\|_F \leq \epsilon_b$, with small $\epsilon_A, \epsilon_b$, we have $d'_{p_\delta} \approx 0$ and hence $s_p \approx s_{p_0} + d_{p_\delta}$. This approximation enables us to estimate the global minimizer of (4) by a sequential scheme: Keeping $p_0$ fixed, we compute the increment $p_\delta$ which globally minimizes the convex energy

$$\min_{x \in \Omega} \sum_{x \in \Omega} f_{p_\delta}(x),$$

(6)

subject to $\|A_\delta\|_F \leq \epsilon_A$ and $\|b_\delta\|_F \leq \epsilon_b$, which leads to $p_0 \leftarrow p_0 + p_\delta$ for the next iteration. The energy (6) is convex, since $\phi_{s_p}(y(x); s_p(x))$ is convex in $s$ and $s_{p_0} + d_{p_\delta}$ is linear in $p_\delta$. In our experiments we found that constraining $C_{s_p}$ to be located within a predefined image region increases the robustness. With $b$ as the center of an elliptic shape model $C_{s_p}$, we formulate a convex program that is solved in each iteration and constrains $b$ to stay within a maximum distance $r$ of a preset point $\mu$:

$$\min_{x \in \Omega} \sum_{x \in \Omega} f_{p_\delta}(x) \quad \text{s.t.} \quad \|A_\delta\|_F \leq \epsilon_A, \quad \|b_\delta\|_F \leq \epsilon_b, \quad \|s_{p_0} + d_{p_\delta} - \mu\| \leq r.$$  

(7)

The feasible set of (7) is convex because the constraints are determined by convex functions (norms) of affine mappings. These constraints are different to [12], where 3-D tubular shapes were enforced for vessel segmentation in 3-D CT data and location constraints were not used. In contrast, we avoid fitting an elliptical model when no such structure is present.

The global solution of (7) is found by primal-dual interior-point methods. Such methods solve the KKT conditions

$$2 \left[ A_\delta \ 0 \ b_0 \ b_0 + b_0 - \mu \right] \lambda + \sum_{x \in \Omega} \begin{bmatrix} \partial_{s_{p_0}} & \partial_{b_\delta} & \partial_{d_{p_\delta}} \end{bmatrix} f_{p_\delta}(x) = 0,$$

$$\text{Diag}(\lambda) \left[ \epsilon_A^2 - \|A_\delta\|^2 \quad \epsilon_b^2 - \|b_\delta\|^2 \quad r^2 - \|s_{p_0} + d_{p_\delta} - \mu\|^2 \right] = 1/t$$

for the primal-dual variable $p_\delta, \lambda$ with an adaptively chosen $t$ repeatedly [13]. We employed the solver [14], which uses a Newton-like scheme to solve these conditions. This second-order optimization scheme is significantly faster than using a first-order method to solve (7) directly. The initialization $p_\delta = 0$ is always feasible if $\|s_{p_0} - \mu\| \leq r$ is assured. We incrementally refine the estimate $p_0$ by solving (7), until the increment $\|p_\delta\|$ becomes smaller than $\epsilon_{\text{min}}$ or $\tau_{\text{max}}$ iterations are reached. This is outlined in Algorithm 1 which determines the solution $p_\ast = \{A_\ast, b_\ast, c_\ast\}$.

Algorithm 1: Sequence of convex programs.

input: $\Omega, g, \tau, \sigma, \mu, \tau_{\text{max}}, \epsilon_{\text{min}}$

$$\begin{array}{ll}
A_0, b_0, c_0 & \leftarrow 1 - \frac{1}{2}, \mu, 1; \\
\text{for } i \leftarrow 1 \text{ to } \tau_{\text{max}} \text{ do} & \\
\quad p_\delta & \leftarrow \text{solution of (7) using the initialization } p_\delta = 0; \\
\quad \text{if } \|p_\delta\| < \epsilon_{\text{min}} \text{ then break; } \\
\quad p_0 & \leftarrow p_0 + p_\delta; \\
\quad \text{return } p_0;
\end{array}$$

2.2. Selection of Image Regions

Our approach presented above determines the shape of a cell nucleus within an image region $\Omega$. Since the shape of $C_{s_p}$ is not restricted to ellipses by the constraints of (7), the result of Algorithm 1 suffices

$$A_\ast < 0 \quad \land \quad c_\ast > 0$$  

(8)

if and only if an elliptical structure is present in $\Omega$. Thus, our aim is to determine one region $\Omega$ for each cell nucleus in the
image, where falsely detected regions without cell nuclei are tolerable, because using \( \theta \) for testing allows identifying and discarding empty regions in a simple but reliable manner.

