

디지털 영상 세포 측정법에 기반한 세포핵의 3차원 정량적 분석

김태윤[†], 최현주^{**}, 최흥국^{***}

요 약

암세포 조직 영상 분석에서 유효한 특성값 추출은 암세포 등급별 분류를 위한 중요한 과정이다. 본 논문에서는 디지털 영상 세포 측정법 기반 세포핵의 3차원 정량적 분석 방법을 제안한다. 먼저 공초점 현미경을 사용하여 신세포암의 각 등급별 3차원 볼륨 데이터를 획득하고, 지도학습 방법을 기반으로 슬라이스 영상의 화소의 컬러 특성값을 이용하여 세포핵을 분할했다. 세포핵의 3차원 가시화를 위해, 윤곽선 기반 표면 렌더링과 3차원 텍스처 사상 방법을 이용한 볼륨 렌더링을 수행했다. 이후 세포핵의 3차원 형태학적 특성값을 정의하고 추출했다. 어떠한 3차원 특성값이 진단 정보로 유용할 것인가를 평가하기 위해, 분산 분석을 이용하여 각 등급 간 3차원 특성값의 통계적 유효성을 분석했다. 마지막으로 추출한 특성값을 2차원 특성값과 비교하고 상관관계를 분석했다. 그 결과, 세포핵 등급과 3차원 형태학적 특성값 간의 유효한 통계학적인 차이를 확인했다. 제안한 방법은 정확한 진단과 예후 추정을 위한 새로운 등급 결정 시스템 개발을 위한 기반 연구로 활용될 수 있는 가능성을 보여주었다.

3D Quantitative Analysis of Cell Nuclei Based on Digital Image Cytometry

Tae-Yun Kim[†], Hyun-Ju Choi^{**}, Heung-Kook Choi^{***}

ABSTRACT

Significant feature extraction in cancer cell image analysis is an important process for grading cell carcinoma. In this study, we propose a method for 3D quantitative analysis of cell nuclei based upon digital image cytometry. First, we acquired volumetric renal cell carcinoma data for each grade using confocal laser scanning microscopy and segmented cell nuclei employing color features based upon a supervised learning scheme. For 3D visualization, we used a contour-based method for surface rendering and a 3D texture mapping method for volume rendering. We then defined and extracted the 3D morphological features of cell nuclei. To evaluate what quantitative features of 3D analysis could contribute to diagnostic information, we analyzed the statistical significance of the extracted 3D features in each grade using an analysis of variance (ANOVA). Finally, we compared the 2D with the 3D features of cell nuclei and analyzed the correlations between them. We found statistically significant correlations between nuclear grade and 3D morphological features. The proposed method has potential for use as fundamental research in developing a new nuclear grading system for accurate diagnosis and prediction of prognosis.

Key words: Digital Image Cytometry(디지털 영상 세포측정법), 3D Feature Extraction(3차원 특성값 추출), 3D Visualization(3차원 가시화)

1. 서 론

진단병리학에서 등급화(grading)란 암세포들의 비정상 정도에 따라 암의 진행정도를 결정하기 위한 하나의 분류 체계이며 등급화가 중요한 이유는 암의 확실적인 성장률과 전이 경향에 관한 정보들 제공하여 임상 치료방법 결정에 중요한 역할을 하기 때문이다.

1932년, Hand와 Broder에 의해 신세포암(renal cell carcinoma)의 등급 별 분류가 환자의 생존율과 관련이 있다고 보고된 이후 다양한 등급 분류 시스템이 제안되었으나[1] 1971년, Skinner 등이 세포핵의 특성과 생존율 사이에 상관관계가 있다는 것을 발표하면서[2], Fuhrman 등에 의해 세포핵의 특성값을 이용한 등급 분류 시스템이 제안되었는데 암의 진행 정도에 따라 4가지 등급으로 분류한다[3].

그러나 현재 Fuhrman 등급 분류 시스템을 기반으로 하는 육안에 의한 신세포암의 등급 결정은 동일한 등급 분류 시스템을 사용함에도 불구하고 서로 다른 분석자 사이에 의견이 같지 않을 수도 있고, 또한 같은 분석자도 반복 분석했을 경우 서로 다른 결과를 나타낼 수 있다. 따라서 이러한 차이를 줄이고 객관적인 분석을 하기 위해 컴퓨터를 이용한 디지털 영상 세포 측정법(digital image cytometry)이 연구되고 있다. 디지털 영상 세포 측정법은 컴퓨터를 기반으로 세포의 특성을 측정하는 방법으로, 비연속적으로 이산화 된 화소(pixel)로 구성된 2차원 디지털 영상을 이용하게 된다. 각 화소의 색상은 이산화 된 수에 의해 기술될 수 있는데, 흑백 영상의 경우, 각 음영들은 특정 화소의 밝기 값, 또는 흑백 정도에 대응하는 수로 표현된다. 또한 컬러 영상의 경우 RGB 밴드별로 독립적으로 이산 흑백 영상으로 표현되며, 동화상은 각 시간 프레임 단위로 이산 영상으로 구성할 수 있다[4].

