THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Gradient Watershed Transform Based Automated Cell Segmentation for THG Microscopy Medical Images to Detect Skin Cancer

S. Lavanya
UG Student, Department of CSE, SNSCT, CBE, India
B. Naveen Kumar
UG Student, Department of CSE, SNSCT, CBE, India
R. Obuliraj
UG Student, Department of CSE, SNSCT, CBE, India
S. Dhanalakshmi
Assistant Senior Professor, Department of CSE, SNSCT, CBE, India

Abstract:
The conventional cell segmentation by manual selection of cells takes several minutes to segment and analyze one image and may suffer from fatigue of human eyes and subjective judgments by medical staff resulting in poor results. In this paper, a computer-aided design for automatic cell segmentation and NC ratio analysis was proposed. The proposed cell segmentation approach, is based on the watershed-based approach and the concept of convergence index filter for automatic cell segmentation. Image thresholding is the simplest method for high-speed cell segmentation, but it produces good results only for images with high contrast between object and background. Watershed transform another common technique, uses the concept of morphological image processing and considers contextual information to produces table segmentation results. Experimental result shows better result when compare with existing one.

Key words*: Cell segmentation, convergence index filter, nuclear-to-cytoplasmic ratio (NC ratio), third harmonic generation (THG), watershed transform. support vector machine (SVM)*

1. Introduction

Conventional biopsy procedures involve invasive tissue removal from a living subject, followed by time-consuming and complicated processes, so noninvasive *in vivo* virtual biopsy, is highly desired which possesses the ability to obtain extensive tissue images without removing tissues, is highly desired. Some sets of *in vivo* virtual biopsy images provided by well volunteers were processed by the future cell segmentation approach, which is based on the watershed-based approach and the concept of convergence index filter for automatic cell segmentation. Investigational results suggest that the proposed algorithm not only reveals high accuracy for cell segmentation but also has remarkable potential for noninvasive analysis of cell nuclear-to-cytoplasmic ratio (NC ratio), which is important in predicting early symptoms of diseases with abnormal NC ratios, such as skin cancers during clinical diagnosis. The method of cell segmentation is preferred in the field of medical research to aid the analysis of biomedical images by delineating round objects and obtaining information about objects, size, area, or shape to locate their positions or measure useful properties. Several algorithms for cell segmentation have been available. High-speed cell segmentation is achived using Image thersolding, but it produces good results only for images with high contrast between object and background. Watershed transform, another common technique uses the concept of morphological image processing and considers contextual information to generate stable segmentation results. Fragment merging and marker-controlled strategies are usually utilized for avoiding over-segmentation of watershed transform.

We concentrate on the design and development of an algorithm that automatically separates and segments individual cell from THG virtual biopsy images, whose segmentation results improve on our existing work. This will finally improve the efficiency and accuracy of the critical part of virtual biopsy procedures. In depth theoretical discussion and experimental setup for THG microscopy can be found in previous works.

2. Related Work

2.1. Medical Imaging

In the clinical context, medical imaging is generally equated to radiology or "clinical imaging" and the medical practitioner responsible for interpreting (and sometimes acquiring) the images is a radiologist. Diagnostic radiography designates the technical aspects of medical imaging and in particular the acquisition of medical images. The radiographer or radiologic technologist is

usually responsible for acquiring medical images of diagnostic quality, although some radiological interventions are performed by radiologists. While radiology is an evaluation of anatomy, nuclear medicine provides functional assessment.
As a field of scientific investigation, medical imaging constitutes a sub-discipline of biomedical engineering, medical physics or medicine depending on the context: Research and development in the area of instrumentation, image acquisition (e.g. radiography), modeling and quantification are usually the preserve of biomedical engineering, medical physics and computer science; Research into the application and interpretation of medical images is usually the preserve of radiology and the medical sub-discipline relevant to medical condition or area of medical science (neuroscience, cardiology, psychiatry, psychology, etc.) under investigation. Many of the techniques developed for medical imaging also have scientific and industrial applications.

