

Segmentation of Cells with Partial Occlusion and Part Configuration Constraint Using Evolutionary Computation

Masoud S. Nosrati and Ghassan Hamarneh

Medical Image Analysis Lab., Simon Fraser University, BC, Canada
{smn6,hamarneh}@sfu.ca

Abstract. We propose a method for targeted segmentation that identifies and delineates only those spatially-recurring objects that conform to specific geometrical, topological and appearance priors. By adopting a “tribes”-based, global genetic algorithm, we show how we incorporate such priors into a faithful objective function unconcerned about its convexity. We evaluated our framework on a variety of histology and microscopy images to segment potentially overlapping cells with complex topology. Our experiments confirmed the generality, reproducibility and improved accuracy of our approach compared to competing methods.

1 Introduction

Histology and microscopy image analysis plays a crucial role in studying diseases such as cancer and in obtaining reference diagnosis (e.g. biopsy histopathology). Automatically segmenting cells in such images is one of the preliminary steps toward automatic image analysis and computer-aided diagnosis. In spite of recent advances in segmenting cells based on some homogeneity and smoothness characteristics, segmenting complex cells with a non-homogeneous appearance (with multiple internal regions) remains challenging. This problem becomes even more challenging when these complex cells overlap. Previous works addressed cell overlapping, for single-region cells, using post-processing [14,15,11] (e.g. finding connected components and using parameter sensitive morphological operations [11]). However, cells in histology and microscopy images typically consist of multiple regions (e.g. membrane, nucleus, nucleolus), each with a unique appearance model (intensity, color or texture) and unique geometric characteristics (e.g. cell size and shape prior). Furthermore, well defined spatial interactions usually exist between different regions of a cell (e.g. membrane contains nucleus, and nucleus contains nucleolus). Most existing methods have only considered simple structured cells and ignored their complex composition [1,3,2].

There are many types of priors that benefit the segmentation of spatially-recurring cells with appearance inhomogeneity along with cell-overlapping. Many state-of-the-art image segmentation methods are formulated as optimization problems, which are capable of incorporating multiple criteria (or priors) as energy terms in the objective function and examining the relative performance of different solutions. Incorporating several energy terms enables us to describe

the problem in more detail and thus obtain a more accurate formulation. On the other hand, adding more terms to the objective function generally makes it more complicated and harder to optimize.

In this work, we opt for ensuring the objective function is flexible enough (even if it is nonconvex) to accurately capture the intricacies of the cell segmentation problem. To optimize such objective function and to deal with imminent problems like initialization and local optima, we adopt a global optimization evolutionary computation method, genetic algorithm (GA), which can attain solutions close to the global optimum, does not require Euler Lagrangian or energy gradient calculations, is generally parallelizable, and allows for arbitrarily complex objective functions. Our framework allows us to leverage a variety of expert knowledge or priors by adding them as additional terms in the objective function without being overly concerned about convexification. Finally, to deal with the spatially recurring aspect in cell segmentation, we use genetic algorithms with *tribes* [13] to obtain multiple distinct solutions for our framework.

2 Problem Formulation

Given an n -channel 2D image $I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^n$, the goal is to segment the objects of interest (cells) in I . We represent the boundary of each object (or each part of a multi-region object) by $\mathbf{X}_i \in \Omega$, where i indicates the i^{th} part/region. Next, we review the useful priors in microscopy images that we can leverage.

