

퍼지 추론 규칙을 이용한 자궁 경부진 핵 인식

김 광 백*, 송 두 헌**

Nucleus Recognition of Uterine Cervical Pap-Smears using Fuzzy Reasoning Rule

Kwang-Baek Kim *, Doo Heon Song **

요 약

본 논문에서는 자궁 경부 세포진 영상에서 세포핵의 후보 영역과 핵을 추출하기 위해 현미경 400배율 확대 사진을 획득하는 과정에서 훼손된 컬러 영상을 복원하기 위한 방법으로 Lighting Compensation을 적용하여 영상을 보정한다. 그리고 배경 영역과 세포핵 영역을 구분하기 위해 영상의 R,G,B 영역의 히스토그램의 분포를 이용하여 배경을 제거한다. 배경이 제거된 영상을 그레이 영상으로 변환 한 후, 히스토그램 명암도의 값을 이용하여 세포핵 영역과 세포질을 분류하여 세포핵 영역을 추출한다. 그리고 Kapur 방법을 적용하여 세포핵 영역의 엔트로피 누적 확률을 구한 후, 영상을 이진화 한다. Kapur 방법이 적용된 이진화 영상에서 세포핵 영역의 중심과 주위 화소를 비교하는 3×3 마스크를 적용하여 영상의 미세한 잡음을 제거 한 후, 8 방향 윤곽선 추적 알고리즘을 적용하여 최종적으로 세포핵 영역을 추출한다. 추출된 세포핵의 영역을 분류 및 인식하는 과정으로 세포의 외각의 방향성 정보, 핵의 크기, 그리고 면적 비율의 특징을 이용하여 퍼지 소속 함수를 설계한 후, 소속 함수의 소속도를 구하고 퍼지 추론 규칙을 적용하여 자궁 경부 세포진 영상에서 정상 세포핵 및 암종 세포핵을 인식한다.

Abstract

In this paper, we apply a set of algorithms to classify normal and cancer nucleus from uterine cervical pap-smear images. First, we use lightening compensation algorithm to restore color images that have defamation through the process of obtaining 1×400 microscope magnification. Then, we remove the background from images with the histogram distributions of RGB regions. We extract nucleus areas from candidates by applying histogram brightness, Kapur method, and our own 8-direction contour tracing algorithm. Various binarization, cumulative entropy, masking algorithms are used in that process. Then, we are able to recognize normal and cancer nucleus from those areas by using three morphological features - directional information, the size of nucleus, and area ratio - with fuzzy membership functions and deciding rules we devised. The experimental result shows our method has low false recognition rate.

▶ Keyword : uterine cervical cytodiagnosis, morphologic features, cancer nucleus, Kapur method, 8-direction contour tracing algorithm

• 제1저자 : 김광백

• 접수일 : 2008. 3. 27, 심사일 : 2008. 4. 1, 심사완료일 : 2008. 5. 24.

* 신라대학교 컴퓨터정보공학부 교수 ** 용인송담대학 컴퓨터게임정보과 교수

1. 서론

자궁경부암은 우리나라 여성에게 발생하는 전체 악성종양 중 유방암, 위암, 대장암, 갑상선암에 이어 5위에 이르는 발생률을 보이며 현재 여성 생식기 암중에서는 가장 발생 빈도가 높은 암이다. 하지만 다른 암과는 달리 암으로 전이되기 이전 단계인 조기 전암 단계에서 정기적으로 암 검진만 한다면 충분히 조기 발견과 치료가 가능하다. 최근에는 의학의 발달로 인해 여러 가지 검진법으로 조기 발견률이 상당히 높은 편이다[1]. 자궁경부암에 대한 검진으로는 세포 도말 검사(pap smear), 골반 검사, 질경 검사, 인유두종 바이러스 검사(HPV-PCR) 등이 있다. 이 중 조기 전암 단계에서의 검진은 세포 도말검사(pap-smear)가 주로 쓰인다[2]. 세포 도말검사(pap-smear)는 자궁 경부 주변의 여러 곳에서 세포조직 표본을 채취하여 염색한 후, 현미경으로 이상 징후를 검사하는 방법으로 상당히 정확한 검진 결과를 가진다. 하지만 높은 암 발생 빈도에 비해 의사나 병리학자가 하루에 검진 할 수 있는 양은 제한되어 있다.

최근의 자궁 경부진 세포 추출 및 인식에 관련된 연구로 영역 확장 기법과 퍼지 RBF 네트워크를 이용한 자궁 경부진 세포 추출 및 인식 방법과 Kapur 방법을 이용한 자궁 경부진 세포 추출 및 인식 방법 등이 제안되었다. 영역 확장 기법과 퍼지 RBF 네트워크를 적용한 세포 추출 및 인식 방법은 세포 영역을 효율적으로 분할하기 위해 영역 확장 기법을 적용한 후, 세포 영역에서 세포핵의 Hue값을 추출하여 퍼지 RBF 네트워크에 적용하여 정상 세포와 비정상 세포를 분류 및 인식하였다. 이 방법은 세포의 형태학적 특징, 즉 구조적 변화와 비정상 세포의 색조 변화 특징 등을 분석하지 않고 세포핵의 Hue값을 퍼지 RBF 네트워크에 적용하여 정상 세포와 비정상 세포를 분류 및 인식하였으므로 대세포 편평 암종 세포, 상피내암 암종 세포 등 다양한 자궁 경부진 세포 추출 및 인식에 적용할 수 없는 문제점이 있다[3]. 기존의 Kapur 방법을 이용한 자궁 경부진 세포 추출 및 인식 방법은 세포진 영상에서 배경과 세포 영역을 분할하는 전처리 과정에서 세포핵의 정보가 손실되어 정상 세포와 암세포를 분류하는 과정에서 정확성이 낮아 세포 감별의 위음성률 및 위양성률이 높아지는 문제점이 있다[4].

