

혈액영상에서 병리진단을 위한 적혈구 세포의 자동분류에 관한 연구

김 경 수*, 김 동 현**

A Study on Automatic Classification System of Red Blood Cell for Pathological Diagnosis in Blood Digital Image

Kyung-Su Kim*, Dong-Hyun Kim**

요 약

의학분야에서 컴퓨터는 병원에서 발생하는 각종 업무데이터의 전산화에서 진단을 위해 사용하는 검사 의료기기들의 자동화, 그리고, 각종 의학영상들을 디지털화하여 처리하는 단계까지 활발하게 활용되고 있는 실정이다. 이러한 시점에서 본 논문에서는 병원의 임상병리과에서 늘어나고 있는 혈액검사를 자동화하기 위한 것으로 혈구영상으로부터 적혈구를 분석하여 정상세포를 비롯한 비정상세포를 16부류로 나누어 분류하였다. 이를위해 UNL푸리에 특징과 불변 모멘트 알고리즘을 사용하여 세그먼트된 적혈구 영상으로부터 특징을 추출하고 이를 인식하기 위한 다단계 신경망을 구축하였다. 실제 임상에서 10명의 환자를 대상으로 실험한 결과 검사자가 참조가능 형태의 결과를 얻을 수 있었다.

Abstract

In medical field, the computer has been used in the automatic processing of data derived in hospital, the automation of diagonal devices, and processing of medical digital images. In this paper, we classify red blood cell into 16 class including normal cell to the automation of blood analysis to diagnose disease. First, using UNL Fourier and invariant moment algorithm, we extract features of red blood cell from blood cell image and then construct multi-layer backpropagation neural network to recognize. We proof that the system can give support to blood analyzer through blood sample analysis of 10 patients.

* 메디정보

** 순천청암정보대학 정보처리학과 교수

논문접수 : 98. 12. 28. 심사완료 : 99. 2. 13.

I. 서론

최근 외국은 물론 국내에서는 정보통신망의 빠른 확산과 영상처리기술 및 컴퓨터 그래픽기술의 발전으로 인해 개인병원이나 종합병원 등에서 생성되는 화상들을 디지털화하여 관리 및 처리하려는 노력들이 활발히 진행되고 있다[2][7].

현재 점점 증가세를 보이고 있는 건강검진등으로 인해 각 병원의 임상병리실에서는 혈액검사의 요구가 많아지고 있어 임상병리사들의 현미경 접안을 통한 검사 작업의 수가 날로 늘어나고 있는 실정이다. 따라서, 환자들과 건강 검진자들에게 혈액검사의 효용성을 높이고 보다 정확하고 신속한 검진결과를 제공하기 위해서는 지금과 같이 수동적으로 처리하는 것보다는 전문가의 고도의 경험이 필요치 않은 부분에 대해서는 진단 자동화를 도입할 필요성이 있다. 또한 기존의 혈구세포에 대한 단순한 카운트 기능을 해주는 시스템은 전적으로 수입에 의존하고 있고 아주 고가여서 일반 개인병원에서는 도입할 수 없는 상태이다. 이를 위해 본 논문에서는 영상처리기법을 이용하여 현미경으로부터 획득된 혈구영상에서 적혈구세포를 분리하고 정상세포를 포함한 여러 종류의 비정상세포를 분류하였다. 이를 통해 임상병리사는 물론 의학도 그리고, 혈액 전문가들의 혈액 분석을 지원할 수 있는 진단지원 시스템의 구축 토대를 마련하였다.

II. 혈구세포

말초혈액은 액체성분인 혈장(plasma)과 적혈구, 백혈구, 혈소판 같은 세포성분으로 구성되어 있다. 이

모든 세포성분은 미분화된 간세포인 혈구조각아구에서 유래된다. 이 초기 원시세포들은 조혈 조직내에서 혈구아구(haemocytoblast)로 분화되고 분지계(clone)로 나누어져 적혈구계, 골수계 그리고 거핵구계열의 세포로 성숙되는데 아래 그림은 혈구조각아구에서 분화되어 가는 순서를 나타내 주고 있다[9][10].