2.2. Image Segmentation
Segmenting the nucleus and cytoplasm of leukocytes from bone marrow images is a very difficult task, as the images show heterogeneous staining and high-cell population as shown in figure below. Some segmentation techniques such as thresholding, edge detection, pixel clustering, and growing regions have been combined to extract the nucleus and cytoplasm of leukocytes. These techniques could be applied as the images showed uniform backgrounds and high contrast that appropriately defined the objects of interest. In this paper an approach is been proposed in which a segmentation algorithm based on color and texture pixels features is performed that can work in bone narrow images showing heterogeneous staining.
Finding dots, i.e. small round regions, in an image is a frequently encountered task in medical and scientific research. The dots could be microscopic views of cochlea hair cells, epithelial A549 cells, HEK293T embryonic kidney cells or silicon wafer defects. Counting these dots, locating them, and measuring their intensity are important for understanding hearing mechanism, cancer development, or material properties. Given the large number of dots in each image and the large number of images in these applications, it is essential to have a computer vision algorithm which extracts these dots automatically. Finding dots is a challenging segmentation problem.
These microscopic images often have poor imaging quality large intensity variation and extensive occlusion and conjunction between dots Even to the human eye, while fuzzy hair cells do pop out, low-contrast A549 and kidney cells need scrutinizing, and conjoined silicon pits require thinking to separate them. Our goal is to develop an algorithm that is capable of finding dots in all these types of images
Image segmentation is conventionally formulated as separating regions of homogeneous features such as intensity and texture However, the difference of features in adjacent regions is not always large enough to separate them completely, creating gaps along region boundaries. A common remedy for completing the gaps is to include in the formulation a prior term which favours a segmentation with smooth boundaries All these traditional formulations of segmentation view regions as solid entities occupied by pixels, and boundaries as abstract lines taking up zero space in-between. While precise region delineation is useful for image manipulation or object recognition and grasping, it is not necessary for our dot and many other applications. Where boundaries should be located is flexible, so long as the core of each dot is retained in the region.
Much of cell biology experimental research is based on microscopy image analysis of cell culture. In many cases the use of different fluorescence dies or proteins is used to enable the collection of multivariate images which contain information on different aspects of each cell. Although analysis of such images can be performed manually, it is time consuming, exhausting and prone to human error, requiring frequent repetitions to validate results. These factors motivate the development of automatic cell analysis tools, to identify each individual cell and extract relevant cell characteristics. This process involves the separation of each individual cell from all others cells and from the background.
Multivariate imaging is widely used in cell microscopy to obtain separated cell nucleus and cytoplasm information by using different fluorescence markers, originating two image channels. Additional fluorescence markers (and channels) can be used to express other phenomena or objects such as parasites or different nucleic acid concentrations

3. Existing System
In existing system, Traditional techniques use image content like color, texture and gradient to represent images has been used. In order to deal with large scale data, mainly two kinds of indexing systems are used. Automatically detected tissue cells have been shown promising in different applications recently. Traditional NC ratio analysis, however, is performed by manual selection of cells from each THG image and subsequent calculation of their volume ratio. In addition to being time consuming, this procedure causes fatigue of the medical personnel and hence inconsistent accuracy of results caused by, for example, visual error. Watershed transform another common technique, uses the concept of morphological image processing and considers contextual information to produce stable segmentation results. The existing system cannot compare the abnormal ratios and normal ratios accurately.

4. Proposed System
We use gradient watershed transform based on the contextual information of pixels in images for nuclei segmentation. The main idea of watershed transform is to identify regional minima in original images, which are frequently the regions of nuclei to be segmented. First, we calculate the gradient map to show up the nuclei boundary. The homogeneous regions in the gradient map match to the regional minima in the input image. We renounce from performing nuclei detection directly on the gradient map because its subsidiary regional minima or noise would cause over-segmentation. Blob detection must be performed at the beginning of nuclei detection Instead, nuclei initialization with marker-controlled strategy. For nuclei initialization, potential nuclei and their corresponding cell boundaries can be obtained to help design a filter to remove the subsidiary regional minima from the gradient map. The filtered gradient map would then be subject to nuclei detection using watershed transform, yielding segmented nuclei. Nuclei validation must also be considered, using a shape descriptor to eliminate the outliers and conserve the valid nuclei.

Of course, if we have user interaction or some prior information in the analysis procedure, the results will be persuasive for medical diagnosis. For cytoplasm segmentation, cytoplasm initialization would be performed based on the information of valid nuclei to obtain potential cytoplasms and corresponding refined cell boundaries. Then, a local filter is designed to be used in the cytoplasm detection stage to obtain segmented cytoplasm. Finally, the evaluated NC ratios are obtained in the NC ratio evaluation stage. The Existing system do not provide the accuracy to predict the symptoms of cancer.

