HHS Public Access

Author manuscript

Med Image Comput Comput Assist Interv. Author manuscript; available in PMC 2016 December 04.

Published in final edited form as:

Med Image Comput Comput Assist Interv. 2015 October ; 9351: 332–339. doi: 10.1007/978-3-319-24574-4_40.

Fast Cell Segmentation Using Scalable Sparse Manifold Learning and Affine Transform-approximated Active Contour

Fuyong Xing[1,2] and **Lin Yang**[1,2]

[1]Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

[2]J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, USA

Abstract

Efficient and effective cell segmentation of neuroendocrine tumor (NET) in whole slide scanned images is a difficult task due to a large number of cells. The weak or misleading cell boundaries also present significant challenges. In this paper, we propose a fast, high throughput cell segmentation algorithm by combining top-down shape models and bottom-up image appearance information. A scalable sparse manifold learning method is proposed to model multiple subpopulations of different cell shape priors. Followed by a shape clustering on the manifold, a novel affine transform-approximated active contour model is derived to deform contours without solving a large amount of computationally-expensive Euler-Lagrange equations, and thus dramatically reduces the computational time. To the best of our knowledge, this is the first report of a high throughput cell segmentation algorithm for whole slide scanned pathology specimens using manifold learning to accelerate active contour models. The proposed approach is tested using 12 NET images, and the comparative experiments with the state of the arts demonstrate its superior performance in terms of both efficiency and effectiveness.

1 Introduction

Effective and efficient cell segmentation of pancreatic neuroendocrine tumor (NET) is a prerequisite for quantitative image analyses such as Ki67 counting. Many state-of-the-art approaches [11, 4, 16, 10] have been applied to cell/nucleus segmentation on specific medical images. In order to handle partial occlusion, shape prior models have been introduced to improve touching cell separation [2, 14] and liver segmentation [17].

However, it is inefficient to exploit the aforementioned shape prior models, which are not adaptive to large data sets, to fast segment thousands of cells in whole slide scanned specimens. In addition, it is necessary to learn multiple subpopulations of shape priors to handle shape variations. In this paper, we propose a high throughput and large-scale cell segmentation algorithm by combing high-level shape priors and low-level active contour models. The main contributions are: 1) A scalable sparse manifold learning algorithm to model multiple cell shape priors; 2) A novel affine transform-approximated active contour model that dramatically accelerates the shape deformation.

2 Methodology

An effective cell segmentation framework combining shape prior models and image appearance information is presented in [14]; however, it requires to solve one associated partial differential equation for each contour within each iteration and therefore is not suitable to handle a large number of cells in whole slide scanned images. In this paper, we propose a novel idea by assuming that there exists an affine transformation between any two similar cell shapes, and approximate shape deformation using the affine transformation instead of solving all computationally expensive Euler-Lagrange equations. In addition, since the shapes of cells on pancreatic NET images lie on a low-dimensional manifold due to the limited number of constraints of shape control (see Figure 1), we present a scalable sparse manifold learning algorithm for cell shape modeling, which can efficiently determine the shape memberships and perform the shape inference. In our algorithm, similar shapes are effectively grouped into the same cluster by taking advantage of the manifold geometry structure to allow the affine approximation for fast shape deformation in each cluster.

In the training stage, the cell shapes aligned with Procrustes analysis [7] are utilized to learn multiple subpopulations of shape priors using sparse manifold clustering and embedding. A deep convolutional neural network (CNN) [6] is trained with small image patches for shape initialization. In the testing stage, the CNN is exploited to generate probability maps with a sliding window on images, and initial cell shapes are obtained by applying an H-minima transform [5] to the maps, one per cell. These shapes deform towards cell boundaries with the affine transform-approximated deformable model. Meanwhile, shape inference and membership update are achieved by using the scalable manifold learning based on the learned shape repositories. The proposed approach alternately performs shape deformation and inference until the active contours converge.