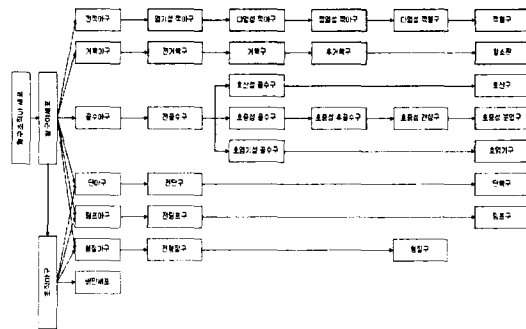
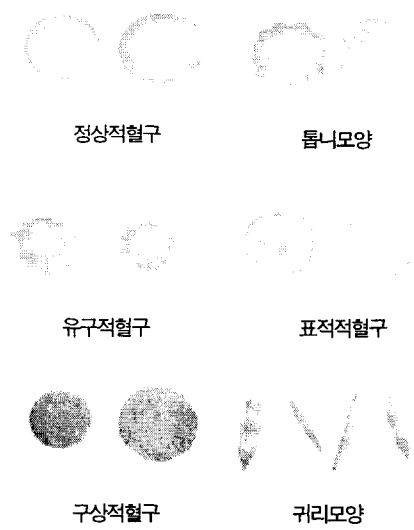


그림 1. 혈액세포의 조혈조직
Fig1. Hematopoietic tissues of blood cells

본 논문에서는 혈구 세포중 정상세포를 포함한 비정상 세포를 16종류로 분류하고 나머지 다른 형태의 적혈구에 대해서는 그 부류에 속하는 여러 변형들로 인해 자동 분류가 불가능하기 때문에 본 연구에서는 자동인식을 하기위한 부류로 포함하지 않았다[9]. 16부류에 대한 적혈구 세포는 아래 그림2와 같다.



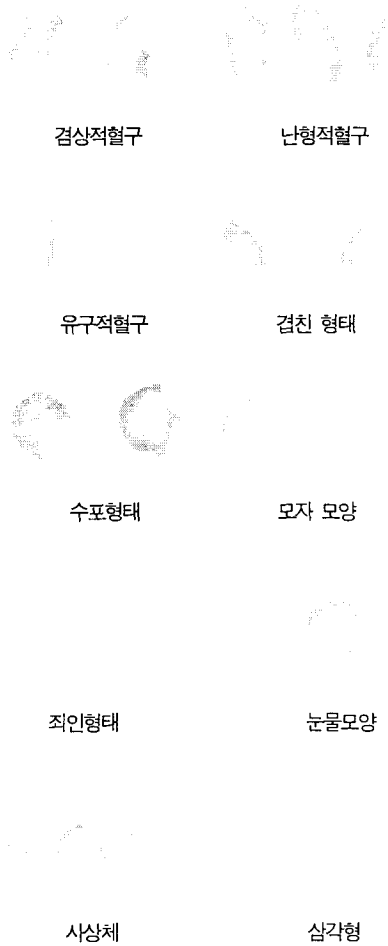


그림2. 적혈구의 형태
Fig2. Shape of red Blood Cells

Ⅲ. 혈구영상데이터 획득

본 논문에서는 주로 성숙된 적혈구 영상을 대상으로 적혈구의 병적 상태를 분류하고자 하였다.

