DOI: 10.3745/KIPSTB.2010.17B.3.239

중요도 맵과 단계적 영역병합을 이용한 백혈구 분할

김 자 원[†] · 고 병 철[††] · 남 재 열[†††]

요 약

혈액 세포 영상에서 백혈구는 환자의 건강상태를 파악하는데 중요한 정보를 제공하며, 이를 통해 다양한 질병을 조기에 예측할 수 있다. 따라서 질병의 조기 예측을 위해 혈액세포에서 백혈구의 분리는 매우 중요한 단계이다. 본 논문에서는 중요도 맵과 단계적 영역 병합을 이용하여 혈액 세포 영상에서 백혈구를 자동으로 분할하는 기법을 제안한다. 백혈구 세포 영역은 염색물질에 의해 주변에 비해 두드러진 색상, 질감 정보를 가짐으로 이를 기반으로 중요도 맵(Saliency Map)을 만든다. 이를 통해 세포 영상에서 두드러진 영역을 찾아 sub-image를 분리하고, 각 sub-image에서 mean-shift 알고리즘을 적용하여 영역 클러스터링을 수행한다. Mean-shift 적용 후 얻은 클러스터들에 대해 단계적 영역 병합 알고리즘을 적용하고, 최종적으로 백혈구 핵으로 판단되는 단일 클러스터를 얻을 수 있다. 본 논문에서 제안한 방법은 혈액 세포 영상을 사용하여 테스트한 결과 71%의 핵 검출 정확도를 보였으며, 기존의 다른 알고리즘보다 뛰어난 성능을 나타내었다.

키워드 : 중요도 맵, Mean-shift, 단계적 영역 병합, 백혈구 분할

Leukocyte Segmentation using Saliency Map and Stepwise Region-merging

Ja-Won Gim[†] · ByoungChul Ko[††] · Jae-Yeal Nam[†††]

ABSTRACT

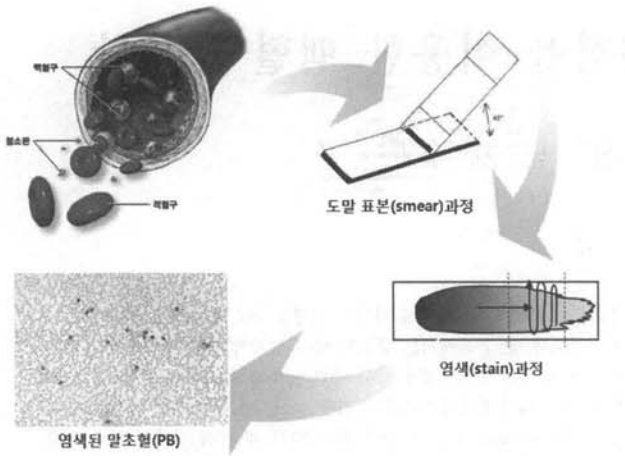
Leukocyte in blood smear image provides significant information to doctors for diagnosis of patient health status. Therefore, it is necessary step to separate leukocyte from blood smear image among various blood cells for early disease prediction. In this paper, we present a saliency map and stepwise region merging based leukocyte segmentation method. Since leukocyte region has salient color and texture, we create a saliency map using these feature map. Saliency map is used for sub-image separation. Then, clustering is performed on each sub-image using mean-shift. After mean-shift is applied, stepwise region-merging is applied to particle clusters to obtain final leukocyte nucleus. The experimental results show that our system can indeed improve segmentation performance compared to previous researches with average accuracy rate of 71%.

Keywords : Saliency Map, Mean-shift, Stepwise Region-merging, Leukocyte Segmentation

1. 서 론

혈액 세포 영상에서 백혈구의 비율은 환자의 건강상태를 파악하는데 중요한 정보를 제공하며, 이를 통해 백혈병과 같은 질병을 초기에 예측할 수 있다[1]. 혈액을 이용한 질병 예측을 위해 개발된 자동혈구 측정검사(Automatic Complete Bood Cell Count: ACBC) 방법은 임상검사실에서 시행하는 다빈도 혈액 검사의 하나로 각종 혈구 세포들의 수를 측정하여 질병 진단에 필요한 기본적인 정보를 제공한다. 일반적으로 현재의 혈액관련 질병진단 과정은 (그림 1)과 같이 혈액으로부터 도말표본과정과 염색과정을 거쳐 혈액 세포 영상을 생성하고 자동혈구측정검사 항목 중 백혈구감별계수(White Blood Cell Differential Count)를 통해 전체 백혈구 수에 대한 호중구(Neutrophil), 호산구(Eosinophil), 호염기구(Basophil), 단핵구(Monocyte), 림프구(Lymphocyte)의 분포율을 구함으로써 여러 병증을 판단하는 기초정보로 사용한다. 대부분의 백혈구 감별계수 검사는 자동혈구분석기를 통해 이루어지고 있으면 이 분석기는 레이저 다각도 편광산란 분리법과 형광 측정법의 원리를 이용하여 혈구를 측정한다. 하지만, 백혈구나 혈소판 관련 분석 결과는 현미경을 통한 말초혈액도말표본의 육안 관찰이 필요한 경우가 많으며, 실제로 자동혈구분석기에서 보통 전체 검사의 20% 이상이 불감지결과로 나타나며 이러한 경우 별도 확인 작업의 대상이 된다 [2].

※ 본 연구는 지식경제부 지방기술혁신사업(RTI04-01-01) 지원으로 수행되었습니다.
† 준 회 원 : 계명대학교 컴퓨터공학과 석사과정
†† 정 회 원 : 계명대학교 컴퓨터공학과 조교수(교신저자)
††† 정 회 원 : 계명대학교 컴퓨터공학과 교수
논문접수 : 2010년 2월 5일
수 정 일 : 1차 2010년 3월 19일
심사완료 : 2010년 3월 19일



(그림 1) 백혈구 분석을 위한 말초혈액 염색 과정

이 경우 혈액 세포내에 포함된 백혈구의 개수를 세는 작업은 병리학자에 의해 직접 수행되며, 1개의 혈액 세포 영상에서 모든 백혈구를 찾아서 각각의 백혈구의 종류를 판단하고, 각 종류의 백혈구의 수를 세는 절차로 이루어진다[3]. 하지만, 환자 한 명의 진단을 내리기 위해서는 수백여장에 달하는 혈액 세포 영상의 분석이 필요하며, 이를 사람이 수작업으로 하는 것은 많이 시간이 소요될 뿐 아니라 분석가의 경험 정도에 따라 정확도가 낮을 수 있다. 이를 효율적으로 개선하기 위해 혈액세포 영상 내에서 백혈구를 자동으로 분리해 내는 연구가 진행되고 있다[1],[3-7].

Liao[1] 등은 스케일 스페이스 필터링 (Scale-Space filtering)과 워터쉐드(Watershed) 클러스터링을 통한 새로운 세포 분할 방법을 제안했다. 이 방법은, red, green, blue 색상 중 green 색상의 히스토그램에 스케일 스페이스 필터를 적용하고, 이를 통해 히스토그램 계곡 점의 임계값을 얻고, 이 값을 통해 이미지를 이진화하여 핵을 검출하였다. 또한, 워터쉐드 알고리즘을 적용하여 백혈구의 세포질을 분할하여, 90%이상의 높은 분할 정확도를 보였다. 하지만 이 경우, 하나의 색에 의존하여 분할을 수행하였기 때문에 백혈구의 핵을 다른 색으로 염색할 경우에는 정확도가 낮아진다.