2.1 Scalable Sparse Manifold Learning for Shape Prior Modeling

In our model, cell shape $x \in R^{2p\times 1}$ is represented by $p = 60$ landmarks following the rules in [14]. We propose to model shape priors by clustering training shapes into multiple subpopulations considering the intrinsic dimensionality of the manifold. The sparse manifold clustering and embedding (SMCE) [8] can robustly achieve simultaneous clustering and dimensionality reduction. However, SMCE is a transductive algorithm that is not able to handle out-of-sample data, and it requires a computational complexity of $\mathcal{O}(N^3)$ to solve the optimization problem over N new testing shapes. In this paper, we efficiently extend SMCE to handle out-of-sample data and update shape clusters via sparse encoding-based shape inference. Specifically, we first apply SMCE to the limited-size training data and obtain multiple subpopulations of cell shapes, then perform sparse encoding for each new shape, and finally assign the new shapes to corresponding clusters. In this scenario, the runtime computational complexity can be reduced from $\mathcal{O}(N^3)$ to $\mathcal{O}(N \cdot M^2)$ where M is the number of training shapes with $M \ll N$.

Shape Prior Modeling via Manifold Learning—SMCE formulates an optimization problem based on sparse representation to allow simultaneous clustering and embedding of

data lying in multiple manifolds. Given a set of aligned training cell shapes $\{x_i\}_{i=1}^{M}$, SMCE solves the following problem

$$\min_{c_i} \gamma \|Q_i c_i\|_1 + \|X_i c_i\|_2, s.t. 1^T c_i = 1, \forall i, \quad (1)$$

where $Q_i \in R^{(M-1)\times(M-1)}$ represents the proximity, and it is a positive-definite diagonal matrix with the j-th diagonal element equal to $\frac{\|x_j - x_i\|_2}{\sum_{k\neq i}\|x_k - x_i\|_2}$, such that the points near x_i receive low penalties. $X_i \in R^{2p\times(M-1)}$ is the normalized shape matrix and the j-th column is $\frac{x_j - x_i}{\|x_j - x_i\|_2}$, $j \neq i$. The γ is the sparsity weight, and $\mathbf{1}^T c_i = 1$ ensures translation invariance. Based on the solution to (1), we can build a similarity graph whose nodes represent the data points [8], and the manifold clustering is achieved by applying spectral clustering to the graph.

Shape Inference and Cluster Update—After the manifold clustering, we can obtain multiple subpopulations of shape priors $\{\phi_k \in R^{2_p \times M_k}\}_{k=1}^{K}$. The original shapes whose embedding vectors lie in the same manifold are similar to each other and form a shape repository/cluster, which is used to perform runtime shape inference for cell shapes assigned to this cluster in the testing stage. It is challenging to efficiently perform shape inference and determine the memberships of out-of-sample data considering the intrinsic dimensionality. Fortunately, any Lipschitz-smooth function defined on a smooth nonlinear manifold can be effectively approximated by a globally linear function with respect to local coordinate coding [15], and the time complexity of sparse encoding completely depends on the much lower intrinsic dimensionality. This indicates that each shape can be sufficiently represented by its coding based on its neighbors. Therefore, we propose to achieve shape inference and cluster update in a unified manner via sparse encoding. Specifically, given the learned shape repository $\Phi = [\phi_1 \ldots \phi_K] = [\psi_1 \ldots \psi_M] \in R^{2p\times M}$, we perform runtime shape inference

$$\min_{\{\alpha_i\}} \sum_{i=1}^{N} (\|\upsilon_i - \Phi\alpha_i\|_2 + \lambda \sum_{j=1}^{M} |\alpha_i| \|\psi_j - \upsilon_i\|^2), s.t.\ \mathbf{1}^T \alpha_i = 1,\ \forall_i. \quad (2)$$

This local coordinate coding converts the difficult nonlinear learning into a linear problem. With the locality constraint in (2), we project each shape to its local coordinate system and solve a smaller linear system for shape inference [13]

$$\min_{\{\alpha_i\}} \sum_{i=1}^{N} \|\upsilon_i - \Phi\alpha_i\|_2 + \lambda \|d_i \odot \alpha_i\|^2, s.t.\ \mathbf{1}^T \alpha_i = 1, \forall i, \quad (3)$$

where d_i measures the similarity between υ_i and ψ_j's in Φ, and $\odot$ is the element-wise multiplication. Equation (3) has an analytical solution and it can be solved efficiently. For

each shape v_i that belongs to cluster k, it selects its neighbors for sparse encoding such that only training shapes in ϕ_k are used for shape inference. Therefore, the nonzero components of the solution $\hat{\alpha}_i$ are grouped in the k-th segment of $\hat{\alpha}_i$ (see Figure 2), and we can determine the label of v_i as

$$\text{label}(v_i)=\max_k\{\|\hat{\alpha}_{i1}\|_0,\|\hat{\alpha}_{i2}\|_0,\ldots,\|\hat{\alpha}_{ik}\|_0,\ldots,\|\hat{\alpha}_{iK}\|_0\}, \quad (4)$$

where the $\hat{\alpha}_{ik}$ corresponds to the code for cluster k, and $\|\cdot\|_0$ is l_0-norm.