To determine suitable image regions, we employed a multiscale blob detector [15]. We used a very conservative threshold for the detection, since falsely detected blobs (and the corresponding fitted shapes) can be reliably discarded by \( \theta \).

Densely located cell nuclei were handled by applying the detector not directly to the image \( g \), but to the image \( g - \frac{\|\nabla g\|}{\max_x\|\nabla g(x)\|} \), where \( \|\nabla g\| \) is the gradient magnitude of \( g \). The image is then partitioned into distinct regions by assigning each pixel \( x \) to the \( i \)-th blob, for which the squared Mahalanobis distance \( \|x - \mu_i\|^2 \mathbf{D}_i (x - \mu_i) \) to the center \( \mu_i \) of the \( i \)-th blob is minimal. We used the matrix

\[
\mathbf{D}_i = -\mathbf{H}_i / (\sigma_i^2 \lambda_{\text{min}} (\mathbf{H}_i)), \tag{9}
\]

where \( \mathbf{H}_i \) is the Hessian of the image \( g \) at \( \mu_i \) and \( \lambda_{\text{min}} (\mathbf{H}_i) \) is the lowest eigenvalue of \( -\mathbf{H}_i \). Blobs with \( \lambda_{\text{min}} (\mathbf{H}_i) = 0 \) were not considered. The denominator in (9) normalizes the scale information in \( \mathbf{H}_i \) with respect to the blob scale \( \sigma_i \). Using the Hessian of an image for analyzing local structure is not new (e.g., [16]). However, in our approach we use the Mahalanobis distance and the normalization in (9). Fig. 2 shows that using the Euclidean distance tends to split cell nuclei into multiple regions (first row), the Mahalanobis distance without normalization is prone to falsely-detected blobs (second row), and using the Mahalanobis distance with normalization \( \theta \) performs best.

For each determined region \( \Omega \), we estimated the globally optimal fit of our shape model using Algorithm 1. The parameters \( \mu \) and \( r \) were chosen as the center and radius \( \sqrt{2}r \) of the blob, the threshold \( \tau \) was determined by Otsu thresholding of the region \( \Omega \). We used \( \sigma = 0.025 \), \( r_{\text{max}} = 50 \), \( r_{\text{min}} = 0.1 \), \( \epsilon_\Omega = 20 \), and \( \epsilon_{\text{br}} = 1 \) in all our experiments. For every fitted model passing condition \( \theta \), we further examined the eigenvalue decomposition \( \mathbf{A}_i = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top \) and computed the ellipse half-axes \( [l_1, l_2] = \sqrt{\epsilon_\Omega} (\mathbf{A})^{-1/2} \mathbf{U} \). We discarded those results additionally, for which the ellipse area \( \pi \cdot \|l_1\| \cdot \|l_2\| \) or circularity \( \frac{4\pi (l_1 l_2)^{1/2}}{\pi l_1 l_2} \) significantly differed from the means of the same image, and merged ellipses with sufficient overlap into single objects.

First, we studied the convergence properties of our method, which is based on a second-order optimization scheme. Using the image shown in Fig. 3a, we compared our method with the first-order scheme in [17], which was previously used for vessel segmentation in 3-D CT data, but omitting the tubular shape constraints. Our algorithm converged in 0.6 seconds (after five iterations there were hardly any changes), achieving a Dice coefficient of 84% (see Fig. 3b). In comparison, the scheme in [12] required already 3.3 seconds for one iteration (Dice coefficient: 80%) and did not terminate within 100 seconds (using the same parameter settings as for our method).

