

Wavelet 변환에 기반한 유방 종양 세포 조직 영상의 분류

황해길[°], 최현주^{*}, 최익환^{*}, 윤혜경^{**}, 최홍국^{*}
인제대학교 정보컴퓨터공학부^{*} 인제대학교 백병원 해부병리학교실^{**}
{seaload, hjchoi, kmiya}@mitl.inje.ac.kr {pathyoon,cschk}@ijn.inje.ac.kr

Classification of Breast Tumor Cell Tissue Section Images Based on Wavelet Transform

Hae-Gil Hwang[°], Hyun-Ju Choi^{*}, Ik-Hwan Choi^{*}, Hye-Kyoung Yoon^{**}, Heung-Kook Choi^{*}
Dept. of Infomation & Computer Engineering, Inje University^{*}
Dept. of Pathology, College of medicine, Inje University^{**}

요 약

본 논문은 유방질환 중에서 Duct(관)에 발생하는 유방 종양을 benign(양성종양)/DCIS(Ductal Carcinoma In Situ)/NOS(Invasive ductal carcinoma)로 자동 분류하기 위한 분류방법을 제안한다. 분류기 생성에서 가장 중요한 단계인 특징 추출단계에서는 wavelet 변환을 적용하였으며, wavelet 변환의 각 depth에 따라 분류기를 생성하여, depth와 생성된 분류기의 분류 정확도와의 상관관계를 비교·분석하였다. 현미경 100배 배율과 400배 배율의 유방 질환 영상을 1, 2, 3, 4 단계(depth)의 wavelet 변환을 적용한 후, 분할된 서브밴드에서 GLCM을 이용하여 질감 특징(Entropy, Energy, Contrast, Homogeneity)을 추출하여, 이 특징값들을 조합하여 판별분석에 의해 분류기(classifier)를 생성한 후, 분류 정확도를 검증하였다. Benign/DCIS/NOS를 분류하려면 최소 3단계 이상의 wavelet 변환을 적용해야 하고, 400배 배율 영상보다는 100배 배율의 영상이 더 나은 결과를 보였다.

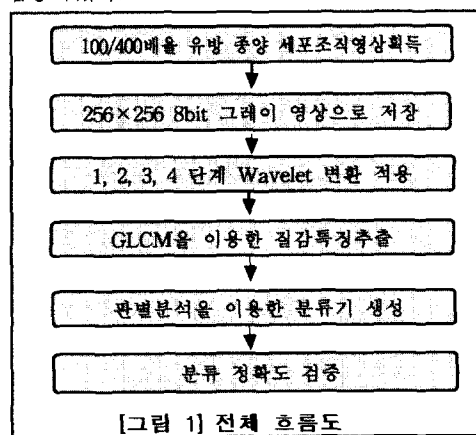
1. 서 론

유방암은 서구여성에게 있어서 가장 빈발하게 발생하는 악성 종양이며, 최근 우리 나라 여성에게도 위암에 이어 두번째로 많은 악성 종양이다. 매년 여성에게 발생하는 위암이나 자궁암은 발생율이 동일하거나 감소하는 추세이나 유방암은 매년 발생율이 급격히 증가하고 있는 추세이다. 이에 따라 유방암에 대한 연구도 활발히 이루어지고 있는데, 유방암 진단 방법으로는 유방촬영술, 초음파 유방촬영술, 유방 CT, 유방 MR imaging을 이용한 방사선학적 방법과 세포조직학적 방법이 있다. 특히 현미경을 통해 획득한 세포조직영상은 질병의 유무나 암의 진행정도를 파악하기 위한 매우 중요한 요소이다[1].

유방 질환의 세포조직영상에서 세포핵의 크기, 방향, 세포핵 사이의 거리 등에 대한 규칙성과 불규칙성은 암의 진행 정도를 판단하는데 있어 중요한 요인이므로, 그 특성을 분석하고 분류하기 위해서는 질감 분석이 필요하다[2]. 질감분석 방법 중 텍스처 정보에 기반한 방법에는 Gray Level의 변환 정도에 따른 Co-occurrence Matrix(GLCM)을 이용한 영상의 통계적 텍스처 분석 방법[3], wavelet 변환에 기반한 텍스처 특징을 분석하는 방법[4] 등이 제안되고 있다.

