2009 Ninth International Conference on Intelligent Systems Design and Applications

Algorithm Selection for Intracellular Image Segmentation based on Region Similarity

Satoko Takemoto and Hideo Yokota
RIKEN
2-1, Hirosawa, Wako, Saitama, 351-0198, Japan
{satoko-t, hyokota}@riken.jp

Abstract—**This paper deals with the problem of intracellular image segmentation. Our goal is to propose an algorithm selection framework that has the potential to be general enough to be used for a variety of intracellular image segmentation tasks. With this framework, an optimal algorithm suited to each segmentation task can be selected automatically by our proposed evaluation criteria derived from region similarity of image features and boundary shape. Furthermore, using our framework, we can rank different algorithms, as well as define each algorithm's parameters. We tested our prototype framework on confocal microscope images and showed that application of these criteria gave highly accurate segmentation results without missing any biologically important image characteristics.**

Keywords-image segmentation; segmentation evaluation; pattern classification; live cell imaging

I. INTRODUCTION

Image segmentation is a process that divides an image into multiple regions corresponding to the components pictured in the image. Segmented regions provide us with identification between the component of interest and the other components. In other words, the segmentation process is one of the first steps in analyzing the component of interest from acquired images. Numerous segmentation algorithms have been proposed [1-2], but most approaches have been developed for a specific application and cannot be generalized for other segmentation tasks and be effective at the level of human competence. Researchers have had to face the difficult duty of choosing the most suitable algorithm for a given task while at the same time facing increasing numbers of images that need to be analyzed.

Recently, some segmentation methods based on algorithm selection have been proposed as cited in [3-5]. In most of these, an appropriate algorithm for each task is automatically selected according to a segmentation performance analysis so as to have a computer emulate human capabilities. That is, when a certain task is specified, a researcher is able to judge which algorithm is suitable to the task, at least to some degree. To emulate similar capabilities on a computer for numerous tasks, various evaluation measures have been proposed for selecting optimal algorithms [5-11].

Algorithm selection is needed in intracellular image processing

Segmentation plays an important role in intracellular image processing because of recent advances in confocal microscopes and relevant imaging systems. These advances have led to the acquisition of an enormous number of high-resolution digital images of living cells that need to be analyzed. It is nearly impossible for researchers in a wide variety of subject areas to analyze the vast quantity of accumulated images, and an ability to quantify the microstructural properties of intracellular components from the acquired images is desperately needed [12-13].

Studies in intracellular image processing have only just begun, and advanced segmentation algorithms, such as those developed for medical image processing, have not generally been applied. In addition, the low signal-to-noise ratios that occur in microscopic imaging with ultra-high sensitivity have greatly hindered progress in intracellular image processing. The noise in the images makes it more difficult for researchers to identify the region of interest and try to obtain quantitative data on the organelles of great value to cell biologists, such as the volume, shape, and dynamics of intracellular structures.

Although some automated techniques for segmentation of cell images have recently been published [14-17], their targets have been relatively simply shaped components, for example, nuclei, so that the segmentation goal is not so difficult. Moreover, recent significant improvements in image acquisition of living cells require the appropriate segmentation algorithm to be flexible enough to accommodate time-variable changes of targets. In fact, no single algorithm is considered to be good for all time-lapse images, and selecting an optimal algorithm for the variety of images is a tedious task for researchers in this field. Thus, computerized selection of the optimal algorithm should prove a useful and efficient solution to this problem. However, almost none of the previous work about the algorithm selection is general enough to be applied to the various tasks required for high-level image segmentation in realistic scientific studies. Therefore, a great deal of time and labor is still necessary to accurately segment intracellular images.

The contributions of our paper

We propose a new higher level evaluation function of algorithm selection on intracellular images that is based on

978-0-7695-3872-3/09 $26.00 © 2009 IEEE
DOI 10.1109/ISDA.2009.205

the combination of two criteria inspired by human judgment. Generally, one criterion (e.g., number of error pixels, feature similarity, and true/false or positive/negative rates between the segmented region and its ground-truth) is used as an evaluation measure (e.g. [5]). In each case, the segmented region is evaluated only computationally, and it is not clear whether the segmentation has actually satisfied its intended purpose.