4.1. Nuclei Segmentation

Nuclei segmentation is performed using gradient watershed transform with marker-controlled strategy, blob detection, and consideration of shape descriptors to obtain accurate segmented nuclei. Nuclei segmentation is a crucial stage in the proposed cell segmentation algorithm because the subsequent stage of cytoplasm segmentation references valid nuclei, which can be thought of as the initial condition of the optimization problem of whole cell segmentation in order to guide the algorithm in finding a feasible solution with high performance.

4.1.1. Watershed Transform With Marker-Controlled Strategy:

Watershed transform which considers contextual information in an image and identifies the regional minima, is chosen here for nuclei segmentation. In observation of images, nuclei with lower intensity locate in the regional minima of images, which are highly correlated with homogeneous regions in the gradient map. Therefore, we utilize the gradient map with prior information for the boundaries of nuclei to help to extract and analyze the almost uniform and round nuclei of adjacent cells from the background. To have a robust gradient map, we use the Sobel operator and Roberts cross-gradient operator to calculate gradient maps in the horizontal, vertical, and positively sloped diagonal, and negatively sloped diagonal directions.

Nuclei initialization would prevent over-segmentation by adopting the marker-controlled strategy followed by the technique of minima imposition existing in morphological image processing. Two kinds of markers are needed: those that roughly mark the locations of potential nuclei and those that mark their corresponding cell boundaries. These two kinds of markers are defined respectively as following

- Internal markers: The groups of connected pixels inside each region where the potential nucleus is to be segmented.
- External markers: The groups of connected pixels ideally relative to the boundary of each cell.

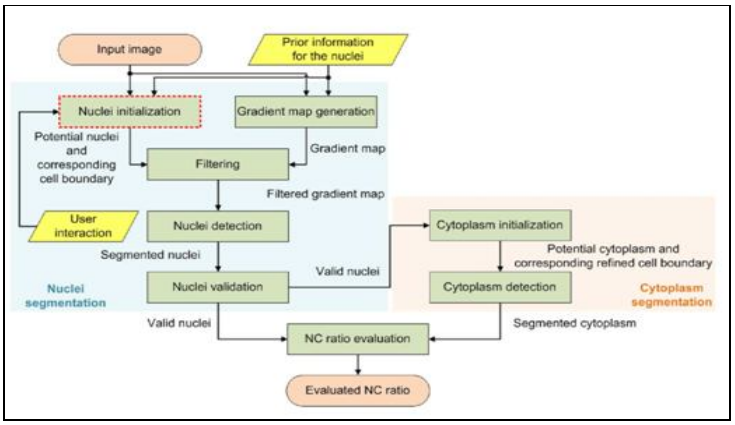

Figure 1: Block diagram of cell segmentation and NC ratio analysis

4.1.2. Nuclei Initialization

The objective of nuclei initialization is to obtain the potential nuclei and their corresponding cell boundary using blob detection followed by outlier removal as well as distance transform. The potential nuclei are marked with internal markers, and the corresponding cell boundaries are marked with external markers. Internal markers must meet the following criteria:

- Regions of internal markers must be surrounded by pixelsof higher intensity.
- All pixels in each region of internal markers should form a connected component with homogeneous intensity.
  To determine the internal markers, blob detection is performed with prior information of the nuclei to be segmented.

4.2. Cytoplasm Segmentation

For cytoplasm segmentation, a convergence index filter with intuitive parameter setting is adopted based on valid nuclei .Since it considers gradient vectors instead of intensity of images, a convergence index filter is suitable for low-contrast and noisy microscopy images. Such a filter makes unnecessary the pre-processing to enhance contrast and remove irregular noisein biomedical images, while preserving the information needed for clinical diagnosis.

In addition, parameter setting for convergence index filter is intuitive for user or medical staff to adjust easily without technical details of algorithmic processing according to some fundamental information of input biomedical images to be analyzed like size or shape of cells. One thing worth mentioning is that convergence index filters consider contextual and locality information to have confident segmentation results and reduce uncertainty of segmentation resulting from low contrast and noise in images.

Several assumptions should be made for using convergence index filtering to segment cytoplasms:

- The shape of cells and their corresponding nuclei are all convex regions.
- Each cytoplasm and its corresponding nucleus are almost concentric, and the gradient vector of each pixel belonging to the cytoplasm and its corresponding nucleus has a trend to point toward the same cell center.