2.2 Fast Active Contour Model

Local Repulsive Deformable Model—For shape deformation, we introduce an edge detector into the Chan-Vese model [3] to better separate cells from the background. To efficiently model the interaction among shapes, for each cell we calculate the repulsion only from its nearest neighbors [14]. Formally, for image I with N cells, the locality-constrained Chan-Vese model can be described as

$$\begin{aligned}\min_V \lambda_1\sum_{i=1}^{N}\int_{\Omega i}(I(\mathbf{v})-h_i)^2d\mathbf{v}+\lambda_2\int_{\Omega b}(I(\mathbf{v})-h_b)^2d\mathbf{v}\\ +\lambda_3\sum_{i=1}^{N}\int_0^1 e(v_i(s))ds+\omega\sum_{i=1}^{N}\sum_{j\in G_i}\int_{\Omega_i\cap\Omega_j}1d\mathbf{v}+\mu\sum_{i=1}^{N}|v_i|,\end{aligned} \quad (5)$$

where Ω_i and Ω_b represent the regions inside v_i and outside all the contours, respectively, h_i (or h_b) denotes the average intensity of Ω_i (or Ω_b), $e(v_i(s))$ is the edge detector and chosen as $-\|\nabla I(v_i(s))\|^2$ ($s\in[0,1]$ is the parameter for contour representation), $|v_i|$ denotes the length of v_i, and G_i represents the neighbors of v_i. The original model [3] requires to solve N associated Euler-Lagrange equations of (5) for shape deformation ($i = 1, \ldots, N$)

$$\frac{\partial v_i}{\partial t}=\left|\frac{\partial v_i}{\partial s}\right|\mathbf{n}_i[-\lambda_1(I-h_i)^2+\lambda_2(I-h_b)^2-\lambda_3\nabla e(v_i)-\omega\sum_{j\in G_i}o_j(v_i)+\mu\kappa(v_i)], \quad (6)$$

where $\mathbf{n}_i$ is v_i's normal unit vector, and $o_j(v_i)$ is the indicator function: $o_j(v_i) = 1$ if $v_i\in\Omega_j$, otherwise 0. $\kappa(v_i)$ is the curvature. Solving (6) for each cell in each iteration is extremely computationally expensive in whole slice scanned images.

Affine Transform Approximation—In order to avoid solving a large number of Euler-Lagrange equations, we propose to deform shapes using (6) for only a few cells and approximate all the other contour evolvements using affine transforms. Since similar cell shapes are grouped into the same clusters, we assume in each cluster there exists a certain affine transform between any two shape instances. Considering a cluster with N_k cells, $\{v_i=[x_1^i\cdots x_p^i y_1^i\cdots y_p^i]^T\in R^{2p\times 1}\}_{i=1}^{N_k}$, we assume that v_i, $i\neq 1$, is created from v_1 via affine transformation $z_i\in R^{6\times 1}$:

$$v_i = S z_i,\ S = \begin{bmatrix} x^1 & y^1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x^1 & y^1 & 0 & 1 \end{bmatrix},\ z_i = \left[a_{11}^i a_{12}^i a_{21}^i a_{22}^i t_1^i t_2^i \right]^T,\ \forall i, \quad (7)$$

where $x^1 \in R^{p\times 1}$ and $y^1 \in R^{p\times 1}$ represent the first and second half of v_1. Therefore, we have $V=[v_1 v_2 \dots v_{N_k}]=S[z_1 z_2 \dots z_{N_k}]$ in each shape cluster.

