

다 해상도 분석을 이용한
X-선 유방 촬영 영상에서의
침상돌기를 보이는 병변 검출에
관한 연구

연세대학교 대학원

의공학과

조 정 진

다 해상도 분석을 이용한
X-선 유방 촬영 영상에서의
침상돌기를 보이는 병변 검출에
관한 연구

지도 김 동 윤 교수

이 논문을 석사 학위논문으로 제출함

2004년 12월 일

연세대학교 대학원

의공학과

조 정 진

조정진의 석사 학위논문을 인준함

심사위원 _____인

심사위원 _____인

심사위원 _____인

연세대학교 대학원

2004년 12월 일

감사의 글

항상 자신에게 무엇인가 부족한 것만 같은 목마름에 시작한 대학원 2년의 시간이 벌써 지나가고, 그 동안 많은 분들의 지도와 도움으로 작은 결실을 맺게 되었습니다. 본 논문이 나오기까지 아낌없는 격려와 지도를 해주신 분들에게 이 글로 감사의 마음을 전하고자 합니다. 부족한 저에게 끊임없는 관심과 지도로 학문의 길로 이끌어주시고 성실히 가르쳐주신 김동윤 교수님께 깊은 감사드립니다. 또한 바쁘신 와중에도 논문에 많은 가르침과 도움을 주신 정보통계학과의 하은호 교수님과 아프신 와중에도 많은 지도를 해주신 김법민 교수님께 감사드립니다. 학부 때부터 지금까지 가르침과 따뜻한 애정으로 지켜봐 주신 윤형로 교수님, 이경중 교수님, 이윤선 교수님, 윤영로 교수님, 김영호 교수님, 신태민 교수님, 김한성 교수님, 김경환 교수님께 감사드립니다.

배움의 길을 먼저 걸어 간 선배로써 늘 모범적인 모습으로 후배들을 이끌어 주신 김동선 선배님, 신동규 선배님과 박순만 선배님, 이경무 선배님께 감사드리며, 가까운 곳에서 많은 시간을 함께 하며 훌륭한 대학원 동기가 되어준 영운과 다양한 개성을 가진 용환, 성원, 석우 에게도 감사의 마음을 전하고 싶습니다. 학부부터 대학원 까지 함께 배우는 동기이면서 또한 좋은 친구이기도 한 상원, 기수 그리고 언제나 든든한 친구인 주신, 령희, 서영, 근화, 설규, 승민, 수지, 믿음직한 선배인 재원오빠, 수일오빠에게 고마운 마음을 전하고 싶습니다. 또한 같이 대학원으로 진학하여 많은 격려와 정보를 나누었던 96 동기들과, 선후배님들께도 감사의 마음을 전하며, 지금까지 많은 조언과 관심을 주신 최승욱 이사님, 노문종씨, 하광씨에게도 감사를 드립니다. 그리고 대학원 생활 동안 힘들고 어려울 때 아낌없는 조언과 도움을 주었던 친구에게 다시 한번 고마움을 전하고 싶습니다.

마지막으로 언제나 든든한 지원자이자 안식처인 사랑하는 아빠, 엄마, 큰이모, 멋진 동생들 유진, 갑진, 하라, 혜미, 두진, 용천에게 진심으로 깊은 감사를 드립니다.

2004년 12월

조정진 올림

차 례

그 림 차 례	iii
표 차 례	v
국 문 요 약	vi
제 1 장 서 론	1
제 2 장 X-선 유방 촬영술	4
2.1 X-선 유방 촬영술의 역사	4
2.2 X-선 유방 촬영 영상과 유방암의 특징	4
2.2.1 유방의 해부학적 구조와 유방 촬영 영상	4
2.2.2 유방 촬영 영상에 나타난 유방암의 특징	6
2.2.3 침상 돌기를 보이는 병변	7
제 3 장 침상 돌기를 보이는 병변의 예상 지역 검출	8
3.1 다 해상도 분석	8
3.1.1 다 해상도 분석	8
3.1.2 다 해상도 복원	11
3.1.3 적용	13
3.2 Histogram Specification	15
3.2.1 기본 원리	15
3.2.2 적용	18
3.3 Maximum Likelihood Segmentation	19
3.3.1 Segmentation	20
3.3.1 Maximum Likelihood Estimation	21
3.3.2 Maximum Likelihood Segmentation	23
제 4 장 침상 돌기를 보이는 병변의 특징 파라미터 추출 및 병변 분류	29
4.1 텍스쳐를 이용한 특징 추출	29
4.1.1 Graylevel Co-occurrence Matrix	30
4.2 Analysis Local Oriented Edge를 이용한 특징 추출	32
4.3 기타 특징 추출	33
4.3.1 통계를 이용한 특징 추출	34
4.3.2 병변의 형태를 이용한 특징 추출	35
4.4 로지스틱 회귀 분석을 이용한 병변 분류	35
4.4.1 로지스틱 회귀 모형	36
4.4.2 병변을 대표하는 특징 파라미터 선별	36

제 5 장	실험 및 고찰	37
5.1	데이터 준비	37
5.1.1	데이터 습득	37
5.2	실험 결과 분석	38
5.2.1	병변 예상 영역 검출	39
5.2.2	침상 돌기를 보이는 병변 검출 및 FROC 분석	43
제 6 장	결 론	48
참 고 문 헌		49
ABSTRACT		52

그 림 차 례

그림 2.1 : 기본적인 유방해부와 조직구조	5
그림 2.2 : X-선 유방촬영술에서 보이는 구조	5
그림 2.3 : 종괴의 모양과 경계	6
그림 2.4 : 유방 촬영 영상에 나타난 다양한 종괴	7
(a) 악성 유방암인 침상 돌기를 보이는 병변	
(b) 양성 유방암의 종괴	
그림 3.1 : 피라미드 영상 구조	8
그림 3.2 : 다 해상도 분석의 시스템 흐름도	9
그림 3.3 : 파라미터 a 에 따른 다양한 가우시안 커널	10
그림 3.4 : 다 해상도 분석 영상	11
(a) 가우시안 피라미드	
(b) 라플라시안 피라미드	
그림 3.5 : 다 해상도 복원의 시스템 흐름도	12
그림 3.6 : 다 해상도 복원의 예	12
그림 3.7 : 다 해상도 분석을 이용한 침상 돌기를 보이는 병변의 영상	14
(a) J 레벨의 병변, $600 * 800$	
(b) J-1 레벨의 병변, $300 * 400$	
(c) J-2 레벨의 병변, $150 * 200$	
그림 3.8 : $a = 0.4$ 의 가우시안 커널	14
(a) 가우시안 커널의 계수	
(b) 가우시안 커널을 이용한 2-D 저역 통과 필터	
그림 3.9 : 다양한 영상의 히스토그램 분포	16
그림 3.10 : 가우시안 확률 밀도 함수	16
그림 3.11 : Histogram Specification 과정	17
(a) $T(r_k)$: 입력 영상의 확률 밀도 함수	
(b) $G(z_k)$: 가우시안 분포를 따른 확률 밀도 함수	
(c) s_k 를 기준으로 한 (a) 와 (b) 의 매핑	
그림 3.12 : 침상 돌기를 보이는 병변의 Histogram Specification 적용 예	19
그림 3.13 : 영상과 히스토그램 분포	20
그림 3.14 : 최적의 임계 값 설정을 위한 히스토그램 분석의 예	21
그림 3.15 : Maximum Likelihood Estimation	23

(a) 가우시안 확률 분포를 가지는 n 개의 데이터 분포	
(b) Likelihood	
(c) Log likelihood 와 Maximum Likelihood	
그림 3.16 : 이상적인 경우의 병변과 주변 조직의 분류	23
그림 3.17 : 병변과 주변 조직을 포함한 영상의 히스토그램 분포	24
그림 3.18 : 임계 값(2500)을 적용한 병변의 영상 예 I	26
(a) 침상 돌기를 보이는 병변	
(b) 병변 예상 영역	
(c) 병변 외 영역의 예상 영역	
그림 3.19 : 임계 값(2300)을 적용한 병변의 영상 예 II	26
(a) 병변 예상 영역, (b) 병변 외 영역의 예상 영역	
그림 3.20 : Log likelihood의 Maximum Likelihood	27
그림 3.21 : Maximum Likelihood Segmentation을 적용하여 검출한 병변	28
(a) 실제 병변 영역	
(b) 검출된 병변 영역	
그림 4.1 : 다양한 텍스쳐	30
그림 4.2 : GCM 의 예	31
(a) 4×4 영상 , (b) 0° 방향 , (c) 90° 방향 , (d) 45° 방향	
그림 4.3 : 유방 촬영 영상에서 조직의 방사 형태의 예	32
(a) 침상 돌기를 보이는 병변	
(b) 정상 조직	
그림 4.4 : 병변과 정상 조직의 방사 형태	32
(a) 방사 방향, (b) 방사 방향의 분산	
그림 5.1 : 침상 돌기를 보이는 병변의 검출 과정	39
그림 5.2 : 다 해상도 상에 나타난 침상돌기를 보이는 병변의 돌기 변화	41
(a) $N \times N$, (b) $\frac{N}{2} \times \frac{N}{2}$	
(c) $\frac{N}{4} \times \frac{N}{4}$, (d) $\frac{N}{8} \times \frac{N}{8}$, (e) $\frac{N}{16} \times \frac{N}{16}$	
그림 5.3 : 침상 돌기를 포함한 영상의 Log Maximum Likelihood	42
(a) Log Likelihood, (b) 확대	
그림 5.4 : J-2 레벨에서 검출된 병변 예상 영역	43
그림 5.5 : 훈련 데이터의 FROC	46
그림 5.6 : 테스트 데이터의 FROC	46