- **Shape:** When an object has a specific geometrical shape (e.g. circle, ellipse, rectillipse, etc.) we model it by shape parameters such as $\mathbf{b} = \{\text{radius, major axis, eccentricity, etc.}\}$. When no clear geometrical representation exists, we model a shape (e.g. i^{th} region's shape) by its statistical (from m training samples) and vibrational properties as $\mathbf{X}_i \approx \bar{\mathbf{X}}_i + \mathbf{S}_i^c \mathbf{b}_i$, where $\bar{\mathbf{X}}_i$ is the average of a set of pose-normalized training shapes and $\mathbf{S}_i^c = \mathbf{S}_{stat} + \beta \mathbf{S}_{vib}$ is the combined (statistical $\mathbf{S}_{stat}$ and vibrational $\mathbf{S}_{vib}$) covariance matrix [8,4], $\beta \propto 1/m$ is the balancing parameter and $\mathbf{b}_i = (b_i^1, \dots, b_i^t)^T$ is a vector of shape parameters. We use the Mahalanobis distance to measure the validity of a novel shape $\mathbf{X}_j$ by $F_i^{sh}(\mathbf{X}_j) = e^{-\sqrt{(\mathbf{X}_j - \bar{\mathbf{X}}_i)^T (\mathbf{S}_i^c)^{-1} (\mathbf{X}_j - \bar{\mathbf{X}}_i)}}$.
- **Appearance:** Histology/microscopy images typically have different discriminative **color** channels, $\mathbf{c} = \{c_1, \dots, c_q\}$, where c_i is a color channel, e.g. R, G, B, etc. Further, cells (and their constitutive regions) might also have different discriminative **texture**, $\mathbf{t} = \{t_1, \dots, t_r\}$, where t_i is a texture channel, e.g. multi-scale Gabor or Haar-like features. To leverage cell appearance (color+texture), we concatenate $\mathbf{c}$ and $\mathbf{t}$ into a regional appearance vector $\mathbf{r}$ calculated within inner and outer bands around $\mathbf{X}_j$, $\Omega_{in,d}^j$ and $\Omega_{out,d}^j$, with thickness d . This band-localization is important since cells can contain inner parts (e.g. nucleus) and can be adjacent to other objects (e.g. other cells), both of which can pollute the regional appearance measures if a band is not used. In addition, by using an inner versus outer band, we are encoding the boundary polarity (e.g. dark to bright). We define the appearance fitness function for object i as

$$F_i^{ap}(\mathbf{X}_j) = \frac{1}{2} \left(\frac{1}{|\Omega_{in,d}^j|} \int_{\Omega_{in,d}^j} p(\mathbf{x} \in O_i) d\mathbf{x} + \frac{1}{|\Omega_{out,d}^j|} \int_{\Omega_{out,d}^j} p(\mathbf{x} \in B_i) d\mathbf{x} \right) \quad (1)$$

where $p(\mathbf{x} \in O_i)$ and $p(\mathbf{x} \in B_i)$ are the probabilities of a given pixel $\mathbf{x} \in \Omega_{in,d}^j \cup \Omega_{out,d}^j$, belonging to object i (O_i) and its background (B_i), respectively, and are estimated by training a random forest (RF) consisting of N_b binary decision trees. To segment an R -region object in I , $R + 1$ patches within R different regions of the object plus background are selected (i.e. regions $\mathcal{L} = \{0, \dots, R\}$) to train the RF. After training, for each pixel $\mathbf{x}$, the feature channels, $\mathbf{r}(\mathbf{x})$, are propagated through each tree resulting in the probability $p_j(\mathbf{x} \in k | \mathbf{r}(\mathbf{x}))$, for the j^{th} tree, where $k \in \mathcal{L}$. These probabilities are combined into a forest's joint probability $p(\mathbf{x} \in k | \mathbf{r}(\mathbf{x})) = \frac{1}{N_b} \sum_{j=1}^{N_b} p_j(\mathbf{x} \in k | \mathbf{r}(\mathbf{x}))$ to determine the probability of $\mathbf{x}$ belonging to class k . Note that $O_i, B_i \in \mathcal{L}$.