그러나 2차원 영상 기반 분석 방법의 문제점은 조직의 절편화 과정으로 인해 세포핵의 일부분만이 포함된 얇은 조직 절편 영상을 이용하므로 본질적으로

3차원 구조를 가지는 세포핵의 정확한 크기와 형태에 대한 특성을 정량화하기에는 한계를 가지고 있다는 점이다. 예를 들어, 타원형의 세포핵이 절편 위에 가로로 놓여있으면 측정된 세포핵의 크기는 크게 나타나며, 이와는 반대로 같은 세포핵임에도 불구하고 절편 위에 세로로 놓여있으면 상대적으로 세포핵의 크기는 작아지게 된다. 즉, 세포핵의 놓인 위치와 절단 각에 따라 세포핵의 크기와 형태가 다르게 측정되어지는 것이다. 그러므로 2차원 영상을 기반으로 하는 특성값의 분석은 정확한 결과를 도출하기 힘들다.

이러한 문제들은 물리적인 절편화 과정 없이 연속적인 2차원 슬라이스들을 얻을 수 있는 공초점 현미경(confocal laser scanning microscopy)을 이용해 해결할 수 있다. 그 이유는 공초점 현미경에서 획득한 볼륨 데이터로부터 3차원 재구성과 3차원 특성값의 정량화가 가능하기 때문이다.

본 연구는 신세포암 세포핵 영상의 컴퓨터 기반 등급분류 시스템 개발을 목적으로 하며, 이를 위해 공초점 현미경 영상과 디지털 영상 세포 측정법에 기반을 둔 세포핵의 3차원 정량적 분석 방법을 제안한다.

본 논문은 다음과 같이 구성되어 있다. 먼저 2장에서는 관련된 기존 연구들과 문제점을 간단히 소개한다. 3장에서는 세포핵의 정량적 측정을 위해 제안된 연구 방법들을 소개하고, 4장에서는 실제 임상 데이터를 이용한 측정 결과와 제안된 방법의 유효성을 평가한다. 마지막으로 5장에서는 간단한 고찰과 함께 마무리 한다.

2. 관련연구

디지털 영상 세포 측정법의 다양한 알고리즘을 이용해 실제로 임상에 적용하기 위한 연구가 지속적으로 이루어져 왔으나 자동화와 정확성을 고려한 최적의 방법 개발에 대한 필요성은 여전히 높은 실정이다.

Lockett와 Herman 등은 디지털 영상 세포 측정법

* 교신저자(Corresponding Author) : 최홍국, 주소 : 경남 김해시 어방동 인제대학교 (621-749), 전화 : (055)320-3437, FAX : (051)322-3107, E-mail : cschk@inje.ac.kr
접수일 : 2006년 12월 4일, 완료일 : 2007년 5월 2일

† 준회원, 인제대학교 전산학과
(E-mail : liminus@paran.com)

** 준회원, 부산대학교 의학전문대학원 BK21 고급의료인력양성사업단 (E-mail : hjchoi@pusan.ac.kr)

*** 중신회원, 인제대학교 전산학과

* 이 논문은 2006년도 정부재원(학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2006-311-D00840).

을 이용해 전립선암 조직의 형광 염색된 세포핵의 자동 검출방법 연구에서, 군집된 세포핵에 대해 핵의 분할 경로를 추정하고 각각의 독립된 세포핵의 형태를 평균화하기 위해 영역 근접도를 조사했다[5]. 그 결과, 군집된 경우와 독립된 경우 모두 90% 이상의 높은 검출율을 얻었다. 그러나 다해상도 영상 분석과 중첩된 개체의 수가 많은 50 μ m 이상의 두꺼운 조직에 대해서는 방법의 한계성을 나타내고 있다. Sears와 Duckworth 등은 28개의 파라미터를 추출하고 다변량 통계 분석을 통해 채장암 세포 표본에 대한 판별 도구로서의 영상 세포 측정법의 잠재력을 확인했다[6]. 그러나 파라미터의 수로 인한 복잡도가 높은 편이며, 유효성 검증에 초점을 맞추어 실제 임상에 활용도는 낮은 편이다. Nederlof와 Flter 등은 형광 현미경으로부터 획득한 형광 교잡법(fluorescence in situ hybridization; F.I.S.H) 영상의 신호를 정량화하기 위한 연구를 수행했다[7]. 형광 교잡법은 개별 세포로부터 특정한 DNA나 RNA의 서열을 검출하기 위해 사용되는 방법이다. 이 연구에서는 밝기 값이 고르지 못한 배경으로부터 신호를 정확하게 분할해 내기 위해 간단한 필터가 적용되었는데 95% 이상의 높은 정확도를 보여주었다. 그러나 현미경 대물렌즈의 선택에 따라 신호의 민감도가 달라지므로 최적의 필터를 적용하기 쉽지 않은 단점이 있다. 또한 기존 연구들이 지닌 공통적인 문제점은 분석 과정에 여전히 전문가의 육안 검사에 의존하는 정도가 크다는 점이다.

공초점 현미경을 이용하여 3차원 데이터로부터 어떻게 보다 객관적인 정보를 얻을 것인가에 대한 연구 또한 다양하게 이루어지고 있다. Lockett 등은 3차원 세포핵 데이터에서 핵의 분할을 위한 대화형 알고리즘을 연구하고 이를 두꺼운 조직 내부의 핵 영역에 적용했다[8]. Adiga 등은 3차원 조직병리 영상의 분할을 위해 영역기반 분할 방법을 연구했으며[9], Jeroen 등은 3차원 데이터의 윤곽선 기반 분할 알고리즘을 제안했다[10].