Dorini[4] 등은 Scale-Space Toggle 연산자를 사용하여 백혈구 이미지를 단순화 한 후, 모폴로지 침식 연산과 워터쉐드 방법을 이용해 핵을 검출해냈다. 그리고, 적혈구보다 큰 디스크모양의 구조 요소를 사용한 모폴로지컬 오프닝 연산과 이전에 분할한 핵 영역을 종합하여, 대략적인 모양의 세포질을 검출했다. 하지만, 이 알고리즘의 경우, 핵과 세포질 검출시에 그레이 영상을 이진 영상으로 변환하여 사용하기 때문에 분할 정확도가 낮을 것으로 예상되며, 논문에서 성능 평가 결과를 언급하지 않고 있다.

Herbert[5] 등은 K-means 클러스터링을 사용하여 혈액 세포 영상에서 백혈구를 분할했다. 이들이 제안한 방법에서는 수작업으로 분할한 백혈구와 적혈구의 색조, 채도, 명암도 값의 분포를 이용하여 모든 픽셀의 백혈구 확률을 계산하고, 이를 백혈구 분할에 이용하여 정확도를 높였다. 하지만, 이 방법에서는 초기 백혈구의 위치를 수작업에 의해 선택해주어야 하는 문제점이 있다.

Nipon[6]은 백혈구의 그레이 이미지에 Fuzzy C-means 클러스터링 알고리즘을 적용하여 핵과 세포질을 각각 분할했다. 하지만, 오직 그레이 이미지에 대해서만 분할작업을 수행하였으며, 이 때문에 핵과 세포질의 명암도가 비슷한 경우, 에러율이 높다는 단점이 있다.

Mira[7, 11] 등은 세포 영상에서 군집세포와 독립세포를 분류하고, 군집세포를 각각의 단일 세포로 분리하였다. 여기서는, 가우시안혼합모델과 최대우도함수를 사용하여 입력 영상을 이진화하고, 전경 영역에 대해 형태학적 특징을 추출하여 독립세포와 군집세포를 분류하였다. 군집세포의 경우, 명암도 기울기 변환 영상과 워터쉐드 알고리즘을 사용하여 군집세포 영역을 작은 영역으로 나누고, 세포 단위의 마커를 추출하여 마커를 중심으로 단계적 병합 알고리즘을 적용하여 최종적으로 각각의 단일 세포로 성공적으로 분리하였다. 하지만 이 알고리즘의 경우, FISH 세포분할을 위해 설계된 알고리즘이므로 혈액 세포 영상 내의 백혈구 세포에 대해서는 낮은 감지율을 보여준다.

위에서 살펴본 관련 연구들은 백혈구를 비교적 효과적으로 분리하고 있지만, 전체 영상을 대상으로 분할 작업을 실시함으로 적혈구나 혈소판과 같은 영역에 대해서도 분할 작업이 이루어져 속도 면에서나 정확도 면에서 성능이 떨어지는 단점이 있다. 특히, 혈액 세포 영상은 크기가 일반 영상에 비해 매우 크므로 실시간 처리를 위해서는 모든 세포 영역을 대상으로 하지 않고 목표가 되는 백혈구 영역만을 찾아내고 이 영역에서만 핵을 분할해 내는 알고리즘이 필요하다.

본 논문에서는, 이러한 단점을 해결하기 위해 백혈구 세포가 가지는 특성을 이용하여 세포 후보 영역을 결정하고 이 후보 영역 안에서만 분할 작업을 실시한다. 우선, 백혈구 세포 영역은 염색물질에 의해 주변에 비해 두드러진 색상, 질감 정보를 가짐으로 이를 기반으로 중요도 맵(Saliency Map)을 만든다. 이후 중요도 맵을 통해 세포 영상에서 두드러진 영역을 찾아 백혈구 세포 후보 영역으로 판단하고 이를 sub-image로 분리한다. 분리한 sub-image에서 mean-shift를 수행하고, 이를 통해 얻은 결과 이미지에서 영역 클러스터간의 병합을 수행 하고 최종적으로 병합된 클러스터를 백혈구 핵이라고 판단한다.

본 연구의 구성은 다음과 같다. 2장에서는 입력된 혈액 세포 영상으로부터 색상, 질감 특징을 추출하여 중요도 맵을 생성하고, 이를 통해 전체 혈액 세포 영상에서 백혈구가 존재하는 영역인 sub-image만을 분리한다. 분리된 sub-image별로 mean-shift 알고리즘을 적용하여 각 영역별로 클러스터링을 수행한다. 3장에서는 각 클러스터의 중요도 맵 평균값, 영상 공간에서의 특징 등을 고려하여 클러스터간의 단계적 영역 병합을 수행하고 최종적으로 백혈구 핵이라고 판단되는 영역을 검출한다. 또한, 4장에서는 기존의 다른 백혈구 분할 알고리즘과의 비교실험을 통해 본 논문에서 제안한 기법의 우수성을 입증하고, 마지막으로 5장에서는 제안된 기법에 대한 전체적인 정리와 함께 본 논문을 마무리한다.

2. 중요도 맵과 Mean-Shift를 이용한 백혈구 세포 영상 분할

2.1 중요도 맵 생성

염색된 세포 영상에서 백혈구 세포 영역은 주변의 적혈구 세포 및 혈소판에 비해 영역의 크기가 크며, 자주색을 가지고, 분화단계와 백혈구의 종류에 따라 1개 혹은 여러 개의 핵으로 분리되어 나타난다. 이러한 백혈구 세포의 특징을 이용하여 세포내의 백혈구 영역만을 분리해 내기 위해 본 논문에서는 Ko[8] 등이 제안한 중요도 맵(Saliency map)을 이용하여 후보 백혈구 영역을 검출한다.

중요도 맵은 색상, 밝기, 방향성에 기반을 두고, 주변 픽셀과 중앙 픽셀의 값 차이를 통해 이미지에서 두드러진 부분을 찾아내는 알고리즘이다. 하지만, 실제 혈액 세포 영상에서는 백혈구 이외에 적혈구도 배경과 두드러지는 밝기 값을 가지므로 본 논문에서는 밝기성분을 중요도 맵 생성에서 제외하였다. 본 장에서는 이 방법을 응용하여, 혈액 세포 영상에서 백혈구를 포함하는 sub-image를 분리해내며, 또한 mean-shift 후 영역 병합 단계에서 병합을 위한 최소 임계값으로 사용한다.

2.1.1 색상 특징 맵

먼저, 색상 특징 맵을 생성하기 위해 인간이 지각하는 색상과 유사한 색상모델인 CIE L*a*b*를 사용한다. 여기서는 a*, b* 색상 모델만 사용하며, L*은 밝기 특징 맵 생성에 쓰이지만, 본 본 논문에서는 밝기 성분을 중요도 맵 생성에서 제외하였다. 백혈구 핵 영역은 적혈구에 비해 두드러지게 red와 blue 색상이 강함으로 a*b*를 사용함으로써 대비 정도가 강한 색상 특징 맵을 얻을 수 있다. 각 a*, b* 이미지는 원본 세포 영상의 1/2크기로 다운 샘플링 한 후, 각 이미지에 11× 11, 13× 13의 두 필터를 적용한다. 필터의 크기는 본 논문에서 사용한 400X300 크기의 영상을 대상으로 실험한 결과에서 가장 좋은 성능을 보여준 값으로 결정했으며, 영상의 크기가 1/2로 줄거나 2배 커질 경우 필터의 크기도 영상의 크기에 비례하여 각각 1/2과 2배로 크게 적용하면 된다.