- **Edge:** Since boundaries of cells and their parts exhibit appearance discontinuities, we incorporate edge information in the image by defining the following edge fitness term $F^{ed}(\mathbf{X}_j) = \frac{1}{|\mathbf{X}_j|} \oint_{\mathbf{X}_j} e^{-g(\mathbf{X}_j)}$, where $g(\cdot) = 1/(\epsilon + \lambda)$, λ is the maximum eigenvalue of the structure tensor $J^T J$ (generalizes scalar field gradients to those of vector fields), where J is the Jacobian matrix of the weighted feature channels, $\mathbf{w}^T \mathbf{r}$, and the vector $\mathbf{w}$, resulting from training the RF, is the importance of each feature channel in discriminating inside versus outside of an object (i.e. maximizes boundary edge response).
- **Pose:** Each cell has a specific size, orientation and position in the image. Given the training data, we estimate the average area ($\bar{A}$) and the principal orientation ($\bar{\theta}$) of cells and use them for imposing constraints on the solutions. We use a cell's centroid, $\mathbf{p} = (p^x, p^y)$, to specify its position.
- **Topological constraints:** In addition to the geometrical and appearance properties of an object (color, texture, edge, shape and pose), in multi-region objects, meaningful topological relationships typically exist between different object's regions, e.g. regions *contain/exclude* others. To enforce containment and exclusion between two regions, e.g. $\mathbf{X}_j$ is contained in $\mathbf{X}_i$, or, $\mathbf{X}_i$ and $\mathbf{X}_j$ are excluded from one another, the following constraints are imposed:

$$D(\mathbf{x}_j) \begin{matrix} \overset{\text{contain}}{\geq} \\ \underset{\text{exclude}}{\leq} \end{matrix} 0 \quad , \quad \forall \mathbf{x}_j \in \mathbf{X}_j \quad , \quad D(\mathbf{x}) = SDM(\mathbf{X}_i) \quad , \quad (2)$$

where $SDM(\mathbf{X}_i)$ is the signed distance map of $\mathbf{X}_i$ and is positive inside and negative outside $\mathbf{X}_i$. Eq. (2) is a general constraint for convex and non-convex shapes. However, for convex shapes, as we typically have in microscopy images, we adopt the following simplification for containment: $\|\mathbf{p}_j - \mathbf{x}_i\| - \|\mathbf{p}_j - \mathbf{x}_j'\| \geq 0$, and exclusion: $\|\mathbf{p}_j - \mathbf{x}_i\| - \|\mathbf{p}_j - \mathbf{x}_j'\| \leq 0$, $\forall \mathbf{x}_i \in \mathbf{X}_i$, for faster computation, where $\mathbf{x}_j' \in \overrightarrow{\mathbf{p}_j \mathbf{x}_i} \cap \mathbf{X}_j$ and $\mathbf{p}_j = (p_j^x, p_j^y)$ is the spatial position of $\mathbf{X}_j$.

- **Inter-part adjacency:** In biomedical applications, the minimum (d^{min}) and maximum (d^{max}) distances between two adjacent regions of an object are sometimes known. Bounding the minimal and maximal distances between two

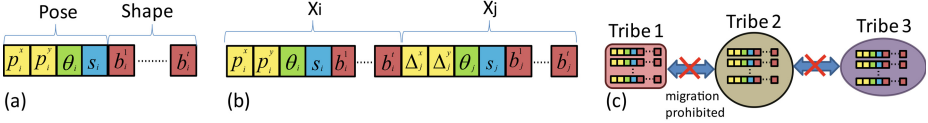


Fig. 1. Chromosome structure for (a) a single-region and (b) a two-region object. The position of the second region, (Δ_j^x, Δ_j^y) , is computed relative to the first object's position. (c) Tribe-based GA. No migration is allowed between tribes.

adjacent boundaries (e.g. i and j) from below and above, respectively, prevents segmentation leakage and improves the results. We impose these constraints by:

$$\min(f_{ij}, f_{ji}) \geq d_{ij}^{min}, \quad \max(g_{ij}, g_{ji}) \leq d_{ij}^{max}, \quad (3)$$

where $f_{ij} = \min_{\mathbf{x}_i \in \mathbf{X}_i} \min_{\mathbf{x}_j \in \mathbf{X}_j} \|\mathbf{x}_i - \mathbf{x}_j\|$ and $g_{ij} = \max_{\mathbf{x}_i \in \mathbf{X}_i} \min_{\mathbf{x}_j \in \mathbf{X}_j} \|\mathbf{x}_i - \mathbf{x}_j\|$. For efficiency, we only calculate and restrict f_{ij} and g_{ij} (not f_{ji} and g_{ji}).