For each iteration t, we randomly select only one shape in each cluster and solve its corresponding equation (6). After this selected shape deformation, we update the temporary positions of the other shapes. To preserve the cell shape information at iteration $t-1$, we apply a weight η to the update at iteration t

$$V^t=(1-\eta)V^{t-1}+\eta S^t Z^{t-1}, \quad (8)$$

where $V^t=[v_1^t v_2^t \dots v_{N_k}^t]$ and $Z^{t-1}=[z_1^{t-1} z_2^{t-1} \dots z_{N_k}^{t-1}]$. Thereafter, we perform shape inference using (3) to get the final contour positions $\hat{S}^t$ and $\hat{V}^t$ of each cluster k, and update the affine transform matrix as

$$Z^t=((\hat{S}^t)^T \hat{S}^t)^{-1} (\hat{S}^t)^T \hat{V}^t. \quad (9)$$

In next iteration $t+1$, we repeat this procedure by alternatively performing shape deformation and inference based on results obtained at iteration t. The membership of each shape is dynamically updated to ensure that it is assigned to a correct cluster. The affine transformation approximation in (8,9), which deals with a small-size matrix inverse (6×6), is much faster than solving N Euler-Lagrange equations within each iteration. In order to finetune the final segmentation, we can deform all contours using (6) only in last iteration.

3 Experiments

The proposed approach is extensively tested on 12 pancreatic NET whole slide scanned TMA discs, which are captured at $20\times$ magnification and contain cells ranging from around 4200 to 17600. The 13-layer CNN model [6] is trained with about 1.3 million image patches with size $32\times 32\times 3$ (half positive and half negative). We have tested different number of training cells for shape prior modeling and observed no significant variations on the performance when the training size is larger than 1000, and thus in total 1395 cells are randomly selected. The algorithm is coded using Matlab on a PC of Intel Xeon CPU with 12 cores and 128 GB RAM. We empirically set $\gamma = 10$ in (1), $\lambda = 0.005$ in (3), $\eta = 0.2$ in (8), and $K = 6$ shape clusters. The Chan-Vese model is relatively insensitive to parameters, which are $\lambda_1 = \lambda_2 = 10$, $\lambda_3 = 0.2$, $\omega = 2.5$ and $\mu = 1$ in (6).

Figure 3 shows the segmentation results using our method on two whole slide scanned TMA discs with size of 3882×3882, where thousands of cells are accurately segmented one-by-

one with shape preserving. More importantly, the average of running time for one whole slide with about 6300 cells is around 251 seconds. We randomly crop 12 image patches of size around 1300×800 from the whole slide images for quantitative analysis. The patches contain 300 to 1200 cells with large shape variations, and the ground-truth of each cell boundary is all manually annotated by pathologists for comparison. We compute the Dice similarity coefficient (*DSC*) [18] to measure the pixel-wise segmentation accuracy, and compare the proposed method with the *full active contour* (FAC), which does not use affine transform approximation but solves (6) for each cell within each iteration, and the transductive SMCE [8] for shape inference. Figure 4 shows that our method can produce competitive performance as the other two in terms of the accuracy with a much lower running time (around 14(Proposed),107(FAC),39(Transductive) seconds on one patch with 465 cells). Figure 4 also shows the running time (15 iterations) with respect to the number of cells, in which the proposed method exhibits the strong scalability and is significantly faster than the other two. The more cells we need to segment, the more advantages we gain using our method.

In Figure 5, we illustrate the comparative segmentation results on one zoomed-in image patch using mean shift (MS), isoperimetric (ISO) [9], marker-based watershed (MWS), graph-cut and coloring (GCC) [1], repulsive level set (RLS) [12], and the proposed approach. It is clear that the proposed method provides best results. Table 1 summarizes the comparative performance between the proposed method and the state of the arts with multiple metrics including DSC, Hausdorff distance (*HD*), and mean absolute distance (*MAD*) [18]. As one can tell, our approach provides the best performance in terms of the mean and standard deviation of the metrics. Table 1 also lists the running time of each algorithm, which demonstrates that our method produces best performance. Except that MS and GCC are implemented in C++, all the others are coded with Matlab. Watershed (MWS), due to its simplicity, is the fastest but with poor segmentation accuracy.

4 Conclusion

We propose a fast cell segmentation approach using scalable sparse manifold learning and affine transform-approximated active contour model, which exhibits outstanding scalability and can efficiently handle large scale images (whole slide scanned specimens) for high throughput analysis using a standard PC machine without parallel computing involving complex image partitioning and stitching.

References

1. Al-Kofahi Y, Lassoued W, Lee W, Roysam B. Improved automatic detection and segmentation of cell nuclei in histopathology images. TBME. 2010; 57(4):841–852.
2. Ali S, Madabhushi A. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. TMI. 2012; 31(7):1448–1460.
3. Chan TF, Vese LA. Active contours without edges. TIP. 2001; 10(2):266–277.
4. Chang H, Han J, Spellman PT, Parvin B. Multireference level set for the characterization of nuclear morphology in glioblastoma multiforme. TBME. 2012; 59(12):3460–3467.
5. Cheng J, Rajapakse JC. Segmentation of clustered nuclei with shape markers and marking functions. TBME. 2009; 56(3):741–748.

