

외산 HCS장비와 호환 가능한 세포 분석 시스템 개발의 기초 연구

나예지*, 이상준**, 왕창원*, 정화영*, 민세동*
 순천향대*, 선문대**

Basic research of foreign HCS equipment compatible with the cell analysis system

Ye-Ji Na*, Sang-Jun Lee**, Chang-Won Wang*, Hwa-Young Jeong*, Se-Dong Min*
 Soonchunhyang University*, SunMoon University**

Abstract - 본 논문은 외산 HCS장비에서 획득한 세포 이미지를 기반으로 다기능 분석 소프트웨어 개발을 위한 기초 연구로서 세포의 외곽선 검출 알고리즘을 제안한다. 기존 체인코드 알고리즘의 탐색 횟수를 줄여 ROI 검출 속도 및 전체 시스템의 성능을 높이고자 하였다. 기존 알고리즘과 제안한 알고리즘을 C언어로 구현하여 실제 영상에 실험해 본 결과 정확한 세포의 ROI 검출뿐만 아니라 속도 향상의 효과를 확인하였다.

세포 분석을 위한 시스템은 소프트웨어적으로 세포영상에서의 잡음을 제거하기 위한 전처리, 이진화, 윤곽선 추출 알고리즘 등의 개발이 요구된다. 아래 그림. 1에 본 시스템의 처리 절차를 나타내었다.



〈그림 1〉 시스템 처리 과정

1. 서 론

최근 제약 및 생명공학 분야에서는 세포 현상에 대한 연구와 신약 개발을 위해 초고속 이미지 분석(HCS, high contents screening) 장비가 많이 사용되고 있다. 연구자들은 HCS 장비를 통해 수 백개에서 수 천장의 세포를 촬영하며 동시에 세포 분석을 수행한다. 이로 인해 방대한 양의 생물학적·화학적 정보를 획득할 수 있으며 획득한 정보를 바탕으로 부적절한 화합물을 배제하여 효과적인 약을 개발할 수 있다는 장점이 있다. 하지만 정확한 세포 분석을 위해서는 숙련된 연구원이 수작업으로 영상을 해석하여야 하고, 이는 수없이 발생하는 영상을 해석하기에는 많은 시간과 노동력을 소비하는 단점이 있다. 또한 장기간 촬영 시간 실험에만 HCS 장비가 독점되는 상황이 발생하며 분석의 정확도 면에서도 전문가마다 상이할 수 있다. 따라서 세포학적 분석을 위한 S/W 개발의 필요성이 대두되고 있다. 국내에서는 HCA(high content analysis) 관련 기술을 보유하고 있으나 실제 분석을 위한 응용프로그램의 개발은 시도되지 않은 것으로 파악된다. 또한 외산 HCS 장비의 경우에는 고가로 판매되고 있어, 많은 소규모 연구실의 사용자들에게는 활용도가 적다.

본 연구의 최종 목표는 외산 HCS 장비에서 획득한 이미지를 기반으로, 실제 장비의 여러 기능들을 탑재한 세포 분석 소프트웨어를 개발하는 것이다. 여러 가지 분석 모듈은 기본적으로 세포의 Counting, Area, Intensity 정보를 포함하며, 이러한 정보를 얻기 위해서는 우선적으로 정확한 ROI 검출이 필요하다. 또한 한 실험에서 발생하는 수많은 이미지 분석을 위해 높은 처리량은 물론 빠른 속도로 ROI를 검출하는 것이 중요하다. 이에 따라 1차적으로는 최종 목표인 세포 분석 시스템 개발을 위한 최적화된 외곽선 검출 알고리즘을 도출하고자 한다.

세포의 외곽선 탐색을 통해 세포를 배경 이미지로부터 분리할 수 있고, 세포의 크기 및 형태를 추정할 수 있다. 기존의 외곽선 추적 알고리즘(contour tracing algorithm)으로 체인코드 알고리즘이 있다. 체인 코드는 Freeman이 제안한 경계 추출 알고리즘으로 미리 정의된 길이와 방향을 가지는 직선 성분들이 연결된 열로서 경계를 표현하는 방식이다. 체인 코드의 주요 단점은 잡음이나 불완전한 분할로 인해 경계를 따라 생긴 작은 교란에 의해 경계 추출에 불필요한 변화가 발생한다는 점이다. 또한 이웃 픽셀 검사를 적용적으로 하지 못해 탐색에 많은 시간이 소요된다는 단점이 있다.