따라서 본 논문의 목적은 세포 도말검사(pap smear)에 사용되는 자궁 경부진 세포핵의 특징 연구를 통해 자동화된 보조 진단 도구로 효과적인 세포 검진과 세포 감별의 위음성률 및 위양성률을 낮추는데 있다. 의사와 병리학자들이 검사

에 사용하는 세포 도말 영상은 400~1000 배율이다. 본 논문에서는 400 배율 자궁 경부진 중층편평 세포 영상을 대상으로 정상 및 암종 세포핵의 형태학적 정보의 특징을 분석하여 퍼지 소속 함수를 설계하고 소속도를 구한 후, 추론 규칙을 적용하여 자궁 경부진 세포 영상에서 정상 세포핵 및 암종 세포핵을 인식한다.

II. 세포핵 영역 추출 과정

본 논문에서 제안하는 자궁 경부진 세포핵 영역 추출 과정은 <그림 1>과 같다.

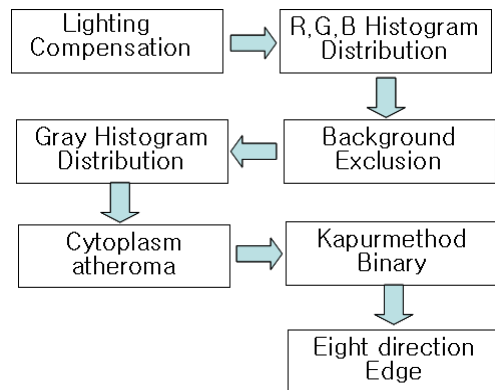


그림 1. 세포핵 영역 추출 과정
Fig 1. Nucleus area extraction process

2.1 Lighting Compensation을 이용한 영상 보정

본 논문에서는 자궁 경부 세포진 영상에서 세포핵의 후보 영역을 추출하기 위해 현미경 400배율 확대 사진을 획득하는 과정에서 훼손된 컬러 영상을 보정하는 과정을 거친다. 영상을 보정하기 위한 방법으로 Lighting Compensation을 적용한다. 보정된 영상은 <그림 2>와 같다.

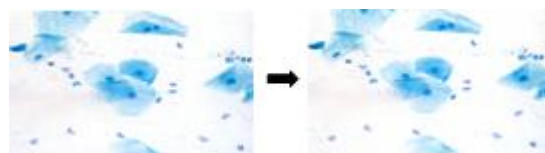


그림 2. 컬러 정보가 보정된 영상
Fig 2. Image after color information restored

2.2 히스토그램 분포를 이용한 배경 제거

Lighting Compensation을 적용하여 컬러 영상이 보정된 영상에서 R,G,B의 히스토그램 분포를 이용하여 배경 영역을 제거한다. 일반적으로 영상에서 R,G,B 분포의 밀도가 가장 높은 영역이 배경이기 때문에 R,G,B의 히스토그램 분포를 분석하여 상위 10%의 분포를 가지는 영역을 배경 영역으로 간주하여 배경 영역을 제거한다. <그림 3>은 R,G,B 각각의 히스토그램의 분포와 배경이 제거된 영상이다.

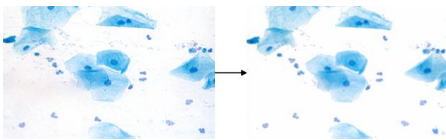


그림 3. 각각의 R,G,B 히스토그램 분포와 R,G,B 분포를 이용한 배경 제거 결과 영상
Fig 3. Image after background removal by R, G, B histogram distribution

2.3 명암도 값을 이용한 세포질과 핵의 분리

배경 영역이 제거된 영상을 그레이 영상으로 변환한 후, 영상의 명암도 히스토그램을 분석하고 명암도 값을 이용하여 세포질 영역과 세포핵 영역을 분리한다. 일반적으로 세포의 핵을 둘러싸고 있는 세포질은 세포의 핵에 비해서 낮은 명암도 값을 가진다. 따라서 본 논문에서는 그레이 영상에서 명암도 히스토그램을 분석하여 영상의 상위 5%의 명암도를 가지는 영역을 핵 영역, 그리고 그 이하의 명암도를 가지는 영역을 세포핵을 둘러싼 세포질 영역으로 판단하고, 세포질 영역을 제거하여 세포핵 영역을 추출한다. <그림 4>는 그레이 영상과 명암도의 분포이다.