6. Cireşan, DC.; Giusti, A.; Gambardella, LM.; Schmidhuber, J. Mitosis detection in breast cancer histology images with deep neural networks. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. MICCAI 2013, Part II, LNCS. Vol. 8150. Springer Berlin; Heidelberg: 2013. p. 411-418.
7. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models-their training and application. CVIU. 1995; 61(1):38–59.
8. Elhamifar E, Vidal R. Sparse manifold clustering and embedding. NIPS. 2011:55–63.
9. Grady L, Schwartz EL. Isoperimetric graph partitioning for image segmetentation. TPAMI. 2006; 28(1):469–475.
10. Irshad H, Veillard A, Roux L, Racoceanu D. Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. RBME. 2014; 7:97–114.
11. Kong H, Gurcan M, Belkacem-Boussaid K. Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. TMI. 2011; 30(9):1661–1677.
12. Qi X, Xing F, Foran DJ, Yang L. Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. TBME. 2012; 59(3):754–765.
13. Wang J, Yang J, Yu K, Lv F, Huang T, Gong Y. Locality-constrained linear coding for image classification. CVPR. 2010:3360–3367.
14. Xing, F.; Yang, L. Robust selection-based sparse shape model for lung cancer image segmentation. In: Mori, K.; Sakuma, I.; Sato, Y.; Barillot, C.; Navab, N., editors. MICCAI 2013, Part III, LNCS. Vol. 8151. Springer Berlin; Heidelberg: 2013. p. 404-412.
15. Yu K, Zhang T, Gong Y. Nonlinear learning using local coordinate coding. NIPS. 2009:1–9.
16. Zhang, C.; Yarkony, J.; Hamprecht, FA. Cell detection and segmentation using correlation clustering. In: Golland, P.; Hata, N.; Barillot, C.; Hornegger, J.; Howe, R., editors. MICCAI 2014, Part I, LNCS. Vol. 8673. Springer International Publishing; 2014. p. 9-16.
17. Zhang S, Zhan Y, Dewan M, Huang J, Metaxas DN, Zhou XS. Towards robust and effective shape modeling: Sparse shape composition. MedIA. 2012; 16(1):265–277.
18. Zhou X, Huang X, Duncan JS, Yu W. Active contours with group similarity. CVPR. 2013:2969–2976.

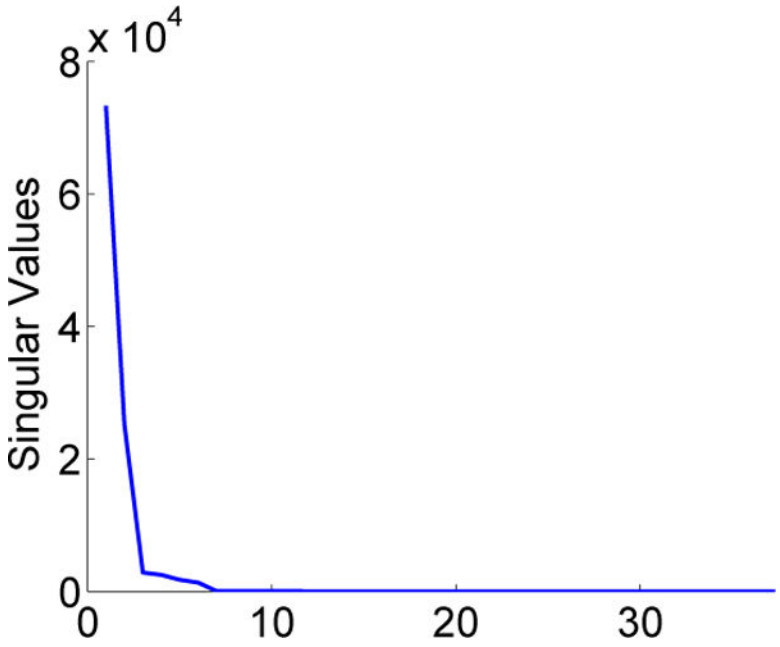

Fig. 1.

The singular values of cell shapes. The first 6 singular values are nonzero and significantly larger than the rest, and thus cell shapes actually lie in a union of subspaces with dimension around 6.