In contrast, our evaluation function consists of similarity measurement of the combination of intensity-based image features and boundary shape between the segmented region and the ground-truth, because cell biologists pay attention to these two features when evaluating a segmented region. Furthermore, previous studies have suggested that such a combination will improve the evaluation quality [7, 9-10]. Our evaluation function is also able to rank different kinds of algorithms, as well as to define each algorithm's parameters.

The rest of the paper is structured as follows. In Section 2, we describe the proposed framework for algorithm selection and its methodology. Experimental results and an evaluation of the proposed framework on confocal microscope images with ground-truth are presented in Section 3. Finally, a conclusion is offered in Section 4.

II. ALGORITHM SELECTION FRAMEWORK

Many possible solutions must be considered when establishing a segmentation algorithm for a specific application that satisfies a user's intention. In many cases, intracellular components are represented by unique image features and can be distinguished from each other, even from background. Here, we first focus on the segmentation techniques implemented by pattern classification technique that can classify image features. When performing segmentation, the computer first calculates *N*-dimensional intensity-based image features and classifies them into multiple classes in the *N*-dimensional feature space (see Fig. 1). Each class is ideally associated with one component, such as an organelle inside the cellular images. In the case of supervised classification, the feature distribution of each class is specified by a user who has knowledge of the segmentation target, and the classification rule is generated by its distribution. The specified class is conducted by means of the user's manual segmentation. According to the classification rule, the computer is able to automatically classify the new inputs that are calculated from the unsegmented images. As a result, target segmentation can be achieved by selecting only the pixels that have the feature classified as the target class.

However, the segmentation algorithm implemented by this classification technique is not general enough, because there will be large differences in segmentation results depending on the algorithm selected. That is, the results are greatly influenced by the type of feature used and the classification rules. Furthermore, suppose that the selected algorithm is optimal for one segmentation task but not for a different segmentation task. To solve this problem, we propose a new framework that can select an optimal algorithm that satisfies user's intention, taking the segmentation task into consideration. Here, "algorithm" means the set including the feature space constructed by the extracted image feature, the classification rule, and their parameter specification. Our framework selects the algorithm that can extract the target region with the highest level of accuracy by means of similarity measurement between the supervised region and the region extracted by some given algorithms. That selection is conducted by the following selection criterion.

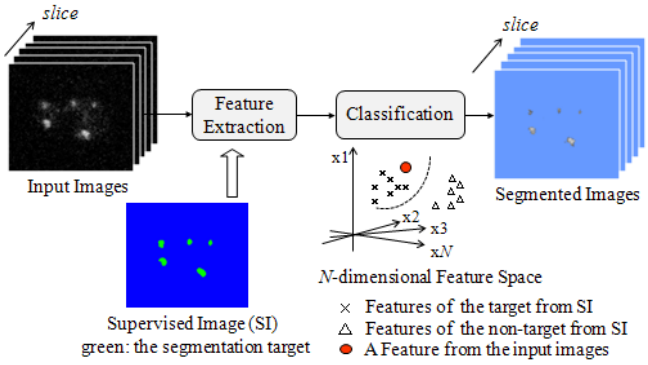


Figure 1. A segmentation approach based on pattern classification theory. In this approach, the user specifies the region of the segmentation target.

A. Selection Criterion

The similarity between the user-supervised region and the automatically segmented regions generated by given algorithms is used as a criterion for selecting the optimal algorithm. The similarity is measured in the intensity-based image features and shape distances between the two regions. In our evaluation function, the algorithm that produces the minimum distance is considered optimal for a segmentation task. That is, the user can obtain the most accurate segmentation result by using the selected algorithm to segment a target that has similar characteristics to the supervised region. This is true because, if a highly accurate identification is achieved for a feature distribution with a certain classification function, the function is also applicable to a similar feature distribution.