- **User interaction:** User interaction is another useful prior. This prior can be applied on the boundary and/or the region of an object by providing corresponding seed points, $\mathbf{s}_i^b$ and $\mathbf{s}_i^r$, and force the solution to satisfy the following constraints $\mathbf{s}_i^r \in \Omega_{in,d=\infty}^i$ and $\mathbf{s}_i^b \in \Omega_{in,d=\epsilon}^i \cup \Omega_{out,d=\epsilon}^i$.

Fitness Function: The overall fitness function (for an R -region cell) is constructed by integrating all above mentioned information as

$$F_{total}(\mathbf{X}) = \sum_{i=1}^R \left(F_i^{sh}(\mathbf{X}) + F_i^{ap}(\mathcal{M}_i(\mathbf{X})) + F_i^{ed}(\mathcal{M}_i(\mathbf{X})) \right), \text{ subject to} \quad (4)$$

$$\text{geometry :} \quad |b_i^j| \leq 3\sqrt{\lambda_i^j} \quad |\theta_i - \bar{\theta}_i| \leq 3\sqrt{\lambda_i^\theta} \quad |\text{Area}(\mathbf{X}_i) - \bar{A}_i| \leq 3\sqrt{\lambda_i^A}$$

$$\text{user interaction :} \quad \mathbf{s}_i^r \in \Omega_{in,\infty}^i \quad \mathbf{s}_i^b \in \Omega_{in,\epsilon}^i \cup \Omega_{out,\epsilon}^i$$

$$\text{topology :} \quad \text{eq.(2)} \quad \text{adjacency :} \quad \text{eq.(3)}$$

where $\mathbf{X} = \overline{\mathbf{X}} + \mathbf{S}^c \mathbf{b}$, $\mathcal{M}_i(\mathbf{X}) = s_i \mathcal{R}_i \mathbf{X} + \mathcal{T}_i$ is a similarity transformation with rotation $\mathcal{R}$, scaling s , and translation $\mathcal{T}$, λ_i^j is the j^{th} eigenvalue of i^{th} region's covariance matrix and λ_i^A and λ_i^θ are the area and orientation variance of i^{th} region, respectively, obtained from the training data.

To find the best fit for such a complex fitness function (4), we adopt GA as a global optimization tool. Although GA does not strictly guarantee the global solution, our results confirm the ability of this approach to accurately segment the spatially-recurring, multi-region cells with partial overlap. In GA, each individual solution is represented by a chromosome consisting of several genes (Fig. 1(a)). The first four genes describe the individual's pose information, p_i^x and p_i^y are the spatial position of i^{th} region and θ_i , s_i and $b_i^1, \dots, b_i^t$ are its orientation, scale and shape parameters.

Encoding Multi-region Object's Information into GA: For simplicity and to conserve space, here we consider a two-region object scenario. Assuming

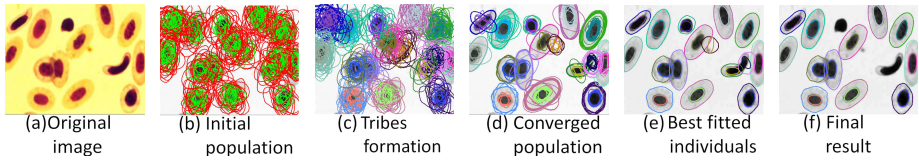


Fig. 2. Fish blood cells segmentation

a cell consists of two regions: $\mathbf{X}_1$ and $\mathbf{X}_2$, where $\mathbf{X}_1$ contains $\mathbf{X}_2$, we compute $\bar{A}$ and $\bar{\theta}$, as well as the shape parameters, $\mathbf{b}$, for $\mathbf{X}_1$ and $\mathbf{X}_2$, separately, as described before. We represent the cell while encoding the interaction between its regions by concatenating the two chromosomes of $\mathbf{X}_1$ and $\mathbf{X}_2$. However, the position of $\mathbf{X}_2$, (p_2^x, p_2^y) , is computed relative to $\mathbf{X}_1$'s position, (p_1^x, p_1^y) , and its corresponding genes are replaced by Δ_x and Δ_y (Fig. 1(b)). Δ_x and Δ_y allow $\mathbf{X}_2$ to move in small distances around its relative position to (p_1^x, p_1^y) . The average relative distance between $\mathbf{X}_1$ and $\mathbf{X}_2$ as well as limits on Δ_x and Δ_y are learned from the training data.