The assumptions above enable information about valid nuclei to be used to detect size and shape of cytoplasm for each cell. It depicts two adjacent cells whose convex regions contain their boundaries of cytoplasm in black and nuclei in red. Green arrows represent the distribution of gradient vectors around cytoplasmic boundaries that point toward the corresponding cell center.

4.2.1. Cytoplasm Initialization

First, we need to determine two constraints, and , which represent the minimum distance from the inner boundary and the maximum distance from the outer boundary to the center of the support region on each orientation, respectively. These two constraints are adaptive to the shape and position of each valid nucleus. The boundary of each valid nucleus can be thought of as a minimum boundary of cytoplasm on that cell that makes sure that the candidates of cytoplasmic region are not present in the nuclear region.For determination of , we adopt the concept of distance transform to generate the refined distance map of valid nuclei to delineate the maximum boundary of potential cytoplasmic region for each cell. Therefore, we make the robust and confident restriction of boundary of potential cytoplasmic region with adaptive constraints and for each cell,which can not only avoid overlap of segmented cells but alsoresolve issues of cell segmentation of adjacent cells with in distinct boundaries in multivariate biomedical images.

4.2.2. Cytoplasm Detection

After defining R_{min} and R_{max} for each cell, the support region of the proposed local filter can be thought of as the union of N line segments , which represent cytoplasmic width radiating from cell center on each orientation. In addition, the definition of the convergence index of the gradient vector is almost the same as that of
the sliding band filter but some constraints about variable distance between cell center and the outer boundary of the support region on each orientation are different.
The output of the proposed local filter applied to the pixel of interest P of Cartesian coordinate (x,y) is defined as the average of convergence indices on N line segments in2-D discrete space.

$$\mathrm{Filter}(x,y) = \frac{1}{N}\sum_{j=0}^{N-1}\mathrm{Filter}_j(x,y)$$

With

$$\mathrm{Filter}_j(x,y) = \frac{1}{r_j - R_{\min}^j}\max_{R_{\min}^j < r_j \le R_{\max}^j}\sum_{i=R_{\min}^j+1}^{r_j}\cos\theta_i^j$$

$$\cos\theta_i^j = \cos\left(\theta_j - \phi(i,\theta_j)\right)$$

$$\theta_j = \frac{1}{N}2\pi j,\ where\ j = 0,1,2,\cdots,(N-1)$$

$$\phi(i,\theta_j) = \tan^{-1}\left(\frac{G_y(i,\theta_j)}{G_x(i,\theta_j)}\right)$$

4.3. Cell Size and Nc Ratio Evaluation

Cellular size and nuclear size are indicators not only of the developing status of some diseases but of skin aging and other quantifiable physical factors. For example, cellular and nuclear size in the layers of basale cells in forearm skin has been found to increase with age.
The NC ratio, which is defined as the volume ratio of nucleus to cytoplasm, is commonly used in diagnosis. A protocol has been developed to obtain accurate NC ratios. Although the NC ratio is defined as a volume ratio, it can be approximated by an area ratio of nucleus to cytoplasm.

4.4. SVM Algorithm

It is used to train normal and abnormal tissue images.This algorithm is used to obtain higher accuracy.It is used to compare the values and to predict normal and abnormal ratios.

5. Experimental Results

Automatic cell segmentation and NC ratio evaluation were performed using the proposed algorithm on about 600 THG virtual biopsy images of the stratum basale layer of human forearm skin from 31 healthy volunteers. The evaluated NC ratios and cell sizes were discussed and interpreted by a dermatologist and a radiologist. Most of the cells were segmented accurately. Moreover, user interaction with medical doctors or medical staff can be adopted to exclude mistakenly segmented cells

further with the profiles that record the information of each segmented cell or select specific cells of interest to enhance the performance of cell segmentation and NC ratio evaluation in clinical diagnosis.