People generally focus on the characteristics of a region when evaluating a segmented region. We consider that image features derived from the pixel intensities and boundary shape of the segmented region are the most important characteristics noticed. We defined S^g as the user-supervised region and $S=\{S_a, a\in A\}$ as the segmented regions by given algorithms in a plane (or a space). The similarity R_A between those two regions can be calculated as follows:

$$R_A = \frac{1}{dist(S^g, S)} = \frac{1}{dist(\mathbf{X}^g, \mathbf{X}_A) + dist(\mathbf{P}^g, \mathbf{P}_A)}, \qquad (1)$$

where $\mathbf{X}=(x_1, x_2, ..., x_N)$ represents the *N*-dimensional image features and $\mathbf{P}=(p_1, p_2, ..., p_n)$, $(p_j \in C^N)$ represents the spatially discrete shape features. That is, R_A is defined

as a linear combination of similarity measurement of $dist(\mathbf{X}^g, \mathbf{X}_A)$ and $dist(\mathbf{P}^g, \mathbf{P}_A)$. We can define a selected algorithm a_i as follows:

$$a_i = \arg \max_{0<i\le k} R_{a_i}, \quad (2)$$

where k is the number of given algorithms. The feature derived from pixel intensity, such as texture statistics or local correlation, is set to $\mathbf{X}$. In our framework, we measure $dist(\mathbf{X}^g, \mathbf{X}_A)$ by using the Bhattacharyya distance, which is an approximate measurement between two statistical distributions. The shape feature $\mathbf{P}$ and $dist(\mathbf{P}^g, \mathbf{P}_A)$ are described in the next subsection.

B. Discrete description of boundary shape

We use the set of boundary points obtained by sampling a sequential boundary to describe the shape of the boundary. A complex autoregressive model [18] is applied to these boundary points, and this leads to a stable description invariant to translation, rotation, and scale of patterns. Each boundary point is represented by a complex number $z_j = x_j + i\,y_j$ (see Fig. 2). Thus, we have a sequence of complex numbers that is defined by a linear combination of boundary points. For this sequence, the complex autoregressive model of order m is defined as follows:

$$\hat{z}_j = \sum_{k=1}^{m} b_k z_{j-k}, \quad (3)$$

where $\{b_k\}_{k=1}^{m}$ is defined by minimizing the mean squared error of $\varepsilon^2(m) = E_j\left|\hat{z}_j - z_j\right|^2$. We calculate the distance between the two boundaries $z^{(n)} \in C^N, (n \in \{1, 2\})$ as follows:

$$Db(1,2) \equiv \sqrt{\sum_{k=1}^{m}\left|b_k{}^{(1)} - b_k{}^{(2)}\right|^2}. \quad (4)$$

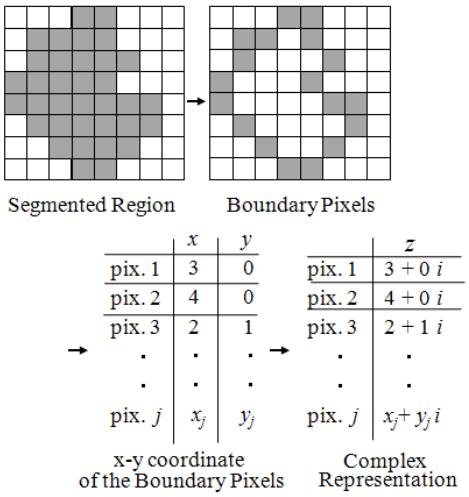


Figure 2. Schematic diagram of the boundary shape description.

That is, the distances represented in (4) are defined as a Euclidean distance of each coefficient b_k represented in (3).

According to these equations, for example, the distance between boundary shape S_0 and its deformed shape S_1 is 52.99 and that between S_0 and its deformed shape S_2 is 36.78 (see Fig. 3). The difference between S_0 and S_2 is less than that between S_0 and S_1, so we can tell that the boundary shape of S_2 is more similar to the shape of S_0 than is the boundary shape of S_1. We use this similarity measure to evaluate whether the automatically segmented region is similar to the supervised region.

III. VALIDATION ON CONFOCAL MICROSCOPE IMAGES

We implemented our selection framework and tested it on intracellular images. The target of segmentation was to extract only the Golgi apparatus region from botanical yeast images. Fig. 4a shows the image used (taken under a confocal microscope), and Fig. 4b shows the user-specified target region. In this validation, we evaluated whether the selected algorithm was able to extract the target region with a high degree of similarity from the viewpoint of the above-mentioned selection criterion.

