

# 자궁경부암 세포 영상 분할을 위한 Thresholding 기법

김재륜\*, 하진영\*, 김백섭\*\*, 김호성\*\*\*

\*강원대학교 컴퓨터·정보통신공학부, \*\*한림대학교 컴퓨터공학부, \*\*\*성신여대 전산학과

## The thresholding method for cervical cell image segmentation

Jae Ryun Kim\*, Jin-Young Ha\*, Baek-sop Kim\*\*, Ho Sung Kim\*\*\*

\*Division of Computer, Information & Communications Engineering, Kangwon National University

\*\*Department of Computer Engineering, Hallym University

\*\*\*Department of Computer Science, Sungshin Women's University

### 요 약

본 논문은 자궁경부암 검사를 위한 전처리 과정인 자궁경부암 세포 영상분할 문제 연구의 결과이다. 자궁경부암 세포 영상은 배경과 세포질 및 세포핵의 구별이 어렵다. 게다가 자궁경부암 검사 시스템은 짧은 시간동안 많은 영상을 처리해야 하기 때문에, 영상의 분석 속도가 빠르고 강력한 영상 분할 기법이 필요하다. 이를 위하여 우리는 thresholding 기법을 연구하였다. 먼저 세포영상의 각 화소의 명암의 분포를 조사하여 히스토그램을 구하였다. 히스토그램은 0~255 사이에 존재하게 되는데, 0~255의 전 영역에 존재하기보다는 그중 일부분에만 존재한다. 우리는 히스토그램이 존재하는 영역을 백분율로 나누고 세포핵 및 세포질이 존재하는 영역의 분포를 구하여 global threshold를 찾았고, 이를 기준으로 각 점을 thresholding 할 때에 주위의 평균값을 보정값으로 두어 local thresholding을 수행하였다. 결과 영상은 핵의 영역을 탐색하기 위한 seed로 사용하기에 적합하다.

## 1. 서 론

자궁암이란 자궁 경부의 상피에 발생한 악성 종양이다. 자궁경부암 검사를 위하여 피험자의 자궁 경부에서 세포를 채취한 뒤, 이를 염색한 표본을 해부병리학 분야의 전문가가 현미경을 통해 세포와 핵의 형태를 육안으로 관찰함으로써 자궁암검사가 이루어진다[1].

본 연구의 목적은 자궁 경부에서 채취된 세포 영상을 분석하여, 피험자의 암 발병 여부를 자동으로 판별하는 시스템을 제작하는 것이다.

세포영상은 채취된 세포를 슬라이드 글라스에 펼쳐놓은 뒤 현미경과 카메라를 이용하여 얻어진다. 우리가 사용하는 영상은 그 배율이 100배 및 400배이며, 본 논문에서 제안한 방법은 100배의 영상에 적용되었다.

자궁경부암 세포영상은 배경, 세포질, 세포핵 및 염증세포로 분할할 수 있다. 그중 전문가가 영상을 분석할 때 중요하게 여기는 특징중의 하나는 세포핵 및 세포질의 특징이다. 그렇기 때문에, 영상분석의 처음 과정으로 영상을 배경, 세포핵 및 세포질로 분할하는 과정이 필요하다. 그러나, 세포영상 자체가 전문적인 지식이 없는 사람이 판별하기에는 어려울 정도로 복잡하며, 그 경계가 분명치 않기 때문에, 이것이 자궁경부암 세포영상분할 문제에 있어서 커다란 걸림돌로 작용한다.

영상의 분할 기법은 이미 기존의 연구가 많이 진행되어 있다[2,3]. 때문에, 기존의 방법들을 우리의 연구에 적용시킬 수 있다. 다만, 자궁경부암 검사 시스템이 요구하는 영상분할 기법은 짧은 시간동안 대단히 많은 영상을 처리할 수 있는 성능을 가지고 있어야 하기 때문에, 이 문제에 최적화 된 기법이 필요하다. 속도의 제약성 때문에 우리가 본 연구에서 고려한 영상분할 기법은 thresholding 기법이다. Thresholding 기법은 단순하면서도 빠르고 꽤 좋은 성능을 발휘하는 영상분할 기법이다[4,5].

2절에서는 관련 연구에 대해서 살펴보고, 3절에서는 우리가 제안하는 thresholding 기법에 대해서 살펴보고, 4절에서는 그 결과를 살펴본다.