1. 전처리

혈액영상을 자동 분석·진단을 하기 위해서는 먼저

현미경으로부터 혹은 혈액 영상을 스캔한 영상을 입력 받는다. 본 논문에서는 현미경이 부착된 컬러 CCD 카메라로부터 400배 확대된 입력영상을 프레임② ③ 그 래버를 통해 캡처한 영상을 사용하였다. 입력③ 된 ④ 영상에 대해 가장 먼저 하게되는 처리는 혈④ 구를 분 리하기 위한 전처리로서 본 논문에서는 임계치를 결정 하여 이진화 영상으로 변환시켰다. 이 때 혈액영상 이 진화를 위해 적용한 알고리즘은 기존의 이진화 알고리 즘 중 본 영상에 가장 좋은 성능을 보인 원래의 그레 이 영상과 이진화된 영상과의 차이를 나타내는 퍼지 척도를 이용한 방법을 이용하였다[4]. 아래 그림3과 4 는 입력영상과 이진화된 영상을 보여주고 있다.



그림3. 입력영상
Fig3. Input Image



그림4. 이진화된 영상
Fig4. Binarized Image

다음 단계에서는 세그먼트된 혈구세포들에 대해 레이블링을 하게 된다. 아래 그림5는 각각의 혈구세포들이 레이블링되어 최소경계사각형으로 표시된 결과를 나타내고 있다[5].

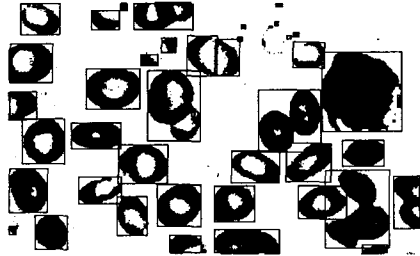


그림5. 레이블링된 영상
Fig5. Labeled Image

레이블링 단계를 거친 후 각각 분리된 혈구세포들 중에서 적혈구, 백혈구, 혈소판, 혈장 등으로 분리하는 단계가 필요하다. 본 논문에서는 인식할 대상이 적혈구이기 때문에 적혈구를 분리해야 하는데 일반적으로 400배 확대된 영상 시야내에서는 백혈구가 가장 큰 용적을 차지하고 있고 거의 2개 이하이기 때문에 백혈구를 분리해낼 수 있고 혈장은 말초혈액에서는 볼 수 없고 혈소판의 경우도 특수한 경우를 제외하고 적혈구나 백혈구에 비해 아주 작은 크기의 형태를 가지는 것을 볼 수 있다(그림3의 위 중간 부분). 그러나, 위와 같이 적혈구세포를 분리해 냈다 하더라도 혈액 샘플 슬라이드를 만들 때 두 개 이상의 세포들이 겹쳐지는 경우가 생기거나 혹은 병적인 원인으로 인해 여러 개의 적혈구 세포들이 뭉치게 되는 적혈구 응집(agglutination)과 연전현상(rouleaux formation)이 있는 경우에 적혈구 하나를 분리하는데 문제가 발생하게 된다. 이 경우 일반적으로 전문가가 뭉쳐있는 각각의 혈구세포들의 이상유무를 경험적으로 판단하여 분석을 하거나 혹은 분석에서 제외시키게 된다. 따라서, 자동진단을 위해서는 각각의 혈구세포들을 분석하기 위해 뭉쳐있는 혈구세포들을 분리하는 작업이 필수적이다. 본 논문에서는 위와 같은 혈구 영상은 보다 신뢰성 있는 분석을 위해 기각함으로써 전문가에게 일임하는 형태를 취했다.

2. 특징 추출

다음 단계에서는 혈구세포의 분류를 위해 각각의 혈구세포들에 대한 특징을 추출하는 단계가 필요하다. 본 논문에서는 혈구세포들의 형태특징을 추출하기 위

해 UNL 푸리에 방법을 사용하였고, 외부 형태가 정상과 같고 내부 형태가 다른 정상적혈구, 구상 적혈구, 표적적혈구, 유구 적혈구들은 모멘트 방법을 사용하여 특징을 추출하였다. 그리고, 정상적혈구와 난형적혈구의 경우 정면에서 보았을 때 세포의 중심부가 창백한데 중심부의 창백한 부위가 적혈구 직경의 1/3 정도일 때 정상인데 이의 양에 따라 저색소성과 과색소성 적혈구로 나누어진다. 본 논문에서는 적혈구 세포의 분류를 위해 UNL 푸리에 특징과 모멘트방법을 사용하였다.