각 필터에서 중심점과 주변 점들과의 차이를 통해, 각색상의 특징 맵인 $C(c,s)$를 구성한다. 주변 점들과의 차이가 클수록 색상 맵에서 밝게 나타난다.

$$\overline{C} = \frac{1}{4}\left(\sum_{c\in\{a^*,b^*\}}, \sum_{s\in\{11\times 11, 13\times 13\}} C(c,s) \right) \tag{1}$$

각 색상 a*, b*에 대해 각각의 필터 11× 11, 13× 13을 적용한 4개의 색상 특징 맵 $C(c,s)$의 합을 정규화하여 최종적인 색상 특징 맵($\overline{C}$)을 생성한다.

2.1.2 방향 특징 맵

이미지에서 두드러진 영역은 두드러진 질감 정보를 가지고 있다. 질감 정보를 통해 객체를 식별하기 위해 웨이블릿 변환을 이용하고, 이를 통해 방향 특징 맵을 구성할 수 있다. 1단계 웨이블릿 변환 후, 수평방향(LH), 수직방향(HL), 대각방향(HH)의 방향성을 가지는 3개의 방향 이미지를 얻을 수 있다.

$$\overline{O} = \frac{1}{6}\left(\sum_{c\in\{HH,HL,LH\}} \sum_{s\in\{11\times 11, 13\times 13\}} O(c,s) \right) \tag{2}$$

세 개의 방향 이미지에 대해 2.1.1에서와 마찬가지로 11× 11, 13× 13의 두 필터를 적용하여 주변 점들과의 차이를 통해 각 방향 이미지의 방향 특징 맵($O(c,s)$)을 구성한다. 이렇게 구해진 6개의 방향 특징 맵을 합하고 정규화하여 최종적인 방향 특징 맵($\overline{O}$)을 얻을 수 있다.

2.1.3 특징 맵 결합

(그림 2)(b)부터 (그림 2)(c)와 같이 입력 영상에 대해 생성된 색상, 방향 특징 맵은 하나의 (그림 2)(d)와 같이 중요 특징 맵(Saliency map)으로 결합되어야 한다. 하지만, 영상마다 백혈구의 크기 및 염색 색상 정도가 다르기 때문에 각각의 특징 맵은 다른 가중치가 적용 되어야 한다. 그림 2에서 보는 것과 같이 색상 특징 맵이 배경과 가장 두드러진 특징을 보이므로, 본 논문에서는 각 맵의 밝기 비율에 따라 가중치를 0.2, 0.8로 결정하고 이를 식(3)에 적용하여 최종적인 중요도 맵 C_m를 생성하였다.

$$C_m = \overline{O}(i,j)\times w_1 + \overline{C}(i,j)\times w_2 \tag{3}$$

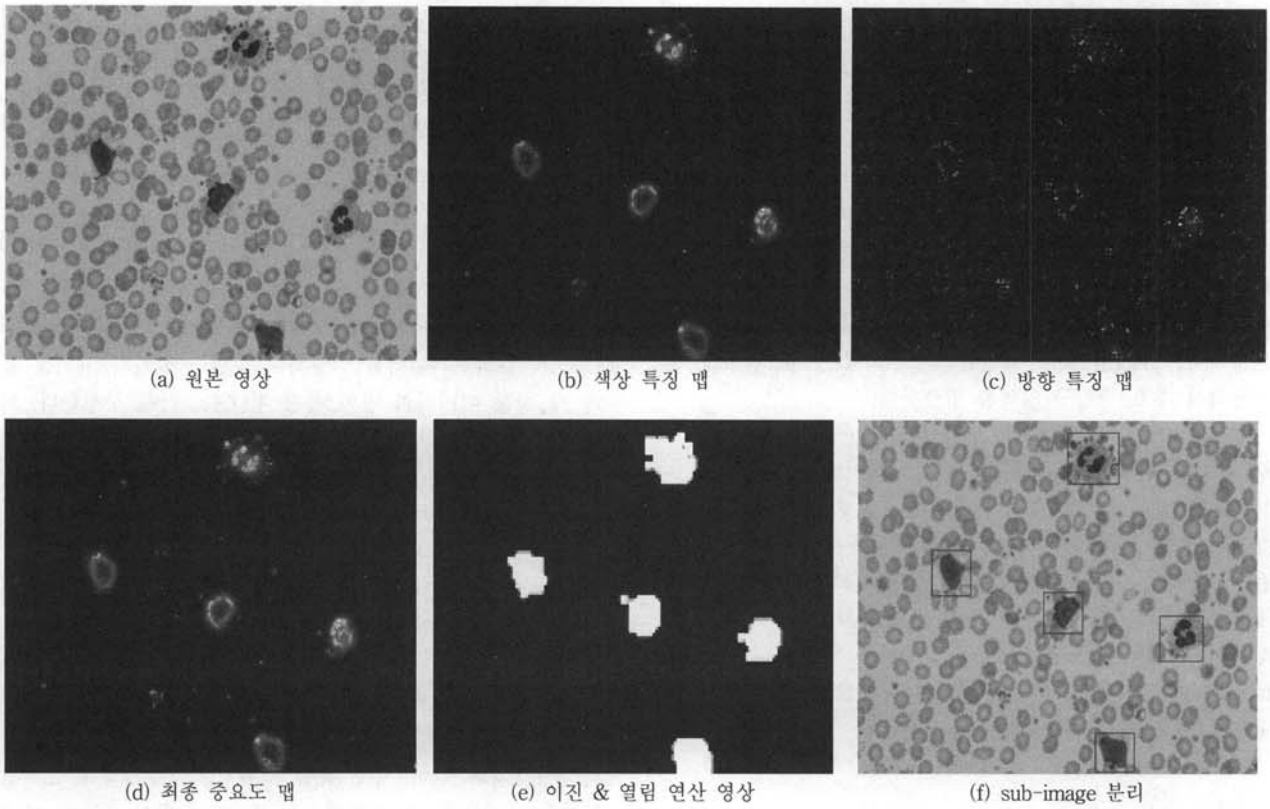
2.1.4 Sub-image 분리

일반적으로, 혈액 세포의 영상에서 백혈구는 적혈구에 비해 상대적으로 크지만, 경우에 따라 영상에 단 한 개의 백혈구도 포함하고 있지 않은 경우도 존재한다. 이러한 영상에 대해서 핵 추출의 절차를 모두 수행하게 되면, 효율성이 떨어지므로 혈액 세포 영상에 대해 후보 백혈구 영역을 먼저 찾고, 후보 백혈구 영역이 존재하는 영상에 대해서만 후속 과정을 수행하는 것이 효율적이다.

본 논문에서는 sub-image의 분리를 위해서 중요도 맵과 Otsu[9]의 이진화 알고리즘을 사용했다. 중요도 맵에서 밝은 색으로 나타나는 부분일수록 다른 영역에 비해 두드러지는 영역이 된다. 따라서 중요도 맵에 대해 Otsu의 이진화 알고리즘을 적용하여 sub-image로 분리할 부분과 배경으로 간주할 부분을 결정한다. 이진화 된 영상에서 모폴로지 열림(opening) 연산을 수행하여 잡음을 제거하고 이후, 영역 레이블링 과정에서 (그림 2)(e)와 같이 작은 영역들은 제거한다. 최종적으로 남은 영역에 대해 X-Y 투영을 실시하여 각각의 영역들에 대한 최소 sub-image를 얻게 된다.