The test segmentation was first conducted for the entire group of given multiple algorithms; therefore, there was the same number of segmentation results as algorithms. Next, for all the segmentation results, we calculated the intensity-based image features inside the segmented region and described the region's boundary shape numerically by the previously mentioned methods. At the same time, we calculated the intensity-based image features inside the supervised region and described the boundary of the supervised region. Finally, we computed the similarity between the supervised region and each automatically segmented region by (1).

Although there are numerous intensity-based image features, we used the two types of image features associated with each pixel as a prototype in this validation: normalized pixel intensity and texture-statistics inside the localized region in which each pixel is centrally positioned. The latter is calculated by (5):

$$X_{pq} = \sum\sum m^p n^q f(m,n), \quad (5)$$

where m and n are the x-y coordinates inside the image, and $f(m,n)$ is the localized region consisting of a 5×5 set of pixels. These calculated features are equivalent to moments, and in this experiment we calculated the normalized moment of order 2 around (m, n) as the second image

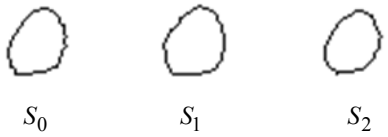
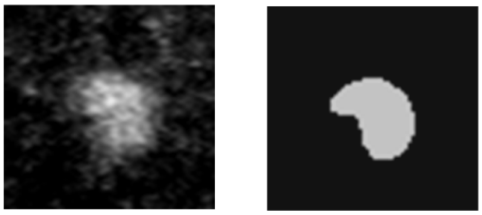


Figure 3. Examples of a boundary shape (S_0) and two deformed shapes (S_1 and S_2).



(a) Original Image (b) User-specified Image

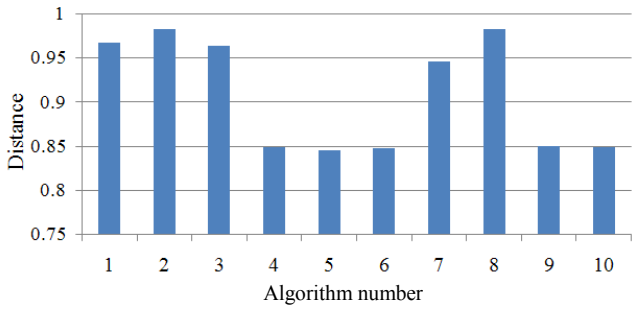
Figure 4. Experimental images. The gray region in (b) is the supervised segmentation target.

feature.

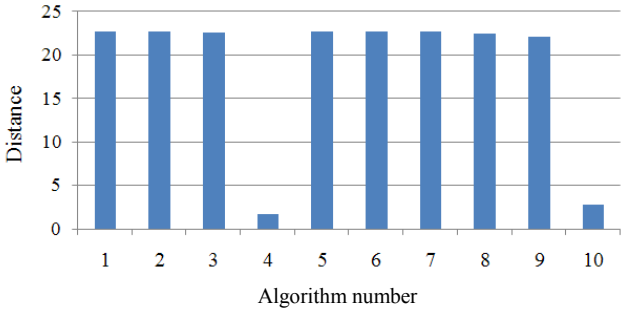
The Support Vector Machine (SVM) [19] and Approximate Nearest Neighbor (ANN) [20] were defined as classification rules in this validation, and a parameter had to be set for each classification rule. We defined three types of parameter settings (P1-P3) related only to the kernel functions in SVM and two types of parameter settings (P4 and P5) related only to the number of nearest neighbors in ANN. The combination of features, classification rules, and their parameters produced the 10 segmentation algorithms shown in Table 1. In the table, F1 shows the feature derived from pixel intensity, F2 shows the feature derived from texture-statistics, M1 is SVM, and M2 is ANN.

Fig. 5a shows the feature distribution distance between the supervised region and each segmented region for each algorithm. Similarly, Fig. 5b shows the shape distance between them. The similarities computed by equation (1) and the performance ranking of each algorithm are shown in Table 2. These results indicate that the segmented region of A4 was most similar to the supervised region. Therefore, we regard A4 as the optimal segmentation algorithm, not only for this task but also for a similar task.