따라서 본 연구에서는 개선된 체인코드의 방법으로 세포의 외곽선을 추출하는 방식을 제안하고, 이를 영상에 적용하여 ROI 검출에 소요되는 시간과 연산량을 단축시켜 전체 시스템의 성능에 긍정적인 영향을 끼치고자 한다.

2. 본 론

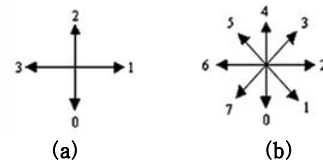
2.1 세포 분석 시스템 소개

본 연구에서는 몇 가지 영상처리 알고리즘을 적용하여 세포분할과 세포학적 분석에 사용되는 중요 정보를 추출해 내는 프로그램을 구현하였다. 본 시스템을 개발하기 위한 환경으로 운영체제는 Windows 7 64bit OS를 사용하였고, 개발 툴은 NI(National Instrument)사의 LabWindows CVI 2013 환경과 OpenCV 2.4.9버전을 사용하였다. 프로그램의 사용 언어는 C로 코딩하였다.

본 시스템에서는 외산 HCS 장비 중 Cellomics사의 ArrayScan 장비로 촬영하여 얻은 512X512 크기의 16BIT 그레이 영상을 입력받아 영상의 잡음을 제거를 위한 가우시안(Gaussian) 필터를 적용한 후, 오투(Otsu)의 방법으로 이진화 처리를 하였다. 배경과 세포가 0과 1의 값으로 이진화된 Binary영상에 본 연구에서 제안한 방법으로 세포의 외곽선을 추출하였다. 이를 통해 세포의 개수를 counting하고, 세포의 Area와 Intensity값을 구하도록 하였다.

2.1.1 세포 외곽선 추출 알고리즘

체인코드 알고리즘은 이진영상을 입력으로 하여 위에서 아래로, 좌에서 우로 탐색하여 처음 외곽선이 시작되는 위치를 찾는다. 각 외곽선 성분들의 4-연결성 또는 8-연결성에 기반을 두고 시계 방향이나 반시계 방향으로 경계를 따라 모든 화소 쌍을 연결하는 선분에 대해 각 방향 벡터를 할당한다.



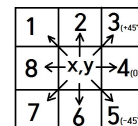
〈그림 2〉 프리만 체인코드의 방향 벡터 (a) 4-연결성, (b) 8-연결성

$$n' = (n + 5) \& 7 \quad (\text{식 1})$$

일반적으로 외곽선의 정확한 추출을 위해 8-연결성을 주로 사용한다. 위 알고리즘은 외곽선이 시작되는 화소 점의 위치로부터 다음에 있을 수 있는 8개의 인접 방향에 대하여 시계방향(또는 반시계방향)으로 화소를 탐색한다. 방향 벡터 0~7은 전체 360°의 단위원으로 생각했을 때 서로 45°씩 차이가 나며, 식 1에 의해 다음에 탐색할 방향 벡터는 이전 추적 방향의 135°를 벗어나지 않는다. 이전 위치에서 135° 이상의 방향 성분들은 이미 윤곽선 여부를 조사했기 때문이다. 이는 이전 방향의 135° 방향부터 정해진 한 방향으로 이웃 픽셀들을 순회하며 외곽선 여부를 조사해야 한다는 것을 의미한다. 그러나 실질적으로 본 시스템의 세포영상에서 세포 외곽선 방향성분의 변화는 ±45°를 크게 벗어나지 않는다. 이와 같은 점을 고려하여 본 연구에서는 이전 방향에 적응적인 체인코드 알고리즘을 제안한다.