2.2 Mean-Shift를 이용한 영역 분할

입력 영상에 대해 sub-image가 분리되고 나면 sub-image



(a) 원본 영상 (b) 색상 특징 맵 (c) 방향 특징 맵

(d) 최종 중요도 맵 (e) 이진 & 열림 연산 영상 (f) sub-image 분리

(그림 2) 원본 영상 (a)에 대한 2가지의 특징 맵(b, c)을 이용해 생성한 중요도 맵(d)과 Otsu 알고리즘으로 이진, 모폴로지 열림 연산 후 영상(e), 중요도 맵으로 sub-image를 선택(e)한 예

영역에 대해서만 mean-shift가 적용된다. Herbert[5]과 Nipon[6]은 백혈구 영역 분할을 위해 K-means 클러스터링과 Fuzzy C-mean 클러스터링 방법을 사용하고 있지만, 클러스터의 개수를 의미하는 K와 C값이 백혈구의 염색정도에 따라 매우 달라 질 수 있으므로 정확한 백혈구 분할을 어렵게 만드는 요인이 된다.

Mean-shift[10]는 비모수적 클러스터링 기술로 클러스터의 수에 대한 사전 지식이 없이도 클러스터링이 가능하다. 또한, 클러스터의 모양에 구애 받지 않기 때문에, 모양이 일정하지 않은 백혈구 세포의 분할에 적합하다. 영상 분할을 위한 mean-Shift 알고리즘은 색상 공간에서 영상 픽셀의 색상 분포를 나타내는 확률밀도함수의 지역적 극대값(local maxima) 찾는 방법이다. 먼저, 주어진 공간 커널 안에서 색상 값들 중 현재 커널의 중심 색상과 유사한 색상분포를 갖는 픽셀들의 평균위치와 색상공간에서의 평균값을 식(4)를 이용하여 계산한다. 수식 (4)에서 평균값 m_h는 결국 이동할 mean-shift 벡터를 의미한다.

$$m_h(x^t)=\frac{\sum_{i=1}^{n} x_i g\left(\left\| \frac{x^t-x_i}{h} \right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\| \frac{x^t-x_i}{h} \right\|^2\right)}-x^t \qquad (4)$$

여기서, x는 커널의 중심을 의미하고 x_i는 커널내의 한 점, h는 커널 윈도우의 크기, g는 커널 함수를 의미한다.

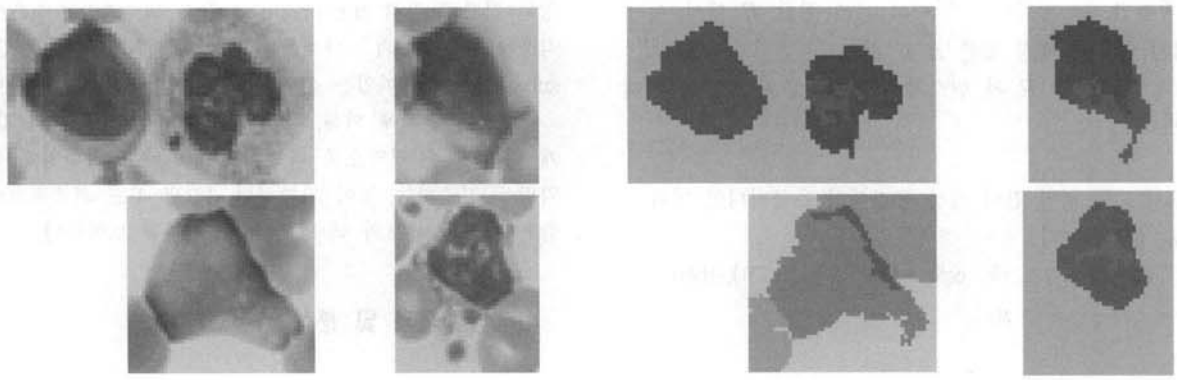
이후, 수식(5)를 이용하여 새로운 점 x^{t+1}로 커널을 이동하여 다시 같은 작업을 시행하고 위와 같은 작업을 윈도우 커널 내에 극대 값이 존재하는 점으로 수렴할 때까지 반복한다.

$$x^{t+1}=x^t+m_h(x^t) \qquad (5)$$

이 극대점에서의 색상 값을 현재 공간위치의 픽셀 값으로 변경하면 공간영역에서 컬러 값을 균일하게 하는 영상 분할 효과를 낸다.

본 논문에서는 sub-image에 적합한 mean-shift의 수행을 위해 영상 공간을 위한 커널 크기는 20, 컬러 공간에서의 커널 크기는 15, 그리고 한 클러스터 당 최소 영역 크기를 25로 설정하고 제안된 논문[10]에서 사용한 L*u*v* 컬러 공간에서 mean-shift를 수행하였다.

(그림 3)(b)와 같이, mean-shift는 혈액 세포 영상에서 각각의 색상 값에 따라 해당 색상의 중심점을 기준으로 하나의 클러스터로 결정한다. 하지만 sub-image 안에는 세포질과 배경 및 적혈구가 일부분 포함되어 있으며, 세포핵 또한 염색정도에 따라 여러 영역으로 나뉘어져 나타나므로 이들을 구분하고 병합 할 수 있는 추가적인 알고리즘이 필요하다.



(a) 원본 영상 (b) mean-shift 적용 영상

(그림 3) Sub-image에 대해 mean-shift를 적용한 예

3. 백혈구 핵 영역 병합

Sub-image 영역 안에서 mean-shift에 의한 영역 분할 작업이 이루어진 후, 각 분할된 영역으로부터 배경과 적혈구 및 세포질 영역을 제거하고 백혈구 핵 영역만을 추출하기 위한 알고리즘이 필요하다. 이를 위해 본 논문에서는 이전에 생성된 중요도 맵과 컬러 공간상의 거리를 이용하여 단계적 영역 병합 알고리즘을 제안한다. 단계적 영역 병합 알고리즘은 두 단계로 구성된다. 첫 번째 단계에서는 분할된 영역들 중 영역병합의 기준으로 사용될 Seed 영역 집합을 선정하는 작업을 진행한다. 두 번째 단계에서는 완성된 Seed 영역 집합 주변 영역들을 대상으로 병합조건을 테스트하고 조건을 만족하면 핵 영역으로 병합하고 그렇지 않을 경우 배경으로 병합한다. 이 과정에 대한 자세한 설명은 다음과 같다.