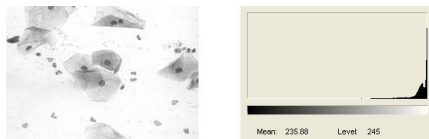


그림 4. 그레이 영상과 히스토그램 분포
Fig 4. Gray image and its histogram distribution

2.4 Kapur 방법을 이용한 세포핵 영역 이진화

세포질 영역이 제거된 자궁 경부진 세포핵 영상을 이진화하기 위하여 Kapur 방법을 적용한다[4,5]. Kapur 방법은 영상의 히스토그램을 획득한 후, 0~255사이의 임의의 명암

값 k 를 정하여 두 구간으로 나눈 후, 식(1)과 (2)를 이용하여 C_0 와 C_1 의 합이 최대가 되는 지점 k 를 threshold 값으로 설정하여 영상을 이진화 하는 방법이다. 세포핵 영역의 엔트로피 C_0 와 핵 주위의 비관심 영역의 엔트로피 C_1 의 계산은 각각 식(3)과 (4)와 같다. 여기서 p_i 는 i 번째 명암도 값을 의미하며, w_k 는 0부터 k 까지의 누적률이다.

$$Class C_0 : \frac{p_0}{w_k}, \frac{p_1}{w_k}, \dots, \frac{p_k}{w_k} \dots\dots\dots (1)$$

$$Class C_1 : \frac{p_{k+1}}{L-w_k}, \frac{p_{k+2}}{L-w_k}, \dots, \frac{p_{L-1}}{L-w_k} \dots\dots\dots (2)$$

두 클래스의 누적 엔트로피는 식(3)과 (4)로 명시된다.

$$H(C_0) = - \sum_{i=0}^k \frac{p_i}{w_k} \ln \frac{p_i}{w_k} \dots\dots\dots (3)$$

$$H(C_1) = - \sum_{i=k+1}^{L-1} \frac{p_i}{L-w_k} \ln \frac{p_i}{L-w_k} \dots\dots\dots (4)$$

두 누적 엔트로피의 합 $\psi(k)$ 는 식(5)와 같다.

$$\begin{aligned} \psi(k) &= H(C_0) + H(C_1) \\ &= \ln \sum_{i=0}^k \frac{p_i}{w_k} + \ln \sum_{i=k+1}^{L-1} \frac{p_i}{L-w_k} - \frac{\sum_{i=0}^k p_i \ln p_i}{\sum_{i=0}^k p_i} - \frac{\sum_{i=k+1}^{L-1} p_i \ln p_i}{\sum_{i=k+1}^{L-1} p_i} \\ &= \ln \sum_{i=0}^k \frac{p_i}{w_k} + \ln \sum_{i=k+1}^{L-1} \frac{p_i}{L-w_k} - \frac{\sum_{i=0}^k p_i \ln p_i}{\sum_{i=0}^k p_i} - \frac{\sum_{i=k+1}^{L-1} p_i \ln p_i}{\sum_{i=k+1}^{L-1} p_i} \\ k_{opt} &= Max(\psi(k)) \dots\dots\dots (5) \end{aligned}$$

여기서 k_{opt} 은 두 클래스 w_k 과 $L-w_k$ 의 엔트로피 합이 최대인 값을 의미한다. 이 k_{opt} 을 threshold 값으로 설정하여 영상을 이진화 한다.

Kapur 방법이 적용된 영상에서 3×3 마스크를 이용하여 미세 잡음을 제거한다. 3×3 마스크는 중심 화소의 주위 8 방향을 스캔한 후, 흰색 픽셀이 검정색 픽셀보다 많이 존재할 경우, 중심화소를 흰색으로 대체함으로써 영상에 존재하는 미세 잡음을 제거한다. 3×3 마스크의 구조는 <그림 5>와 같다.

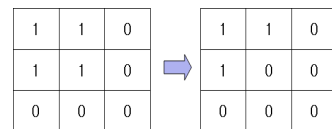


그림 5. 3×3 마스크 구조
Fig 5. 3×3 mask structure

미세 잡음이 제거된 영상에서 8 방향 윤곽선 추적 알고리즘을 적용하여 정상 세포핵과 암종 세포핵을 추출한다. 그러나 <그림 6>과 같이 자궁 경부진 증층 편평 세포 영상에서는 백혈구와 림프구들이 세포핵과 같이 군집화 현상을 이루고 있는 경우가 있다.

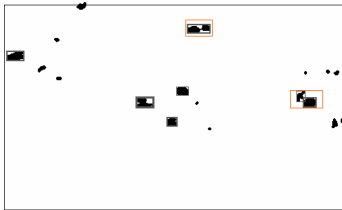


그림 6. 세포핵의 군집화 영상
Fig 6. Nucleus cluster image

본 논문에서는 정상 세포핵의 형태학적 특징 정보를 분석하기 때문에 백혈구와 림프구에 의해서 군집화 된 세포핵의 영역에서는 정확한 핵의 특징 정보를 추출할 수 없다. 따라서 군집화 된 영역에서 백혈구와 림프구를 관심영역에서 제외하기 위해, 추출된 세포핵의 평균을 이용하여 평균 보다 30% 이상의 크기를 가진 영역을 군집화된 영역이라고 판단하고 군집화된 영역 내에서 평균 보다 30% 이하의 크기를 가진 영역이 림프구와 백혈구 영역이므로 군집화 영역에서 제거한다. kapur 방법을 적용하여 이진화된 영상에서 정상 세포핵과 암종 세포핵을 추출한 결과와 kapur 방법에서 threshold 값을 설정하는 과정을 <그림 7>로 나타내었다.