2.1 UNL 푸리에 방법

UNL(Universidade Nova de Lisboa) 푸리에 방법은 푸리에 서술자(Fourier descriptor)의 확장 개선된 알고리즘으로 객체로부터 특징을 추출하기 위해 아래의 그림과 같이 크게 두 단계로 실행된다 [1][8]. 첫 번째 단계에서는 이진 곡선 패턴의 객체를 가지고 있는 영상이 카치교 좌표계(Cartesian coordinate system)에서 극 좌표계(polar coordinate system)로 변환된다. 이를 위해 이상데이터인 영상에 대해 해석적 곡선 방정식이 구하고 UNL 변환을 거친 후 극 좌표계 상에 곡선 방정식이 놓이게 된다. 따라서, UNL 변환은 원래의 패턴에 대해 이동 및 회전연산에 대해 불변인 특성을 갖게 된다.

또한, UNL변환의 순환 이동(cyclic shift)은 원래 영상을 회전한 결과를 얻을 수 있다. 이러한 과정은 두 번째 단계에서 이차원 푸리에 변환을 통해 생략할 수 있다. 즉, 푸리에 변환의 스펙트럼은 원래 인식할 객체 패턴에 대한 회전에 영향을 받지 않는다. 결과적으로 이러한 UNL 푸리에 특징은 인식 객체에 대한 회전, 이동, 축소 및 증가에 영향을 받지 않는다.

$\Omega(t)$ 를 n개의 화소 $z_i=(x_i, y_i)$ 로 구성된 이산 객체, $O=(O_x, O_y)$ 를 객체의 중심, M을 중심좌표 Ω 로부터 모든 화소 z_i 로의 대 유클리디언 거리라고 가정하자. 객체를 구성하는 두 인접화소 $z_i=(x_i, y_i)$ 와 $z_j=(x_j, y_j)$ 사이의 라인 세그먼트 $z_{ij}(t)$ 로 구성된 이산 객체 $U(\Omega(t))$ 의 UNL 변환은 다음과 같은 식에 의해 직교 좌표계로부터 극좌표계로의 매핑하는 결과가 된다.

$U : ((0, 1) \rightarrow R^2) \rightarrow ((0, 1) \rightarrow R^2)$
 여기서, $\forall 1 \dots n \quad t \in [0, 1]$

$$U(z_{ij}(t)) = \zeta_{ij} = (E_{ij}(t), \theta_{ij}(t))$$

$$= \left(\frac{\|z_i + t(z_j - z_i) - O\|}{M}, \right.$$

$$\left. \text{atan} \left(\frac{y_i + t(y_j - y_i) - O_y}{x_i + t(x_j - x_i) - O_x} \right) \right)$$

$i(x, y)$ 를 이산 객체 $\Omega(t)$ 를 나타내는 2차원 영상, $f(R, \theta)$ 를 $\Omega(t)$ 의 UNL 변환을 나타내는 2차원 영상이라고 가정하면 영상 $f(R, \theta)$ 의 정규화된 이산 푸리에 스펙트럼은 다음과 같다.

$$UFF(u', v') = \frac{\|F(f(R, \theta))\|}{F(0,0)}$$

$$= \frac{\|F(u, v)\|}{F(0,0)}$$

UNL 푸리에 특징을 추출하기 위해 세그먼트된 적혈구 세포의 경계선을 구하고 64×64 크기로 스케일링을 행한 후 각각에 대해 모두 76개의 UNL 푸리에 특징을 추출하였다.

2.2 모멘트 방법

모멘트 방법은 2차원 형태를 기술하는데 많이 사용되고 있는 알고리즘으로 어떤 함수 $f(x, y)$ 의 차수 (p, q) 에 대한 2차원 모멘트는 다음 식과 같이 정의된다[3].

$$m_{pq} = \int \int x^p y^q f(x, y) dx dy$$

첫 번째 모멘트 μ_{00} 는 m 에 의해 명시된다.