3.1 Seed 영역 집합 결정 단계

$S=\varnothing$ (S : *seed 집합*)

$R=\{r_1 \cdots r_n\}$ (R : *분할된 영역 집합*)

[단계 1] 중요도 맵 C_m와 R 의 한 영역 r_i의 픽셀 수 N_r을 이용해 각 영역 r_i에 대한 중요도 맵 값의 평균(μ_r) 계산

$$\mu_r = \frac{1}{N_r}\sum_{i\in r} C_{m_i} \tag{6}$$

[단계 2] 각 영역의 중요도 맵 평균값에 따라 조건 (7)을 만족하는 집합 R 의 N개의 영역을 Seed 영역으로 선정하고 r_i는 집합 R 에서 제거

$$\begin{aligned} &\text{If } (\mu_r > th_1), \\ &\text{then, } \{S \ni r_r,\ r_i \notin R\} \end{aligned} \tag{7}$$

[단계 3] 만약 Seed 집합 S 에 속하는 영역이 없을 경우, sub-image의 중심점과 R 의 원소 영역들 간의 거리가 최소이고 일정크기 이상인 영역을 Seed영역으로 선정

Seed영역이 하나도 선정되지 못했을 경우 최소 근접한 영역을 Seed 영역으로 선정

C_O_S : sub-image의 중심

$SDist$: C_O_S와 영역 중심 간의 유클리디언 거리

$Ascen_Dist=\{r_i,\ldots,r_n\}$: $SDist$ 값이 작은 순서로 재 정렬된 영역집합

```
i=0;
while (i<n && S=∅)
{
  If(Ascen_Dist[i].Size > th2)                    (8)
  then, {S ∋ Ascen_Dist[i], Ascen_Dist[i] ∉ R}
  i++;
}
if(S=∅) S ∋ Ascen_Dist[0];
```

3.2 핵 영역 병합 단계

$S=\{s_1 \cdots s_m\},\ R=\{r_1 \cdots r_n\},\ S\cap R=\varnothing$

$B=\varnothing$: 배경 영역집합

[단계 1] if($S=\varnothing$) Stop
else Next Step

[단계 2] R 집합의 한 영역 r_i가 sub-image의 가장자리와 만나는 길이를 B_{r_i}라 했을 때, 조건에 따라 r_i를 배경집합으로 선정하고 집합 R 에서 제거

N_{r_i}: 영역r_i의 전체 픽셀 수

If $(B_{r_i}/N_{r_i} > th_3)$ then,

$\{B \ni r_i,\ r_i \notin R\}$

[단계 3] S 의 s_i와 R 집합의 한 영역 r_i가 아래의 조건

을 만족하면, 두 영역을 병합하고 r_i는 집합 R 에서 제거, 만족하지 않으면 배경 집합 B 로 선정

이 과정을 집합 R 의 원소가 더 이상 존재하지 않을 때까지 반복

$CDist$: 두 영역 간의 색상 공간상의 유클리디언 거리
while ($R \neq \{\varnothing\}$)
If ($CDist(s_i, r_i) \leq th_4$ && $SDist(s_i, r_i) \leq 0$), then,
$\{s_i = s_i \cup r_i \ , \ r_i \notin R\}$
else
$\{B \ni r_i, \ r_i \notin R\}$

[단계 4] 최종 S 집합의 원소를 백혈구 핵으로 선언

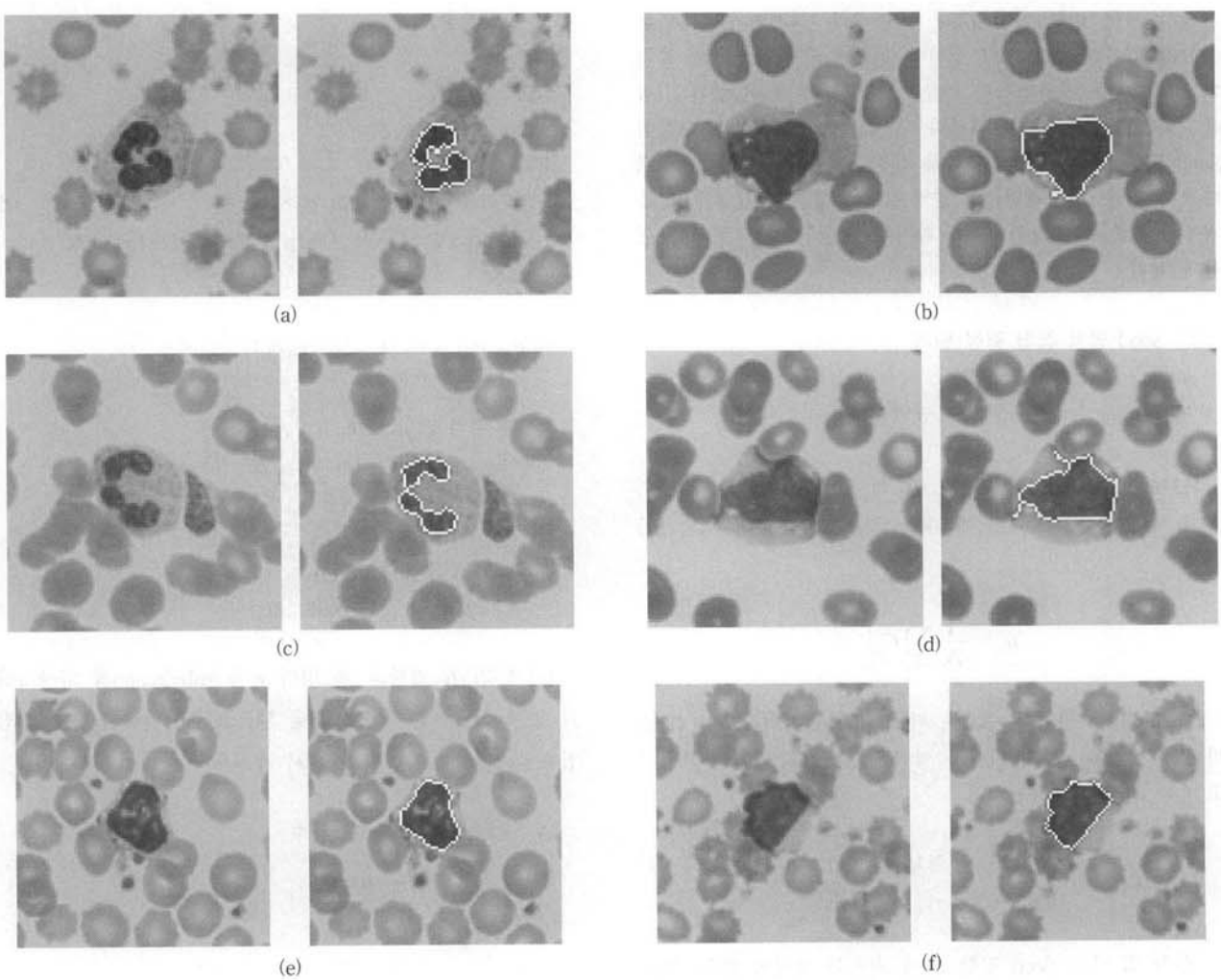
병합단계 중 초기 seed 선택에서 사용된 임계값 th_1은 실험적으로 150으로 설정되었으며 th_2는 sub-image 크기의 5%로 역시 실험적으로 결정된 값이다. 이 두 값은 영상의 크기 비율이나 중요도 맵 생성 시 사용되는 필터의 크기에 따라 비례하여 달라질 수 있다. 본 논문에서 th_3와 th_4는 고정된 값으로 각각 sub-image 크기의 5%, 그리고 10000으로 적용하였다. 위의 단계를 거쳐 영역 병합이 끝나면 sub-image에 남아있는 seed를 최종적인 핵이라고 판단한다.

위 병합 단계에 의해, 핵 영역이 (그림 4)(a),(c)와 같이 2개 이상의 부 영역으로 나누어져 있더라도 정확하게 핵 영역만을 검출하는 것이 가능하다. (그림 4)는 최종적으로 병합된 백혈구 세포의 핵 영역 병합 결과를 보여준다.