본 논문은 유방 질환 중에서 Duct(관)에 발생하는 유방 종양을 benign(양성종양)/DCIS(Ductal Carcinoma In Situ-관상피 내암종)/NOS(Invasive Ductal Carcinoma-

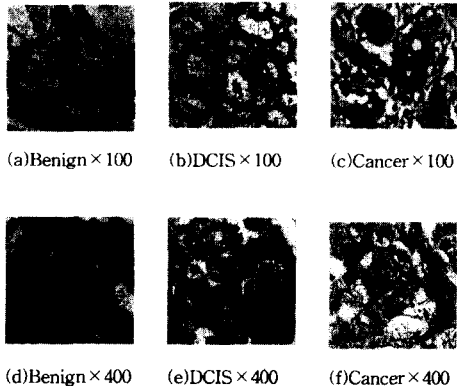
침윤성 내암종)로 자동 분류하기 위해 현미경 100배와 400배 배율의 유방 종양 세포 조직 영상을 각 그룹(benign/DCIS/NOS)에서 20개씩, 총 60개의 영상을 1, 2, 3, 4단계의 wavelet 변환을 적용한 후, 분할된 서브밴드에서 GLCM을 이용하여 질감 특징(Entropy, Energy, Contrast, Homogeneity)을 추출하였다. 이 특징값들을 조합하여 판별 분석에 의해 분류기(classifier)를 생성한 후, 총 60개의 영상 중에서 30개는 training data set으로 나머지 30개는 test data set으로 사용하여, 분류 정확도를 검증하였다.



[그림 1] 전체 흐름도

2. 재료와 이미지 획득

본 연구에 사용된 Breast 세포조직영상들은 1999년과 2000년 유방 질환 환자의 슬라이드로 인제대학교 백병원 해부병리학교실에서 제공받았다. 양성 종양(benign) 슬라이드와 DCIS 슬라이드, NOS 슬라이드를 CCD 칼라 비디오 카메라(sony)가 연결되어 있는 Olympus 광학 현미경을 통해 병리 전문가가 흥미있어 하는 영역을 현미경 100배와 400배 배율에서 640×480 크기의 칼라 이미지로 획득하였다. 256×256 크기의 8bit Gray 영상으로 저장하였다.



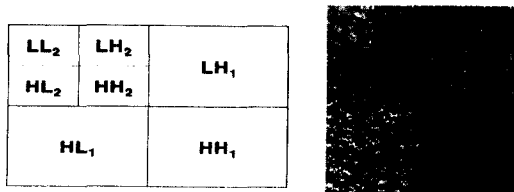
[그림 2] 유방 종양 세포 조직 영상

3. Wavelet 변환

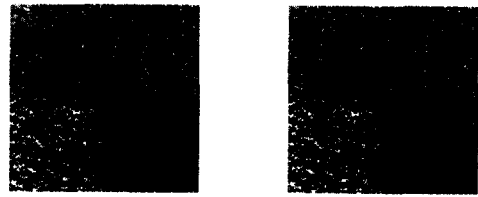
Wavelet 변환은 최근 들어 신호 처리나 패턴 인식 분야에서 가장 많이 쓰이고 있는 변환 방법 중 하나이다. 이 변환 방법은 공간 영역과 주파수 영역에서 지역적 특징과 다중해상도(Multiresolution Analysis)을 잘 나타낸다는 특성으로 인해 질감 정보를 얻고자 할 때 많이 쓰이고 있다.

영상 신호는 2차원 신호의 한 형태이므로 wavelet 필터를 적용할 때는 두 단계를 거친다. 우선 영상의 행(row)방향으로 wavelet 필터를 적용하고 열(column)방향으로 wavelet 필터를 적용한다.

그림 2는 2-채널 필터로 2단계 wavelet 변환한 후 이미지의 분할된 영역이다. LL_2 는 해상도가 반으로 줄어들고 원 영상의 블러링한 영상을 얻게되고, LH_1, LH_2 는 원 영상수평성분의 에지탐색이 적용된 영상을 얻게 되고, HL_1, HL_2 는 수직성분이, HH_1, HH_2 는 대각선 성분의 에지탐색이 적용된 영상을 얻게 된다.



