

히스토그램 분포 분류를 통한 효율적인 세포 이미지 분할 시스템

조미경*

An Efficient Segmentation System for Cell Images By Classifying Distributions of Histogram

Migyung Cho*

Department of Media Engineering, Tongmyong University, Busan 608-711, Korea

요 약

세포 분할 작업은 세포 이미지의 배경으로부터 세포 영역을 추출하는 작업으로 배양과정에 있는 살아있는 세포를 이미지화하여 분석하는 바이오 이미징 분야에서 기초적인 작업들 중 하나이다. 선명한 이미지의 경우 바이모달 히스토그램 분포를 가지므로 Otsu와 같은 전역임계값 알고리즘을 이용하여 쉽게 세포분할 작업을 수행할 수 있지만 희미한 이미지의 경우는 정확한 세포 분할을 하기가 어렵다. 본 논문에서는 입력된 세포이미지의 히스토그램을 분석하여 히스토그램 분포에 따라 분류한 후 바이모달 분포를 가지는 이미지의 경우 전역임계값 알고리즘을 적용하고 유니모달 분포를 가지는 이미지의 경우 영역을 분할하여 부분 영역별로 다른 임계값을 적용하는 세포 분할 시스템을 개발하였다. 실험결과 제안한 시스템은 바이모달 분포를 가지는 세포이미지 뿐만 아니라 유니모달 분포를 가지는 세포 이미지에 대해서도 정확한 세포 분할 작업을 수행하였다.

ABSTRACT

Cell segmentation which extracts cell objects from background is one of basic works in bio-imaging which analyze cell images acquired from live cells in cell culture. In the case of clear images, they have a bi-modal histogram distribution and segmentation of them can easily be performed by global threshold algorithm such as Otsu algorithm. But In the case of degraded images, it is difficult to get exact segmentation results. In this paper, we developed a cell segmentation system that it classify input images by the type of their histogram distribution and then apply a proper segmentation algorithm. If it has a bi-modal distribution, a global threshold algorithm is applied for segmentation. Otherwise it has a uni-modal distribution, our algorithm is performed. By experimentation, our system gave exact segmentation results for uni-modal cell images as well as bi-modal cell images.

키워드 : 분할, 세포 이미지, 히스토그램, 유니모달 분포, 바이모달 분포

Key word : Segmentation, Cell images, Histogram, Uni-modal distribution, Bi-modal distribution

접수일자 : 2013. 10. 15 심사완료일자 : 2013. 11. 11 게재확정일자 : 2013. 11. 25

* **Corresponding Author** Migyung Cho(E-mail:mgcho@tu.ac.kr, Tel:+82-51-629-1214)
Department of Media Engineering, Tongmyong University, Busan 608-711, Korea

Open Access <http://dx.doi.org/10.6109/jkice.2014.18.2.431>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

줄기 세포의 움직임을 분석 및 관찰하기 위해 공초점이나 DIC 현미경 등을 이용하여 배양 중에 있는 세포들을 촬영하여 세포이미지를 획득한 후 세포 이미지를 분석하기 위한 연구들이 활발히 진행되고 있다. 현미경으로부터 세포이미지를 촬영할 때 환경에 따라 선명한 이미지를 획득하기도 하지만 잡음이 많고 명암 대비가 낮은 이미지를 얻기도 한다.

세포분석을 위한 첫 번째 단계는 배경으로부터 세포 영역을 분할하는 작업이다. 잡음이 없고 명암 대비가 선명한 세포 이미지들은 바이모달 분포(bi-modal distribution)나 멀티모달 분포(multi-modal distribution)를 가지므로 전역 임계값 알고리즘과 같은 영상처리 분야에서 일반적으로 사용하고 있는 분할 알고리즘을 적용하여 정확하게 분할 작업을 수행할 수 있다[1-2].