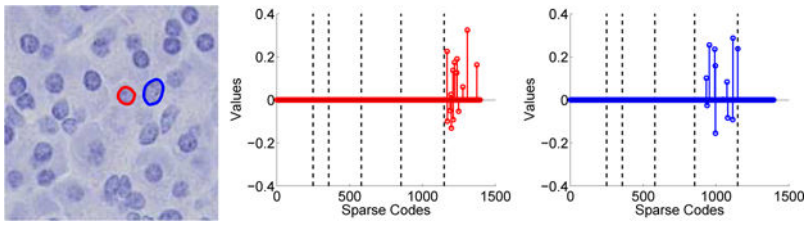

Fig. 2.

Left: Two cells with different shapes in one NET image. **Middle:** The sparse codes corresponding to the red cell, where the black dash vertical lines separate different shape clusters. **Right:** The sparse codes corresponding to the blue cell.

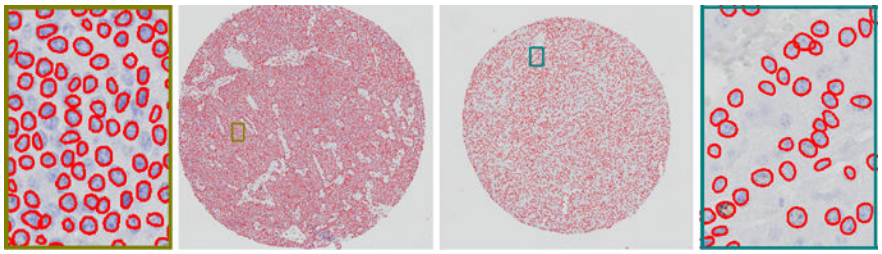

Fig. 3.
Segmentation of whole slice scanned TMA discs with the proposed algorithm.

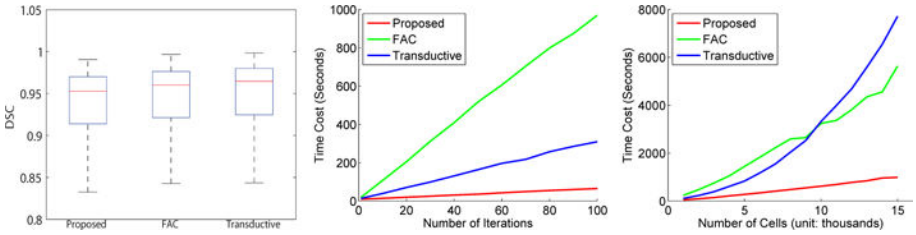

Fig. 4.

Performance with the proposed method, full active contour (FAC), and transductive learning. **Left:** Dice similarity coefficient. **Middle** and **Right:** The running time with respect to the number of algorithm iterations and cells, respectively.

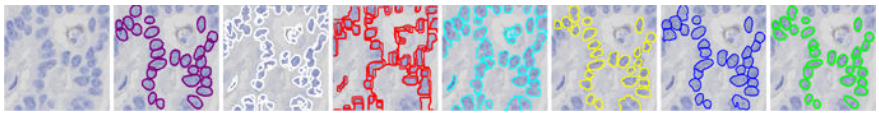

Fig. 5.

Comparative segmentation using different methods. From left to right: original images, ground truth, MS, ISO [9], GCC [1], MWS, RLS [12], and the proposed. MWS, RLS, and the proposed use the same initialization, and the lymphocytes are discarded on purpose. Note that cells touching image boundaries are ignored.

Table 1

Comparative pixel-wise segmentation accuracy

	MS	ISO	GCC	MWS	RLS	Proposed
DSC	0.62 ± 0.26	0.49 ± 0.26	0.58 ± 0.26	0.81 ± 0.12	0.63 ± 0.20	0.91 ± 0.11
HD	9.59 ± 9.78	16.74 ± 19.38	7.94 ± 7.90	4 88 ±. 3.39	7.83 ± 5.30	2.21 ± 3.05
MAD	5.63 ± 4.69	10.80 ± 9.84	5.55 ± 4.29	2.66 ± 1.84	5.02 ± 3.53	1.67 ± 2.04
Runtime†	10.07*	139.1	32.96*	3.07	317.2	36.34

†Runtime unit: seconds.

*MS and GCC are coded with C++, and others with Matlab.

Med Image Comput Comput Assist Interv. Author manuscript; available in PMC 2016 December 04.