4. 실험 결과 및 분석

본 논문에서 제안하는 방법의 실험은 펜티엄 PC (CPU: dual core 3.40GHz)와 Windows 기반의 Visual Studio 2008을 이용하여 수행하였다. 실험은 연세대 세브란스병원 진단의학과에서 제공한 800× 600 사이즈의 말초 혈액 영상 표본 데이터를 사용하였다. 실험 영상 안에는 87개의 백혈구 세포를 포함한다. 본 논문에서는 영상 분할의 속도를 높이기 위해 원본영상을 400× 300 크기로 줄여 사용하였다.

먼저, 제안 기법의 sub-image 분리 정확도를 측정하기 위해, 다음과 같이 4가지의 분류 성능 측정방법을 이용하여

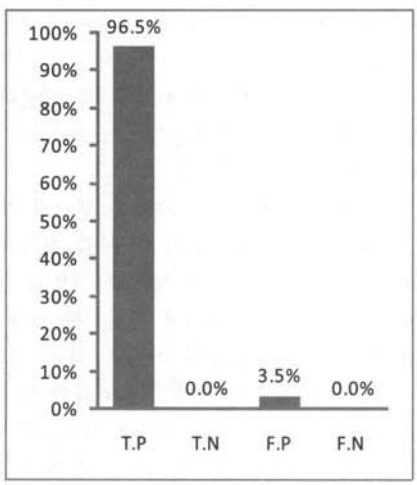


(그림 4) 백혈구 세포 핵 영역 병합 결과 예 (왼쪽: 원본, 오른쪽: 병합결과)

테스트 하였다.

- True Positive(TP) : 실제 백혈구 영역을 실험에서도 백혈구 영역으로 분리하는 경우
- True Negative(TN) : 실제 백혈구가 아닌 영역을 실험에서도 백혈구가 아닌 영역으로 분리하는 경우
- False Positive(FP) : 실제 백혈구가 아닌 영역을 실험에서는 백혈구 영역으로 분리하는 경우
- False Negative(FN) : 실제 백혈구 영역을 실험에서는 검출하지 못한 경우

(그림 5)에서 보는 바와 같이 True Positive(TP)는 96.5%, True Negative(TN)는 0%, False Positive(FP)는 3.5%, False Negative(FN)는 0%로 나타났다. 특히 True Positive의 경우, 95%를 넘는 우수한 정확도를 보여주었으며, 백혈구가 아닌 영역을 sub-image로 분리한 경우를 의미하는 False Positive(FP)가 3.5%로 오검출율이 매우 낮고, 백혈구 영역을 sub-image로 분리하지 못한 경우를 의미하는 False Negative(FN)는 0%로 Missing이 발생하지 않음을 알 수 있다.



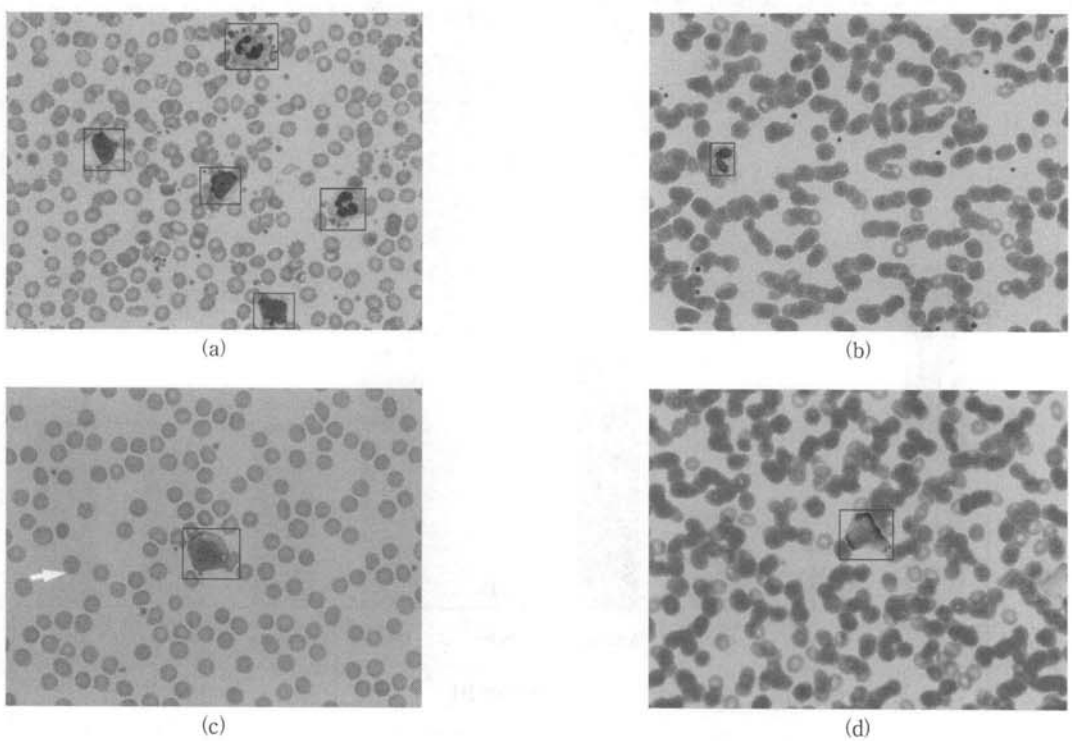
(그림 5) sub-image 분리 성능 평가 결과

Sub-image 분리 결과를 (그림 6)에 나타내었다. 그림 6에서 보는 바와 같이 혈액 세포 영상에서 sub-image를 정확하게 분리해 내는 것을 알 수 있다. 특히, (그림 6)(a)와 같이 여러 개의 백혈구가 존재하는 이미지나 (그림 6)(b)와 같이 아주 작은 크기의 백혈구도 정확하게 찾아내는 것을 알 수 있다. 또한, (그림 6)(c)와 같이 화살표와 같은 잡음이 있고 백혈구가 흐릿한 경우 (그림 6)(d)와 같이 적혈구끼리 뭉쳐있어 주변은 진하고 백혈구는 흐릿한 경우에도 정확하게 백혈구만을 찾아 sub-image를 분리 해 내는 것을 확인하였다.

추가적으로, 제안된 기법의 백혈구 핵 분할 성능 평가를 위해서, 제안된 기법에 의해 추출된 백혈구 핵 영역을 Herbert 등[5]이 제안한 기법, Mira 등[7]이 제안한 기법을 이용한 분할 방법과 비교하였다.

평가를 위해 각 기법의 분할 결과는 전문가가 판단하여 손으로 분할한 ground-truth를 제안된 알고리즘의 분할결과와 비교하여 저분할(Under segmentation), 과분할(Over segmentation) 오차 비율을 평가하였다.

