Proceedings of the 2013 International Conference on Machine Learning and Cybernetics, Tianjin, 14-17 July, 2013

ADAPTIVE MARKER-BASED WATERSHED SEGMENTATION APPROACH FOR T CELL FLUORESCENCE IMAGES

GE FAN[1], JIAN-WEI ZHANG[1*], YONG WU[1], DONG-FA GAO[2]

[1]School of Computer Science & Engineering, South China University of Technology，Guangzhou 510006, China
[2] School of Informatics, Guang Dong University of Foreign Studies, Guangzhou 510420, China
E-MAIL: 644604918@qq.com, jwzhang@scut.edu.cn, 422837083@qq.com, gaodf@mail.gdufs.edu.cn

Abstract:

It is intractable problem to segment the fluorescence image of T-cells with different sizes, irregular shape, and severe overlapping by conventional marker-based watershed segmentation. In this paper, Adaptive Marker-controlled Watershed method（AMWS） will be proposed. The Otsu strategy firstly is performed twice in a row to capture as many T-cells as possible. Then based on T-cells' roundish shape, an improved strategy to obtain markers adaptively is present using the evaluation of the segmentation result. This strategy is able to mark the single cell and the overlapping cells accurately. It avoids the ineffectiveness of ultimate erosion which is due to different sizes of cells. The experimental results show that the proposed strategy in this paper can effectively avoid over-segmentation and under-segmentation thus improves both accuracy and robustness of the segmentation.

Keywords:

Marker-based watershed; Distance reconstruction; T cell fluorescence image

1. Introduction

T cell is one of the largest numbers and the most complicated classes of lymphocytes, and plays a key role in the human adaptive immune system. A number of researches on the evolution of T cells and their relationship with some cancers have been done based on the T tell images. However, the user-interaction in the research process is a tedious, time-consuming and error-prone. Therefore, utilizing the computer technology to recognize, analyze the biological mechanism of T cells has become an important research topic. Since the performance of the cell analysis largely depends on the precision of cell segmentation, it is of great importance to study accurate, automatic and reliable segmentation approaches.

There exist a variety of segmentation algorithms for cell images. Clustering method[1] is mainly used in the absence of prior knowledge about the image, where the segmentation is based on the target intensity, texture and other features. Presented clustering approaches include: K-means clustering[2], fuzzy C-means clustering[3], and etc. Mathematical morphology[4-5] extracts or processes the image by using structural elements, in order to achieve segmentation and identification. Segmentation methods based on differentiation[6] consider that there exist sudden changes of the intensity on the boundary pixels. The common differential operators include Robert, Sobel, Prewitt and Canny, and so on. Segmentation methods based on models, in particular Active Contour Method in cell segmentation has been widely used. Zimmer[7] used the Parametric Active Contours in segmentation and tracked of cells. For multi-target image, Zhou[8] adopted CV(Chan-Vese) model, which used a multiphase Level Set Method for cell image segmentation, to achieve more accurate segmentation result.

Through the analysis and the research on the segmentation algorithms above, we can find that there are several problems in the existing cell image segmentation algorithms. The first, human intervention is required to get a good segmentation results more or less. In many cases, we have to adjust the parameters manually in order to deal with different images, that is to say, the generality of the algorithms is unsatisfying. The second, there is lack of criteria to evaluate the segmentation results, so we have no effective mechanism to guide the improvement of the segmentation methods using the quality of the segmentation result. The third, it is impossible to get a satisfied segmentation results by using any simple segmentation algorithm, because of the complexity of processing images. It is a main reason why we have to find a way to comprehensively utilize more than one kind of segmentation methods.

T cell images processed in our experiments are from Department of Immunology, University of Texas Southwestern Medical Center. One example is showed by Fig.1. We can see that 1) The intensity is non-uniform. This can be seen not only in the whole image, but also sometimes

978-1-4799-0260-6/13/$31.00 ©2013 IEEE

within a same cell. 2) The cells are irregular in shape. Although most T cells seem like a circle, some are prolate or irregular. Besides, the size of cells varies. 3) The overlapping is serious. Overlapping between two, three or even more cells can be found in many locations. This is the main difficulty in cell segmentation.