일반적으로 영상처리 분야에서 사용하는 분할 알고리즘은 의약 영상 분야에서 널리 구현되어 있으며 최근 사용하는 분할 알고리즘들은 워터세드, 보르노이 기반 알고리즘(Voronoi-based algorithm), 히스토그램 기반 클러스터링, 임계값 알고리즘 등이 있다[3-7]. 그런데 이러한 분할 기법들은 이미지의 품질이 좋은 경우 정확한 분할 작업을 기대할 수 있다. 명암대비가 낮고 잡음이 많은 세포이미지의 경우 이러한 알고리즘을 이용하여 정확한 분할 결과를 얻기 힘들다.

명암대비가 낮고 잡음이 많은 세포이미지들의 히스토그램을 분석해 보면 대부분 유니모달 분포(uni-modal distribution)를 가진다[2,8]. 이처럼 유니모달 분포를 가지는 이미지를 위한 분할 알고리즘들도 제안되었는데 [8-9] 대체로 영역 분할과 적응적 임계값 알고리즘 등을 이용하며 전역 임계값 알고리즘에 비해 비교적 계산 비용이 많이 소요되는 분할 기법이 요구된다. 또한 유니모달 분포를 가지는 이미지들에 대한 분할 기법을 바이모달 분포나 멀티모달 분포를 가지는 품질이 좋은 이미지에 적용하면 원하는 결과를 얻지 못하는 경우가 발생한다. 따라서 효율적인 분할 작업을 위해서는 세포 이미지의 품질을 먼저 분석하여 각 이미지에 맞는 분할 기법을 적용하는 것이 바람직하다.

본 연구에서는 다양한 품질의 세포 이미지들에 대한 정확한 분할 작업을 위해 강도 히스토그램의 분포 분류를 통한 자동화된 세포 분할 시스템을 제안하고 개발하

였다. 제안된 시스템은 입력된 세포 이미지에 대한 히스토그램 분포를 분석하여 세포 이미지가 어떤 히스토그램 분포(바이모달, 멀티모달, 유니모달 분포)를 따르는지를 먼저 결정한다. 만약 바이모달 분포나 멀티모달 분포를 따르면 대표적인 전역 임계값 알고리즘인 Otsu 알고리즘으로 분할 작업을 수행하고 유니모달 분포를 따르면 본 연구에서 제안하는 영역 분할과 지역 임계값 알고리즘을 이용한 분할 기법을 적용하여 세포 분할 작업을 수행한다.

II. 세포 분할 시스템 구조와 기능

본 논문에서 제안하는 세포 이미지 분할 시스템의 구조는 그림 1과 같다.

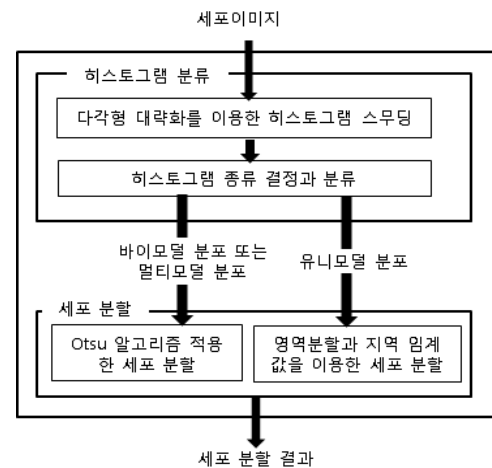


그림 1. 제안된 세포 이미지 분할 시스템
Fig. 1 Proposed segmentation system for cell images

분할 시스템은 히스토그램 분류 단계와 세포 분할 단계로 나누어진다. 히스토그램 분류 단계에서는 그레이스케일 세포 이미지를 입력으로 받아 강도 히스토그램을 얻은 후 다각형 대략화 기법을 적용하여 히스토그램을 대략화(스무딩)하는 작업을 수행한다. 이 작업은 들쭉날쭉한 히스토그램의 끝 모양을 완만하게 만들기 위해 수행하며 자세한 방법은 참고문헌 2을 참고하기 바란다. 그림 2는 명암대비가 비교적 높은 세포 이미지의 히스토그램과 참고문헌 2에서 제안한 방법으

로 스무딩한 결과로 생성된 대략화된 히스토그램을 보여준다.