저분할은 ground-truth 영역에 대해 각 기법이 분리하지 못한 영역의 비율을 나타내고, 과분할은 각 기법이 분리한 영역 중 ground-truth 영역이 아닌 오차 영역을 나타내며,



(그림 6) sub-image 분리 결과 예

각각의 오차율을 구하는 식은 식(9)와 같다.

$$S_U = \frac{M-(M\cap S)}{S_M},\ S_O = \frac{S-(M\cap S)}{S_S} \tag{9}$$

여기서, M 과 S 는 ground-truth 영역과 각 분할 기법을 이용하여 추출한 영역을 의미하고, S_M과 S_S는 추출된 영역의 전체 크기(pixel)이다. S_U는 두 영역의 해당 좌표 픽셀을 비교하여 사람이 선택한 영역 중에서 각 분할 기법에 의해 추출되지 못한 영역의 비율인 저분할을 나타내고 S_O는 두 영역의 해당하는 좌표 픽셀을 비교하여 각 분할 기법에 의해 추출된 전체 영역 중에서 사람이 선택하지 않은 영역의 비율인 과분할을 의미한다. 수식 (10)과 같이, 두 오차율을 이용하여 평균 분할 정확도를 구할 수 있다.

$$AVG_S = 1-(S_U \oplus S_O)\times 100 \tag{10}$$

AVG_S는 높은 값을 가질수록 적은 오차, 높은 정확도를 나타내고, 반대로 낮은 값을 가질수록 큰 오차율, 낮은 분할 신뢰도를 나타낸다.

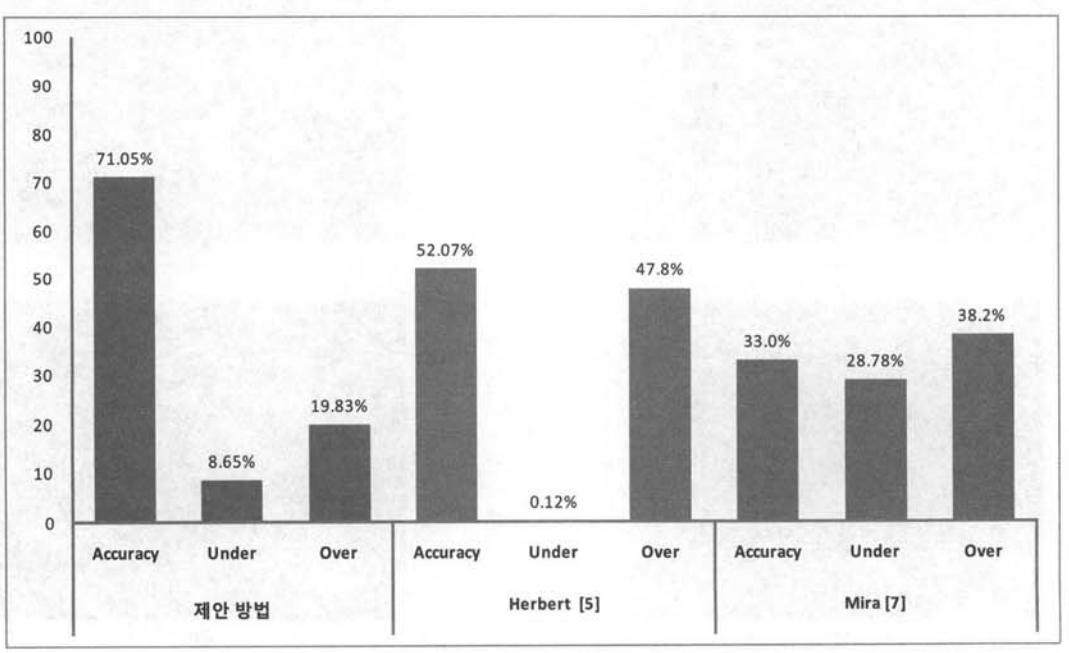
(그림 7)에 백혈구 분할의 비교 실험 결과를 나타내었다.

위의 그림에서 보는 바와 같이, Herbert 등의 알고리즘의 정확도 52%, Mira 등의 알고리즘의 정확도 33%에 비해, 제안된 기법은 71% 이상의 정확도를 나타내었다. 또한, 과분할 오차율이 비교 알고리즘은 각각 47%, 38%인데 반해 제안된 기법의 과분할 오차율은 19%로 낮은 오차율을 보였다. 저분할 오차율의 경우, 제안된 기법은 8%, Herbert 등의 알고리즘은 0%, Mira 등의 알고리즘은 28%로, 제안 기법은 Herbert 등의 알고리즘 보다는 높은 오차율을 보였지만 Mira 등의 알고리즘에 비해 매우 낮은 오차율을 보여주었다.

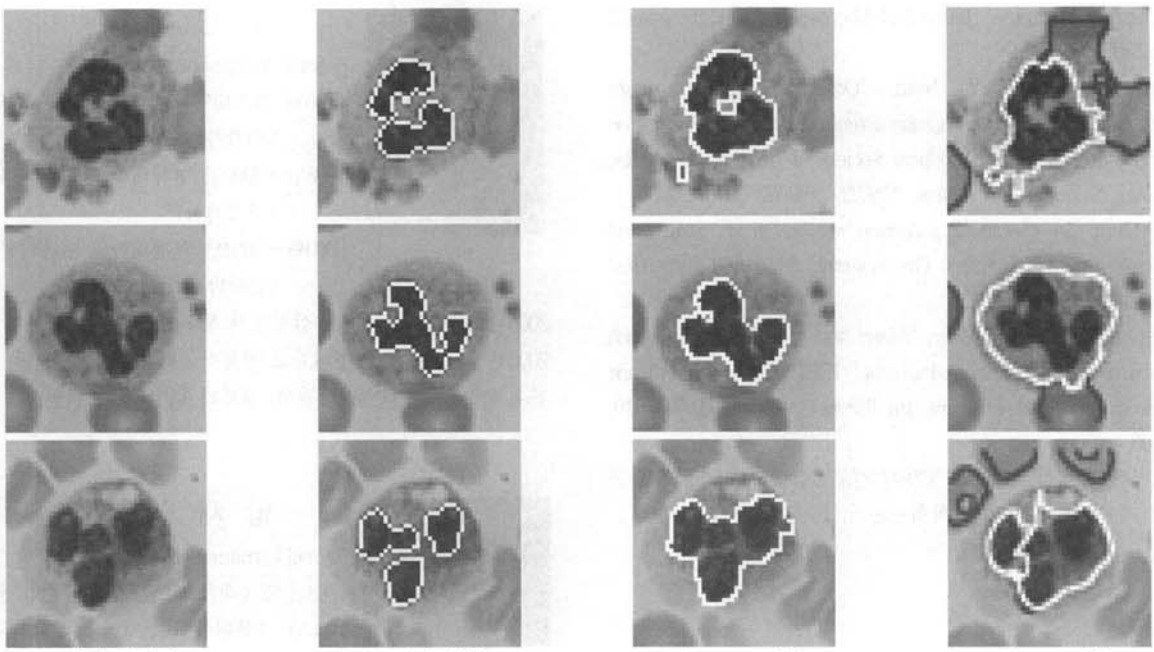
Herbert 등은 K-means 클러스터링 알고리즘을 적용한 후에 별도의 영역 병합과정이나 비핵 영역 제거 과정을 거치지 않기 때문에 주변에 핵과 유사한 색상의 잡음이 있는 경우를 핵 영역으로부터 제거 할 수 없고, 다핵인 경우 핵과 핵 사이의 세포질 공간도 핵으로 판단하는 경우가 많아 과분할 오차율이 제안 기법보다 높게 나타남을 확인하였다. 하지만, 저분할 에러율은 제안된 기법보다 Herbert 등의 알고리즘이 더 낮게 나타났다. 이는 제안된 기법은 mean-shift 클러스터링을 수행하는 동안 클러스터당 최소 영역 크기를 정해두고 이 조건을 만족하지 못하는 클러스터의 경우, 다른 클러스터에 병합하게 되는데, 이때 핵의 일부분이 다른 클러스터로 옮겨가게 된다. 하지만 Herbert 등의 알고리즘의 경우, 이미지의 모든 픽셀에 대해 독립적으로 백혈구 확률을 계산하기 때문에 백혈구 핵 픽셀이 다른 클러스터에 포함되는 경우가 상대적으로 드물어 더 낮은 저분할 오차율이 나타남을 확인하였다.