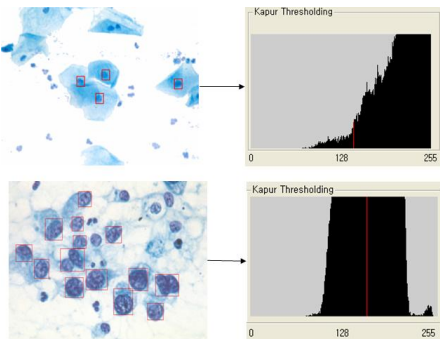


그림 7. 정상 세포의 핵과 암종 세포의 핵 추출
Fig 7. Extraction of nucleus from normal and cancer cell

III. 세포핵의 특징 추출

일반적으로 암종 세포의 핵은 정상 세포의 핵에 비해 크고, 다양한 형태를 보이며, 뚜렷한 핵소체를 가진다. 또한 암종 세포핵은 이염색질(heterochromatin)의 증가 때문에 거친 염색질을 보인다. 실제로 암종 세포핵은 이와 같은 특징 때문에 자궁경부암종 등의 세포학적 진단이 가능하다(6). 본 논문에서는 이러한 세포핵의 형태학적 특징을 이용하여 핵 외곽의 방향성 정보와 면적 비율, 핵의 크기 정보를 이용하여 정상 세포핵과 암종 세포핵을 분류하고 인식한다. 방향성 정보는 핵 외곽의 거친 정도를 수치상으로 나타낸 것이며, 면적 비율은 세포핵의 염색질의 짙은 정도와 세포핵이 서서히 파괴되면서 생기는 공포의 비율이 낮은 정도를 나타낸다. 공포의 비율과 면적 비율의 수치는 반비례 관계가 있다. 또한 암종 세포핵의 대부분은 정상 세포의 핵보다 상대적으로 크기가 큰 특징을 가진다. <그림 8>은 세포핵의 특징을 이용하여 분류하는 과정이다.

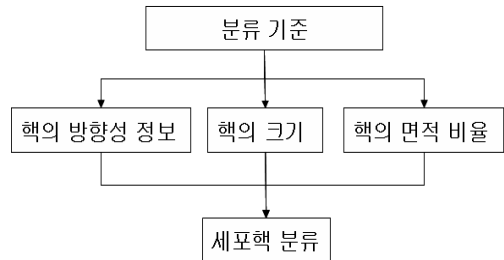
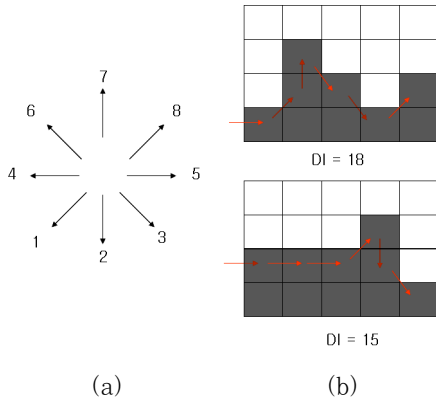


그림 8. 세포핵의 특징을 이용한 핵 분류 과정
Fig 8. Nucleus classification process with morphological features

3.1 세포핵의 방향성 정보

본 논문에서는 8 방향 윤곽선 추적 알고리즘의 진행 방향에 따라 가중치를 적용하여 이전 방향의 가중치와 다음 방향의 가중치의 차를 계산하여 계산된 값을 모두 더한 방향성 수치를 세포핵 인식을 위한 첫 번째 특징으로 적용한다. 여덟 방향의 각각의 가중치는 <그림 9>와 같다. <그림 9>의 (a)는 방향에 따른 가중치이고, (b)는 각 영상에서 외곽선을 추적하면서 가중치를 합한 수치인 DI의 한 부분이다. (b)에서 들리는 값이지만 외곽이 상대적으로 더 거친 위쪽 영상의 가중치 합이 더 큰 것을 확인할 수 있다.