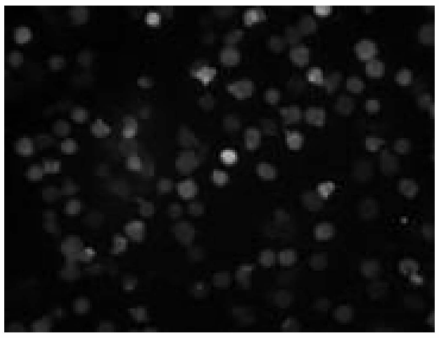


Figure 1: A T-cell fluorescence image.

In this paper, an adaptive and automatic segmentation method for T cell fluorescence images is proposed based on marker-based watershed algorithm. It improves the strategies about markers determination and post-processing in order to obtain more accurate results.

2. Deficiency of existing marker-based Watershed algorithm

There are two key steps when apply the watershed transform to segment cellular image: One is watershed algorithm and the other is marker selection.

The intuitive idea about Watershed algorithm comes from the concept of topography. The grey value of a pixel is interpreted as its altitude. Each local minimum and the range of influence is regard as a basin and the watershed line correspond to the boundary of the adjacent basins. Several methods[9-11]of the watershed transform have been presented, among which Meyer's flooding watershed is one of the most popular methods[12].

Since marker-controlled watershed only works on the premise that the extracted markers really represent the true objects, it is important to detect markers accurately. Ultimate erosion[13] and distance reconstruction[14-16] are two classic methods used to mark the markers.

However, we get poor results when we applied distance reconstruction to the T-cell. Although, distance reconstruction can overcome over-segmentation, it causes under-segmentation, which is show in Fig2.

Our aim is to develop a fully automatic method for segmentation of T-cells in the fluorescence image, which can achieve much more robust and precise results.

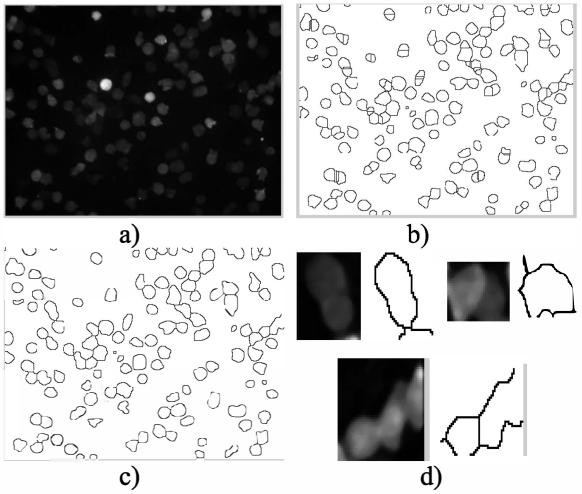


Figure.2 Segmentation results and under-segmentation.

a)Original T-cell image. b)Results with ultimate erosion. c)Results with distance reconstruction. d)Cases of under-segmentation.

3. Proposed algorithm

The details of our Adaptive Marker-Controlled Watershed application for T cell segmentation is illustrated in Fig 3.

The proposed segmentation scheme consists of three modules.

1.Preprocessing module. Some of the basic image operations and morphological processing are took into account, in order to eliminate noises and enhance the image. An improved iterative Otsu thresholding method is used in this module.

2.AMWS module. In this module, we first apply distance transform, gray scale reconstruction to the preprocessed, enhanced image, and carry out the first time watershed. By making the statistical analysis about the sizes of the regions, we obtain a fitting radius range according to the normal distribution of radius. Note that, The T cells segmented should belong to the range. The radius can be used for the selection of markers which obtained by the ultimate erosion. Then we get a new radius distribution. According to previous estimation, iteration should stop when the variation of the radius is less than a threshold.

3.Post-process module. After iterations, we establish evaluation criteria for the segmentation result to merge some small regions or split the big regions, and obtain the final result.