Fig. 6 shows the target regions segmented automatically by using each algorithm; it is clear that several results include isolated regions other than the target region. In that case, we calculated the distance on the basis of only the largest region. For comparison, we also show a binarization result provided by the Otsu method [21] as A11. Because the original image was extremely noisy, the binarization result contained false positive errors. Algorithm A4,



(a)



(b)

Figure 5. (a) Distance between the results of Ai and the supervised region for the image features. (b) Distance between the results of Ai and the supervised region for boundary shape.

however, was not affected by the noise and achieved a highly accurate segmentation.

If we had used only the criterion derived from image features, A4, A5, A6, A9, and A10 could have been selected as the optimal algorithm. If we had used only the criterion derived from boundary shape, A4 or A10 could have been selected. However, as can clearly be seen in Fig. 6, over-segmentation occurs in A9 and A10. Because we used combination criteria in the evaluation function in equation (1), we avoided the risk of choosing a suboptimal algorithm.

In addition, although A4-A6 appear to be similar to each other in Fig. 6, there is a large difference between the

TABLE I. EXPERIMENTAL ALGORITHMS, A1-A10

method	feature	classification rule	parameter
A1	F1	M1	P1
A2	F1	M1	P2
A3	F1	M1	P3
A4	F2	M1	P1
A5	F2	M1	P2
A6	F2	M1	P3
A7	F1	M2	P4
A8	F1	M2	P5
A9	F2	M2	P4
A10	F2	M2	P5

TABLE II. PERFORMANCE RANKING OF THE ALGORITHMS

Algorithm Number	Normalized Similarity	Performance Ranking
A1	1.48	8
A2	1.73	10
A3	1.41	7
A4	-3.05	1
A5	-0.52	3
A6	-0.49	5
A7	1.13	6
A8	1.70	9
A9	-0.50	4
A10	-2.91	2

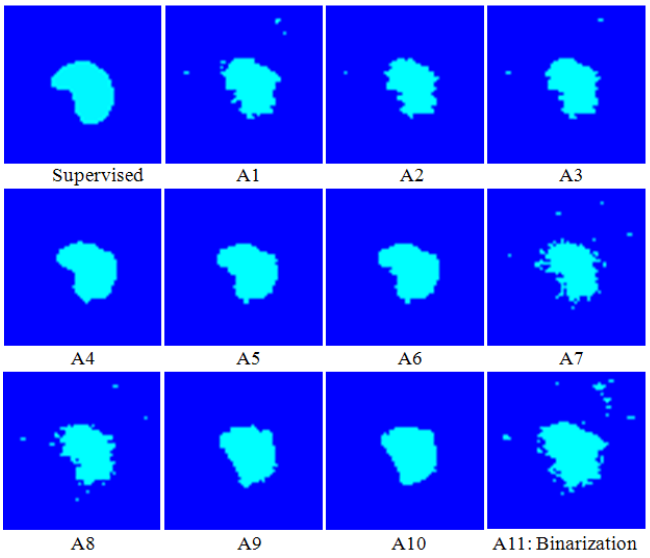


Figure 6. Segmentation results for all the algorithms. The result of A4 was determined to be the optimal algorithm.

boundary shape of A4 and those of A5 and A6 (see Fig. 5b). In the segmented images of A4-A6, the center-left of each segmented region clearly has a larger boundary change than the other regions. Although false-negative error occurs in that region in A5 and A6, A4 achieved an accurate segmentation reflecting the boundary of the supervised region (see Fig. 6). Our evaluation function did not miss the difference between these results, which appears to be biologically important. Even if the differences were trivial, however, the evaluation framework was able to select the optimal algorithm to reflect the user's intention.

Our segmentation framework assumes that images having similar characteristics will show similar good segmentation results. To validate this concept, we conducted a follow-up experiment. Fig. 7 shows six sequential images (in depth) taken by a confocal microscope of the marked Golgi apparatus from botanical yeast. In fact, the image shown in Fig. 4 was cropped from this set of images. Therefore, the segmentation target inside these six images should be similar to that of the previous experiment. We implemented an automatic segmentation of these six images by using algorithm A4, which had been selected as the optimal algorithm. As can clearly be seen in Fig. 8, the target region (that is, the Golgi apparatus) was correctly segmented from these very noisy images. The cell biologist who provided the supervised region evaluated these results and determined that they were correct.