2.2 개선된 체인코드 알고리즘

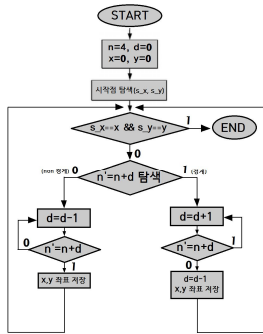
제안한 알고리즘에서는 이전 방향의 외곽선 여부에 따라 다음 탐색 방향을 선택하도록 하였다.



〈그림 3〉 제안한 체인코드의 방향 벡터

$$n' = n + d \quad (\text{식 } 2)$$

제안한 알고리즘에서는 현재 위치 좌표 (x, y)에서 우선적으로 탐색할 방향이 이전에 이동했던 방향이다. 이 값을 조사하여 1(경계)이면 이전 방향에서 +45°방향의 화소 점을 조사하고, 0(Non-경계)이면 이전 방향에서 -45°방향의 화소 점을 조사한다. 본 알고리즘에서는 식 2와 같이 다음 탐색 방향 n'은 이전 방향벡터 n에 d의 값을 더한 것으로 식을 표현했다. 변수 d(direction)의 값은 이전 방향 n의 외곽선 여부에 따라 ±45° 방향으로 이동 방향을 결정한다. 이와 같이 설정한 이유는 촬영된 세포의 형태학적 특징이 타원에 가까운 모양을 가진 데이터가 대다수이기 때문이다[1,2]. 따라서 이전 방향에서 다음 탐색 방향으로의 각도 변화가 크지 않을 것이다. 아래에는 제안한 외곽선 탐색 알고리즘의 흐름도를 나타내었다.

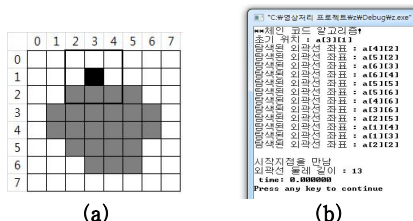


〈그림 4〉 제안한 외곽선 탐색 알고리즘의 흐름도

원 영상에서 초기 경계 화소 점 s_x, s_y를 발견하면 이 점으로부터 4번 방향(0°)의 인접 화소를 먼저 탐색한다. 4번 방향에 경계가 있을 때는 경계 값이 없을 때까지 5방향, 6방향, 7방향 등의 +45°방향을 순차적으로 탐색하여 다음 이동 방향을 찾는다. 경계가 없을 때에는 경계 값이 탐색될 때까지 4방향에서 상대적으로 -45°방향인 3방향, 2방향, 1방향, 8방향 등의 순서대로 탐색을 한다. 진행 방향에 경계선에 해당하는 화소가 존재하면 해당 화소로 이동하여 재탐색을 하며, 본 과정은 재귀함수로 구현하였다.

2.2.1 기존 알고리즘과의 비교

제안한 알고리즘과 기존 체인코드 알고리즘의 비교를 위해 C언어로 직접 구현을 하였다. 먼저 0과 1의 값만을 가진 그림. 2(a)의 8x8 binary 영상으로 비교 실험을 해 보았다. 그림. 2(b)는 두 알고리즘을 적용하여 얻은 실행 결과 화면이다.



〈그림 5〉 (a) 8x8 입력 영상, (b) 알고리즘 실행 결과 창

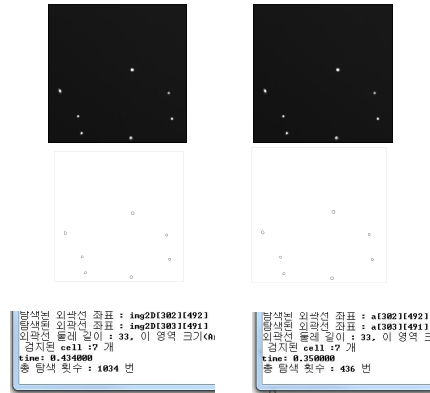
세포의 첫 시작점부터 연결된 외곽선 성분이 차례대로 검출되는 것을 확인할 수 있다. 위 영상에 적용된 기존 알고리즘과 개선된 알고리즘의 수행 속도는 0.00초가 걸렸으며, 보다 정확한 속도 차이를 계산하기 위해 50x50의 binary 영상을 입력영상으로 하여 다시 실험해 보았다. 결과적으로 기존 알고리즘의 수행 속도는 0.156초로 제안한 알고리즘의 속도 0.046초보다 약 0.11초가 더 걸리는 것으로 나타났다. 또한 이웃 픽셀의 총 탐색 횟수는 기존 알고리즘은 581회, 제안한 알고리즘은 236회로 2배 이상의 차이를 나타냈다.