## 2. 관련 연구

세포영상에서 특정한 세포핵이나 세포질 등의 객체를 추출하기 위하여 가장 먼저 해주어야 할 과정이 영상분할 과정이다. 영상분할 과정을 통하여 추출된 객체들로부터 특징 추출이 이루어지게 되므로, 영상분할은 매우 중요한 과정이라 할 수 있다.

영상분할 기법은 이미 기존에 많은 연구가 진행되어 있고, 많은 기법들이 존재한다. 대표적인 기법들로 들 수 있는 것이 thresholding 기법, edge based 기법, region growing 기법, split and merge 기법들이

1) 본 연구는 보건복지부의 선도기술·의료공학기술개발사업의 연구비 지원을 받았음

다[3].

이들 중에서 가장 기본이 되고 간단한 방법이 thresholding 기법이다. Thresholding 기법이란 영상을 특정한 threshold 값을 기준으로 그 값보다 큰 값을 가진 화소과, 그렇지 않은 화소로 나누어 분할하는 기법이다. 이 기법에서 중요한 부분은 threshold 값을 찾아내는 방법이며 이 방법에 따라 thresholding 기법은 여러 가지로 나눌 수 있다. 단순히 gray level 값을 이용한 thresholding 기법이나, 히스토그램을 이용한 방법, entropy를 이용한 방법들이 존재한다[4-6].

### 3. Global / Local Thresholding 기법

적당한 threshold 값을 찾기 위해 우리가 적용한 알고리즘은 다음과 같다. 첫째 히스토그램을 구한다. 둘째 히스토그램에 대하여 스무딩 과정을 거친다. 셋째, 초기 threshold값을 찾는다. 넷째, threshold값을 보정한다. 다섯째, erosion과 dilation 과정을 거쳐 최종 이미지를 얻어낸다. 각 과정에 대해서 자세히 살펴보겠다.

#### 3.1. Global Thresholding

##### 3.1.1. 히스토그램

영상은 화소로 이루어진 2차원 배열이며, 각 원소는 명암의 크기를 나타내는 값을 가지고 있다. 이 값의 분포를 조사하여 그래프로 나타낸 것이 히스토그램이다. 동일한 명암값이 많으면 그 부분의 히스토그램은 높게 나타나고, 그렇지 않을 때에는 낮게 나타난다. 세포 영상은 세포 핵, 세포질, 배경으로 나눌 수 있는데, 이상적인 경우의 히스토그램은 명암값이 세포핵, 세포질, 배경부분에 많이 분포하여 3개의 마루를 가진 모양을 하고 있다. 그러나, 우리가 사용하는 세포영상에서는 이러한 이상적인 히스토그램을 기대하기 어렵다.

##### 3.1.2. 히스토그램 평활화

히스토그램에서 3개의 마루가 나타나는 것이 이상적이지만, 실제 세포영상에서 히스토그램을 구해보면, 아주 작은 여러 개의 골과 마루가 나타난다. 따라서, 히스토그램을 직접 사용하기는 어려우며, 적당한 처리를 해주어 threshold 값을 찾기 쉽게 해주어야 한다. 우리가 적용한 방법은 히스토그램 평활화이다.

사용한 히스토그램 평활화 방법은 다음과 같다. 자신을 포함하여 이전의 2개와 이후의 2개를 합쳐서 모두 5개의 값에 weight를 주어 자신의 값을 재 할당해 준다.

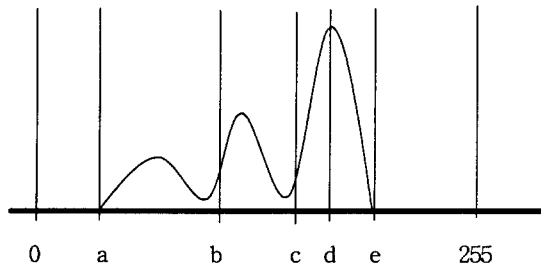
$$iB = (v_1 \times 1 + v_2 \times 2 + v_3 \times 3 + v_4 \times 2 + v_5 \times 1) / 9$$

##### 3.1.3. 초기 Threshold 값 찾기

초기 threshold 값을 찾기 위하여 여러 이미지를 분석해 본 결과 세포핵과 그 외의 것의 경계는 히스토그램에서 67%로 나타나고 세포질과 배경과의 경계는 92%에서 나타난다는 것을 알 수 있었다. 여기에서 백분율의 기준은 <그림 1>을 보면 알 수 있듯이, 히스토그램이 존재하는 구간에 대한 백분율이다.