IV. Conclusion

We proposed a novel framework for intracellular image segmentation based on effective algorithm selection. Selection is conducted by measurement both of similarities of intensity-based image features and of boundary shape between the user-supervised region and the automatically

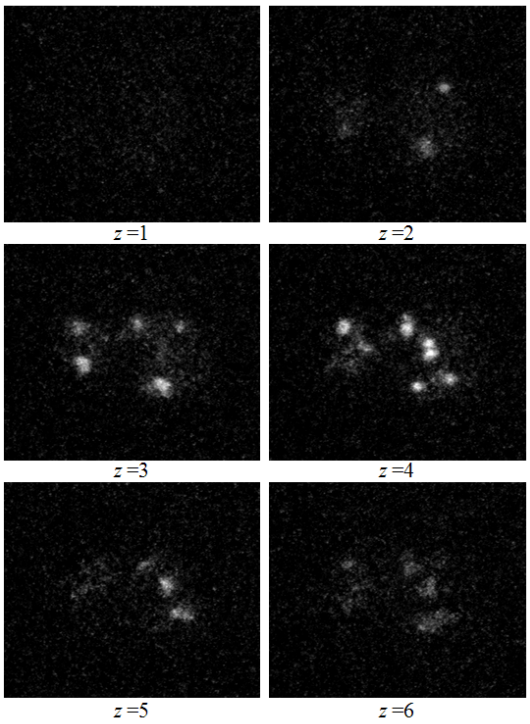


Figure 7. Living cell images of botanical yeast with marked regions of Golgi apparatus. *z* indicates the depth position of each image.

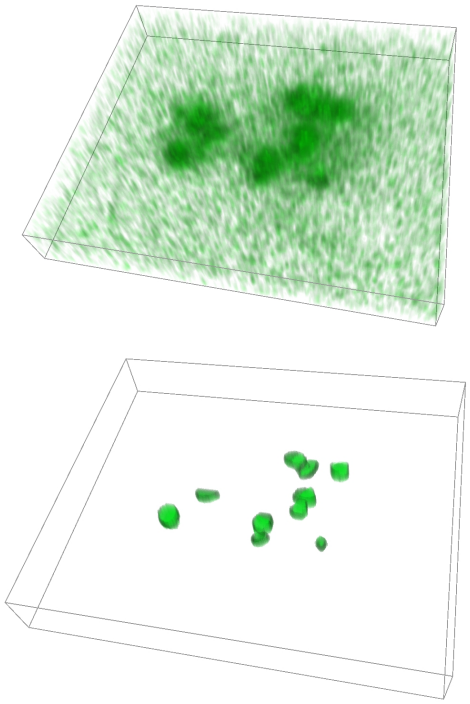


Figure 8. Living cell images of botanical yeast with marked regions of Golgi apparatus (above). The automatic segmented region as a Golgi apparatus (below). These figures were visualized by the volume-rendering method.

segmented regions generated by the given pattern classification techniques. Our framework assumes that the algorithm, which has powerful segmentation ability for a test image, will show good segmentation results for other similar images. That is, our framework can select an optimal algorithm to segment a region that has similar characteristics to the user-supervised region, even from many images. Furthermore, as shown in the experiment, our framework can rank different algorithms and define the parameters of each algorithm.

The evaluation function presented here is versatile, but further investigation may reveal other functions that are better able to reflect a user's intention. In addition, our framework needs to be expanded to be able to better represent image features and boundary shape, and it should include more classification rules and a greater variety of parameters. We tested only two types of features and two types of classification rules as a prototype framework. These types of improvements will lead to segmentation that will have the necessary generality to conduct the various segmentation tasks required by researchers. As a result, we believe that researchers will be released from a labor-intensive and troublesome task and able to concentrate on the accumulation of valuable data.

ACKNOWLEDGMENT

This work was supported in part by Strategic Programs for R&D (President's Discretionary Fund) of RIKEN and Grants-in-Aid for Scientific Research of Japan (19800062 and 20113007). The cell images were taken at the Molecular Membrane Biology Laboratory of RIKEN. Part of the results of the calculations was performed by using the RIKEN Super Combined Cluster System (RSCC).