Cell Index	Cell Area (pixels)	Nuclear Area (pixels)	Cytoplasmic Area (pixels)	NC Ratio	Position (row, column)
1	924	229	695	0.329496403	(476 , 16)
2	649	160	489	0.327198364	(289 , 36)
3	901	228	673	0.338781575	(453 , 66)
4	1076	316	760	0.415789474	(422 , 91)
5	836	201	635	0.316535433	(451 , 91)
...	...	...	...	...	...
49	862	222	640	0.346875	(143 , 364)
50	972	319	653	0.488514548	(242 , 374)
51	943	296	647	0.457496136	(442 , 370)
52	722	217	505	0.42970297	(452 , 394)
53	1226	303	923	0.328277356	(222 , 398)
54	1221	408	813	0.501845018	(460 , 419)
55	1357	400	957	0.417972832	(267 , 428)
56	1341	452	889	0.508436445	(193 , 464)
57	1323	195	1128	0.17287234	(143 , 467)
58	1066	265	801	0.330836454	(157 , 490)
Total	56250	15732	40518	-	-
Average	969.8276 (pixels)	271.2414 (pixels)	698.5862 (pixels)	0.388271879	-
Average	53.2742 (μ m^2)	14.8997 (μ m^2)	38.3745 (μ m^2)	-	-

Table 1: Profile of Each Segmented Cell

6. Conclusion

In this work, a watershed transformation for automatic cell segmentation and NC ratio analysis was proposed. SVM algorithm is used to train normal and abnormal tissue images.The experimental results show that the proposed method provides objective segmentation results with high efficiency and consistent accuracy of NC ratio value. Moreover, the determination of NC ratios of skin cells using the proposed automatic algorithm is more intention and forceful than that using manual approach, and hence medical doctors can diagnose potential diseases without the power of any subjective factor, such as the subjective judgment of analyzer and the fatigue of the medical personnel. It is the faster and efficient way to predict skin cancer in tissue cells by comparing the evaluated values. Better experimental result has been achieved when compare with the existing technique. Future work in this paper may involve developing the algorithm using SVM to
detect cancer in any tissue images not only in skin related tissue images.

7. References

1. C.-K. Sun, "Higher harmonic generation microscopy,"Adv.Biochem.Eng./Biotechnol., vol. 95, pp. 17–56, 2005.
2. S.-Y. Chen, S.-U. Chen, H.-Y. Wu, W.-J. Lee, Y.-H. Liao, and C.-K. Sun, "In vivo virtual biopsy of human skin by using noninvasive higher harmonic generation microscopy," IEEE J. Sel. Topics Quantum Elec-tron., vol. 16, no.3, pp. 478–492 2010.
3. S.-Y. Chen, H.-Y. Wu, and C.-K. Sun, "In vivo harmonic generation biopsy of human skin," J. Biomed. Opt., vol. 14, no. 6, p. 060505, 2009.
4. M.-R. Tsai, S.-Y. Chen, D.-B. Shieh, P.-J. Lou, and C.-K. Sun, "In vivo optical virtual biopsy of human oral mucosa with harmonic generation microscopy," Biomed. Opt. Expr., vol. 2, no. 8, pp. 2317–2328, 2011.
5. Y.-H. Liao, S.-Y. Chen, S.-Y. Chou, P.-H. Wang, M.-R. Tsai, and C.-K. Sun, "In vivo harmonic generation biopsy for quantitative evaluation in chronological aged skin keratinocytes," Biomed. Opt. Expr., vol. 4, no. 1, pp. 77–88, 2013.
6. M.-R. Tsai, Y.-H. Liao, and C.-K. Sun, "Differential diagnosis of pig-mented skin lesions based on harmonic generation microscopy," pre-sented at the Photonics in Dermatology and Plastic Surgery Conf., Pho-tonics West, San Francisco, CA, USA, 2013, paper 8565-6.
7. M.-R. Tsai, D.-B. Shieh, P.-J. Lou, C.-F. Lin, and C.-K. Sun, "Charac-terization of oral Squamous Cell Carcinoma based on higher-harmonic generation microscopy," J. Biophoton., vol. 5, no. 5–6, pp. 415–424, 2012.
8. E. Bernardis and S. X. Yu, "Finding dots: Segmentation as popping out regions from boundaries," in Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Jun. 2010, pp. 199–206.
9. P. Quelhas, M. Marcuzzo, A. M. Mendonca, and A. Campilho, "Cell nuclei and cytoplasm joint segmentation using the sliding band filter," IEEE Trans. Med. Imag., vol. 29, no. 8, pp. 1463–1473, 2010.
10. H.-H. Lin, M.-R. Tsai, C.-F. Chen, S.-Y. Chen, Y.-H. Liao, G. G. Lee, and C.-K. Sun, "Cell segmentation and NC ratio analysis of Third Har-monic Generation virtual biopsy images based on marker-controlled gradient watershed algorithm," Proc. IEEE Int. Symp. Circuits and Sys-tems, pp. 101–104, 2012