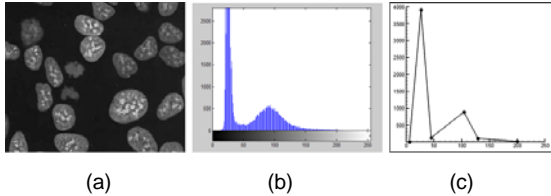


그림 2. 세포이미지와 대략화된 히스토그램:(a)세포이미지 (b) 히스토그램 (c)대략화된 히스토그램

Fig. 2 A cell image and its approximated histogram: (a) A cell image (b) Its histogram (c)Its approximated histogram

그림 2 (a) 세포이미지는 잡음이 없고 명암대비가 높은 이미지이다. 따라서 (b)의 히스토그램도 바이모달 분포를 가진다. 히스토그램이 어떤 분포를 따르는지 자동적으로 결정하기 위해 다각형 대략화 기법을 적용하여 대략화한 결과가 그림 2 (c)이다. 대략화된 히스토그램을 이용하여 입력된 세포 이미지가 유니모달분포를 가지는지 아니면 바이모달 혹은 멀티모달 분포를 가지는지 결정하게 된다. 이것을 결정하는 자세한 방법은 참고문헌 2를 참고하기 바란다.

세포분할시스템의 두 번째 단계인 세포분할 단계에서는 히스토그램의 분포 종류에 따라 바이모달 분포인 경우 Otsu 알고리즘을 적용하여 세포 분할을 수행하고 유니모달 분포를 가질 경우 3장에서 제안하는 방법에 의해 세포 분할을 수행한다.

III. 영역분할과 지역 임계값 추출

영상처리 분야에서도 유니모달 분포를 가지는 이미지의 배경으로부터 객체를 분할하는 것은 어려운 문제로 알려져 있다[8-9]. 이는 유니모달 분포에서는 최적의 임계값을 설정하는 것이 매우 어렵기 때문이다. 임계값 설정은 크게 전역 임계값 알고리즘과 지역 임계값 알고리즘으로 나눌 수 있다. 조명의 불균형으로 전체적으로 선명하지 못하거나 잡음이 심한 이미지의 경우 전역 임계값 알고리즘보다 지역 임계값 알고리즘을 적용하는 것이 바람직하다.

본 장에서는 영역 분할과 지역 임계값 알고리즘을 이

용한 세포 분할 방법을 제안한다. 조명의 불균등은 이미지의 중간영역과 가장자리 영역의 밝기를 다르게 만들며 광학 현미경의 잡음도 영역에 따라 잡음의 정도가 다르다. 본 논문에서는 이러한 것에 착안하여 이미지를 몇 개의 부분영역으로 나누어 각 부분 영역에 대한 임계값을 구하고, 구해진 임계값의 특징에 따라 최종 임계값을 적용하는 분할 알고리즘을 제안한다.

영역 분할은 그림 3과 같이 분할할 영역의 개수 n 을 사용자로부터 입력받아 수직 방향과 수평 방향으로 균등하게 분할한다. 그림에서 $width$ 는 입력된 세포 이미지의 넓이를 $height$ 는 높이를 나타낸다. 마지막 행과 열의 부분 영역의 높이와 넓이는 그림 3에서 보는 것처럼 전체 이미지의 높이와 넓이에 따라 나머지 다른 부분영역의 높이와 넓이 값과 다를 수 있다.