그러나 공초점 현미경 데이터를 이용한 대부분의 연구들은 주로 두꺼운 조직 표본의 볼륨 데이터로부터 핵을 분할하기 위한 연구에 초점을 맞추고 있다. 그리고 기존의 세포핵 3차원 분석 방법은 많은 시간이 걸리며 진단 데이터를 쉽게 해석할 수 없다. 더욱이 여러 연구들을 통해 암의 등급화를 위해 2차원

기반 디지털 영상 세포 측정법과 비교하여 3차원 기반 방법의 상대적인 유용성은 평가 되었지만 어떠한 방법들이 더 우수하거나 떨어지는가에 대한 지표가 없는 실정이다.

따라서 본 논문에서 제안하는 3차원 특성값 기반의 새로운 세포핵 등급화 시스템은 기존연구의 문제점을 줄이고, 세포핵의 3차원 특성값에 대한 기준을 개발할 수 있도록 할 것이다.

3. 연구방법

3.1 영상획득

실험을 위해 연세대학교 병리학교실로부터 획득한 8개의 신세포암 조직을 포르말린(10% neutral-buffered formalin)으로 고정시킨 뒤 파라핀으로 포매했다. 20 μ m 절편으로 잘라진 각 조직들을 RNase A(final concentration, 0.5mg/ml)를 포함하고 있는 Propidium Iodide(PI)로 염색한 후, 형광매체 위에 고정했다(DAKO, USA). 이 신세포암 조직들을 공초점 현미경(Leica Microsystems Ltd. Mannheim, Germany)을 이용해 630배 배율로 영상화했다. 또한 3차원 렌더링을 위해, 0.4 μ m의 간격으로 연속적인 2차원 광학 절편을 가장 위 표면에서 시작해 가장 아래까지 확장해 나가면서 볼륨 데이터로 재구성했다. 각 볼륨들은 50장의 슬라이스로 구성되었으며, 각 슬라이스들은 512 \times 512 해상도의 24비트 영상이었다.

3.2 세포핵 분할을 위한 분류기 생성

지도학습(supervised learning) 알고리즘을 기반으로 세포핵을 분할하기 위한 분류기를 생성하기 위해 마우스를 이용해 세포핵과 배경 두 영역 별로 화소들을 각각 200개씩 선택했다. 선택된 화소들 중 200개는 분류기의 학습을 위해 사용했으며, 나머지 200개는 분류기를 테스트하기 위해 사용했다. 대부분의 컬러 공간들은 크게 주 공간(primary space), 인식 공간(perceptual space), 독립 축 공간(independent axis space) 그리고 휘도-색도 공간(luminance-chrominance space)의 4개의 주 그룹으로 분류할 수 있는데, 주 공간 내의 RGB 공간으로부터 선형 혹은 비선형 변환을 사용하여 다른 종류의 컬러 공간으로 변환할 수 있다[11]. 이에 따라, 학습 데이터 집합의

각 화소들로부터 추출한 RGB 값을 이용해 HIS, I₁I₂I₃ 그리고 La*b* 값으로 변환했다[12-14]. 변환된 각 컬러 공간은 각 그룹의 대표적 컬러공간으로 총 12개의 컬러 특성값(RGB, HIS, I₁I₂I₃, La*b*)들을 추출했다.

세포핵을 분할하기 위한 분류기의 성능을 최대화하기 위해서는 특성값 선별 과정이 필요하다. 특성값 선별방법은 최선(optimal)의 방법과 차선(sub-optimal)의 방법으로 나눌 수 있다. 전자의 경우, 이용 가능한 모든 특성값들의 조합을 찾는 검색을 수행한다. 그러나 이 방법은 시간 복잡도가 높기 때문에 고차원 문제에 대해서는 사용할 수 없다. 차선의 방법은 비록 최적의 성능을 보장하지는 못하지만, 대부분의 경우에 최선의 방법에 거의 근접한 결과를 보이며, 속도 면에서는 더 우수하다. 차선의 접근방법은 일반적으로 크게 순차적 전진(sequential forward), 후진(backward), 단계별(stepwise) 선택법의 세 가지로 나뉜다. 본 논문에서는 특성값 선별의 방법으로 순차적 단계 선택 방법을 사용했다. 이 방법은 전진 선택법과 후진 선택법을 결합한 방법으로 각 단계마다 추가와 제거를 반복하면서 진행해 나간다[15]. 특성값 집합의 중요도를 평가하기 위해 선별 기준이 필요하게 되는데 이를 위해 윌크스의 람다(Wilks' Lambda) 값과 각 분류기의 오분류율을 사용했다. 윌크스의 람다 값은 클래스 간의 공분산 행렬과 전체 공분산 행렬의 행렬식의 비율이다. 이것은 식(1)과 같이 표현될 수 있다.

$$A^* = \frac{|W|}{|B+W|} \tag{1}$$

여기서 $W = \sum_i \sum_j (x_{ij} - \mu_i)(x_{ij} - \mu_i)'$ 가 되며

$B + W = \sum_i \sum_j (x_{ij} - \mu)(x_{ij} - \mu)'$ 가 된다. 만일 일변화된 분산의 비율이 너무 낮을 경우, 아무런 효과가 없다는 가설은 기각되며 우수한 특성값 집합일수록 작은 윌크스의 람다 값을 가진다. 표 1에서 보자. 바와 같이 본 논문에서는 다변량 분산분석의 결과, I₂R을 분류기를 생성하기 위한 최적의 특성값으로 선

별했다.

최적의 특성값이 선별된 후, 베이즈 결정 이론(Bayesian decision theory)의 최소 오류율(minimum error rate)을 기반으로 분류기 생성을 위한 판별함수 g_i 의 집합을 정의했다.