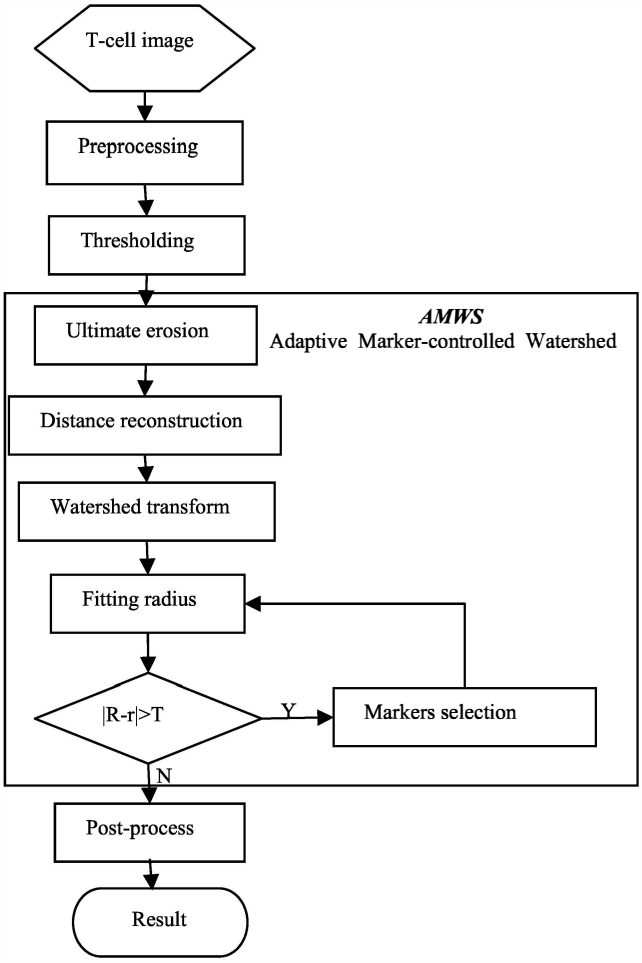


Figure 3 Experimental process design

3.1 Preprocessing

The preprocessing module has two goals. One is to eliminate the impact of the noise in the fluorescence images, the other is to enhance the contrast between the object and the background. The basic steps of the preprocessing module are described as follows.

1)Top-hat transformation. This is composite operation that provides a solution to solve the uneven gray distribution of the background or objects in the image.

2)Median filtering. Under certain conditions, median filtering preserves edges while removing noise. Also, median filter can smooth out the jagged edges of binary image effectively.

3)Thresholding. If the threshold is not set correctly, many T cells are eliminated. The only goal of choosing setting threshold is to get as many t cells as possible in the binary image.

4)Filling holes. Due to the great intensity variations in some cells, thresholding may lead some holes in a few cells. We fill them.

5)Morphological opening. Opening operation removes small objects that are smaller than the structural element (noise) from the foreground, as well as smoothes out the jagged edges of T cells.

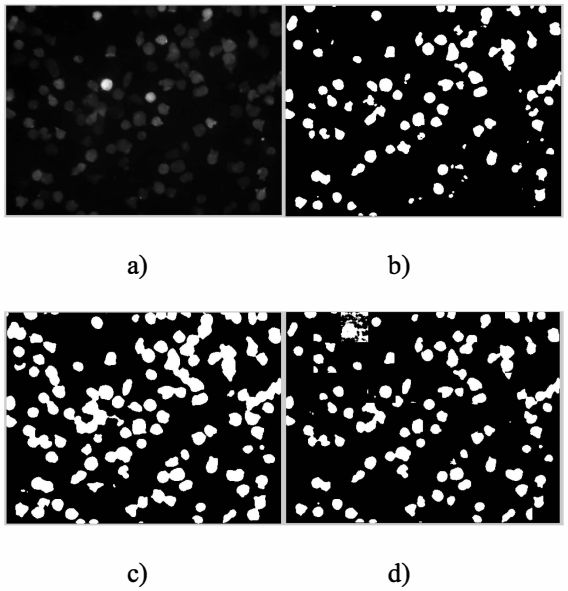
In step 3), the thresholding module uses an improved iterative Otsu thresholding method, that is applying the Otsu method n times in a row to the image. In our experiment, we set n=2. The iterative Otsu threshold can be naturally formulated as:

$$b_n \leftarrow Otsu(f_n) \tag{1}$$

$$f_{n+1} \leftarrow f_n(b_n) = 0 \tag{2}$$

$$B^n = \bigcup_{i=1,n} b_i \tag{3}$$

where f_n refers to the image before the nth iteration while the initial value f_1 is the original image. $b_n \leftarrow Otsu(f_n)$ is the procedure to obtain the binary image b_n by thresholding f_n with Otsu algorithm. $f_n(b_n) = 0$ means those pixels of which the gray value is 1 are set to 0, that is to say, we temporarily set the foreground as the background to segmentation the reminder of the image. The equation (11) shows that the final binary image is the union of the first n binary images. Fig. 4 lists the results of some thresholding methods applied to a T-cells image.



a) b)

c) d)

Figure.4 Comparision between several methods.
a) Original T-cell image. (b) Otsu thresholding. c) Fuzzy c-means. d) The proposed iterative Otsu thresholding method algorithm..

From Fig.4 we can find that iterative Otsu thresholding and Fuzzy c-means threshold can all obtain most T cells. Nevertheless because the fuzzy c-means methods can only segment the brighter region of T cells, it fails to capture the whole areas of the T cells with low intensity, so that they are easily eliminated as noises. The iterative Otsu thresholding method can achieve better result, as well as correct markers.

3.2 AMWS: Adaptive Marker-controlled Watershed Segmentation

After preprocessing, we obtain the binary image, in which, the foreground pixels are set 1, the others are 0. The detailed procedure of AMWS is as follows.

1.Invert the binary image, perform the distance transform, then compute distance image reconstruction, invert the result, and apply Watershed.

$$D = Dist(\sim B) \tag{4}$$

$$R = \rho_D(D-1) \tag{5}$$

$$S = WS(\sim R) \tag{6}$$

Here, we denote $\sim$ as the inversion operation. Let B denote the binary image, D the distance image, R the reconstructed distance image and S the rough segmentation result.

2. Analyze the segmentation result *S*, find out region set *S'*, in which the areas of all regions are in the predetermined size range. Fit the sizes of these regions to a normal distribution, obtain the mean value A_N, then calculate the radius *r*.

$$A_M = Med(S) \tag{7}$$

$$S' = \{S \mid 0.5 * A_M \le S \le 2 * A_M\} \tag{8}$$

$$A_N = Norm(S') \tag{9}$$

$$r = \sqrt{A_N / \pi} \tag{10}$$

where $Med(S)$ computes the median value of the set S. S' is the subset of S ,in which the sizes satisfy two times or one half of the median value. $Norm(S')$ is to fit the areas of S' to a normal distribution and return the mean value.

3. Perform the ultimate erosion to the binary image B to get a rough marker set M. Because of the roundish characteristics of T cells, the two markers with small distance between them should be considered to belong to the same cell. We get the foreground markers M_{fg} by making use of the radius r obtained in the last step to select markers. The detailed steps are as follows:

// $Seed_i$ is the marker set in the i-th iteration
Initiate the marker set $Seed_0 = \Phi$;
for each connected component $B_i \in B$ {
  ultimate erode B_i and obtain the markers M_i
  for each marker $M_i^j \in M_i$ {
    Compute the distance between M_i^j and each element S_i^k in $Seed_i$
    $d_j^k = \| M_i^j - S_i^k \|$;
    if $d_j^k \ge 2r$, M_i^j is added to $Seed_i$;
  }
}
Output the foreground $M_{fg} = \cup Seed_i$.