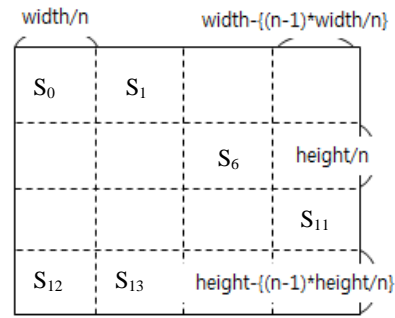


그림 3. 영역 분할 예
Fig. 3 An example of area partition

주어진 이미지를 그림 3과 같이 크기 n 에 대하여 나누는 다음, Otsu 알고리즘을 이용하여 나누어진 각 부분 영역에 대한 임계값 T_i 를 구한다. 크기 n 이 4인 경우 전체 이미지는 16개의 부분영역으로 나누어지므로 T_0 부터 T_{15} 까지의 임계값을 계산하게 된다. 구해진 부분영역들에 대한 임계값들을 이용하여 전체 이미지에 대한 임계값을 구하는 방법은 알고리즘 1과 같다.

알고리즘 1. 부분영역 임계값을 이용한 세포분할 알고리즘

- [단계1] 모든 부분영역 S_i 에 대해 Otsu 알고리즘을 이용하여 임시 임계값 $T_i(0 \leq i \leq n^2-1)$ 를 구한다.
- [단계2] 식1, 식2, 식3, 식4를 이용하여 $T_{min}, T_{max}, T_{diff}, T_{mean}$ 를 구한다.

$$T_{\min} = \min\{T_i : 0 \leq i \leq n^*n - 1\} \quad (1)$$

$$T_{\max} = \max\{T_i : 0 \leq i \leq n^*n - 1\} \quad (2)$$

$$T_{diff} = |T_{\max} - T_{\min}| \quad (3)$$

$$T_{mean} = \frac{\sum_{i=0}^{n^*n-1} T_i}{n^*n} \quad (4)$$

[단계3] 식5를 이용하여 모든 부분영역에 대한 최종 임계값 $ST_i(0 \leq i \leq n^2-1)$ 를 구한다.

$$ST_i = \begin{cases} T_i & \text{if } T_{diff} < \alpha * T_{mean} \\ T_{max} & \text{otherwise} \end{cases} \quad (5)$$

[단계4] 부분 영역 임계값 $ST_i(0 \leq i \leq n^2-1)$ 를 이용하여 대응되는 부분 영역들에 대한 세포 분할을 식6과 같이 수행한다.

$$t_i(x,y) = \begin{cases} 1 & \text{if } f_i(x,y) > ST_i \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

식 6에서 $f_i(x,y)$ 는 부분 영역 S_i 에 속한 x, y 위치의 픽셀 값 강도를 나타낸다.

IV. 실험 및 분석

제안한 알고리즘의 성능을 검증하기 위해, 전역 임계값 알고리즘인 otsu 알고리즘[6]과 지역적으로 적용된 임계값을 적용하는 MAT 알고리즘[8]의 분할 결과와 본 논문에서 제안한 알고리즘의 분할 결과를 비교해 보았다. Otsu 알고리즘은 바이모달 분포를 지니는 이미지의 분할에 대표적으로 사용되는 방법으로 성능이 입증된 알고리즘이다. MAT 알고리즘은 적응적 지역 임계값 알고리즘 중 하나로 특정 픽셀을 중심으로 일정한 영역 b 이내의 픽셀 강도의 평균과 분산을 이용하여 임계값을 구하여 분할하는 방법으로 부분적으로 잡음이 많고 배경 강도가 고르지 못한 이미지에 적용하기 위해 제안된 방법이다.