표 차 례

표 2.1 유방 촬영 상의 종괴의 특징 소견	6
표 3.1 다 해상도 분석 시 데이터 처리 량	13
표 5.1 데이터 구성	38
표 5.2 다 해상도 분석 상의 병변의 크기	39
표 5.3 훈련 데이터를 이용한 병변 예상 영역 검출 결과	40
표 5.4 그림 5.2 영상의 특징들	42
표 5.5 훈련 데이터의 각 해상도에서 선별된 특징 파라미터	44
표 5.6 훈련 데이터의 FROC 분석 결과	47
표 5.7 테스트 데이터의 FROC 분석 결과	47

국 문 요 약

다 해상도 분석을 이용한 X-선 유방 촬영 영상에서의 침상돌기를 보이는 병변 검출에 관한 연구

X-선 유방 촬영술은 유방암 진단을 위해서 사용되는 가장 좋은 방법으로 유방암의 초기 진단에 유용하게 사용된다. X-선 유방 촬영술에서 보이는 대표적인 이상 소견으로 종괴를 들 수 있는데, 종괴는 종종 유선 조직에 부분적으로 덮여진 형태로 나타나기 때문에 검출이 어렵다. 최근 방사선과 전문의가 판독 시 검출하지 못하는 종양이 자동 검출 시스템에서 검출 할 수 있다는 연구 결과들이 발표되고 있다. 컴퓨터 진단 보조(CAD, Computer Aided Diagnosis) 시스템은 방사선과 전문의들이 간과하기 쉬운 병변 의심 영역을 판독 시 알려줄 수 있기 때문에 방사선 전문의들의 많은 관심을 끌고 있다.

본 연구에서는 다 해상도 상에 나타나는 침상 돌기를 보이는 병변을 검출하는 방법을 구현 하였다. 이를 위해 훈련 데이터에 대해 다 해상도 분석과 Maximum Likelihood Segmentation을 적용하여 침상돌기를 보이는 병변으로 의심되는 영역을 검출하였고, 검출된 영역의 데이터의 특성으로부터 49 개의 특정 파라미터를 추출한 후 이에 대해 로지스틱 회귀분석을 이용하여 유용한 특정 파라미터를 선택하였다. 선택된 특정 파라미터를 이용하여 로지스틱 회귀 모형을 추정한 후 이를 테스트 데이터에 적용하여 악성 유방의 소견인 침상 돌기를 보이는 병변을 분류하였다.

본 논문의 검출결과를 J 레벨, J-1 레벨에서 FROC(free response receiver operating characteristic)를 이용하여 분석하였다. 훈련 데이터의 J 레벨, J-1레벨의 해상도에서 TPF(True Positive Fraction)가 90% 일 때 영상 당 평균 FP(False Positive) 는 0.9 와 1 이었고, TPF 가 100% 일 때 영상 당 평균 FP는 1.5 와 1.6 이었다. 테스트 데이터에서는 J 레벨과 J-1 레벨의 해상도에서 TPF가 90% 일 때 영상 당 평균 FP 는 0.8 과 1, TPF 가 100% 일 때 영상 당 평균 FP는 1.3의 결

과를 보였다. 이는 Cheng 등이 J 레벨에서 TPF가 90% 일 때 영상 당 평균 FP는 약 1.4 , TPF가 100% 일 때 영상 당 평균 FP는 약 2.1인 연구 결과 보다 우수한 결과이다.

핵심되는 말 : X-선 유방 촬영술, 유방암, 침상 돌기를 보이는 병변, 다 해상도 분석, 로지스틱 회귀분석, FROC

제 1 장 서 론

유방암은 여성에게는 가장 흔한 질병으로 우리나라에서 매년 발생빈도가 증가하여 2000년 통계에는 위암에 이어 2위를 차지하고 미국의 경우 1991년의 Wallis에 의하면 발생률이 매년 증가하여 년 150,000 명에 이르고 약 45,000 명이 유방암으로 사망하며 실제 여성 10명중 1명꼴로 유방암에 걸릴 확률이 있다고 보고하였다[1]. 유방암을 진단하기 위해서는 X-선을 이용한 유방 촬영술, 초음파 유방 촬영술, 유선조영술, 유방자기공명영상검사, PET(position emission tomography) scan 등이 있는데 가장 간단하며 기본적인 방법이 X-선 유방 촬영술이다[1]. 대부분의 유방암은 원인 없이 발생하여 유방암을 예방한다는 것은 불가능하므로 조기 발견이 중요하다 할 수 있는데, X-선 유방 촬영술을 이용한 검진은 암이 촉진되는 시기보다 2년 일찍 조기 유방암을 발견해 준다는 보고가 있다[2]. X-선 유방 촬영술의 영상에서 관찰되는 대표적인 이상 소견으로 종괴와 석회화를 볼 수 있다. 이들은 모양이나, 분포 상태, 경계선등이 악성 유방암과 양성 유방암에서 각기 다른 특징을 보이므로. 유방 촬영상에 나타난 이상 소견의 특징들을 관찰하여 유방 방사선과 전문의에 의해 악성 혹은 양성의 판독이 이루어진다[3]. 그러나 숙련된 유방 방사선과 전문의에 의하여 판독이 이루어짐에도 불구하고 육안에 의한 판독에 대해 정확성과 효율성에 대한 문제가 제기되었고 이에 따라 디지털 영상 처리를 이용한 컴퓨터 보조 진단(CAD, Computer Aided Diagnosis)에 대한 연구가 활발히 진행되고 있다. 유방암의 진단에 있어서의 컴퓨터 보조 진단은 판독의사에게 참고 의견(second opinion)과 의중 판독(double reading)을 제시하여 판독 오차를 줄일 수 있다는 점에서 각광을 받아 2001년에는 일부 상용 제품이 미국 식품의약국(FDA)의 공인을 받아 판독의의 보조 진단으로 사용되고 있다[6].

컴퓨터 보조 진단에서 유방암의 이상 소견을 자동적으로 검출하기 위해서 악성 유방암의 종괴와 석회화를 정확하게 추출하는 것이 가장 중요하다. 그러므로 이들의 자동적 추출과 양성, 악성을 정확하게 분류하기 위한 연구가 다양하게 진행되고 있으며, X-선 유방 촬영 영상에서 나타나는 종괴와 석회화가 서로 다른 특징

을 보이므로 일반적으로 각 소견의 특징에 맞추어 분리되어 연구되고 있다.

종괴가 있는 병변을 검출하기 위한 연구 방법은 크게 bilateral image subtraction과 single image segmentation 두 가지로 나뉜다. bilateral image subtraction 방법은 양쪽 유방은 해부학적으로 같은 구조를 갖는다는 전제하에 유방 촬영술에 나타난 양쪽 유방의 비대칭성을 비교하여 병변을 검출하는 방법이다. Tabar[14] 와 Windsberg[15] 등은 양쪽 유방 영상에서 같은 해부학적 조직을 가지는 영역을 제거하고 나머지 영역의 비 대칭성을 관찰하는 연구를 하였으며 Giger[13] 등은 유두를 기준으로 두 영상의 Correlation을 이용하여 경계선을 추출하고 두 영상의 비 대칭성을 이용하여 종괴를 포함한 병변을 찾는 연구를 하였다. Yin[16] 등은 nonlinear bilateral subtraction을 이용하여 종괴를 검출하고 크기, 경계선등을 기준으로 악성, 양성을 이용하여 병변을 구분하였다. 두 번째 연구 방법인 single image segmentation은 양쪽 유방의 X-선 유방 촬영 영상 중 각각의 영상에서 종괴를 포함한 병변을 추출하는 방법이다. 이 방법은 주로 악성 유방암의 특징인 침상 돌기를 보이는 종괴를 포함하는 병변과 양성 유방암의 특징인 명확한 경계선을 보이는 종괴 검출에 이용되는 연구 방법으로 Lai[17] 등은 메디안 필터 와 템플릿 매칭을 이용한 검출 방법을 제시하였다. Brazakoviz[18] 등은 fuzzy pyramid linking 과 다 해상도 분석을 이용하여 악성 유방암의 병변과 양성 유방암의 병변을 분류하였으며, Kegelmeyer[7] 등은 Laws 텍스쳐와 ALOE(Analysis Local Oriented Edge)를 이용하여 침상돌기를 보이는 병변을 검출하였다. Youssry[8] 등은 병변 가능 영역을 추출 한 후 COM(Co-occurrence Matrix)를 이용하여 특징을 추출하고 뉴로-퍼지 모델을 이용하여 병변을 분류하였으며, Liu[9] 는 linear phase nonseparable 2-D 웨이블렛 변환을 이용한 다 해상도 분석을 사용하여 ALOE, 평균, 분산 등의 특징을 추출하고, 이진 결정 트리 구조를 이용하여 병변을 분류 하여 FROC(free-response receiver operating characteristic)로 검출 결과를 분석하였다. Kinnard[10] 등은 침상 돌기를 보이는 병변의 후보 영역을 Maximum Likelihood Segmentation을 이용하여 추출한 후 평균과 분산을 비롯한 통계적인 특징, 텍스쳐 등을 이용한 신경망회로로써 실제 병변을 분류하고 FROC를 사용하여 평가하였으며, Kupinski[11]등은 병변에 가우시안 분포를 적용하였을