---

4. Calculate the marker set M_{bg} of the background by skeletonizing method.

5. Perform the watershed-based segmentation. Then we give the evaluation to the segmentation result, and return a new fitting circular radius r'. The evaluation rule adopted here is the same as that in step 2.

6. If $|r' - r| > T$, turn to step 3. Otherwise, output a segmentation result S .

This strategy of the marker selection aims to find exact one marker in a T cell to avoid over-segmentation and under-segmentation.

Based on the analysis of segmentation result in each iteration, we can obtain a radius adaptively to guide next segmentation. This strategy integrates a priori know-ledge to segmentation, making the result more reliable and robust.

3.3 Post-process

The marker selection strategy showed above can effectively solve the problem existing in the distance reconstruction, however it would product a wrong result in case of very irregular shape, showed as Fig.5. At some time in activation of T cells, they might outstretch tentacles, thus present irregular appearance. These cells are often wrongly segmented to two or more, among which, one might be much bigger than the other. In this case, we can use a rule-based approach to merge them. We also use these rules to splitting regions due to under-segmentation that sometimes occurs.

Fig.5 Over-segmentation examples

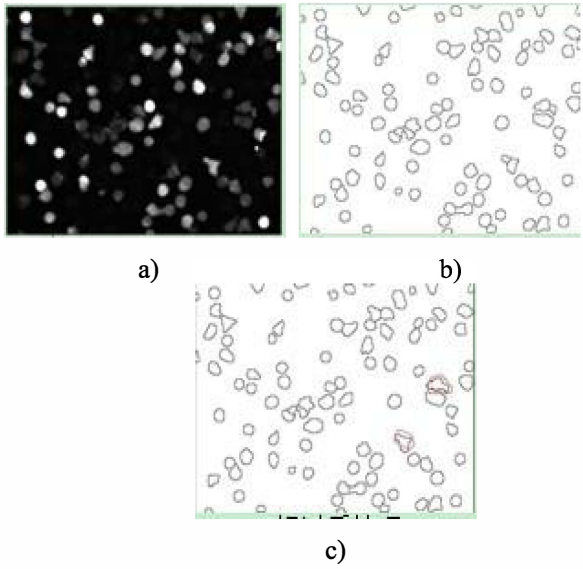


Figure.6 C Improvement by post-processing.

a) Original T cell image. b)Segmentation result without post-processing.
c) Segmentation result after post-processing.

Based on the r' calculated by the last step, we set the minimal area threshold T_A, and from the segmentation result obtained by above-mentioned algorithm select the regions which size is smaller than T_A And A_m denotes the set of these regions. These regions in A_m are small regions that do not seem whole cells. What we should do is decide whether we combine them with the nearest regions from them.

For each small region $A_m^i \in A_m$, if it has some neighbor regions, we select the nearest one to merging it with A_m^i, if it is alone, it should represents a cell, so we believe that the small region is the part of the cell.

Improvement by post-process is illustrated in Fig 6.

4. Experimental results

In this section, we apply our algorithm on the segmentation of fluorescence images of T cells, which come from Department of Immunology, University of Texas Southwestern Medical Center. The experiments are performed on the Matlab 2008a. The dataset of T cells is composed of images of four types of proteins, named Cofilin, Coro¬nin1A, CPalpha1 and EGFP. Each type of protein is imaged at different time. We use "Run" to denote a batch of images to the same culture dish in one biology experiment. Almost 46 images constitute one Run. Currently, the segmentation of each image with 500*385 pixels takes approximately 2–3 seconds on a PC with 1GHz CPU. A few samples of segmentation results are displayed in Fig 7.