다음 그림은 제안한 기법으로 수행한 백혈구 핵 분할 결과 영상과 기존의 다른 알고리즘의 핵 분할 결과 영상이다.

(그림 8)에서 보는 바와 같이, 제안한 방법이 (a)의 입력 영상에 대해 핵만을 가장 정확하게 분할하는 것을 알 수 있다. Herbert 등[5]의 알고리즘도 비교적 핵을 잘 분할하고 있지만 핵 주변에 핵과 유사한 색상의 잡음이 있는 경우 잡음도 핵으로 분할하는 문제점과 다핵인 경우, 핵과 핵 사이의 세포질 공간을 모두 핵으로 인식하는 등의 단점이 있다. Mira 등[7]의 알고리즘은 본래 백혈구의 분할을 위한 알고리즘이 아니라, FISH 영상에서의 단일 세포와 군집 세포를 분류하고 군집 세포를 분리하기 위한 알고리즘이므로, 백혈



(그림 7) 제안 기법과 기존 알고리즘의 평균 분할 정확도 및 분할 오차율



(a) 입력 이미지 (b) 제안 방법 (c) Herbert [5] (d) Mira [7]

(그림 8) 제안 방법과 기존 알고리즘의 백혈구 핵 분할 결과 예

구의 핵 분할에 적용했을 때 낮은 정확도를 보였다.

5. 결 론

본 논문에서는 환자의 혈액 관련 질병의 진단을 위해, 백혈구 감별 계수의 기초가 되는 백혈구 핵 분할을 효율적으로 수행하는 방법을 제안하였다. 본 논문에서 제안한 방법은 핵이 염색되어있다는 특징을 기반으로, 색상, 방향 특징에 기초한 중요도 맵(Saliency map)을 생성하고, 이를 통해 세포 영상에서 백혈구를 포함하는 sub-image를 분리하여, sub-image에 대해서만 백혈구 핵 분할 작업을 수행하였다. Sub-image 분리 후 mean-shift 알고리즘을 적용하고, 이를 통해 얻은 클러스터들에 단계적 영역 병합 과정을 수행하여 최종적으로 단일한 클러스터의 핵을 검출하였다. 실험 수행 결과 sub-image 분리 정확도는 96%이상으로 거의 완벽하게 백혈구가 있는 영역만을 sub-image로 분리함을 알 수 있었고, 비교 실험 수행 결과, 기존의 다른 알고리즘보다 우수한 성능을 보여주었다. 하지만,sub-image내에서 백혈구분할결과 약71%의 성능을 보여줌으로써, 다른 알고리즘보다는 우수하지만 아직도 개선해야할 문제점이 많은 것으로 드러났다. 이를 위해 차후 알고리즘에서는 현재 연구에서 사용된 영상의 두 배 크기인 원본영상을 사용하여 백혈구분할 정확도를 더 높이는데 초점을 둘 것이다.

추가적으로, 현재 개발 중인 세포질의 분할 연구와 분할한 핵 및 세포질의 특징을 통해 백혈구를 분류하는 방법을 제안하고, 이를 통해 가장 가능성이 높은 질병까지 자동 진단하는, 의료 전문가의 진단 보조 시스템으로 개발하고자 한다.

참 고 문 헌

[1] K. Jiang, Q.M Liao and S.-Y Dai, *"A Novel White Blood Cell Segmentation Scheme using Scale-Space Filtering And Watershed Clustering,"* Int. Conf. on Machine Learning and Cybernetics, Vol.5, pp.2820-2825, 2003.

[2] G. Bourner, J. Dhaliwal and J. Sumner, *"Performance evaluation of the latest fully automated hematology analyzers in a large, commercial laboratory setting: a 4-way, side-by-side study,"* LABORATORY HEMATOLOGY, Vol.11, No.4, pp.285-97, 2005.

[3] N. Theera-Umpon, *"Morphological Granulometric Features of Nucleus in Automatic Bone Marrow White Blood Cell Classification,"* IEEE Trans. on Information Technology in Biomedicine, Vol.11, Issue3, pp.787-796, 2007.

[4] L.B Dorini, R. Minetto and N.J Leite, *"White blood cell segmentation using morphological operators and scale-space analysis,"* Computer Graphics and Image Processing, pp. 294-304, 2007.

[5] H. Ramoser, V. Laurain, H. Bischof and R. Ecker, *"Leukocyte segmentation and classification in blood-smear images,"* IEEE Engineering in Medicine and Biology Society 27th Annual Conference, pp.3371-3374 , 2006.

[6] N. Theera-Umpon, *"White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images,"* Lecture notes in Computer Science Vol.3614, pp.787-796, 2005.

[7] M. Jeong, B.C. Ko and J.-Y Nam, *"Overlapping Nuclei Segmentation based on Bayesian Networks and Stepwise*

Merging Strategy," Journal of Microscopy, Vol.235, Issue2, pp.188-198, 2009.

[8] B.C. Ko and J.-Y. Nam, *"Object-of-interest image segmentation based on human attention and semantic region clustering,"* Journal of Optical Society of America A: Optics, Image Science, and Vision, Vol.23, pp.2462-2470, 2006.

[9] N. Otsu, *"A threshold selection method from gray level histogram,"* IEEE trans. On Systems, Man and Cybertics, Vol.9, pp.62-66, 1979.

[10] D. Comaniciu and P. Meer, *"Mean-shift : A Robust Approach toward Feature Space Analysis,"* IEEE Trans. on Pattern Recognition and Machine Intelligence, Vol.24, pp.603-619, 2000.

[11] 정미라, 고병철, 남재열, *"FISH 세포영상에서의 군집세포 분할 기법,"* 한국정보처리학회논문지, 16-B, pp.1-10, 2009.

김 자 원

e-mail : yttrium@kmu.ac.kr

2009년 계명대학교 컴퓨터공학과 졸업(학사)

2009년 3월~현재 계명대학교 컴퓨터공학과 석사과정

관심분야 : 컴퓨터 비전 및 패턴인식

고 병 철

e-mail : niceko@kmu.ac.kr

1998년 경기대학교 전자계산학과 졸업 (이학사)

1998년~2000년 연세대학교 컴퓨터과학 (공학석사)

2000년~2004년 연세대학교 컴퓨터과학 (공학박사)

2004년 3월~2005년 8월 삼성전자 통신연구소 책임연구원

2005년 9월~현재 계명대학교 컴퓨터공학과 조교수

관심분야 : 내용기반 영상검색, 컴퓨터 비전 및 패턴인식

남 재 열

e-mail : jynam@kmu.ac.kr

1983년 경북대학교 전자공학과 졸업(공학사)

1985년 경북대학교 전자공학(공학석사)

1991년 University of Texas at Arlington 전기공학(공학박사)

1985년 5월~1987년 7월 한국전자통신연구소 연구원

1991년 9월~1995년 2월 한국전자통신연구소 선임연구원

1995년 3월~현재 계명대학교 컴퓨터공학과 교수

관심분야 : 영상압축, 영상통신, 멀티미디어 시스템