때의 병변과 병변 외 지역의 확률 분포 변화를 이용한 Probabilistic Segmentation 알고리즘을 사용하여 병변 가능 영역을 찾고 이를 일반적인 영역 검출법에 의한 병변 검출 결과와 비교하는 연구를 제시하였다. 이들 연구들의 궁극적인 목적은 X-선 유방 촬영 영상에서 악성 유방암의 병변을 검출하되 검출된 병변의 TP(True Positive)율은 높이고 영상 당 평균 FP(False Positive)은 낮추는 것이다.

본 연구에서는 악성 유방암으로 분류되는 침상돌기를 보이는 병변을 다 해상도 분석을 이용하여 병변의 특징을 분석함으로써 각종 영상 처리 시 빠른 데이터 계산과 속도 향상을 시도하였다. 또한 병변의 검출 방법으로 실험값을 임계 값으로 선택하는 대신 병변 영역과 병변 외 영역의 확률 분포를 이용한 Maximum Likelihood Estimation 을 적용하여 가장 적절한 임계 값을 선택하고, 추출된 임계 값을 이용하여 침상돌기를 보이는 병변 검출을 위한 병변 예상 영역을 추출하였다. 병변 예상 영역으로부터 실제 병변 지역을 검출하기 위해서 훈련 데이터로부터 통계, 텍스쳐, ALOE(Analysis Local Oriented Edge)를 포함한 49 개의 특징 파라미터를 추출하고 로지스틱 회귀 분석을 통해 병변을 대표하는 적절한 특징 파라미터를 선택하였다. 선택된 병변의 특징 파라미터를 이용한 로지스틱 회귀 모형을 사용하여 병변 예상 영역으로부터 병변과 병변 외 영역을 분류하고, 테스트 데이터를 통해 검증하였다. 그리고 FROC(free response receiver operating characteristic)을 이용하여 분류 결과에 대한 TPF(True Positive Fraction)와 FP(False Positive) per image를 분석하였다. 그리하여 다 해상도 분석을 통한 TP(True Positive)율은 높이되 영상 당 평균 FP(False Positive)는 낮추는 병변 검출 방법 제시하는 것을 목표로 한다.

본 논문의 2장에서는 병변을 추출하기 위해 유방의 해부학적 구조와 X-선 유방 촬영술에 나타나는 병변의 특징을 소개하고 3장에서는 다 해상도 분석과 Maximum Likelihood을 이용한 병변 예상 영역 검출을 설명한다. 4장에서는 검출된 병변 예상 영역으로부터 추출한 특징 파라미터와 로지스틱 회귀분석에 대해 설명하고 5장에서 실제 X-선 유방 촬영 영상으로 구성된 훈련 데이터 과 테스트 데이터를 사용한 실험을 소개하고 병변 분류 결과를 FROC를 통해 분석하였다.

제 2 장 X-선 유방 촬영술

2.1 X-선 유방 촬영술의 역사

X-선 유방촬영술은 1913년 베를린 대학의 외과의사인 Albert Salomon 이 일반 X-선 촬영 장치를 유방검사에 이용함으로써 시작 되었다. 그러나 유방은 다른 기관과 달리 매우 유사한 음영의 연조직으로 구성된 기관으로 높은 대조도의 촬영이 요구되므로 일반 X-선 촬영기를 유방 촬영에 사용하기 위해서 개선이 필요했다. 이후 1965년 프랑스의 Charles Gros 에 의해 전압과 전류 등을 변화시킨 최초의 유방 촬영 전용 장치가 개발 되었으며 1970년에 Dupont 사에서 증감지-필름을 이용한 X-선 유방 촬영 장치를 소개하였다. 2000년에는 전자 기술의 발달과 더불어 미국의 식품의약국(FDA)의 인가를 받은 디지털 유방촬영 장치가 개발 되었고 디지털 유방촬영 장치를 통해 생성된 디지털 유방 촬영 영상은 2000년 11월부터는 고성능 모니터를 통한 직접 판독도 허용되었다[1,3].

2.2 X-선 유방 촬영술내 유방암의 특징

2.2.1 유방의 해부학적 구조와 유방촬영 영상

그림 2.1 은 기본적인 유방해부와 조직 구조를 나타낸 것으로 유방은 크게 지방으로 구성된 피하층(subcutaneous layer)과 유선엽 결합 조직의 유선층(mammary layer) 그리고 유선후층(retromammary layer) 으로 나눌 수 있으며 대부분의 유방암은 유관과 유엽에서 기원한다고 할 수 있다[3,4]. 그림 2.2 는 X-선 유방 촬영술을 이용하여 유방을 촬영하였을 때 보이는 구조를 보인 것이다. 그림2.2 에서 보듯이 유방내의 조직들이 촬영 영상에서 매우 유사한 음영으로 나타남을 알 수 있다.

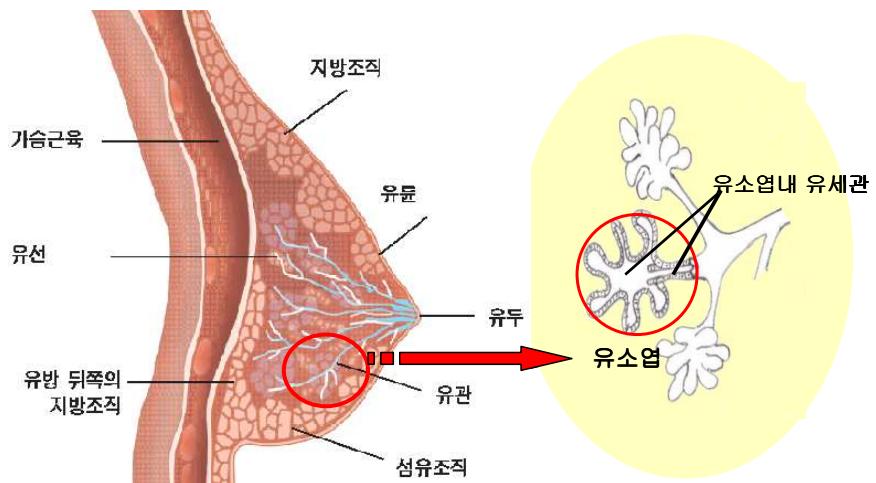


그림 2.1 : 기본적인 유방해부와 조직구조

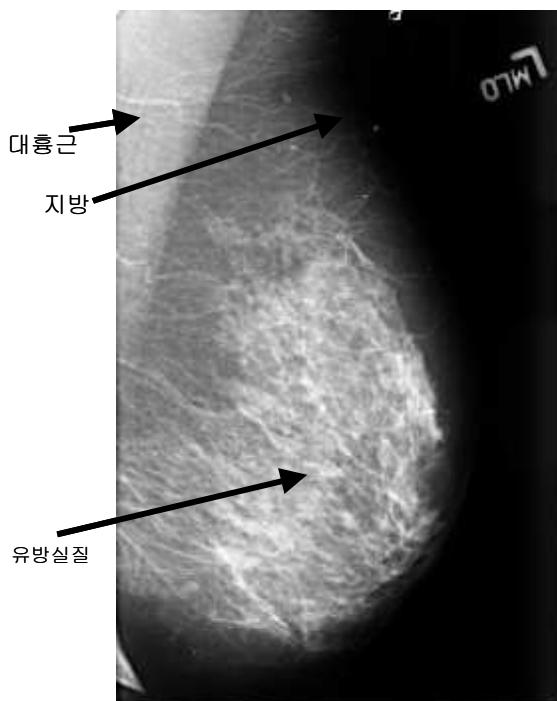


그림 2.2 : X-선 유방촬영술에서 보이는 구조

2.2.2 유방 촬영 영상에 나타난 종괴의 특징

X-선 유방 촬영 영상에서 관찰되는 대표적인 유방암 소견으로 종괴를 볼 수 있다. 종괴는 그 모양이나 경계선 또는 크기 및 성분에 따라 유방 방사선과 전문의에 의해 양성과 악성으로 판단되어지는데 보통의 양성 병변에 해당하는 종괴는 모양이 둥글고 경계가 명확하며 유동성이 있는 반면 악성 병변에 해당하는 종괴는 모양이 불규칙하여 방사형을 띠고 가장자리 경계가 불명확하며 고정되어 있다. 그림 2.3은 유방 촬영 영상에 나타나는 종괴의 모양과 경계에 대한 특징을 보여주는데 종괴를 자동적으로 검출하고자 할 때 중요한 기준이 된다[35].