통계적인 정보를 이용하여 우리는 67%와 92%를 기준으로 thresholding을 수행하기로 했다. 물론 이때 구해지는 threshold 값은 모든 이미지에 대해서 적절하지는 않으며, 때문에 과정 4에서

threshold값 보정을 거치게 된다.



a: 0% 지점 b: 세포핵을 위한 threshold 값  
c: 세포질을 위한 threshold 값  
d: 값이 가장 클 때의 luminance e: 100%지점  
<그림 1> 히스토그램

<그림 1>에 나타난 a와 e가 threshold 값을 찾을 때의 기준선이다. 그런데, 100배의 세포영상의 경우 대부분을 차지하는 영역이 배경이기 때문에, global threshold를 찾을 때 d를 100%로 놓고 사용할 수도 있다. 현재 실험에서는 a와 d를 기준으로 그 백분율을 구했다.

##### 3.1.4. 초기 Threshold 값 보정

세포 이미지에서 핵이나 세포질이 흡수하는 빛의 양은 거의 같으므로, 통계적으로 67%와 92%를 구한다. 그러나 모든 이미지에 대해서 정확한 threshold값을 얻어낼 수 없다. 세포핵, 세포질, 배경의 경계부분에서 골이 나타난다는 가정 하에서 우리는 과정 3에서 얻은 초기 threshold값을 기준으로 좌우로 10%씩 탐색공간을 잡아 threshold값을 변화 시켜 보아서 골을 탐색해내기로 했다. 골을 찾아내면 그 값을 threshold값으로 결정한다.

##### 3.1.5. Morphology Operator

3\*3 마스크 erosion과 dilation과정을 거쳐 최종적인 결과 영상을 얻어낸다. 이 방법은 좀더 부드러운 이미지를 얻어낼 수 있으며, hole filling효과와, 어느 정도의 artifact 제거가 가능하다.

### 3.2. Local Thresholding 기법

보통의 global thresholding 기법은 thresholding value  $t$ 를 구하는 것으로서  $t$ 를 영상 전체에 똑같이 적용하게 된다. 그러나, 이 방법을 그냥 적용하기에는 세포영상의 gray level의 변화가 영역별로 많으므로, 지역적인 정보를 적용한 thresholding 기법이 요구된다. 우리가 local thresholding 기법을 적용하기 위해서 사용한 값들은 영상 전체의 그레이 레벨 평균과 현재  $t$ 를 적용중인 화소근처의 그레이레벨의 평균이다.

$$Th_{New} = Th_{Global} + (AvgWindow - AvgGlobal) \times \alpha$$

- $Th_{global}$  : global thresholding value
- $AvgGlobal$  : 영상 전체의 그레이 레벨의 평균
- $AvgWindow$  : 현재 화소 주위의 사각 띠의 평균
- $Th_{New}$  : local thresholding value
- $\alpha$  : 비율

구하는 방법은 다음과 같다. 먼저 AvgGlobal과 Thglobal을 구한다. AvgGlobal은 세포영상 전체 화소의 그레이 값의 평균을 구하면 되고, ThGlobal은 기존의 thresholding 기법을 이용하여 구한다. AvgWindow는 현재 화소를 중심으로 사각 띠를 먼저 구하는데, 일정 화소 밖의 화소들에 대해서 평균을 구한다.

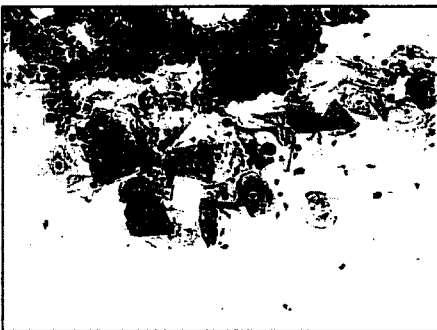
이 local thresholding 기법을 적용하면 특히 세포질과 세포핵이 일 부분에 집중되어 있는 경우, 그 영역에 대해서 핵을 찾아내는데 좋은 성능을 보여준다.