(a) (b)

그림 9. (a) 방향에 따른 가중치
(b) 각 영상에 대한 가중치의 합

Fig 9. (a) Weight by direction
(b) Sum of weights from images

3.2 세포핵에서의 면적 비율

암종 세포핵의 또 다른 특징은 정상 세포와는 다르게 핵의 파괴로 인해 핵 내에서의 공포가 나타난다. 이 특징을 인식에 적용하기 위하여 추출된 단일 세포에 평균 이진화를 적용하여 핵 내에서의 공포가 핵의 전체 영역에서 차지하는 비율을 식 (6)과 같이 계산한다. 정상 세포와 공포가 발생한 암종 세포를 비교한 것은 <그림 10>과 같다.

$$\text{Cell rate} = \text{Cell size} - \text{Cell hole} \dots\dots\dots (6)$$



그림 10. (a) 정상 세포핵 (b) 공포가 생긴 암종 세포핵
Fig 10. (a) Normal nucleus
(b) Cancer nucleus with vacuole

3.3 세포핵의 크기

추출된 세포핵의 크기도 중요한 특징이 된다. 정상 세포의 핵은 암종 세포의 핵보다 크기가 작고, 정상세포는 핵과 세포질의 비율이 대개 1:8~1:9로 나타난다. 그러나 자궁 경부진 암종 세포의 핵은 이형성증과 화생성으로 발전하는 가능성이 90%이상이기 때문에 괴립과 이 염색질의 특징에 의해서 핵의 크기가 커지며, 세포질 또한 거친 형태를 가진다[2].

IV. 세포핵 인식을 위한 소속 함수 설계 및 퍼지 추론 규칙을 이용한 핵 인식

본 논문에서는 추출된 자궁 경부진 정상 세포핵과 암종 세포핵을 인식하는 과정에서 세포핵의 방향성 정보, 핵의 크기, 핵의 면적 비율의 특징을 퍼지 기법에 적용하기 위하여 소속 함수를 설계한다. 본 논문에서 소속 함수의 카테고리는 Normal, Middle, Cancer와 같이 3개로 구성하고 Normal은 N, Middle은 M, Cancer는 C로 표시한다. M의 카테고리는 N과 C구간의 교집합으로 설정한다. 각각의 카테고리에서 계산된 소속도를 출력 소속 함수에 적용하여 퍼지 값을 구한 후에 Max_Min 연산으로 추론한 후, 결과 값을 무게 중심법을 적용하여 비퍼지화를 수행한다. 그리고 비퍼지화를 통해서 최종적으로 정상 세포핵 및 암종 세포핵의 소속도를 계산하여 인식한다. <그림 11>과 같이 핵의 방향성 특징에 대한 소속 함수는 [0, 10]의 소속 구간을 가지며, N_D는 N구간의 방향성이며, M_D는 중간구간의 방향성, C_D는 C구간에 대한 방향성이다. 세포의 방향성 특징은 8 방향 윤곽선 추적 알고리즘의 진행 방향에 의한 방향성 수치의 값에 따라 소속 구간을 구분한다. 정상 세포의 핵은 원형에 가까운 형태를 가지기 때문에 방향성 수치의 값이 적은 구간 [0,5]의 N_D구간에 속하며, 암종 세포의 핵은 이 염색질의 증가 때문에 세포의 핵이 거친 형태를 가지기 때문에 방향성 수치의 값이 커지게 된다. 따라서 암종 세포의 핵의 구간은 소속 구간이 높은 [5,10]까지의 C_D 구간이다. 정상 세포의 핵이라도 월경주기에 따라서 세포가 분열하고 성장하는 과정이 빠르기 때문에 핵의 표면이 거칠어질 수 있다. 이런 경우 정상 세포의 핵 구간 N_D와 암종 세포의 핵 구간 C_D의 중간 구간인 [4,6] 구간을 M_D 구간으로 정의한다.

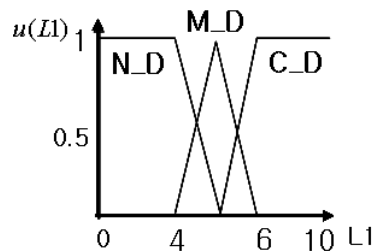


그림 11. 핵의 방향성 특징에 대한 소속 함수
Fig 11. Membership functions by directional information

〈그림 12〉에서 핵의 크기에 대한 소속 함수의 구간 L2는 [0,20]의 구간이다. N_S는 Normal 구간에 대한 크기이며, M_S는 중간 구간에 대한 크기이며, C_S는 Cancer 구간에 대한 크기이다. 정상 세포의 핵의 크기는 균등하고, 일정한 크기를 가지기 때문에 크기의 차이가 적은 [0,10]의 N_S 구간에 속하며, 암종 세포의 핵은 크기가 균일 하지 않고, 여러 가지 다양한 크기를 가지기 때문에 크기의 차이가 큰 [10,20]의 C_S 구간이다. 그러나 정상 세포핵이 분열과 성장하는 과정에서 암종 세포핵의 크기와 비슷한 경우가 있다. 이러한 경우에는 N_S와 C_D의 중간인 M_S로 정의하였고 구간은 [8,12]이다.

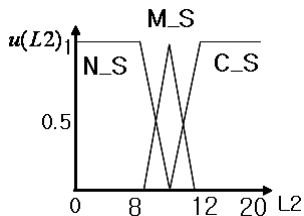


그림 12. 핵의 크기에 대한 소속 함수
Fig 12. Membership functions by the size of nucleus

〈그림 13〉에서 핵의 면적 비율에 대한 소속 함수의 구간 L3은 [0,15]의 구간을 가진다. N_R은 Normal 구간에 대한 면적 비율이고, M_R은 중간 구간에 대한 면적 비율, C_R은 Cancer 구간에 대한 면적 비율이다. 정상 세포의 핵은 세포 핵이 서서히 파괴되면서 생기는 공포가 거의 없기 때문에 공포의 비율이 낮다. 따라서 [0,7.5]의 N_R 소속 구간이며, 암종 세포핵은 공포의 비율이 높기 때문에 C_R의 [7.5,15]의 높은 소속 구간을 가진다. 그러나 정상 세포도 노화 현상으로 공포가 발생할 수 있다. 이러한 경우는 N_R과 C_R의 중간 구간 [6,9]의 소속 구간을 M_R 구간으로 정의한다. 핵의 면적 비율에 대한 소속 함수는 〈그림 13〉과 같다.