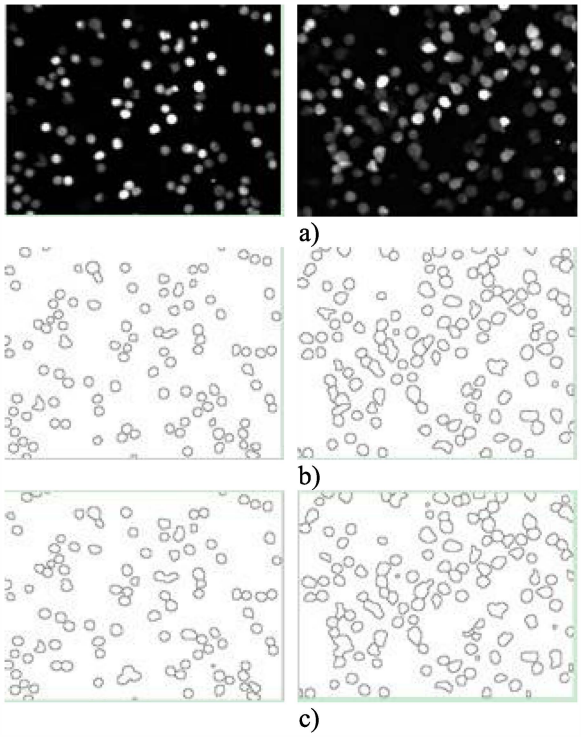


Figure.7 Segmentation results of two algorithms.

a)Original T cell images. b)Results of distance reconstruction based watershed. c)Results of the proposed method.

As seen in Fig7, both the proposed approach and the classical reconstruction-based[14] watershed method can deal with over-segmentation well. As we know, reconstruction-based method cannot control markers. Once the marker does not detect correctly, it will miss an object. In contrast, the proposed method generated more precise and reliable segmentation by avoiding over-segmentation and under-segmentation. The radius that used for markers selection was obtained adaptively based on the T cells of the images. Therefore, we do not need to set the optimal value manually.

Table 1 presents the segmentation results of Cofilin images. We randomly chose 6 batches of images, then we

obtain the accuracy rate by visual inspection of a professional technician in a biological lab. And table 2 lists the segmentation results of four types of protein images in summary. "Other errors" means the others mistakes, such as the mismatched shape, threshold-caused disappearance. The criteria used to evaluate the over-segmented and under-segmented is compare with original t-cell. We also show the precision of traditional marker-based watershed method for comparison.

In a very small number of cases, we obtain wrong results. The first, several cells with significant intensity variance were seriously clustered, so that it is really difficult to separate them. The second, because of the greatly blurred edge of T cells, the gradient information tends to reflect the wrong contour, resulting in the over-segmentation or the very irregular shape of the cell regions in the results. The third, since in some time of activation of T cells, they may become very peculiar. It will often be split incorrectly.

TABLE Ⅰ The statistical information of segmentation result of Cofilin

Number of images	Correctly Segmented	Over-Segmented	Under-Segmented	Other errors	traditional watershed
20	97.8%	1%	1%	0.2%	95%
16	98%	1%	1%	0	97%
22	98.5%	0.8%	0.3%	0.4%	97.5%
25	97.6%	1.2%	0.6%	0.6%	96%
25	97%	0.7%	1.5%	0.8%	97%
30	97.6%	1.2%	1.2%	0	96.6%

TABLE Ⅱ The statistical information of four types of protein images

Protein	Correctly Segmented	Over-Segmented	Under-Segmented	Other errors	traditional watershed
Cofilin	97.7%	1%	0.9%	0.4%	96.6%
Coronin1A	93%	5%	1%	1%	88%
CPalpha1	96.4%	2%	1%	0.6%	90%
EGFP	99%	0.5%	0	0.5%	95%

5. Conclusion

In this paper, we present an automated technique to segment T cells in fluorescent microscopy images, which is capable of accurately separating most of the clustered cells. An adaptive threshold was proposed to accurately recognize T cells, avoiding setting the threshold manually. Additionally, a closed-loop processing is introduced to adaptively find markers, avoiding over-segmentation and under-segmentation. The experimental results show that the proposed method is more accurate and robust than the classical marker-based watershed.

However, when the size and the shape of T cell vary a lot or the cluster is too complicated, the algorithm could fail to detect the correct number of T cells. We should take measures to solve in our future work. First, we expect to use more priors characteristic or inclusion of strong shape constraints for accurate description of T cells, in case of very irregular boundaries. Second, we can integrate more information such as the priori biological knowledge into the segmentation process to guide the segmentation procedure.