#### 4. 실험 및 결과 분석

<그림 2>는 실험에 적용했던 영상 중의 하나이다. 위 부분에 뭉쳐진 세포핵들이 보이고, 영상의 가운데에 다른 종류의 세포들이 존재하는 것이 보인다. 그리고, 대부분의 영역이 배경임을 알 수 있다.



<그림 2> 입력 세포 영상



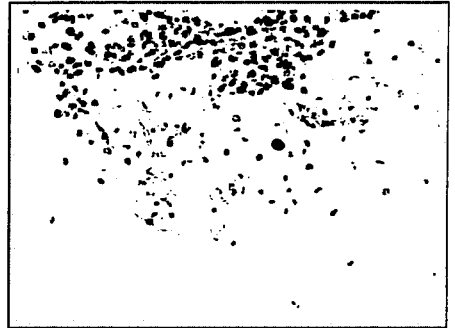
<그림 3> Global thresholding 결과 (핵과 세포질)

먼저 global thresholding 알고리즘을 100배의 세포 영상에 적용시켜보았다. 검은색이 세포핵, 회색이 세포질 영역이다. 세포핵이 뭉쳐진 부분이 잘 분할되지 않았음을 볼 수 있다. 이것은 global thresholding 이 가진 근본적인 한계점이며, 사람이 눈으로 보고 thresholding 할 때에도 나타나는 현상이다.

<그림 3>은 local thresholding 알고리즘을 적용한 뒤의 결과 영상이다. local thresholding은 오직 세포핵에 대해서만 적용시켜 보았다. global thresholding에 비하여 더 좋은 결과를 얻을 수 있음을 알 수 있다. local thresholding은 global thresholding에 비하여 연산속도가 그리 느리지 않으면서도 더 좋은 결과를 얻을 수 있다.

실험 결과 전체적으로 세포영상은 global thresholding 기법만으로

는 분할하기 힘들다는 것을 알 수 있었다. 이것은 명암의 분포가 불균일한에서 오는 것으로 local thresholding 기법을 적용하여 약간 더 향상된 결과 영상을 얻을 수 있었다.



<그림 4> Local thresholding 결과 (핵 영역)

#### 5. 결론 및 향후연구

Thresholding 은 매우 간단한 영상분석 도구이다. 우리는 세포영상을 분할하기 위하여 thresholding 알고리즘을 적용시켜보았으며, 그 결과를 살펴보았다. Global thresholding 및 local thresholding 알고리즘에 대한 실험 결과, thresholding 알고리즘은 속도가 빨라 검사시스템을 제작하기에 적합함을 알 수 있었다. 결과 영상은 핵의 영역을 탐색하기 위한 seed로 사용될 수 있다. 앞으로 결과 영상을 이용한 여타의 다른 방법에 대한 연구가 필요하다고 여겨진다. Thresholding 기법에 대해서는 선형적이지 않은 threshold 평면을 가진 thresholding 기법에 대한 연구가 필요하다.

#### 참고문헌

- [1] 은성경, 박찬모, 박화순, 윤소영, 조민선, 조수연, 김성숙, "영상처리를 이용한 자궁경부 세포진의 자동탐색 방법에 관한 연구," 대한세포병리학회지, 제5권, 제1호, pp. 15-22, 1994.
- [2] Kenong Wu, David Gauthier and Martin D. Levine, "Live Cell Image Segmentation," IEEE Transactions on Biomedical Engineering, vol. 42, no. 1, pp. 1-11, 1995.
- [3] Robert M. Haralick and Linda G. Shapiro, "SURVEY : Image Segmentation Techniques," Computer Vision, Graphics, and Image Processing 29, pp. 100-132, Academic Press, 1985.
- [4] Calum MacAulay and Branko Palcic, "A Comparison of Some Quick and Simple Threshold Selection Methods for Stained Cells," Analyt Quant Cytol Histol, Vol. 10, No. 2, pp. 134-138, 1988.
- [5] Joan S. Weszka, "SURVEY : A Survey of Threshold Selection Techniques," Computer Graphics and Image Processing 7, pp. 259-265, Academic Press, 1978.
- [6] J. R. Parker, Algorithms for Image Processing and Computer vision, John Wiley & Sons, 1997.