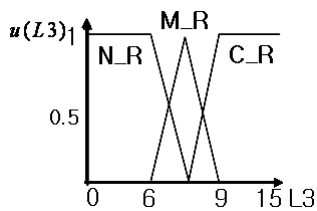


그림 13. 핵의 면적비율에 대한 소속 함수
Fig 13. Membership functions by nucleus area ratio

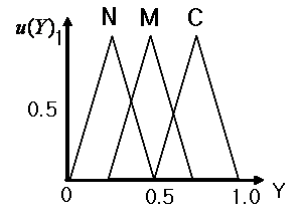


그림 14. 출력 소속 함수
Fig 14. Output membership functions

본 논문에서는 핵의 방향성 정보, 크기, 면적 비율에 따른 소속 함수를 통해서 계산된 소속도는 다음과 같은 추론 규칙에 의해서 추론한다. 핵의 방향성 정보의 소속도 $\mu(L1)$, 핵 크기의 소속도 $\mu(L2)$, 그리고 핵의 면적 비율에 대한 소속도는 $\mu(L3)$ 이다. 〈그림 14〉는 각 소속 함수의 입력에 의한 출력 소속 함수이다. 본 논문에서 추론 한 규칙은 핵의 방향성 정보가 자궁 경부진 세포핵의 인식에 첫 번째 특징으로 적용하기 때문에 방향성 정보의 구간 L1이 정상 세포의 N_D 구간, 암종 세포핵의 C_D구간, 정상 세포핵과 암종 세포핵 구간의 교집합 구간 M_D에 속하는 결과에 따라서 추론 규칙을 구분하고 세 소속 함수에 대한 결과는 Y로 표시하고, 다음과 같이 추론한다.

(1) 핵의 방향성 정보가 Normal구간에 속할 때

If L1 is N_D and L2 is N_S and L3 is N_R then Y is N
 if L1 is N_D and L2 is N_S and L3 is C_R then Y is N
 if L1 is N_D and L2 is N_S and L3 is M_R then Y is N
 if L1 is N_D and L2 is C_S and L3 is N_R then Y is N
 if L1 is N_D and L2 is C_S and L3 is M_R then Y is N
 if L1 is N_D and L2 is M_S and L3 is N_R then Y is M
 if L1 is N_D and L2 is M_S and L3 is C_R then Y is M
 if L1 is N_D and L2 is M_S and L3 is M_R then Y is M

(2) 핵의 방향성 정보가 Cancer 구간에 속할 때

if L1 is C_D and L2 is N_S and L3 is N_R then Y is N
 if L1 is C_D and L2 is N_S and L3 is C_R then Y is C
 if L1 is C_D and L2 is C_S and L3 is N_R then Y is C
 if L1 is C_D and L2 is N_S and L3 is M_R then Y is C
 if L1 is C_D and L2 is C_S and L3 is M_R then Y is C
 if L1 is C_D and L2 is M_S and L3 is N_R then Y is M
 if L1 is C_D and L2 is M_S and L3 is C_R then Y is M
 if L1 is C_D and L2 is M_S and L3 is M_R then Y is M

(3) 핵의 방향성 정보가 Middle 구간에 속할 때

if L1 is \underline{M}_D and L2 is \underline{N}_S and L3 is \underline{N}_R then Y is N
 if L1 is \underline{M}_D and L2 is \underline{N}_S and L3 is \underline{C}_R then Y is N
 if L1 is \underline{M}_D and L2 is \underline{C}_S and L3 is \underline{N}_R then Y is C
 if L1 is \underline{M}_D and L2 is \underline{N}_S and L3 is \underline{M}_R then Y is M
 if L1 is \underline{M}_D and L2 is \underline{C}_S and L3 is \underline{M}_R then Y is M
 if L1 is \underline{M}_D and L2 is \underline{M}_S and L3 is \underline{N}_R then Y is M
 if L1 is \underline{M}_D and L2 is \underline{M}_S and L3 is \underline{C}_R then Y is M
 if L1 is \underline{M}_D and L2 is \underline{M}_S and L3 is \underline{M}_R then Y is M

예를 들어 핵의 방향성 정보 L1이 \underline{N}_D 구간에 속할 때, 핵의 크기 L2가 \underline{N}_S 구간이고, 핵의 면적 비율 구간 L3가 \underline{N}_D 구간이면, 결과 값 Y는 N 구간으로 추론한다. 다른 예로 핵의 방향성 정보 L1이 \underline{C}_D 구간에 속할 때, 핵의 면적 크기 L2가 \underline{C}_S 구간이며, 핵의 면적 비율 L3가 \underline{N}_R 구간이면 결과 값 Y는 C구간으로 추론한다. 추론 규칙을 적용하여 얻어진 소속 구간을 Max_Min 합성 연산으로 추론하는 방법은 식(7)과 같다.