Each object (cell) typically recur in different parts of the image domain. To deal with such spatially recurring aspect of cell segmentation, we use GA with *tribes* to obtain multiple distinct solutions (i.e. cells). In tribes-based GA, the whole population is grouped into several tribes. During the GA evolution and in the gene crossover phase, any two selected parents must be from the same tribe. In fact, tribes are too choosy about who is allowed to join them (Fig. 1(c)); they do not accept any stranger (no migration is allowed) and even children who are not similar to the tribe's population are rejected. This tribes-based GA allows for the desired multiple distinct solutions. We choose the tribes' membership based on the spatial position of each member (cell), i.e. (p^x, p^y) .

Initialization and Implementation: We used 6 channels of colors (RGB+HSV) and 3-channel Gabor features as our regional cues. Gabor filters were calculated in 8 different orientations and 3 different scales and were summed up across orientations to obtain rotational-invariance texture features. For RF, we used $N_b = 50$ binary trees. We randomly spread 10,000 random chromosomes over the image wherever the probability of existing cells obtained from RF is large enough, i.e. $p(\mathbf{x} \in O | \mathbf{r}(\mathbf{x})) > 0.6$ (Fig. 2(b)). d_{ij}^{min} and d_{ij}^{max} were set based on the training dataset. Although our method can handle user interaction, none was used in our experiments. The crossover and mutation rates were fixed to 0.7 and 0.01, respectively, in all of our experiments. Individuals that are within a distance of ℓ pixels from each other establish a tribe (Fig. 2(c)). We implemented our method in MATLAB in a way that all individuals are evaluated simultaneously in parallel. After convergence, Fig. 2(d), the best solution in each tribe is examined (Fig. 2(e)). We use the final fitness measure as a confidence measure, where the user can request displaying e.g. the top 10% confident segmentation. According to our fitness function, the ideal fitness score is 3. In all of our experiments we kept the solutions that are higher than 2.4 (top 20% confident segmentation) as the final solution (Fig. 2(f)).

3 Experiments

In our **first experiment**, we evaluated our method on stained breast cancer tissue images used in ICPR 2010's HIMA contest on 'Counting Lymphocytes on Histopathology Images' [7]. We benchmarked our results against the state-of-the-art methods, including the contest's finalists. We used the centroid of the segmented cells to compare our results against the expert annotated ground truth (GT). Fig. 3(a) quantitatively compares our results against the competing methods. The evaluation criteria are based on the Euclidean distance, d_E , between the GT and segmented lymphocytes, as well as the absolute difference, N , between the true number of cells in GT and detected cells. m and s in Fig. 3(a) are mean and standard deviation, respectively. In Fig. 3(a), Bernardis et al. [2] reported results for different thresholds, ρ , on the same dataset. While for some cases (e.g. $\rho = 2$) they achieved better distance accuracy, d_E , than our method, they found less true cells (bigger N). On the other hand, they obtained better detection rate (smaller N) for $\rho = 5$ but with less accuracy, d_E . We emphasize that their method has been designed for single-region cells only. Fig. 3(b) demonstrates how our method distinguishes the merged cells. Our method can not only detect and segment the single-region cells but also delineate the different boundaries of a multi-region cell.

To further showcase our method, we ran a **second experiment** on another dataset, MICR, consisting of 20 different histology and microscopy images with multi-region cells. Our results in Fig. 4 verify the use of proposed constraints (topology, thickness and shape) as compared to ubiquitous unconstrained image segmentation methods; graph cuts (GC), and constrained methods; Delong and Boykov [5] (DB). While GC is not designed to segment cells, its results show the issues and difficulties involved in segmenting complex, multi-region cells. Delong and Boykov's method incorporates containment constraint in the GC framework, however, their method is unable to segment the targeted objects solely (Fig. 4). Fig. 3(c) quantitatively compares our method with watershed (WS), GC and DB on both HIMA and MICR datasets using Dice similarity coefficient (DSC).