만약 다음의 조건이 만족된다면 분류기는 특성값 벡터 x 를 클래스 ω_i 에 할당한다.

$$g_i(x) = p(\omega_i | x) > g_j(x) = p(\omega_j | x) \quad \forall j \neq i$$

판별 함수는 식(2)와 같이 간단히 정의될 수 있는데, 만약 클래스 ω_i 에 속한다면 $p(x | \omega_i)$ 는 특성값 벡터 x 의 조건부 확률이 되며 $p(\omega_i)$ 는 클래스 i 의 사후 확률이 된다.

$$g_i(x) = \log p(x | \omega_i) + \log p(\omega_i) \tag{2}$$

$p(x | \omega_i)$ 가 다변량 정규분포가 될 경우, g_i 는 식(3)과 같은 판별함수 집합으로 표현된다.

$$g_i(x) = -\frac{1}{2}(x - \mu_i)' \sum_i^{-1} (x - \mu_i) - \frac{d}{2} \log 2\pi - \frac{1}{2} \log |\sum_i| + \log p(\omega_i) \tag{3}$$

여기서 μ_i 와 \sum_i 는 각각 클래스 i 의 분포에 대한 공분산 행렬이 된다. 본 논문에서는 두 개의 클래스들에 대한 공분산 행렬이 동일하다고 가정하고 위의 식을 이용하여 선형 판별함수를 생성했다. 위 식에서 $|\sum_i|$ 와 $(d/2) \log 2\pi$ 는 i 에 독립적이며, 2차식의 확장인 $(x - \mu_i)' \sum_i^{-1} (x - \mu_i)$ 에서 $x' \sum_i^{-1} x$ 역시 i 에 독립적이므로 무시될 수 있다.

따라서 결과적으로 판별함수는 식(4)와 같이 선형으로 표현될 수 있으며, $W_i = \sum_i \mu_i$, $W_{i0} = -(1/2) \mu_i' \sum_i^{-1} \mu_i + \log p(\omega_i)$, 세포핵일 경우 $i = 1$. 배경일 경우 $i = 2$ 가 된다.

$$g_i = W_i^* x + W_{i0} \tag{4}$$

표 1. 단계별 선택 결과

단계	입력	결정계수	F값	확률>F값	윌크스의 람다값	확률<윌크스의 람다값
1	I ₂	0.9125	2064.32	<.0001	0.08752	<.0001
2	R	0.0947	20.62	<.0001	0.0792	<.0001

분류 규칙이 설정된 후, 선택된 특성값들을 이용하여 분류기를 생성했다.

3.3 3차원 레이블링

연결요소 레이블링(connected component labeling)은 전처리 과정의 하나이며 일반적으로 정량적인 분석 전에 수행된다. 본 연구에서는 슬라이스 정보에 기반을 둔 새로운 3차원 레이블링 방법을 사용했다 [16]. 윤곽선 추적을 사용해 2차원의 연결요소 레이블링을 각 슬라이스에 적용한 후에, 레이블링 정보들을 동일한 객체에 속해있는 다음 슬라이스 상의 객체를 찾아내기 위해 사용했다.

그림 1은 슬라이스 정보를 이용한 레이블링 알고리즘의 수행 절차를 설명하고 있다. 예를 들어, k 번째, $k+1$ 번째에 공통으로 속한 객체가 있을 경우, 그 객체 간의 연결 여부를 판단하기 위해 각 객체들의 중심 화소들을 비교했다. 만약 연결되어있다고 판단 되면 양쪽에 동일한 레이블을 할당했다. 연속된 슬라이스 상에 공통의 객체가 존재하지만 각 객체에 구멍이 생겨 중심 화소가 연결되지 않은 경우, 객체의 나머지 영역은 중심 화소를 제외하고 연결여부를 판단했다. 두 객체가 연결되어있을 경우, 역시 동일한 레이블을 할당했고, 그렇지 않은 경우 $k+1$ 번째의 영상의 객체에는 다른 레이블을 할당했다.

3.4 3차원 특성값 추출

3차원 레이블링 후, 각각의 세포핵으로부터 3차원 특성값을 추출하기 위해 등급 결정 시 진단 전문가들의 등급 결정 기준이 되는 핵의 크기, 형태의 다

양성 등 형태학적인 특성에 따라 세포핵의 3차원 특성값들을 다음과 같이 정의했다[17].

부피(volume) : 핵의 전체 복셀 수에 의해 결정된다. 전체 복셀의 수는 세포핵의 표준 단위에 세포핵의 크기가 되는 하나의 복셀 크기를 곱했다.

표면적(surface area) : 3차원에서 세포핵의 표면적은 적어도 하나의 이웃하는 배경 복셀을 가지는 핵에 속해있는 복셀 들의 수로 추정될 수 있다. 그러나 모든 표면의 복셀 들을 찾아내기 위해 연결 연산을 사용하여 모든 복셀 들 사이의 관계가 재계산되어야만 한다. 따라서 본 논문에서는 표면적 측정의 대안으로 헤론의 공식을 이용했다(식(5)). 렌더링 된 표면이 삼각형들로 구성되므로, 세변의 길이가 주어질 경우 직접적으로 삼각형의 면적을 구할 수 있게 된다. 각 삼각형 영역의 면적을 계산한 후에, 이들의 합으로 전체 표면적의 넓이를 얻을 수 있다. 여기서 a, b, c 는 삼각형의 세 변의 길이이며 $s = (a + b + c)/2$ 가 된다.