Acknowledgements

We are grateful to Department of Immunology, University of Texas Southwestern Medical Center, for discussion and provision of T cell datasets. This paper is supported by Guangzhou Science and Technology Support Key Projects (No.2012J4300030) and Guangdong Provincial Science and Technology Plan Projects (No.2011B040200073).

References

[1] SriparnaSaha, SanghamitraBandyopadhyay. MRI Brain Image Segmentation by Fuzzy Symmetry Based Genetic Clustering Technique. IEEE Congress on Evolutionary Computation 2007: 4417-4424

[2] Laura Bradbury and Justin W.L. Wan, A Spectral K-Means Approach to Bright-field Cell Image Segmentation, 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina, August 31 - September 4, 2010

[3] S.shen, W.A.Sandham, M.H.Granat, M.F.Dempsey, J.Patterson, Fuzzy clustering based applications to medical image segmentation, Proceedings of the 25th Annual International Conference of the IEEE EMBS Cancum, Mexico September 17-21,2003

[4] L. Vincent, E.R. Dougherty, Morphological segmentation for textures and particles, E.R. Dougherty (Ed.), Digital image processing methods, Marcel Dekker, New York, NY (1994), pp. 1-472

[5] Yuqian Zhao, Research on Medical Images Processing Theories and Methods Based on Mathematical Morphology, PhD thesis, Central South University 2006.4

[6] ArjanKuijper, Bettina Heise, An automatic cell segmentation method for differential interference contrast microscopy, ICPR 2008: 1-4

[7] Christophe Zimmer, Elisabeth Labruyere,et, Segmentation and Tracking of Migrating Cells in Videomicroscopy With Parametric Active Contours:A tool for Cell-Based Drug Testing, IEEE Transactionson

on medical imaging vol.21, no.10,october 2002

[8] Zhou Y,Kuijper A,He L, Multiphase level set method and its application in cell segmentation, Proceedings of the Fifth IASTED International Conference on Signal Processing,Pattern Recognitionand Applications,2008,134-139

[9] Joe B.T.M Roerdink and Arnold Meijster, The Watershed Transform:Definitions, Algorithms and Parallelization Strategies, FundamentaInformaticae 41 (2001) 187-228

[10] Luc Vincent and Pierre Soille, Watersheds in Digital Spaces:An Efficient Algorithm Based on Immersion Simulations, IEEE Transactions on pattern analysis and machine VOL.13 NO.6 JUNE.1991

[11] Yung-Chieh Lin, Yu-Pao Tsai,Yi-Ping Hungand Zen-Chung Shih, Comparison Between Immersion-Based and Toboggan-Based Watershed Image Segmentation, IEEE Transactions on image processing. VOL.15 NO.3 MARCH.2006

[12] Richard Beare, Gaetan Lehmann, The watershed transform in ITK - discussion and new developments, Department of Medicine,Monash University, Australia, June 9,2006

[13] Luc Vincent, Efficient computation of various types of skeletons, In Procs. Medical Imaging V, SPIE, volume 1445, pages 297-311, 1991

[14] Luc Vincent, Morphological grayscale reconstruction in image analysis applications and efficient algorithms, IEEE Transactions on image processing vol.2,No.2, April.1993

[15] Olivier CUISENAIRE, Distance Transformations: Fast Algorithms and Applications to Medical Image Processing, PhD thesis, Universitecatholique de Louvain(UCL), Belgium, 1999

[16] Luc Vincent& B.Masters, Morphological Image Processing and Network Analysis of Cornea Endothelial Cell Images in Proc. SPIE, vol.1769, Image Algebra and Morph Image Processing III ,San Di¬ego,CA,pp.212-226,July 1992

[17] Nobuyuki Otsu. A threshold selection method from gray-level histograms IEEE Transactions on Systems, Man, and Cybernetics, VOL. SMC-9, NO.1, JANUARY 1979.