Due to the random initialization and evolution in GA, they may not always produce the same result. To examine the reproducibility of the proposed approach, in our **third experiment**, we ran our method 20 times on sample images and monitored the fitness and DSC vs. generation (Fig. 5). The results confirm that our method converges to almost similar results (similar DSC) although we randomly initialized the population at each run. From Fig. 5, the fitness values for the 1st and 2nd cells are lower than the 3rd simply because the first two cells overlap. The variations between different runs can be reduced by increasing the size of initial population but at the cost of computational complexity. **Runtime:** Using non-optimized MATLAB code on standard 2.3 GHz CPU, the running time of our algorithm ranged between 60-300 s/image, which depended primarily on the number of cells per image, which varied between 2-60 for both HIMA and MICR.

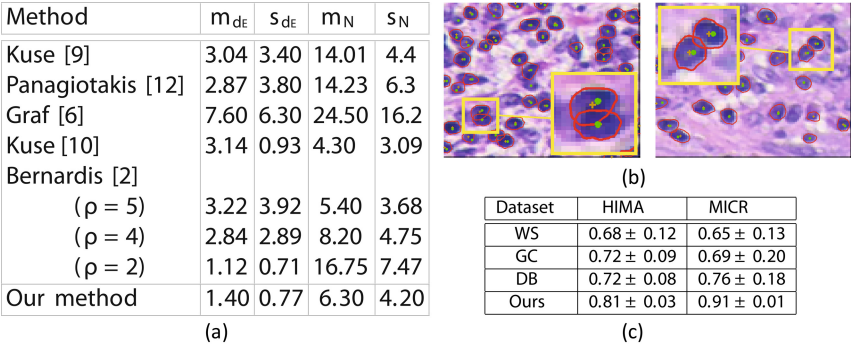


Fig. 3. (a) Comparison against state-of-the-art methods on HIMA dataset [7]. (b) Sample results. Red contours: our segmentation result; small gold + sign: our segmentation centroid; green dots: ground truth. (c) Accuracy comparison with WS, GC and DB methods using DSC (*mean ± std*).

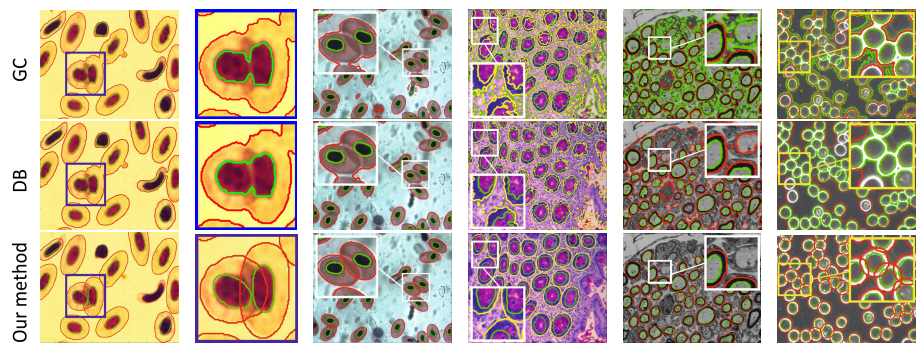


Fig. 4. Sample results on MICR dataset. Note how the proposed method segments only the targeted cells. Same data term was used for all experiments.

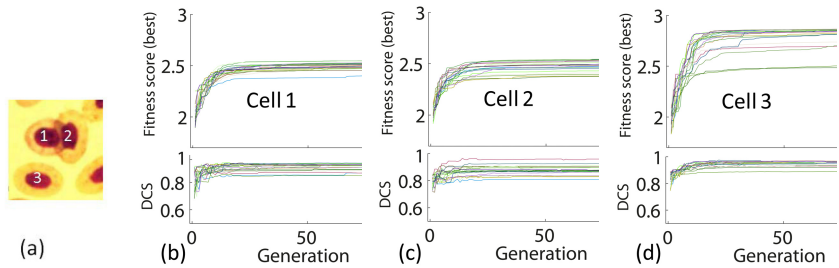


Fig. 5. Sample segmentation shown in Fig. 4. (b-d) Fitness and DSC of the best individual of each tribe vs. generation number for the three tribes (1^{st} , 2^{nd} and 3^{rd}) corresponding to the three cells in (a), for 20 different runs on (a).