$$SA = \sqrt{s(s-a)(s-b)(s-c)} \quad (5)$$

구형도(spherical shape factor) : 이 특성값은 세포핵이 얼마나 구에 가까운 가를 나타낸다. A 가 세포핵의 표면적, V 가 핵의 부피가 된다면, 구형도는 식(6)과 같이 정의 된다.

$$SSF = 36 \cdot \pi \cdot V^2/A^3 \quad (6)$$

3.5 3차원 가시화

세포핵의 형태 및 3차원 구조에 대한 가시화를 위해 윤곽선 기반의 표면 렌더링(surface rendering)과 3차원 텍스처 기반의 볼륨 렌더링(volume rendering)을 사용했다. 표면 렌더링의 경우, 각 슬라이스 영상들에 가우시안-라플라시안(Laplacian of Gaussian; LoG)연산자를 적용해 객체의 윤곽선을 추출하고, 윤곽선 데이터의 개선을 위해 Ryu 등이 제안한 평활화 알고리즘을 적용했다[18]. 또한 각각의 정점들로부터 표면을 생성하기 위해 로프팅 알고리즘을 사용했다. 3차원 텍스처 기반 렌더링을 위해서는 텍스처 맵을 생성한 후, 3차원 텍스처 배열 안에 볼륨 데이터를 적재시켰다. 일단 텍스처가 생성되면 현재의 시점 방향에 대한 영상 평면에 병렬인 슬라이스들을 렌더링 했다. 또한 볼륨 내부의 보간을 위해

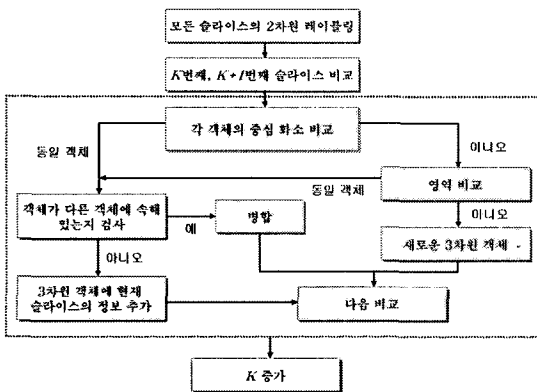


그림 1. 슬라이스 정보를 이용한 레이블링 알고리즘(SIL)

삼선형 보간법을 사용했으며, 볼륨 렌더링 최종 단계인 합성 단계에서, 전향(back-to-front)방식으로 텍스처를 입힌 다각형들을 평면에 알파 블렌딩했다 [19].

볼륨 데이터에서 서로 다른 시각적인 속성들의 하위 구조를 가시화하기 위해, 밝기 값을 적절한 컬러와 불투명도로 사상시키는 진이 함수를 생성했다. 이 함수의 결과 값들은 선형 사상의 형태로 표현되는 컬러 룩업 테이블을 사용하며, 대응 컬러와 불투명도 값으로 표현된다.

4. 연구결과

4.1 세포핵 분할

생성된 분류기를 이용한 실제 영상에서의 세포핵 분할 과정은 다음과 같다. 영상의 행과 열을 따라 탐색해 나가면서, 각 화소로부터 12개의 컬러 특성값을 추출한 후, 각 분류기들의 의사 결정 규칙을 기반으로 각 화소들을 세포핵과 배경으로 분류했다. 예를 들어, 만약 판별 함수가 $g_1(x) > g_2(x)$ 가 되면 분류기는 그 화소에 세포핵 영역으로 밝기 값 255를 할당하고 만약 판별함수가 $g_1(x) < g_2(x)$ 인 경우 반대로 그 화소에 배경영역으로 밝기 값 0을 할당하게 된다. 따라서 결과 영상은 세포핵이 분할된 이진화 영상이 된다. 그림 2는 세포핵이 분할된 결과 영상의 한 예이다.

4.2 3차원 가시화 결과

그림 3은 표면 렌더링의 결과이다. 3차원 가시화를 통해 다양한 각도에서 관찰이 가능하므로 세포핵의 정확한 형태를 확인 할 수 있다. 등급 1(그림 3(a))