$$\mu_R(N_m, C_m, M_m; Y_m) = \max(\min(\mu_N(N_i), \mu_C(C_i), \mu_M(M_i), \mu_Y(Y_i))) \dots \dots (7)$$

Max_Min 연산으로 추론한 결과들 중 소속도가 가장 높은 값을 찾는다. 찾은 값을 무게 중심법을 이용하여 비퍼지화 한다. 비퍼지화는 식 (8)과 같이 계산한다(7).

$$y^* = \frac{\sum \mu(y_i)x_i}{\sum \mu(y_i)} \dots \dots \dots (8)$$

식(8)에서 y^* 값은 비퍼지화의 출력 값이다. 이렇게 구해진 퍼지 값을 비퍼지화를 통해서 구해진 값으로 정상 세포 핵과 암종 세포핵의 평가 범위는 <표 1>과 같다.

표 1. 최종 평가 범위
Table 1. Final region of evaluation

최종 평가	평가 범위
정상세포핵	$0.0 \leq y^* \leq 0.5$
암종세포핵	$0.5 < y^* \leq 1.0$

V. 실험 및 결과 분석

실험 환경은 IBM 호환 기종의 Pentium-IV PC상에서 Visual C++ 6.0으로 구현하였다. 실험 표본은 640×480 크기를 가진 자궁경부 편평 세포진 400 배율 정상 세포 영상 12장과 암종 세포 영상 20장을 대상으로 실험하였다. 자궁 경부진 정상 세포 영상에서 추출된 정상 세포핵의 개수는 총 61개중 59개가 추출되었으며, 암종 세포 영상에서 추출된 암종 세포핵의 개수는 총 138중 131개가 추출되었다. <표 2>는 정상 세포 영상에 대한 세포핵 인식 결과를 나타내었고 <표 3>은 암종 세포 영상에 대한 암종 세포핵 인식 결과를 나타내었다. <표 4>는 자궁 경부진 세포 영상에 대한 세포핵 인식 결과를 나타내었다. <그림 15>는 본 논문에서 제안된 자궁 경부진 추출 및 인식 화면이다.

본 논문에서 정상 세포핵의 경우, 중층 중간층 세포핵과 표면 편평상피 세포핵의 경우는 인식에 실패하는 경우가 발생하였다.

표 2. 정상 세포핵에 대한 인식 개수
Table 2. Recognition result for normal cells

정상세포 영상	인식된 세포핵 개수	총 세포핵 개수
중층 편평상피 세포	12	12
중층 중간층 세포	18	19
원주상피 세포	13	13
표면 편평상피 세포	16	17

표 3. 암종 세포핵에 대한 인식 개수
Table 3. Recognition result for cancer cells

암종 세포영상	인식된 세포핵 개수	총 세포핵 개수
각화성 암종 세포	12	12
비각화성 암종 세포	17	19
화생성 암종 세포	13	13
대세포 편평 암종 세포	15	17
소세포 편평 암종 세포	6	6
미소 침윤성 암종 세포	16	16
투명 세포 선 암종 세포	10	10
상피내암 암종 세포	24	24
이형성암 암종 세포	18	21

표 4. 제안된 방법에 의한 세포핵 인식 결과
Table 4. Result for recognition and extraction by proposed algorithm

세포 기준	총 세포 인식 개수	총 세포 개수
정상 세포핵	59	61
암종 세포핵	131	138

중간층 세포핵의 경우, 월경주기에 따라서 세포의 분열과 성장이 빠르기 때문에 세포핵의 표면이 거칠고, 핵의 크기가 커지며, 핵 표면 내에서 공포가 발생하였다. 따라서 핵의 방향성 정보, 크기, 면적 비율에 대한 소속 함수를 적용하여 추론한 결과, 암종 세포핵으로 인식되었다. 암종 세포핵의 경우에는 비각화성 암종세포, 대세포 편평 암종 세포, 이형성증 암종 세포가 정상 세포가 인식되는 경우가 발생하였다. 비각화성 암종 세포의 경우는 세포핵의 표면이 변하지 않고 크기 또한 정상 세포의 크기와 비슷하며, 공포만이 생겨서 암세포 인식에 실패한 경우이다. 대세포 편평 암종 세포의 경우는 세포핵의 크기만 변하고, 이형성증 암종 세포의 경우는 염색에서 발달한 암종 세포이기 때문에 이 염색질의 변화로 핵의 표면만 거칠어졌다. 따라서 퍼지 추론 기법을 적용한 결과, 정상 세포핵으로 인식하는 문제점이 있었다.