4 Conclusion

Segmenting spatially recurring complex objects consisting of different regions with varying shapes, colors and textures remains a challenging problem in

biomedical image segmentation. Another layer of complexity is added once these multi-region objects overlap. In this paper we showed how to address this complexity holistically by incorporating several intuitive priors into an objective function without being overly concerned about its optimization. The proposed high level priors help us to segment only the targeted objects (cells) in an image. Fully parallelizing the proposed method is in our agenda as future work.

References

1. Ali, S., Veltri, R., Epstein, J.I., Christudass, C., Madabhushi, A.: Adaptive energy selective active contour with shape priors for nuclear segmentation and gleason grading of prostate cancer. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 661–669. Springer, Heidelberg (2011)
2. Bernardis, E., et al.: Pop out many small structures from a very large microscopic image. *MedIA* 15(5), 690–707 (2011)
3. Cheng, L., Ye, N., Yu, W., Cheah, A.: Discriminative segmentation of microscopic cellular images. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 637–644. Springer, Heidelberg (2011)
4. Cootes, T.F., et al.: Combining point distribution models with shape models based on finite element analysis. *Image Vis. Comp.* 13(5), 403–409 (1995)
5. Delong, A., et al.: Globally optimal segmentation of multi-region objects. In: ICCV, pp. 285–292 (2009)
6. Graf, F., Grzegorzec, M., Paulus, D.: Counting lymphocytes in histopathology images using connected components. In: Ünay, D., Çataltepe, Z., Aksoy, S. (eds.) ICPR 2010. LNCS, vol. 6388, pp. 263–269. Springer, Heidelberg (2010)
7. Gurcan, M.N., Madabhushi, A., Rajpoot, N.: Pattern recognition in histopathological images: An ICPR 2010 contest. In: Ünay, D., Çataltepe, Z., Aksoy, S. (eds.) ICPR 2010. LNCS, vol. 6388, pp. 226–234. Springer, Heidelberg (2010)
8. Hamarneh, G., Jassi, P., Tang, L.: Simulation of ground-truth validation data via physically-and statistically-based warps. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 459–467. Springer, Heidelberg (2008)
9. Kuse, M., Sharma, T., Gupta, S.: A classification scheme for lymphocyte segmentation in H&E stained histology images. In: Ünay, D., Çataltepe, Z., Aksoy, S. (eds.) ICPR 2010. LNCS, vol. 6388, pp. 235–243. Springer, Heidelberg (2010)
10. Kuse, M., et al.: Local isotropic phase symmetry measure for detection of beta cells and lymphocytes. *J. Pathol. Inf.* 2 (2011)
11. Mao, K.Z., et al.: Supervised learning-based cell image segmentation for p53 immunohistochemistry. *IEEE TBE* 53(6), 1153–1163 (2006)
12. Panagiotakis, C., Ramasso, E., Tziritas, G.: Lymphocyte segmentation using the transferable belief model. In: Ünay, D., Çataltepe, Z., Aksoy, S. (eds.) ICPR 2010. LNCS, vol. 6388, pp. 253–262. Springer, Heidelberg (2010)
13. Turner, A., et al.: Obtaining multiple distinct solutions with genetic algorithm niching methods. In: Ebeling, W., Rechenberg, I., Voigt, H.-M., Schwefel, H.-P. (eds.) PPSN 1996. LNCS, vol. 1141, pp. 451–460. Springer, Heidelberg (1996)
14. Wu, X., et al.: Embedding topic discovery in conditional random fields model for segmenting nuclei using multispectral data. *IEEE TBE* 59(6), 1539–1549 (2012)
15. Yang, L., Tuzel, O., Meer, P., Foran, D.J.: Automatic image analysis of histopathology specimens using concave vertex graph. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 833–841. Springer, Heidelberg (2008)