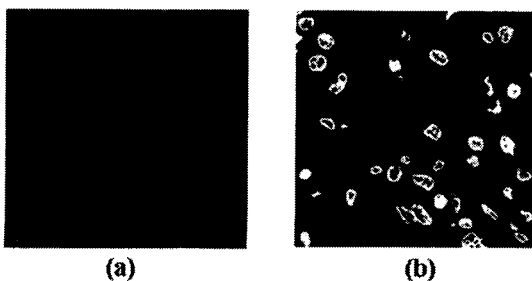


그림 2. 분류기를 이용한 세포핵 분할 (a)원본 영상; (b)결과 이진 영상

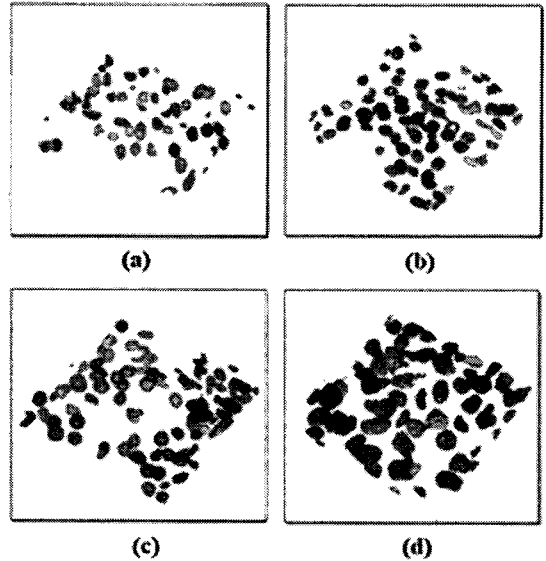


그림 3. 3차원 표면 렌더링의 결과 (a)등급 1; (b)등급 2; (c)등급 3; (d)등급 4

의 경우 약간의 변형이 존재하지만, 대체적으로 일정하고 정렬된 패턴을 보여주고 있다. 등급 2(그림 3(b))에서는 등급 1에 비해 핵의 크기가 커졌음을 쉽게 관찰 할 수 있으며, 등급 3(그림 3(c))에서는 무질서도 현저하게 높아졌음을 알 수 있다. 등급 4(그림 3(d))에서는 무질서의 경향이 가장 강하게 나타나며, 3차원 특성값 측정에서 극단적인 변형이 발생함을 알 수 있다.

그림 4는 동일한 데이터의 볼륨 렌더링의 결과로 각 복셀들의 밝기 값에 따라 서로 다른 컬러를 이용해 세포핵을 표현하고 있다. 이를 통해 3차원에서 핵의 염색질 패턴의 다양성을 찾아 낼 수 있다. 등급 1(그림 4(a))의 경우, 균일한 컬러 분포를 보여주지만 반면, 등급 4(그림 4(d))의 경우, 컬러 분포도의 변화가 큼을 알 수 있다

4.3 데이터 분석 결과

추출된 3차원 특성값을 분석하고 그 중요성을 평가하기 위해 데이터 분석을 수행했다. 어떤 3차원 특성값이 진단 정보에 영향을 주는가와 어떻게 핵 등급별 분류의 정확성을 높일 것인가를 평가하기 위해, 서로 다른 등급들 간의 3차원 특성값의 통계학적인 차이를 분석했다. 서로 다른 등급들 간의 분포의 차이에서 오는 통계학적인 중요도를 결정하기 위해 분

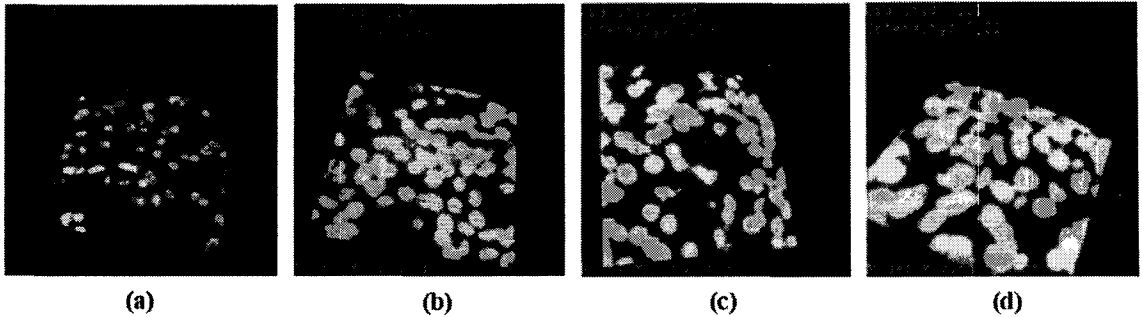


그림 4. 3차원 볼륨 렌더링 결과 (a)등급 1; (b)등급 2; (c)등급 3; (d)등급 4

산 분석을 사용했다. 각 특성값에 대한 분산분석의 결과는 표 2와 같다. 각 테스트의 분산분석의 결과에서 F값은 그룹 간(between) 제곱 합과 그룹 내(within) 제곱합의 비율을 나타낸다. 최적의 특성값은 그룹 내 제곱 합과 비교해 그룹 간 제곱 합이 상대적으로 높은 값을 가지며, 이는 최상의 특성값이 가장 높은 F값을 가짐을 의미한다. 테스트 결과, 등급

별 분류에 있어 부피가 가장 우수한 특성값임을 알 수 있었다.

표 3은 각 등급에 대해 핵의 3차원 특성값들을 비교한 결과이다. 가장 높은 등급의 세포핵의 평균 부피와 평균 표면적은 가장 낮은 등급의 것보다 높게 나타났다. 가장 낮은 등급의 평균 구형도는 가장 높은 등급에서보다 크게 나타났으며, 이는 완전한 구가

표 2. 부피, 표면적, 구형도에 대한 분산분석 결과

	데이터	자유도	제곱합	평균 제곱	F값	확률>F값
부피	그룹 간 제곱합	3	39522658.12	13174219.37	356.54	<.0001
	그룹 내 제곱합	234	8646352.89	36950.23		
표면적	그룹 간 제곱합	3	19760885.03	6586961.69	260.63	<.0001
	그룹 내 제곱합	234	5913874.38	25272.97		
구형도	그룹 간 제곱합	3	2.3249	0.7749	68.13	<.0001
	그룹 내 제곱합	234	2.6618	0.0113		

표 3. 핵의 등급별 부피, 표면적, 구형도에 대한 측정 데이터

		핵의 수	평균±표준편차(μm^3)	최소값	최대값
부피	등급 1	52	352.977±68.589	225.767	495.577
	등급 2	54	523.393±74.704	414.912	783.803
	등급 3	61	641.095±137.575	455.383	978.149
	등급 4	71	1377.258±315.711	840.123	2530.357
표면적	등급 1	52	312.853±49.740	208.319	463.868
	등급 2	54	466.302±64.649	363.104	786.912
	등급 3	61	572.590±104.313	421.228	885.682
	등급 4	71	1056.103±264.934	687.057	1843.806
구형도	등급 1	52	0.481±0.100	0.137	0.716
	등급 2	54	0.332±0.120	0.089	0.630
	등급 3	61	0.272±0.114	0.067	0.546
	등급 4	71	0.212±0.093	0.045	0.422