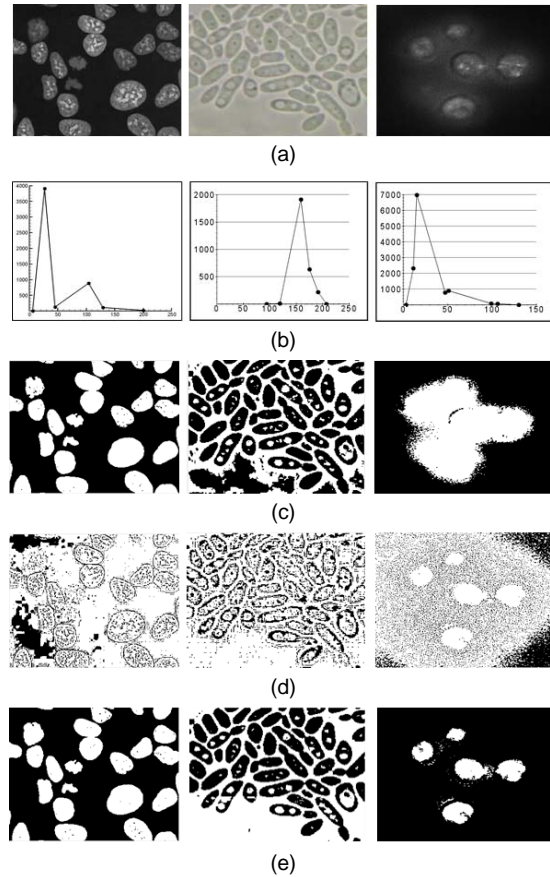


그림 4. 세포분할 결과: (a) 원본세포이미지 (b) 대략화된 히스토그램 (c) Otsu 알고리즘 결과 (d) MAT 알고리즘 결과 (e) 제안된 알고리즘의 결과

Fig. 4 The results of cell segmentation: (a) three original images (b) Their approximated histograms (c) The results of Otsu algorithm (d) The results of MAT algorithm (e) The results of our algorithm

그림 4은 원본 이미지와 대략화된 히스토그램과 세 알고리즘의 분할 결과를 보여 준다. 그림 4 (a)는 실험에 사용된 원본 이미지로 왼쪽 이미지는 잡음이 없고 배경이 선명한 이미지이고 가운데 이미지는 배경이 선명하지 못한 이미지이고 오른쪽 이미지는 잡음이 많은 이미지이다. 그림 4 (b)는 분할시스템의 히스토그램 분류 단계에서 생성된 대략화된 히스토그램의 결과를 보여주며 왼쪽 이미지는 바이모달 분포를 가운데와 오른쪽은 유니모달 분포를 가진다.

그림 4 (c)는 Otsu 알고리즘 분할 결과로 선명한 왼쪽

데이터에 대해서는 정확한 결과를 주지만 그렇지 않은 두 데이터에 대해서는 제대로 된 세포 분할을 해 주지 못했다. 이는 Otsu 알고리즘이 전역 임계값을 사용하기 때문에 배경의 강도의 불균등한 이미지에 대해 강하지 못하기 때문이다.

그림 4 (d)는 적응적 지역 임계값을 사용하는 MAT 알고리즘 결과를 보여준다. Otsu 결과와 반대로 선명한 왼쪽 데이터에 대한 분할 결과는 좋지 않지만 나머지 두 데이터들에 대해서는 Otsu 알고리즘보다 좋은 분할 결과를 주었다. 가운데 이미지의 경우 하단 부분의 세포 영역과 비슷한 강도를 지니는 배경부분도 배경 영역으로 정확하게 분할을 해 준 것을 볼 수 있고 오른쪽 이미지의 경우도 잡음이 많지만 후처리 과정을 거치게 되면 세포 영역을 추출할 수 있다.

그림 3의 (e)는 제안한 분할 알고리즘을 적용한 결과이다. 선명한 데이터뿐만 아니라 명암대비가 낮고 잡음이 많은 데이터에 대해서도 정확한 분할 결과를 주었다. 제안한 알고리즘은 바이모달 분포를 가질 경우 Otsu 알고리즘을 적용하고 그렇지 않는 경우 제안된 알고리즘 1의 식5와 같이 임계값을 설정하여 분할한다. 부분 영역별로 다른 임계값을 사용하기도 하고 부분영역별 임계값의 특징에 따라 전역 임계값을 사용하기도 한다. 따라서 이미지의 특징에 따라 다른 임계값이 설정되므로 잡음과 불균등한 조명으로 인해 유니모달 분포를 가지는 이미지에 대해서도 정확한 세포 분할 결과를 보여 주었다.