(a) 2단계 wavelet 분할영역 (b) 2단계 wavelet



(c) 3단계 wavelet (d) 4단계 wavelet
[그림 3] 웨이블릿 변환

Benign/DCIS/NOS의 해상도 256×256 8bit 그레이 영상에서 각각 20개씩 총 60개의 영상에 대해 1, 2, 3, 4 단계의 wavelet 변환을 적용하고, 분할 된 서브 영역인 LH_1, HL_1, HH_1 -총 3개 영역(1단계), $LH_1, LH_2, HL_1, HL_2, HH_1, HH_2$ -총 6개 영역(2단계), $LH_1, LH_2, LH_3, HL_1, HL_2, HL_3, HH_1, HH_2, HH_3$ -총 9개 영역(3단계), $LH_1, LH_2, LH_3, LH_4, HL_1, HL_2, HL_3, HL_4, HH_1, HH_2, HH_3, HH_4$ -총 12개 영역(4단계)에서 질감 특징값(Entropy, Energy, Contrast, Homogeneity)을 추출한다.

4. GLCM을 이용한 질감 특징 추출

영상에서의 질감 분석은 영상이 가지고 있는 질감의 상호 대비성, 질감의 방향성, 질감 무늬의 규칙성을 분석하여 객관적 수치로 나타내는 것이다. 질감 특징 추출 방법은 세포 조직 영상이 거의 불규칙적인 패턴을 보이므로 통계학적 방법을 사용하였다. Joint Probability Density 함수 $f(i,j,d,\theta)$ 를 통해 Gray level의 Co-occurrence matrix(GLCM)을 생성한 후, 두 화소간의 거리 d 값은 1에 대해, 각 θ 는 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 에 대해 계산하였다. 원 영상의 256 gray level에 대해 모든 두 화소간의 밝기값의 분포를 계산 할 경우, 크기도 클 뿐만 아니라 계산량이 많이 요구되므로, 16 gray level로 축소하여 16×16 GLCM을 생성하여 질감 특징을 추출하였다[5]. 추출한 질감 특징은 아래와 같다[6].

$$\begin{aligned}
 - \text{Entropy} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (PM) \log (Px(i)Py(j)) \\
 - \text{Energy} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (PM)^2 \\
 - \text{Contrast} &= \sum_{n=0}^{N-1} n^2 \left(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} PM \right) \quad (\text{단, } |i-j|=n) \\
 - \text{Homogeneity} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{p[i,j]}{1+n} \quad (\text{단, } |i-j|=n)
 \end{aligned}$$

5. 실험 결과 및 분석

추출한 질감 특징값들에 대한 유의성을 평가하기 위해 통계 패키지인 SAS을 이용하여 판별 분석에 의해 분류기(classifier)를 생성한 후, 분류 정확도를 검증하였다. 판별분석은 집단간의 차이를 의미있게 설명해 줄 수 있는 독립변수들을 찾아내고, 이들의 선형결합으로 판별식을 만들어 내는 것으로, 추정된 판별식이 집단간의 차이를 잘 반영하는지에 대한 검증을 하였다[7].

[표1] Wavelet 단계(depth)에 따른 분류 결과

	B	D	N	%
B(20)	15	1	4	75
D(20)	3	12	5	60
N(20)	3	1	16	80
Total				71.7

(a) ×100 1depth

	B	D	N	%
B(20)	12	8	0	60
D(20)	2	18	0	90
N(20)	0	2	18	90
Total				80.0

(b) ×400 1depth

	B	D	N	%
B(20)	15	5	0	75
D(20)	1	19	0	95
N(20)	0	0	20	100
Total				90.0

(c) ×100 2depth

	B	D	N	%
B(20)	19	1	0	95
D(20)	3	16	1	80
N(20)	1	1	18	90
Total				88.3

(d) ×400 2depth

	B	D	N	%
B(20)	18	2	0	90
D(20)	0	20	0	100
N(20)	0	0	20	100
Total				96.7

(e) ×100 3depth

	B	D	N	%
B(20)	18	1	1	90
D(20)	2	18	0	90
N(20)	0	1	19	95
Total				91.7

(f) ×400 3depth

	B	D	N	%
B(20)	18	1	1	90
D(20)	0	20	0	100
N(20)	0	0	20	100
Total				96.7

(g) ×100 4depth

	B	D	N	%
B(20)	18	1	1	90
D(20)	2	17	1	85
N(20)	0	0	20	100
Total				91.7

(h) ×400 4depth

[표1] Wavelet 단계(depth)에 따른 분류 결과를 살펴보면, 1, 2depth wavelet를 적용한 결과(71.7과 80.0 / 90.0과 88.3) 보다는 3depth wavelet를 적용한 결과(96.7과 91.7)가 좋은 분류를 보였고, 3, 4depth Wavelet을 적용한 결과는 96.7과 91.7로 같은 분류 결과를 보였다. 따라서 benign/DCIS/NOS를 분류하려면 최소 3depth 이상의 wavelet 변환을 적용시켜야 함을 알 수 있었다.