구형도 1의 값을 가지는 경우, 더 낮은 등급의 세포핵이 높은 등급의 경우보다 더 구형에 가까움을 나타낸다. 표의 결과를 통해 등급 진행에 따라 부피와 표면적이 증가함을 관찰할 수 있었다. 양극단에 놓인 등급(등급 1과 4)의 경우, 현저한 차이를 보이므로 쉽게 구별이 가능했지만, 중간 등급(등급 2와 3)의 세포핵들은 두 등급 간에 차이가 작았다. 전반적으로 3차원에서의 등급 간 세포핵의 크기 비교는 통계학적인 차이를 보여주었다.

그림 5는 회귀분석을 통해 동일한 세포핵에 대해 2차원에서 측정된 핵의 특성값(직경, 면적, 주변길이, 원형도)과 3차원에서 측정된 핵의 특성값(부피, 표면적, 구형도)의 관계를 나타낸 결과이다. 이 결과를 바탕으로 3차원 특성값에 대한 2차원 특성값의 영향력을 평가했다. 비교한 결과에서 직경과 면적이 부피와 높은 관련성이 있음을 알 수 있었다(그림 5(a), (b)). Pearson 상관계수는 각각 0.85701과 0.84242였다.

선형 회귀 함수는 $y = -450.454 + 115.706x$ 와 $y = 47.112 + 11.998x$ 이었으며 회귀 방정식에 의해 설명될 수 있는 변동 비율을 나타내는 결정계수 R^2 은 0.7908이었다.

또한 결과에서 주변길이가 표면적과 높은 관련성이 있음을 알 수 있었다(그림 5(c)). 가장 높은 상관

계수는 0.89042였으며, 선형 회귀 함수는 $y = -367.619 + 33.496x$, $R^2 = 0.7930$ 이었다. 그러나 원형도와 구형도 사이에는 의미 있는 관련성을 찾을 수 없었다(그림 5(d)). Pearson 상관계수는 0.27713, 선형 회귀 함수는 $y = -0.0057 + 0.397x$, 그리고 결정 계수 R^2 은 0.0768로 가장 낮은 값으로 나타났다. 이 결과는 지름, 면적, 주변길이가, 3차원 특성 값 추정 시 좋은 지표가 됨을 의미하며, 원형도를 사용하여 세포핵의 구형도를 추정하는 것이 어렵다는 것을 의미한다.

5. 결 론

본 연구에서는 세포핵의 최상의 분할을 위해 어떠한 특성값이 적합한지를 우선적으로 분석했다. 선별된 최적의 특성 값 I₂R을 사용한 세포핵의 분류는 만족할만한 성능을 보여주었다. 제안된 방법은 별도의 컬러 채널 별 분리과정이 필요 없이 직접적으로 세포핵을 분할 할 수 있으므로 기존의 방법에 비해 상대적으로 복잡도가 낮으며 속도도 빠른 장점을 가진다. 등급별 차이에 따른 3차원 특성값의 중요도를 평가하기 위해 수행한 분산 분석 결과, 부피가 가장 우수한 특성값으로 나타났다. 또한 세포핵의 2차원 특성값들로부터 상응하는 세포핵의 3차원 특성값을 추정하는 것이 가능했다. 2차원 특성값 직경과 면적이 부피에 대해 각각 0.85701과 0.84242의 높은 상관계수를 보였으며, 이 결과는 암의 진행에 따른 핵의 등급 변화와 3차원 형태학적 특성값들의 사이에 통계학적으로 높은 연관성이 존재함을 의미한다.

참 고 문 헌

[1] J.R. Hand and A. Broders, "Carcinoma of Kidney : The Degree of Malignancy in Relation to Factors Bearing on Prognosis," *Journal of Urology*, Vol. 28, pp. 199-216, 1932.
 [2] D.G. Skinner, R.B. Colvin, R.C Pfister, and W.F Leadbetter, "Diagnosis and Management of Renal Cell Carcinoma : A Clinical and Pathological Study of 309 Cases," *Cancer*, Vol. 28, No. 5, pp. 1165-1177, 1971.
 [3] S.A. Fuhrman, L.C. Lasky, and C. Limas,