$x' = \mu_{10}/m$, $y' = \mu_{01}/m$ 으로 정하고 차수 (p, q) 의 중심 모멘트는 다음과 같이 정의된다.

$$\mu_{pq} = \sum \sum (x - x')^p (y - y')^q f(x, y)$$

η_{pq} 로 명시된 정규화된 중심 모멘트는 다음과 같은 식에 의해 정의된다.

$$\eta_{pq} = \mu_{pq} / (\mu_{00})^r$$

여기서, $r = \frac{1}{2}(p+q) + 1$, $p+q=2, 3, \dots$

2차와 3차 모멘트로부터 7개의 불변 모멘트가 구해진다. 그 식은 다음과 같다.

$$m_1 = \eta_{20} + \eta_{02}$$

$$m_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$m_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$m_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$m_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$m_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$m_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

형태 기술자로서의 이러한 불변 모멘트는 영상내의 인식대상인 객체에 대한 이동, 회전 축소, 확대 변환에 불변인 특징을 가지고 있다.

불변 모멘트를 구하기 위해 세그먼트된 적혈구 세포의 최소경계사각형을 구한 후 그 영상으로부터 불변 모멘트 특징을 추출하였다.

IV. 실험 및 결과

실험은 병원의 임상병리과에서 얻은 환자의 혈액 샘플과 혈액학 도감등을 활용하여 혈구의 분류를 위해 역전파 학습을 이용한 한 층의 은닉노드를 가진 다단계 신경망을 사용하였다[6]. 신경망은 UNL 푸리에 특징의 입력인 76개의 입력노드, 은닉노드는 120노드, 출력노드는 13노드로 구성되어 있다. 이때 출력이 13

인 이유는 UNL 푸리에 특징은 객체의 인식에 있어 경계선을 이용하게 된다. 따라서, 표적적혈구, 구상적혈구, 유구적혈구의 외곽선 형태는 정상 적혈구와 같기 때문에 이들 혈구들은 모멘트 특징을 이용하여 분류하였다. 학습데이터의수는 16부류의 적혈구에 대해 각각 50개씩 900개의 데이터를 사용하였다.

실제로 병원에서 환자의 혈액을 검사할 때 현미경을 통해 10개의 시야를 보게되는데 본 논문에서는 적혈구 분류 시스템의 성능을 평가하기 위해 실제로 환자10명으로부터 각각 10시야의 영상을 입력받아 적혈구 분석을 행하였다.

임상에서 실제로 접하게 되는 환자의 비정상 혈구 형태들을 여러 종류의 비정상 형태들이 보이지 않고 주로 몇 가지 형태의 세포형태가 주를 이루고 있다. 표1은 본 실험에서 검증을 위해 사용한 환자의 혈액 샘플에 대한 주요 형태를 나타내고 있다.

표2는 각 환자들에 대해 본 시스템이 기각하거나 인식하지 못한 혈구의 형태를 나타내고 있다.

표1과 1에서 볼 수 있듯이 외곽선 형태의 특징이 어느 정도 일관된 정상적혈구, 톱니모양 적혈구, 유구적혈구, 난형 적혈구 등은 모든 환자의 혈구 샘플에서 인식하였다. 그러나, 형태가 일관되지 않은 귀리모양, 겸상적혈구, 수포모양 적혈구들은 환자에 따라 학습 데이터와 차이가 많이 나는 경우 인식하지 못했다. 그리고, 모멘트를 이용하여 정상적혈구와 외곽선 형태가 같은 혈구들중에 정상적혈구와 유구적혈구의 오인식이 발생하였다.