REFERENCES

[1] N.R. Pal and S.K. Pal, "A review on image segmentation techniques," Pattern Recognition, Vol. 26, No. 9, pp. 1277–1294, 1993.

[2] R. Haralick and L. Shapiro, "Image segmentation techniques," Computer Vision, Graphics, and Image Processing, Vol. 29, pp. 100–132, 1985.

[3] Y.J. Zhang, "A survey on evaluation methods for image segmentation," Pattern Recognition, Vol. 29, No. 8, pp. 1335–1346, 1996.

[4] Y.J. Zhang, Advances in Image and Video Segmentation, IRM Press, USA, 2006.

[5] J.S. Cardoso and L. Corte-Real, "Toward a generic evaluation of image segmentation," IEEE Transactions on Image Processing, Vol. 14, No. 11, pp. 1773–1782, 2005.

[6] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," Proceedings of IEE International Conference on Computer Vision, Vol. 2, pp. 416–423, 2001.

[7] X. Jiang, C. Marti, C. Irniger and H. Bunke, "Distance measures for image segmentation evaluation," EURASIP Journal on Applied Signal Processing, Vol. 2006, No. 35909, pp. 1-10, 2006.

[8] J.K. Udupa, V.R. LeBlanc, Y. Zhuge, C. Imielinska, H. Schmidt, L.M. Currie, B.E. Hirsch and J. Woodburn, "A framework for evaluating image segmentation algorithms," Computerized Medical Imaging and Graphics, Vol. 30, pp.75-87, 2006.

[9] S.K. Shah, "Performance modeling and algorithm characterization for robust image segmentation," International Journal of Computer Vision, Vol. 80, No. 1, pp. 92–103, 2008.

[10] R. Cardenes, M. Bach, Y. Chi, I. Marras, R. de Luis, M. Anderson, P. Cashman and M. Bultelle, "Multimodal evaluation for medical image segmentation," Proceedings of 12th International Conference of Computer Analysis of Images and Patterns (CAIP 2007), pp. 229–236, Vienna, August 2007.

[11] S. Cagnomi, A.B. Dobrzeniecki, R. Poli and J.C. Yanch, "Genetic algorithm-based interactive segmentation of 3D medical images," Image and Vision Computing, Vol. 17, No. 12, pp. 881–895, 1999.

[12] R. Eils and C. Athale, "Computational imaging in cell biology," Journal of Cell Biology, Vol. 161, No. 3, pp. 477–481, 2003.

[13] H. Bhaskar and S. Singh, "Live cell imaging: a computational perspective," Journal of Real-Time Image Processing, Vol. 1, pp. 195–212, 2007.

[14] D.P. McCullough, P.R. Gudla, B.S. Harris, J.A. Collins, K.J. Meaburn, M.A. Nakaya, T.P. Yamaguchi, T. Misteli and S.J. Lockett, "Segmentation of Whole Cells and Cell Nuclei From 3-D Optical Microscope Images Using Dynamic Programming," IEEE Transactions on Medical Imaging, Vol. 27, No. 5, pp. 723–734, 2008.

[15] N. Harder, R. Eils and K. Rohr, "Automated classification of mitotic phenotypes of human cells using fluorescent proteins," Methods in Cell Biology, Vol. 85, pp. 539–554, 2008.

[16] G. Cong and B. Parvin, "Model Based Segmentation of Nuclei," IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, p. 1256, 1999.

[17] O. Dzyubachyk, W.A. Cappelen, J. Essers, W. Niessen and E. Meijering, "Energy minimization methods for cell motion correction and intracellular analysis in live-cell fluorescence microscopy," Proceedings of The Sixth IEEE International Symposium on Biomedical Imaging (ISBI'09), Boston, June-July 2009, in press.

[18] I. Sekita, T. Kurita and N. Otsu, "Complex autoregressive model for shape recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 4, pp. 489–496, 1992.

[19] V.N. Vapnik, The Nature of Statistical Learning Theory, New York, Springer Verlag, 1995.

[20] S. Arya, D.M. Mount, N. Netanyahu, R. Silverman and A.Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," Proceedings of 5th ACM-SIAM Symposium, pp. 573–582, Arlington, January 1994.

[21] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transaction on Systems, Man and Cybernetics, No. 9, pp. 62–66, 1979.