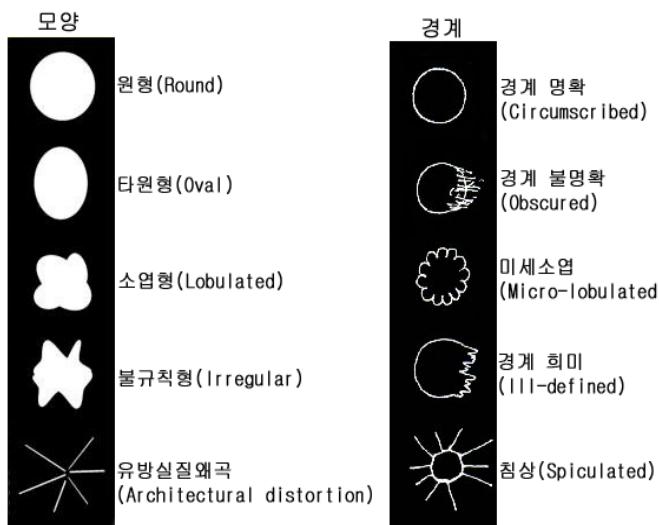


그림 2.3 : 종괴의 모양과 경계

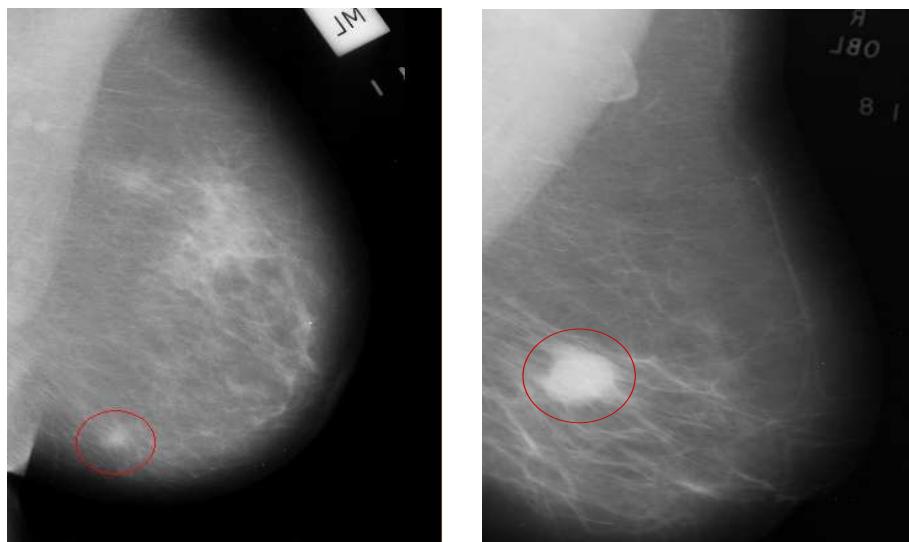
유방 촬영상의 특징 소견	악성 병변
종괴 모양	유방 실질 왜곡형은 병변의 가능성이 높음
종괴 음영	주로 비 균등
종괴 경계	불명확
침상 소견	주로 방사형을 이룸
X-선 Attenuation	주변 조직에 비해 높은 밀도를 가짐

표 2.1 : 유방 촬영상의 종괴의 특징 소견

표 2.1 은 X-선 유방 촬영 영상에 나타나는 악성 유방암의 병변의 특징으로, 자동으로 병변을 분류하고자 할 때 중요한 기준이 되는 특징들을 나타낸 것이다[1].

2.2.3. 침상 돌기를 보이는 병변

유방 촬영 영상에 나타나는 종괴와 관련된 악성 유방암의 소견은 침상 돌기를 보이는 병변(a lesion with spiculated border)[그림 2.4(a)], 경계가 불분명한 종괴(a mass with spiculated border), 미세분엽양 종괴(a mass with microlobulation)[5] 등이 있다. 그림 2.4(a)의 타원형으로 표시된 부분은 침상 돌기를 보이는 병변을 보인 것이며 (b)는 양성으로 판정된 종괴를 나타낸 영상이다. 악성 유방암의 침상 돌기를 보이는 병변은 방사형의 모양과 불분명한 경계선을 가지는 특징을 가지나 양성 유방암의 종괴는 명확한 경계선과 모양을 보임을 알 수 있다[8,9,35].



(a) 악성 유방암인 침상 돌기를 보이는 병변 (b) 양성 유방암의 종괴

그림 2.4 : 유방 촬영 영상에 나타난 다양한 종괴

제 3 장 침상 돌기를 보이는 병변 검출

본 장에서는 X-선 유방 촬영 영상에 나타난 침상 돌기를 보이는 병변을 검출하기 위하여 다 해상도 분석과 Maximum Likelihood Segmentation를 이용한 방법에 대해 설명한다.

3.1 다 해상도 분석

1983년 P. J. Burt 등에 의해 소개된 다 해상도 분석은 신호 또는 영상의 해상도를 변화시키면서 고 주파수 성분과 저 주파수 성분으로 신호 또는 영상을 분리하고, 각 해상도에 나타난 각 주파수의 특징을 분석하는 방법이다[19, 20]. 다 해상도 분석은 해상도가 변화해도 신호 또는 영상의 특징이 일정하게 유지되므로 데이터 압축 등에 이용 할 수 있으며, 각 해상도에서의 고 주파수와 저 주파수 성분을 쉽게 분리해서 볼 수 있어 신호나 영상의 특징 분석을 위한 다양한 분야에서 사용되고 있다.

3.1.1 분석

다 해상도 분석의 가장 기본적인 원리는 영상의 해상도를 일정 비율로 감소시켜 영상의 스케일을 변화 시키는 것이다.

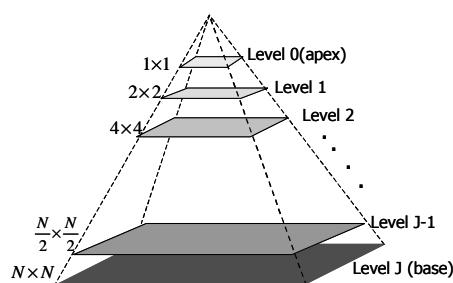


그림 3.1 : 피라미드 영상 구조

그림 3.1 은 J 레벨의 영상을 $J-1, J-2, \dots, 0$ 의 레벨까지 변화시켜 얻는 영상을 파라미드 구조화 시킨 것이다. J 레벨의 $N \times N$ 영상이 레벨이 변화함에 따라 영상의 해상도가 $\frac{1}{2} \times \frac{1}{2}$ 씩 감소함을 알 수 있다. 그림 3.2 는 다 해상도 분석을 하기 위한 시스템 흐름도를 나타낸 것이다. 레벨 J 의 입력 영상에 대하여 저역 통과 필터를 통과 시키고 이를 $\frac{1}{2}$ 로 다운샘플링 함으로써 J-1 레벨의 가우시안 영상을 얻고, 이를 2배로 업 샘플링 한 후 레벨 J의 영상에 대하여 빼기를 하여 J 레벨의 라플라시안 영상을 얻는다.

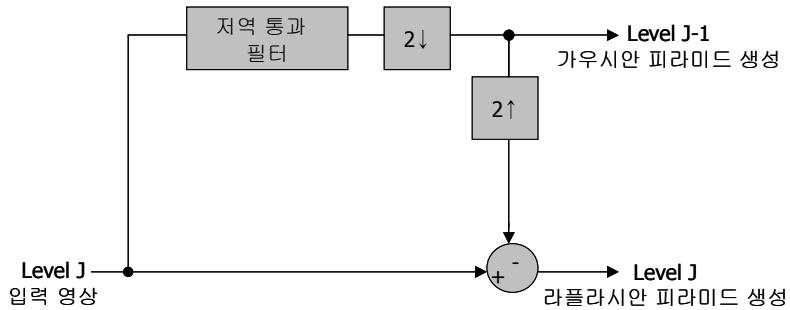


그림 3.2 : 다 해상도 분석의 시스템 흐름도

1) 가우시안 피라미드

J 레벨의 영상이 입력되면 가우시안 커널을 이용한 저역 통과 필터를 거침으로써 저 주파수 성분의 영상을 얻을 수 있다. 얻어진 저 주파수 성분의 영상은 다운 샘플링 단계를 거쳐 가우시안 피라미드의 레벨 $J-1$ 영상(g_0)을 생성한다.