표1. 혈액샘플의 주요모양
Table1. Principal shape of blood sample

환자의 혈액 샘플	주요 혈구 모양
1	정상, 톱니모양
2	구상적혈구, 난원적혈구
3	겸상적혈구
4	난원적혈구, 귀리모양
5	톱니모양, 수포모양
6	모자모양, 분열적혈구
7	표적적혈구, 결정(SC, CC)
8	겹친모양, 분열적혈구
9	정상적혈구, 유구적혈구
10	죄인형태, 겹친형태

V. 결론 및 향후 연구과제

본 논문에서는 혈액영상을 자동으로 분석하여 혈구 세포들의 수를 카운트하고 이상유무를 판별하여줌으로써 병리 진단에 도움을 줄 수 있도록 하는 적혈구 자동인식을 수행하였다. 본 논문에서 수행한 적혈구 인

표2. 혈액샘플중 오인식된 혈구모양과 갯수
Table2. Shape and number of false recognition

환자의 혈액 샘플	실제혈구(오인식된 혈구형태-갯수)
1	정상(유구-4)
2	구상(정상-2), 난원(정상-2)
3	겸상(귀리-5)
4	정상(난원-2)
5	톱니모양(유구-1), 수포(모자모양-1)
6	모자모양(수포-1), 분열적혈구(분류기각-검사자에게 일임)
7	결정(분류기각)
8	분열적혈구(분류기각)
9	모두 인식
10	죄인형태(눈물방울양-1), 겹친형태(표적-1)

식의 결과는 약간의 성능이 개선이 이루어지면 임상병리실에서 충분히 참조 가능한 시스템임을 확인하였다. 그러나, 혈구세포의 변이를 인식하여 병리진단을 위한 작업은 적혈구뿐만 아니라 백혈구의 인식이 필요하다. 따라서, 향후 백혈구에 대해 검사자가 참조 가능하도록 인식알고리즘의 연구가 필요하고 기존의 검사 장비의 결과를 인터페이스 구축을 통해 얻은 결과를 종합적으로 분석할수 있는 진단 전문가 시스템 구축이 필요하다

참고 문헌

- [1] Babu M. Mehtre, Mohan S. Kankanhalli and Wing F. Lee, "Shape Measures For Content Based Image Retrieval : A Comparison", Technical Report95-195-0, Institute Of Systems Science, National University of Singapore, 1995.
- [2] Brent K. Stewart and Steve G. Langer, "Medical Image Databases and Informatics", IEEE Internation Conf. on image processing, Oct 4-7, Chicago, Illinois, pp.29-33, 1998.
- [3] C.C.Chen, "Improved Moment Invariants for Shape Discrimination", Pattern recognition, Vol.26, No.5, pp.683-686, 1993.
- [4] Huang L-K and M-J. J. Wang, "Image thresholding by minimizing the measures of fuzziness", Pattern Recognition, Vol.28, 1:41-51, 1995.
- [5] Ioannis Pitas, Digital Image processing Algorithms, Prentice Hall, 1993.
- [6] James A. Freeman and David M. Skapura, Neural Networks: Algorithms, Applications and programming Techniques, Addison-Wesley Publishing, 1991.
- [7] Michael W. Vannier and John W. Haller, "Biomedical Image Segmentation", IEEE International Conf. on image processing, Oct 4-7, Chicago, Illinois, pp.20-24, 1998.
- [8] T.W.Rauber, "Two-Dimensional Shape description", Technical Report GRUNI-

NOVA-RT-10-94, Universidade Nova de Lisboa, Lisboa, Portugal, 1994.

- [9] 고려의학 출판부편, 혈액학 도감, 고려의학, 1995
- [10] 조한익 편역, 실용 혈액학, 고려의학, 1995.

저자 소개



김경수

1992년 조선대학교 컴퓨터공학과 졸업
 1994년 중앙대학교대학원 컴퓨터공학과 졸업
 1996년~현재 조선대학교 대학원 전자계산학과 박사과정 재학중

1998년~현재 메디정보 대표



김동현

1987년 조선대학교 전산기공학과(공학사)
 1991년 광운대학교전산기공학과(공학석사)
 1998~현재 조선대학교 전자계산학과 박사과정
 1996~현재 순천청암대학 전산정보처리과 교수