100배 배율의 유방 종양의 세포 조직 영상과 400배율의 유방 세포 조직 영상을 비교해 보면 100배 배율의 세포 조직 영상이 더 나은 분류 결과를 보임을 알 수 있다. 이는 현미경 100배 배율 슬라이드에서는 세포 하나 하나 보다는 관(Duct)구조를 중심으로 영상을 볼 수 있고, 400배 배율의 조직 영상에서는 100배 배율과는 반대로 세포 하나 하나를 중심으로 영상을 보게 되므로, 400배율의 영상을 wavelet 변환을 적용하면 암의 Grading을 분석하는데 좋은 결과를 보였고[8], 100배율의 영상을 wavelet 변환을 적용하면 benign/DCIS/NOS를 분류하는데 좋은 결과를 보였다.

6. 결론 및 향후 과제

유방 질환 중에서 Duct(관)에 발생하는 유방종양을

benign/DCIS(Ductal Carcinoma In Situ)/NOS(Invasive Ductal Carcinoma)로 자동 분류하기 위해, 현미경 100배와 400배 배율의 유방 종양 세포 조직 영상을 1,2,3,4단계의 wavelet 변환을 적용한 후, 분할된 서브밴드에서 GLCM을 이용하여 질감 특징(Entropy, Energy, Contrast, Homogeneity)을 추출하였다. 이 특징값들을 조합하여 판별 분석에 의해 분류기(classifier)를 생성한 후, 분류 정확도를 검증하였다. Benign/DCIS/NOS를 분류하려면 최소 3depth 이상의 wavelet 변환을 적용시켜야 하고, 100배 배율의 세포 조직 영상을 이용하는 것이 더욱 나은 결과를 얻을 수 있다.

향후 과제는 wavelet 변환시 LL 영역에서도 특징값을 추출하여 나머지 서브밴드들과 같이 실험을 해 보고, 각 밴드별로 유의한 질감 특징값이 무엇인지 알아 보아야 할 것이다. 그리고 wavelet 변환 후 추출한 질감 특징값과 형태학적 특징값[9]을 조합하여 유방 종양 세포 조직 영상을 분류하는데 최적의 분류기를 생성하여야 할 것이다.

[참고문헌]

- [1] 대한병리학회, "병리학(PATHOLOGY)", 고문사, 2000
- [2] Heung-Kook Choi, Janos Vasko, Ewert Bengtsson, Torsten Jarkrans, Per-Uno Malmstrom, Kenneth Wester, Christer Busch, "Grading of Transitional Cell Bladder Carcinoma by Texture analysis of Histological Section," Analytical Cellular Pathology Vol.6, pp.327-343, 1994
- [3] H. Tamura, S. Mori and T. Yamawaki, "Textures corresponding to visual perception", IEEE TransSyst. Man Cybern. SMC-8(6), 460-473, 1978
- [4] 서상환, 이연숙, 김상균, 김홍식, "웨이브릿 변환에 기반한 색상과 질감 특징을 이용한 이미지 검색 시스템", 1999년도 한국정보과학회학술발표논문집 Vol. 26. No. 2, pp.30-32, 1999
- [5] 최예찬, 김호영, 김선아, 김백섭 "세포영상에서의 텍스처 특징을 이용한 영역 분할", 1999년도 한국정보과학회 봄 학술발표논문집 Vol. 26. No. 1, pp.576-578, 1999
- [6] Robert M. Haralick, K. Shanmugam, Its' Hak Dinstein, "Textural Features for Image Classification," IEEE Trans. On System, Man, and Cybernetics, Vol. SMC-3, No. 6, pp. 610-624, 1973
- [7] Richard A. Johnson, Dean W. Wichern, "Applied Multivariate Statistical Analysis," Prentice-Hall Inc. pp.629-725, 1998
- [8] 최현주, 최홍국, "현미경 영상에서 유효 특성 추출을 위한 알고리즘 비교분석", 제 13회 영상처리 및 이해에 관한 워크샵 발표 논문집, 2001, p21-26
- [9] 황해길, 최현주, 윤혜경, 최홍국, "유방질환 진단을 위한 유방 종양 세포핵의 형태학적 특성 비교 분석" 2001년도 한국멀티미디어 춘계학술발표논문집, Vol. 4, No. 1, pp.91-94