Second, we applied our method to images from three datasets. The first dataset (from [17]) consists of 49 images of Hoechst-stained NIH3T3 cells. Visible artifacts and non-elliptic nuclei shapes hamper the analysis of these images. The other two datasets are training datasets from the ISBI 2013 Cell Tracking Challenge [8], consisting of 51 images of GFP-transfected GOWT1 cells. Both datasets are difficult due to strong image noise and low contrast (e.g., see Fig. 3d).

We used the Dice coefficient and the SEG measure from [8] to evaluate our results. SEG was computed for all images. It is defined as the mean Jaccard similarity index \( J(R) = |R \cap S(R)| / |R \cup S(R)| \) of a ground truth cell nucleus \( R \) and its corresponding segmented object \( S(R) \). If no segmented object corresponds to \( R \), \( S(R) = \emptyset \) is set. For the Dice coefficient, we used all images, which a fully-labeled ground truth was available for (all 49 images for the NIH3T3 dataset and 4 images for each GOWT1 dataset).

For the evaluation with the Dice coefficient, we used the NIH3T3 dataset and both GOWT1 datasets. For NIH3T3, results were previously reported for a convex variational level sets approach [5], which does not use shape information. We performed a comparison with this method and Otsu thresholding, and studied the effectiveness of the location constraint of our method in [7]. The results in Tab. 1 show that the location constraint improves the accuracy significantly. In addition, our method outperforms [5] and Otsu thresholding.

We also assessed the performance of our method on the two GOWT1 datasets using the SEG measure and compared it with three other methods, for which [19] provided results. The first method is KTH [19], which performed overall best for segmentation in the ISBI 2013 Cell Tracking Challenge [8]. The other two methods are a blob detection approach (BLOB) [20] and a deep learning method (CPN) [18]. We also used Otsu thresholding for our comparison. The results in Tab. 2 show that our method performs best on the second
GOWT1 dataset. For the first GOWT1 dataset, our method is second best (somewhat worse than CPN), but significantly more accurate (13.6% better) than KTH, which achieved the best overall result for segmentation in the challenge [8]. Our method is also 7.9% better than BLOB. The low values for Otsu thresholding indicate the difficulties of the datasets, where cell nuclei are easily missed due to low contrast and strong noise. Considering this, and also that BLOB and particularly CPN exploit temporal information by performing joint segmentation and tracking, which our method does not, the results of our method are very competitive.

Table 1. Average Dice coefficients (in %) of our method with and without location constraints (LC) and other approaches. The best-performing method of each dataset is highlighted.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Otsu</th>
<th>Level sets</th>
<th>Ours w/o LC</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIH3T3</td>
<td>40.5</td>
<td>85</td>
<td>86.4</td>
<td><strong>87.4</strong></td>
</tr>
<tr>
<td>GOWT1-01</td>
<td>59.2</td>
<td>–</td>
<td>57.4</td>
<td><strong>63.7</strong></td>
</tr>
<tr>
<td>GOWT1-02</td>
<td>60.4</td>
<td>–</td>
<td>86.7</td>
<td><strong>89.4</strong></td>
</tr>
</tbody>
</table>

Table 2. SEG performance values (in %) of Otsu thresholding, CPN [18], BLOB [20], KTH [19], and our method. The best-performing method of each dataset is highlighted.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Otsu</th>
<th>CPN</th>
<th>BLOB</th>
<th>KTH</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOWT1-01</td>
<td>21.7</td>
<td><strong>85.1</strong></td>
<td>74.2</td>
<td>68.5</td>
<td>82.1</td>
</tr>
<tr>
<td>GOWT1-02</td>
<td>42.5</td>
<td>87.3</td>
<td>90.5</td>
<td>89.4</td>
<td><strong>91.3</strong></td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

In this paper, we have presented a new model-based approach for robust segmentation of cell nuclei in microscopy images. The model is directly fitted to the image intensities by solving a sequence of convex programs. A fast second-order optimization scheme determines the global solution of each convex program. By the choice of the constraints in the convex programs, our approach is intrinsically tolerant to falsely selected image regions. A quantitative comparison with previous methods showed that our approach yields competitive results or outperforms previous methods.

Acknowledgement: This work has been supported by the RTG 1653 of the German Research Foundation (DFG).

5. REFERENCES