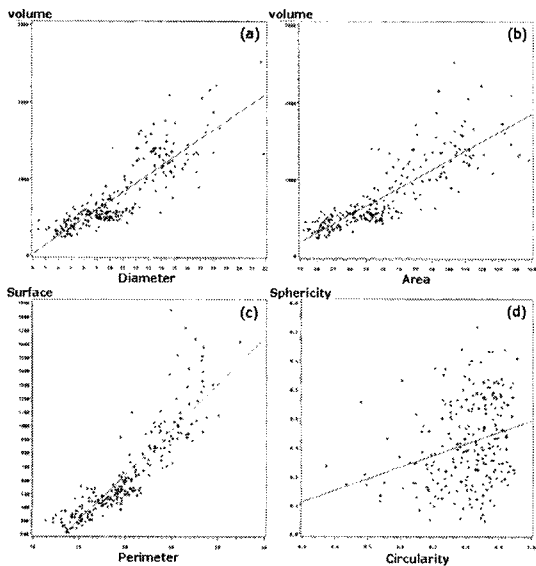


그림 5. 회귀 분석 결과 (a)직경과 부피; (b)면적과 부피; (c)주변길이의와 표면적; (d)원형도와 구형도

- “Prognostic Significance of Morphologic Parameters in Renal Cell Carcinoma,” *American Journal of Surgical Pathology*, Vol. 6, pp. 655-663, 1982.
- [4] C. Wahlby, “Algorithms for Applied Digital Image Cytometry,” PhD thesis, Uppsala University, Sweden, 2003.
- [5] S.J. Lockett and B. Herman, “Automatic Detection of Clustered, Fluorescent stained Nuclei by Digital Image-based Cytometry,” *Cytometry*, Vol. 17, No. 1, pp. 1-12, 1994.
- [6] R.J. Sears and C.W. Duckworth, “Image Cytometry as a Discriminant Tool for Cytologic Specimens Obtained by Endoscopic Retrograde Cholangiopancreatography,” *Cancer Cytopathology*, Vol. 84, No. 2, pp. 119-126, 1998.
- [7] P.M. Nederlof, S. van der Flier, N.P. Verwoerd, J. Vrolijk, A. K. Raap, and H.J. Tanke “Quantification of Fluorescence In Situ Hybridization Signals by Image Cytometry,” *Cytometry*, Vol. 13, No. 8, pp. 846-852, 1992.
- [8] S.J. Lockett, D. Sudar, C.T. Tompson, D. Pinkel, and J.W. Gray, “Efficient, Interactive, and Three-dimensional Segmentation of Cell Nuclei in Thick Tissue Sections,” *Cytometry*, Vol. 31, No. 4, pp. 275-286, 1998.
- [9] P.S. Adiga, and B.B. Chaudhuri, “Region-based Techniques for Segmentation of Volumetric Histopathological Images,” *Computer methods and programs in biomedicine*, Vol. 61, No. 1, pp. 23-47, 2000.
- [10] A.B. Jeroen, A.M. Hielke, T. Paulos, S.P. Lennert, M.P. Neal, P.A.B. Jan, and J. Paul, “Confocal DNA Cytometry : A Contour-based Segmentation Algorithm for Automated Three-dimensional Image Segmentation,” *Cytometry*, Vol. 49, No. 1, pp. 12-21, 2002.
- [11] N. Vandenbroucke, L. Macaire, and J.G. Postaire, “Color Image Segmentation by Pixel Classification in Adapted Hybrid Color Space,” *Computer Vision and Image Understanding*, Vol. 90, No. 2, pp. 190-216, 2003.
- [12] Y. Ohta, T. Kanade, and T. Sasaki, “Color Information for Region Segmentation,” *Computer Graphics and Image Processing*, Vol. 13, pp. 222-241, 1980.
- [13] S.E. Umbaugh, *Computer vision and image processing*, Prentice Hall, New Jersey, 1998.
- [14] R.S. Berns, *Principles of color technology 3rd Ed.* John Wiley&Sons Inc., New York, 2000.
- [15] H. Schulerud, G.B. Kristensen, K. Liestol, L. Vlatkovic, F. Albrechtsen, and H.E. Danielsen, “A Review of Caveats in Statistical Nuclear Image Analysis,” *Analytical Cellular Pathology*, Vol. 16. No. 2, pp. 63-82, 1998.
- [16] H.J. Choi, I.H. Choi, T.Y. Kim, N.H. Cho, and H.K. Choi, “Three-Dimensional Visualization and Quantitative Analysis of Cervical Cell Nuclei with Confocal Laser Scanning Microscopy,” *Analytical and Quantitative Cytology and Histology*, Vol. 27, No. 3, pp. 174-180, 2005.
- [17] P.S. Adiga, “An Integrated System for Feature Evaluation of 3D Images of A Tissue Specimen,” *Analytical Cellular Pathology*, Vol. 24, No. 2-3, pp. 47-55, 2002.
- [18] J.H. Ryu, H.S. Kim, and K.H. Lee, “Contour-based Algorithms for Generating 3D CAD Models from Medical Images,” *International Journal of Advanced Manufacturing Technology*, Vol. 24, No. 1-2, pp. 112-119, 2004.
- [19] M. Meißner, U. Hoffmann, and W. Straßer, “Enabling Classification and Shading for 3D Texture Mapping Based Volume Rendering Using OpenGL and Extensions,” *Proceeding of the 10th IEEE Visualization 1999 Conference(VIS'99)*, pp. 207-214, 1999.



김 태 운

2004년 인제대학교 정보컴퓨터
학과 학사
2006년 인제대학교 전산학과 석
사
2006년~현재 인제대학교 전산학
과 박사과정

관심분야 : 의료영상처리 및 분석,
컴퓨터 그래픽스



최 흥 국

1988년 Linköping University
BS.
1990년 Linköping University
MS.
1996년 Uppsala University Ph.
D.

2004년~현재 인제대학교 컴퓨터
공학부 부교수, 한국 신호처리 시스템 학회 상
임이사
관심분야 : 멀티미디어, 컴퓨터 그래픽스, 영상처리 및
분석



최 현 주

1995년 인제대학교 전산학과 학
사
2001년 인제대학교 전산학과 석
사
2005년 인제대학교 전산학과 박
사
2006년 인제대학교 의생명공학

원 연구교수
2007년~현재 부산대학교 의학전문대학원 BK21고급의
료인력양성사업단 연구교수
관심분야 : 의료영상처리 및 분석, 컴퓨터 그래픽스