① 가우시안 커널

$$\sum_{m=-2}^2 \hat{w}(m) = 1 \quad (3-1)$$

$$w(m, n) = \hat{w}(m)\hat{w}(n) \quad (3-2)$$

식 (3-1) 은 길이 5 의 1-D 가우시안 커널을 나타낸 것이며, 식 (3-2)는 1-D 커널을 이용해서 2-D 가우시안 커널을 표현한 것이다. 가우시안 커널의 각 계수는

$\hat{w}(0) = a, \hat{w}(-1) = \hat{w}(1) = 0.25, \hat{w}(-2) = \hat{w}(2) = 0.25 - a \times 0.5$ 를 이용하여 만족하는 커널 계수를 구성 할 수 있다[25]. 그림 3.3 은 파라미터 a 에 따른 다양한 가우시안 커널을 보여준 것으로 $a = 0.4$ 일때 가장 가우시안 형태와 가까운 커널임을 알 수 있다.

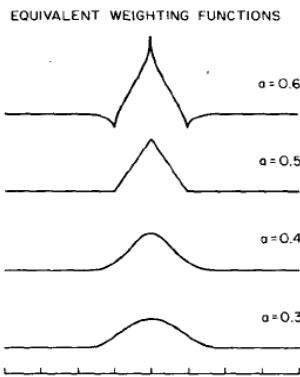


그림 3.3 : 파라미터 a 에 따른 다양한 가우시안 커널

② 가우시안 피라미드 생성

$$g_l(i,j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m,n) g_{l-1}(2i+m, 2j+n) \quad (3-3)$$

level : $0 < l < N, \quad 0 < i < Row, \quad 0 < j < Column$

식 (3-3) 은 가우시안 피라미드를 생성하는 과정을 표현한 식으로 J 레벨의 입력 영상으로부터 생성된 $J-1$ 레벨의 가우시안 영상은 다음 레벨의 입력 영상이 되고, 가우시안 커널과 다운 샘플링을 거쳐 $J-2$ 레벨의 가우시안 영상을 출력하는 과정을 반복함으로써 가우시안 피라미드를 생성함을 표현하고 있다.

2) 라플라시안 피라미드

라플라시안 피라미드는 입력 영상의 고 주파수 성분만을 추출한 것으로, 저 주파수 성분인 가우시안 피라미드의 영상을 확장 시킨 후 입력 영상과의 차이를 구한 것이다.

$$L_l = g_l - EXPAND(g_{l+1}) \quad (3-4)$$

식 (3-4) 는 라플라시안 피라미드를 생성하는 과정을 표현한 식이다.

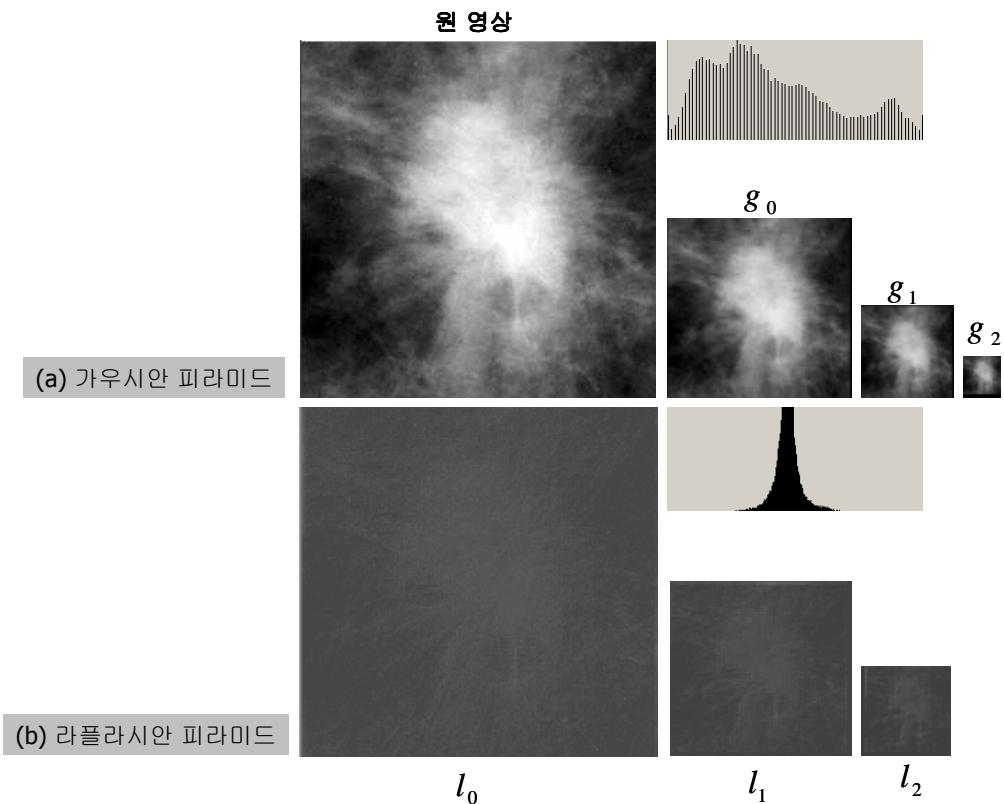


그림 3.4 : 다 해상도 분석영상

그림 3.4 는 다 해상도 분석을 통해 생성된 가우시안 피라미드의 영상(a)과 라플라시안 피라미드의 영상(b)을 보여준 것으로 해상도가 변하여도 각 레벨의 히스토그램 특성이 변하지 않음을 보여준다.

3.1.2 복원

분석 과정을 통해 생성된 가우시안 피라미드와 라플라시안 피라미드를 이용함으

로써 원 영상으로의 복원이 가능하다. 그림 3.5 는 다 해상도 복원의 과정을 나타낸 것으로 J-1 레벨의 가우시안 영상을 확장하여 J 레벨의 라플라시안 영상과 더함으로써 J 레벨의 원 영상을 복원 할 수 있다. 그림 3.6 은 J-3 레벨까지 분석된 가우시안과 라플라시안 영상들을 이용하여 J 레벨의 영상을 복원하는 예이다.