〈그림 6〉 50x50 입력 영상. (a) 기존 알고리즘 실행 결과 창, (b) 제안한 알고리즘 실행 결과 창

제안한 알고리즘을 실험 영상보다 약 100배 큰 실제 512x512 세포영상에 직접 적용해 보았다. 원본 영상에서의 세포 개수 및 형태, 면적이 정확하게 검출됨을 육안으로 확인하였다. 또한 외곽선 검출을 위한 이웃 픽셀 탐색 횟수가 제안 알고리즘에서 436번으로 기존 알고리즘 1034번

의 연산량보다 약 2배 적은 것으로 나타났으며 수행 속도는 0.074초 차이가 났다.



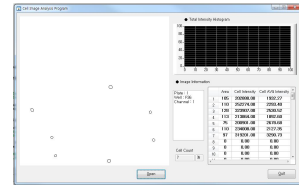
〈그림 7〉 위에서부터 세포 이미지 입력 영상, 외곽선 검출, 연산 비교 (a) 기존 알고리즘의 결과, (b) 제안한 알고리즘 결과

〈표 1〉 두 알고리즘의 비교

	이미지 크기	수행 속도	픽셀 탐색 횟수
기존 알고리즘	50x50	0.156 (s)	581 번
	512x512	0.434 (s)	1034 번
제안한 알고리즘	50x50	0.046 (s)	236 번
	512x512	0.350 (s)	436 번

3.1 개선된 체인코드를 적용한 소프트웨어

개선된 체인코드를 적용한 프로그램의 결과 화면을 그림. 7에 나타내었다. 결과로서 세포 counting, Area, Intensity의 정보를 얻을 수 있었다.



〈그림 8〉 구현된 세포 분석 프로그램

3. 결 론

본 논문에서는 기존의 외곽선 정보 추출 방법에 비해 수행시간이 빠르고 탐색횟수가 적어 메모리 낭비를 줄인 새로운 방법을 제안한다. 실제 장비에서의 ROI 검출 속도는 Evotec사의 Opera장비에서는 24장에 30초가 걸리며, Cellomics사의 Arrayscan장비에서는 150장에 30초가 걸리는 것을 확인하였다. 본 연구에서 구현한 시스템은 1장의 영상에서 ROI 검출 속도가 평균 0.4초가 걸렸으며, 이로 미루어 보아 75장의 검출 속도가 30초가 걸림을 예측할 수 있다. 최종적으로 구현된 알고리즘을 통해 입력영상에서의 세포 영역추출, cell counting 및 Area와 Intensity를 확인할 수 있었다. 앞으로 이러한 정보를 기반으로 세포 주기, 독성, Mitotic Index 등의 생물학적 분석을 위한 알고리즘을 개발하여 다기능 소프트웨어를 제작할 계획이다. 본 시스템은 사용자 신뢰성 검증 및 필드테스트를 필요로 하며, 차후에 검증이 연구되면 병리진단 및 신약개발에 도움이 되는 소프트웨어가 될 것으로 기대된다.

Acknowledgements

이 논문은 2014년 교육부와 한국연구재단의 지역혁신 창의 인력양성사업의 지원을 받아 수행된 연구임(NRF-2014H1C1A1066998).

[참 고 문 헌]

[1] Xiangzhi Bai, Changming Sun, Fugen Zhou, "Splitting touching cells based on concave points and ellipse fitting", Pattern Recognition 42, pp. 2434-2446, 2009.
 [2] Sonal Kothar et al, "Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques," IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009.