V. 결 론

바이오 이미징에서는 광학현미경으로부터 획득한 세포 이미지를 분석하여 살아있는 세포의 변화 과정을 분석한다. 이때 배경으로부터 세포를 분할해 내는 것은 가장 기초적이고 중요한 작업인데 본 논문에서는 세포 이미지의 히스토그램 분포에 따라 세포 이미지를 분류한 다음 세포를 분할하는 시스템을 개발하였다. 바이모달 분포를 가지는 경우 Otsu 알고리즘을 적용하고 유니모달 분포를 가지는 이미지의 경우 영역을 일정한 크기로 분할하여 부분영역에 대한 임계값을 구하여 임계값의 특징을 분석한 후 부분 영역별로 최종 임계값을 결정하는 분할 알고리즘을 제안하고 적용하였다. 제안한 세포

분할 시스템은 선명한 이미지뿐만 아니라 잡음과 불균등한 조명으로 인해 배경이 선명하지 못한 이미지에 대해서도 정확한 세포 분할 작업을 수행해 주었다. 앞으로 더 다양한 세포 데이터들에 대한 실험을 수행하여 제안한 세포 분할 시스템의 성능을 검증하고자 한다.

감사의 글

이 논문은 2012학년도 동명대학교 교내학술연구비 지원에 의하여 연구되었음 (2012A006)

REFERENCES

- [1] Xian Du, Sumeet Dua, "Segmentation of Fluorescence Microscopy Cell Image Using Unsupervised Mining," *The Open Medical Information Journal*, Vol.4, pp.41-49, 2010.
- [2] Migyung Cho, "An Algorithm for Determining Histogram Type of a Cell Image Using Polygonal Approximation," *International Journal of Advancements in Computing Technology*, Vol.5 No.13, pp.229-234, Sept., 2013.
- [3] Thouis R. Jones, Anne Carpenter, Polina Golland, "Voronoi-based segmentation of cells on image manifolds", *Lecture Notes of Computer Science*, Vol. 3765, pp 535-543, 2005.
- [4] Bazi Y, Ruzzzone L, Melgani F., "Image thresholding based on the EM algorithm and the generalized Gaussian distribution", *Pattern Recognition Letter*, Vol. 40, pp.619-634, 2007.
- [5] P.K Sahoo, S. Soltani, A.K.C Wong, "A survey of thresholding techniques", *Computer Vision, Graphics, and Image Processing*, Vol. 41, Issue 2, pp. 233 - 260, 1988.
- [6] Nobuyuki Otsu, "A threshold Selection Method from Gray-Level Histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-9, No. 1, pp62-66, Jan., 1979.
- [7] Migyung Cho and Jaesool Shim "Segmentation Algorithm of Fluorescence Microscopy Cell Images using Area Partition," in *Proceeding of the 3rd International Conference on Convergence Technology 2013*, Chiang Mai, Thailand, pp. 402-403, 2013.
- [8] Feixiang Yan, Hong Zhang, C. Ronald Kube, "A multistage

adaptive thresholding method,” *Pattern Recognition Letters*, vol. 26, pp. 1183-1191, Dec. 2005.

[9] Nicolas Coudray, Jean-Luc Buessler, Jean-Philippe Urban,

“Robust threshold estimation for images with unimodal histograms,” *Pattern Recognition Letters*, vol. 31, pp. 1010-1019, Jan. 2010.



조미경(Migyung Cho)

1998년 2월 부산대학교 전자계산학과(이학박사)
1998년 3월 - 2000년 7월 신라대학교 교양학부
2000년 9월 ~ 2002년 8월 부산대학교 연구교수
2005년 9월 ~ 2006년 8월 워싱턴주립대학교 방문교수
2002년 9월 ~ 현재 동명대학교 미디어공학과 부교수
※관심분야: 알고리즘, 바이오이미징, 바이오시뮬레이션