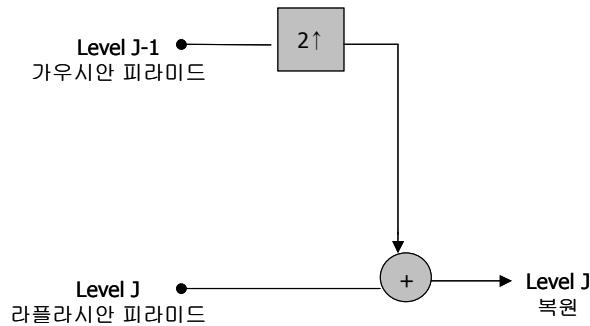


그림 3.5 : 다 해상도 복원의 시스템 흐름도

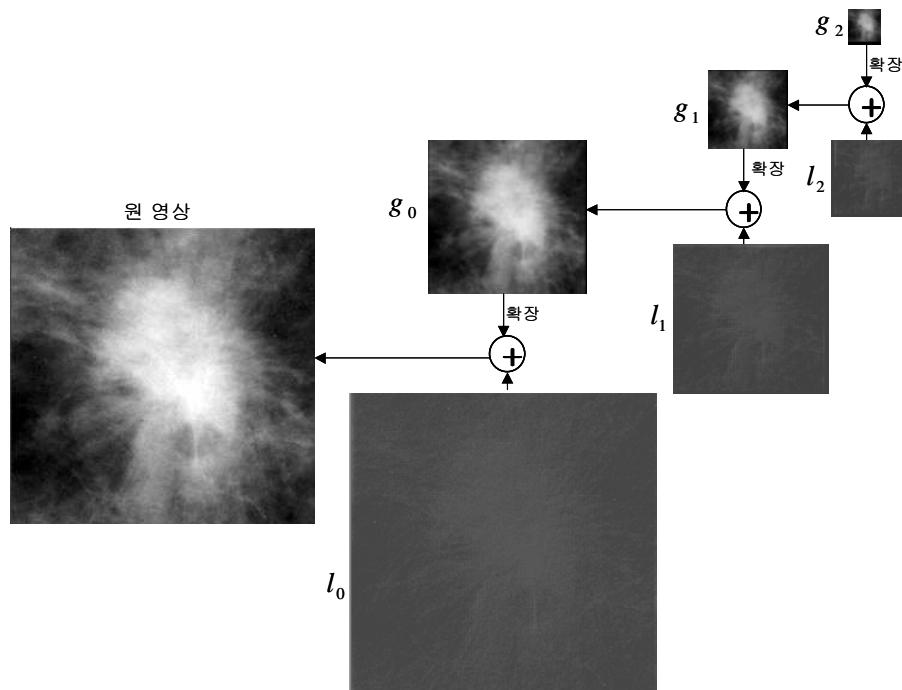


그림 3.6 : 다 해상도 복원의 예

3.1.3 적용

디지털화 된 X-선 유방촬영 영상의 공간 해상도는 약 1000×2000 에서 5000×7000 까지 다양한 크기를 보이는데 약 5000×7000 크기의 디지털 영상 신호 처리 시 많은 계산과 시간이 필요하므로 고성능의 시스템 사양이 요구된다. 그러나 다 해상도 분석은 영상의 해상도가 변해도 영상의 특징을 일정하게 유지하기 때문에, 유방 촬영 영상에 나타난 병변의 특징이 변하지 않으면서 빠른 영상 처리가 가능한 해상도를 찾을 수 있다. 표 3-1 은 5000×7000 의 영상을 다 해상도 분석을 이용하였을 경우 각 레벨에서의 데이터 처리량을 나타낸 것이다.

다 해상도 분석 레벨	데이터 처리량
J	35,000,000
J-1	8,750,000
J-2	2,187,500

표 3.1 : 다 해상도 분석 시 데이터 처리량

침상 돌기를 가지는 병변의 특징은 2.2.3에서 언급하였듯이 방사형의 모양과 불분명한 경계선을 가진다[8,9]. 그러나 침상돌기를 보이는 병변을 다 해상도 분석을 이용하여 해상도를 낮추면 경계선이 좀더 뚜렷해짐으로써 방사형의 모양을 가지는 침상 돌기들 역시 보다 분명하게 보인다[9].

그림 3.7 은 침상 돌기를 보이는 병변을 다 해상도 분석을 통하여 얻는 가우시안 피라미드의 영상을 보인 것으로 J 레벨의 불분명한 경계선과 침상 돌기의 모습[그림 3.7 (a)]이 레벨을 변화시킨 그림 3.7 (b), (c)에서 좀더 뚜렷하게 보이는 것을 알 수 있다. 그러나 다 해상도 분석의 레벨을 변화 시켜 가우시안 영상을 얻는 것은 일정 레벨을 지나면 영상의 고주파수 성분이 필요 이상으로 제거되어 병변의 중요한 특징인 방사형의 돌기가 제거될 수도 있다. 이는 Lee[26] 등에 의해 지적된 것으로 침상 돌기를 보이는 병변의 특징이 유지되는 분석 레벨을 찾는 것이 중요하다.

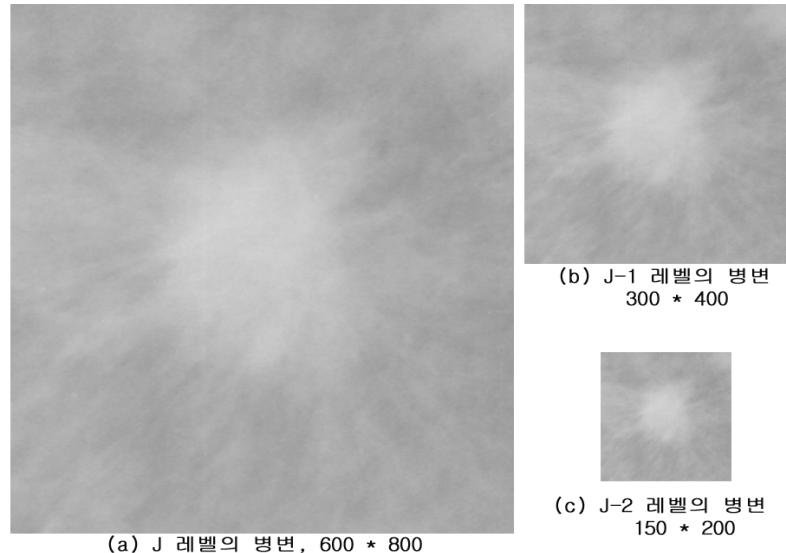
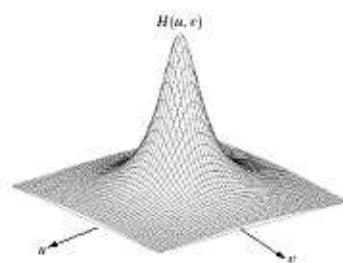


그림 3.7 : 다 해상도 분석을 이용한 침상 돌기를 가지는 병변의 영상

다 해상도 분석 시 사용한 가우시안 커널은 그림 3.3에서 보았듯이 가우시안 형태와 가장 유사한 모양을 보이는 $a = 0.4$ 의 커널을 사용한다[그림 3.8][19].

0.0025	0.0125	0.02	0.0125	0.0025
0.0125	0.0625	0.1	0.0625	0.0125
0.02	0.1	0.16	0.1	0.02
0.0125	0.0625	0.1	0.0625	0.0125
0.0025	0.0125	0.02	0.0125	0.0025

(a) 가우시안 커널의 계수



(b) 가우시안 커널을 이용한 2-D 저역 통과 필터

그림 3.8 : $a=0.4$ 의 가우시안 커널

3.2 Histogram Specification

Histogram 은 $[0 \sim L-1]$ 범위의 픽셀 값을 가지는 영상의 픽셀 분포를 알기 위해 사용한다. 반대로 영상의 Histogram 분포의 형태를 수학적인 식에 의하여 임의로 변형한 후 원 영상의 픽셀 값을 바꾸기도 한다. Histogram Equalization은 원 영상의 히스토그램 확률 밀도 함수를 자신의 확률 분포 함수에 따라 바꿈으로써 새로운 영상의 확률 밀도 함수를 균일하게 하는 방법이며 Histogram Specification은 원 영상의 히스토그램 확률밀도를 구한 후 이를 주어진 확률 분포 함수에 따라 변화시킴으로써 새로운 픽셀 값을 결정하는 방법이다. Histogram Equalization 과 Histogram Specification 은 영상이 전체적으로 밝거나 어두운 경우처럼 픽셀 분포가 한곳으로 몰려 있는 영상에 대해 높을 대조도를 얻을 수 있으므로 다양한 분야에서 사용한다[20].

3.2.1 기본 원리

히스토그램은 영상의 픽셀 분포를 알기 위해 사용한다. 그림 3.9 는 다양한 대조도를 가지는 영상과 해당 영상의 히스토그램 분포를 보여준다. 그림 3.9 의 (a), (b), (c) 의 영상은 너무 밝거나 너무 어두운 영상으로 해당 영상의 히스토그램 분포 또한 한곳으로 몰려 있어 대조도가 낮음을 확인 할 수 있다. 그러나 그림 3.9 의 (d) 는 대조도가 높은 영상으로 해당 영상의 히스토그램 분포가 $[0 \sim L-1]$ 범위의 픽셀 값 전 구간에 걸쳐 고르게 분포 되어 있어 영상의 세부 구조가 분명하게 보임을 알 수 있다. 그림 3.9의 (a), (b), (c) 영상과 같이 히스토그램의 분포가 한곳으로 몰려있어 대조도가 낮은 영상을 (d) 의 영상처럼 히스토그램의 분포를 변형시켜 영상의 대조도를 높이되 히스토그램 분포를 균일한 형태가 되도록 하는 것이다. 그림 3.10 과 식 (3-4) 는 Histogram Specification 적용 시 사용하는 $N(\mu, \sigma)$ 의 가우시안 확률 밀도 함수를 보인 것이다.

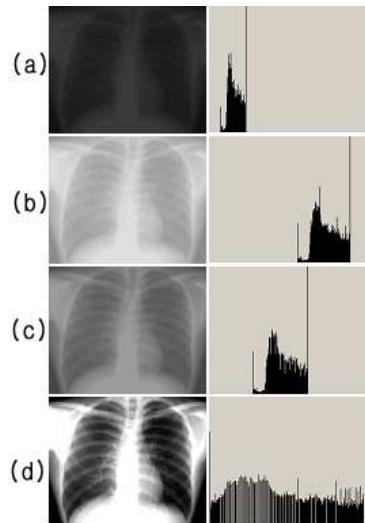


그림 3.9 : 다양한 영상의 히스토그램 분포

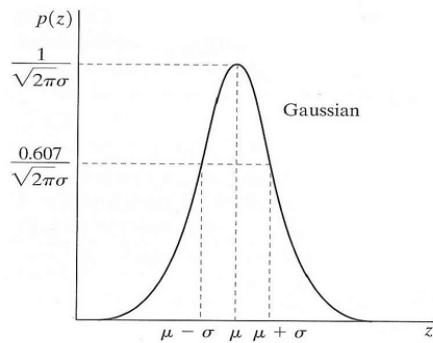


그림 3.10 : 가우시안 확률 밀도 함수

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2}, \quad N(\mu, \sigma) \quad (3-4)$$

가우시안 확률 밀도를 따르는 Histogram Specification 을 적용하기 위한 과정은 아래와 같은 단계를 거친다.