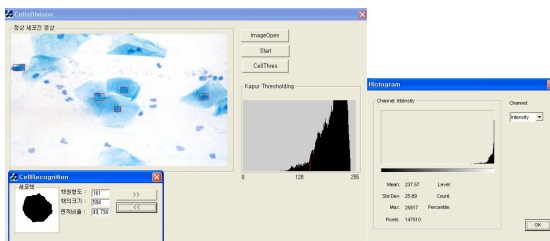


그림 15. 자궁 경부진 추출 및 인식 화면
Fig 15. Screen shot of the extraction and recognition of Uterine Cervical Pap-smears

VI. 결 론

본 논문에서는 세포 도말 검진에 사용되는 자궁경부 세포진 영상에서 세포핵을 추출하기 위해 Lighting Compensation을 적용하여 영상을 보정하였다. 그리고 배경 영역과 세포핵 영역을 구분하기 위해 영상의 R,G,B 영역의 히스토그램 분포를 이용하여 배경을 제거하였다. 배경이 제거된 영상을 그레이 영상으로 변환한 후, 히스토그램의 명암도의 값을 이용하여 세포핵

과 세포질 영역을 분리하였다. 그리고 Kapur 방법을 적용하여 세포핵 영역의 엔트로피 누적 확률을 구하고 이것을 이용하여 영상을 이진화 하였다. Kapur 방법이 적용된 이진화된 영상에서 세포핵 영역의 중심과 주위 화소를 비교하는 3x3 마스크를 적용하여 영상의 미세한 잡음을 제거한 후, 8 방향 윤곽선 추적 알고리즘을 적용하여 최종적으로 세포핵을 추출하였다. 추출된 세포핵을 정상 세포와 암세포로 분류 및 인식하는 과정으로 세포의 외각의 방향성 정보, 핵의 크기, 그리고 면적의 특징을 각각 구하고 구한 특징들에 대해 각각의 퍼지 소속 함수에 적용하여 소속도를 구하고 퍼지 추론 규칙을 적용하여 자궁 경부 세포진 영상에서 정상 세포핵 및 암종 세포핵을 인식하였다. 세포를 분류하고 판별에 적용되는 특징들은 100여 개의 특징이 있다. 하지만 그 중 암 세포의 특징들을 선택하기 어렵고 암 세포의 특징들을 추출하기 위해서는 오랜 시간이 걸린다. 이 중에서 본 논문에서는 크기, 면적, 방향성의 3가지 특징을 퍼지 추론 규칙에 적용한 결과, 정상 세포핵과 암종 세포핵은 199개 세포핵 중에서 190개가 정확히 인식되었고, 9개의 세포핵은 오인식 되었다.

향후 연구 과제는 여성의 월경 주기와 자궁 경부진 세포핵의 성장 과정에서 세포핵의 형태가 변하는 특징을 분석하고, 중층 편평 세포에 한정하지 않고 자궁 경관과 자궁 표면에 대한 다양한 세포핵을 인식할 수 있는 방법에 대해 연구할 것이다.

참고문헌

- [1] 박관규, 권선영, 제갈문영, 한상미, 장은숙, 차순도, 한인숙, "자궁경부암종에서 암 전이의 조기 발견을 위한 면역조직화학 연구," 대한병리학회지, Vol.35, pp.391-400, 2001.
- [2] 은성경, 박찬모, 박화춘, 김성숙, "A Study on an Automatic Detection of Nucleus from Pap-smears by Image Processing," 한국정보과학회, 학술발표논문집 Vol. 20, No. 1, 1993.
- [3] K. B. Kim, D. H. Song, Y. W. WOO, "Nucleus Segmentation and Recognition of Uterine Cervical Pap-Smears, Lecture Notes in Artificial Intelligence, LNAI 4482, pp.153-160, 2007.
- [4] 김광백, "Kapur 방법과 형태학적 특징을 이용한 자궁경부암 세포 추출 및 인식," 한국해양정보통신학회 논문지, 11권 10호, pp.1992-1998, 2007.

- [5] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, "A New Method for Gray Level Picture Thresholding Using the Entropy of the Histogram," Computer Vision, Graphics, and Image Processing, Vol.29, pp.273-285, 1985.
- [6] 김미숙, 정상우, "진염색질-이염색질 경계면에서의 mRNA 합성: 경계면 이론의 제안," 대한병리학회지, Vol.35, pp.93-97, 2001.
- [7] R. Babuska, Fuzzy Modeling For Control, Kluwer Academic Publishers, 1998.

저 자 소 개



김 광 백
 1999년 부산대학교 전자계산학과 (이학박사)
 1997년~현재 신라대학교 컴퓨터정보공학부 교수
 2005년~현재 한국해양정보통신학회 이사 및, 논문지 편집부위원장
 2005년~현재 한국멀티미디어학회 이사 및 논문지 분과 편집위원장 및 편집위원
 <관심분야> Neural Networks, Image Processing, Fuzzy Logic, Medical Imaging and Biomedical System, Support Vector Machines



송 두 현
 1981년 서울대학교 계산통계학과 졸업
 1983년 KAIST 전산학과 졸업
 1994년 미 UC Irvine 전산학과 박사 수료
 2004년~현재 한국멀티미디어학회 부회장 및 한국특허학회 감사
 1997년~현재 용인송담대학 컴퓨터계임정보과 교수
 <관심분야> Image processing, ITS, Machine Learning, Data mining, Game software development process, Internet security