- 1) 입력 영상의 픽셀 값 r_k 에 대한 확률 밀도 함수를 구한다.

$$s_k = T(r_k) = \sum_{j=0}^k P_r(r_j) = \sum_{j=0}^k \frac{n_j}{n} \quad (j, k = 0, 1, 2, \dots, L-1)$$

2) 출력 영상의 픽셀 값 z_k 에 대한 확률 밀도 함수를 가우시안 분포에 따른 확률

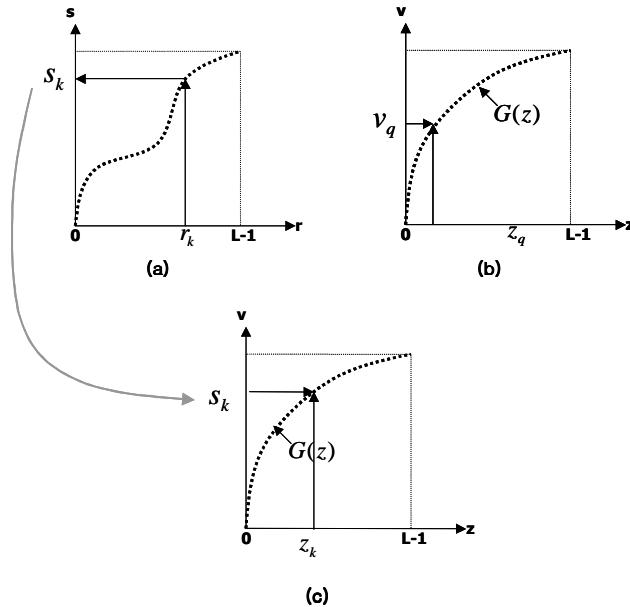
밀도 함수로 매핑하기 위해 가우시안 확률 밀도 함수 $G(z_k)$ 를 구한다.

$$G(z_k) = \sum_{i=0}^k P_z(z_i)$$

$$s_k = G(z_k)$$

3) s_k 를 기준으로 하여 입력 영상의 픽셀 값 r_k 을 출력 영상의 z_k 로 변환한다.

$$z_k = G^{-1}[T(r_k)] = G^{-1}[s_k], \quad (k = 0, 1, 2, \dots, L-1)$$



(a) $T(r_k)$: 입력 영상의 확률 밀도 함수

(b) $G(z_k)$; 가우시안 분포에 따른 확률 밀도 함수

(c) s_k 를 기준으로 한 (a) 와 (b) 의 매핑

그림 3.11 : Histogram Specification 매핑 과정

그림 3.11 은 가우시안 확률 분포를 따르는 Histogram Specification을 적용하기 위한 과정을 그림으로 나타낸 것으로 그림 3.10 의 (c) 와 같이 s_k 를 기준으로 입력 영상의 확률 밀도 함수와 가우시안 확률 밀도 함수를 매핑 함으로써 출력 영상의 확률 밀도 함수를 가우시안의 형태로 바꿀 수 있다.

3.2.2 적용

유방은 다른 기관과 달리 매우 유사한 음영의 연조직으로 구성된 기관으로 높은 대조도의 촬영이 요구되므로 X-선 유방 촬영술은 이러한 유방의 특징을 고려하여 유방 내부의 조직들이 영상에 잘 나타나도록 고안된 특수 촬영 장치이다[1,3]. X-선 유방 촬영 영상은 종괴 등을 포함한 병변이 높은 픽셀 값을 가지는 특징을 가지고 있으나[9,10] 일부 병변은 훈련된 유방 방사선과 전문의 조차 병변과 주변조직을 육안으로 구별하는 것이 어려울 정도로 병변과 병변 외 지역의 대조도가 낮다[7].

실제 병변을 포함한 영상[그림 3.12 (a)]의 히스토그램 분포[그림 3.12 (b)]를 보면 픽셀 값의 분포가 한곳으로 몰려 있음을 확인 할 수 있다. 그림 3.12 (c) 는 (a) 영상을 가우시안 확률 밀도 분포를 따르는 Histogram Specification을 적용했을 때 결과 영상이다. 병변의 불분명한 경계선과 낮은 대조도로 인하여 병변과 주변조직의 구별이 육안으로 어려웠으나 변환 후 대조도가 높아져 육안으로 병변의 형태를 구분 할 수 있으며 변환 영상의 확률 분포 또한 가우시안의 형태로 바뀌었음을 확인할 수 있다[그림 3.12 (d)].

Histogram Specification의 적용으로 인하여 병변과 병변 외 조직의 대조도를 높임으로써 병변의 검출을 보다 쉽게 할 수 있다.

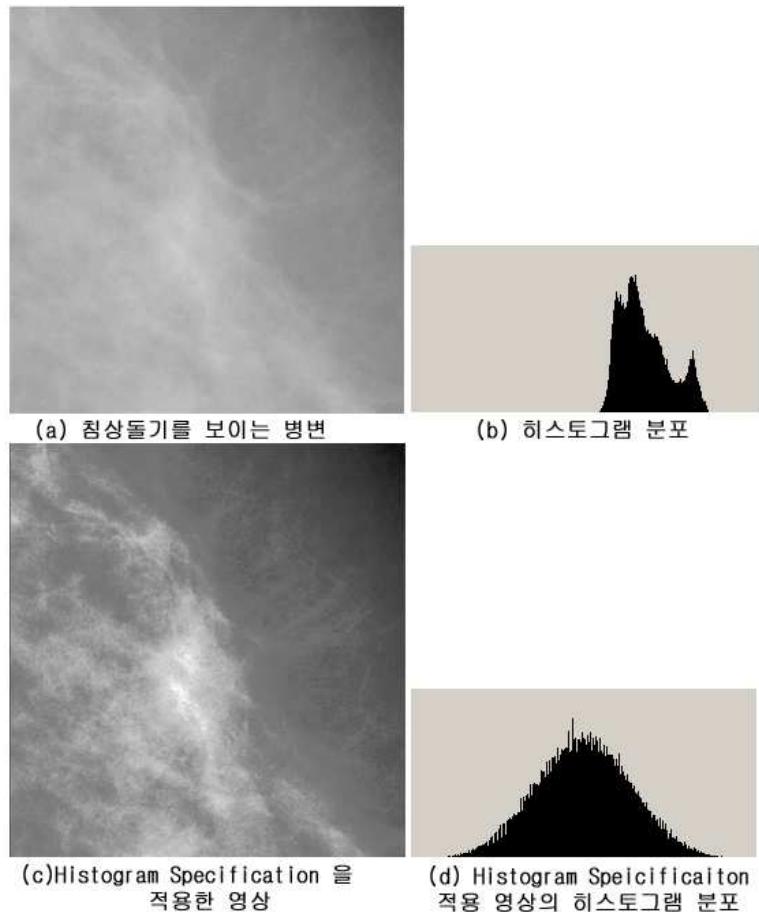


그림 3.12 : 침상 돌기를 보이는 병변의 Histogram Specification 적용 예

3.3 Maximum Likelihood Segmentation

Maximum Likelihood Segmentation 은 병변의 검출에 있어 가장 핵심이 되는 과정으로 주변조직과의 대조도가 낮은 유방 촬영 영상으로부터 자동으로 그리고 정확하게 병변으로 추정되는 영역을 추출하기 위한 방법이다. 이 방법은 Kinnard 등이 X-선 유방 촬영 영상의 양성과 악성 종양을 구별하기 위한 연구의 한 과정으로 제시한 것으로[10,11], 확률 분포에 의해 최적의 Decision Threshold 를 결정하기 위해 Maximum Likelihood Estimation 과 Segmentation 를 결합한 방법이라

할 수 있다.

3.3.1 Segmentation

Segmentation 은 영상으로부터 특정 물체를 추출하는 영상 처리 기법으로, 추출하고자 하는 물체와 배경을 구별 짓는 특징을 찾는 것이 핵심이다. 가장 보편적으로 사용되는 Segmentation 기법은 임계 값을 적용한 Segmentation 이다. 그림 3.13 의 (a) 는 배경과 물체로 이루어진 영상이며 (b)는 해당 영상의 히스토그램 분포를 나타낸다. 배경과 물체의 히스토그램 분포가 각각 다른 분포를 보이므로 두 분포 사이의 값을 임계 값으로 하여, 임계 값보다 작은 픽셀 값을 배경으로 간주하여 물체를 추출 할 수 있다(식3-5). 그림 3.13 (c) 는 임계 값을 적용하여 물체를 추출한 결과를 나타낸 그림이다.

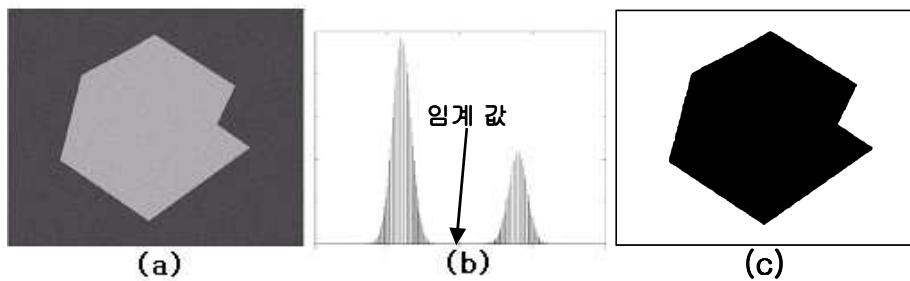


그림 3.13 : 히스토그램 분포와 Segmentation

$$f(x, y) \begin{cases} 0, & f(x, y) \text{ 가 임계값 보다 작을 때} \\ f(x, y), & f(x, y) \text{ 가 임계값 보다 크거나 같을 때} \end{cases} \quad (3-5)$$

영상으로부터 원하는 물체를 추출하기 위한 임계 값을 선택하기 위해서는 추출하고자 하는 물체와 배경의 픽셀 값의 분포를 아는 것이 중요하다. 그림 3.14 의 (a) 는 각 영상에서 배경의 히스토그램 분포와 추출하고자 하는 물체의 히스토그램 분포를 분리시켜 나타낸 것이고 (b)의 영상은 두 분포를 중첩시켜 나타낸 것이다. 그림 3.14의 첫 번째 열의 히스토그램 분포는 배경의 분포와 물체의 히스토그램의 분포가 겹쳐지는 부분이 매우 적다. 배경과 물체의 분리를 위한 최적의 임계 값

(Optimal Threshold)은 두 히스토그램 분포가 겹쳐지는 부분이 가장 적은 곳의 값을 임계 값으로 선택하여, 배경을 물체로 혹은 물체를 배경으로 잘못 분류하는 Segmentation 에러를 가장 적게 하는 값이 될 것이다. 첫 번째 열의 히스토그램 분포에서 물체와 배경 분류를 위한 가장 최적의 임계 값(Optimal Threshold)은 두 분포 사이의 오목한 부분(Conventional Threshold)으로 설정 할 수 있다. Segmentation 에러 또한 최소한으로 발생함을 알 수 있다.

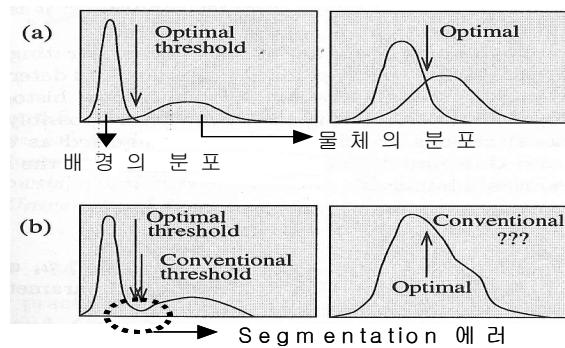


그림 3.14 최적의 임계 값 설정을 위한 히스토그램 분포 분석의 예

그림 3.14의 두 번째 열의 히스토그램 분포는 대부분의 Segmentation 과정에서 볼 수 있는 경우로 최적의 임계 값(Optimal Threshold)을 찾는 과정에 문제가 발생한다. 배경과 물체의 히스토그램 분포를 구분 할 수 있는 오목한 부분(Conventional Threshold)을 발견 할 수 없기 때문에 Segmentation 에러를 가장 적게 하는 임계 값을 찾는 작업이 어려워짐을 알 수 있다[21]. 그림 3.14 의 여러 가지 예에서 볼 수 있듯이 성공적인 Segmentation을 위해서는 물체와 배경을 구분하되 Segmentation 에러는 최소로 하는 최적의 임계 값을 찾는 것이 핵심이라 할 수 있겠다.

3.3.2 Maximum Likelihood Estimation

Maximum Likelihood Estimation 은 $N(\mu, \sigma)$ 의 가우시안 확률 분포를 따르고 서로 독립적인 집단들이 있을 때, 집단들의 데이터를 분류하되 에러가 가장 적게 발생하는 임계값을 찾는 방법이다.

생하는 Maximum Likelihood 를 찾는 방법이다[22].

가우시안 확률 분포를 따르는 n 개의 데이터 분포가 있을 때 이들 데이터 분포의 분류를 위한 Maximum Likelihood를 선택하기 위한 과정은 다음과 같다.

1) 가우시안을 따르는 n 개의 데이터 분포의 확률 밀도 함수가 있다.

$$p_1(z), p_2(z), \dots, p_{n-1}(z), p_n(z)$$

2) n 개의 데이터 분류를 위한 Likelihood , $p(z)$ 를 구하기 위해 각 확률 밀도 함수를 곱한다. .

$$p(z) = \prod_{i=0}^n p_i(z)$$

3) Likelihood 와 같은 의미로, 로그를 취한 각 분포의 확률 밀도 함수를 더하여 Log-likelihood , $\ln(p(z))$ 를 구할 수 있다.

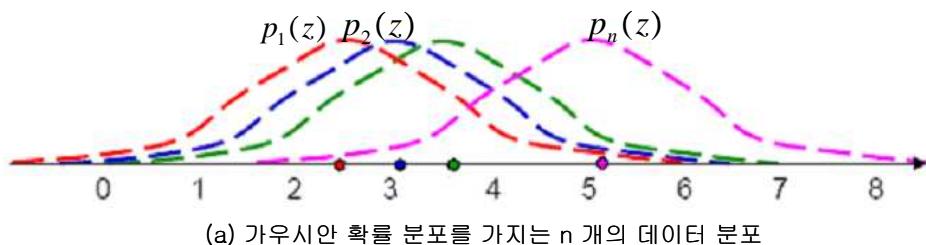
$$\ln(p(z)) = \sum_{i=0}^n \ln(p_i(z))$$

4) log likelihood를 최대로 하는 값을 Maximum Likelihood 로 설정한다. 미분이 0 이 되는 지점을 찾기도 한다.

$$\text{Maximum Likelihood} : \arg\max(\ln(p(z)))$$

$$\text{Maximum Likelihood} : \frac{d}{dz}(\ln(p(z))) = 0$$

그림 3.15 는 Maximum Likelihood Estimation 과정을 나타낸 것으로, (c) 의 그림은 Log likelihood 값을 최대로 하는 값을 보여준다. 최대 값 혹은 미분이 0인 값을 선택하여 데이터를 분류함으로써 분류에 따른 에러를 최소화 할 수 있다.



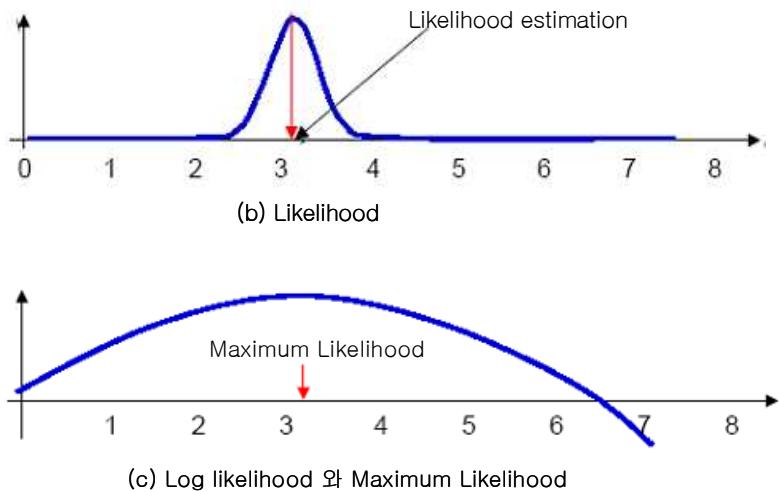


그림 3.15 : Maximum Likelihood Estimation 과정

3.3.3 Maximum Likelihood Segmentation

X-선 유방 촬영 영상에서 병변을 포함하는 영상은 크게 병변과 병변 외 조직으로 구분 할 수 있다. 병변의 검출 시 병변의 확률 밀도 함수와 병변 외 조직의 확률 밀도 함수를 정확히 알 수 있다면 Maximum Likelihood Estimation을 이용하여 Maximum Likelihood를 선택하고 이를 최적의 임계 값(Optimal Threshold)으로 설정하여 병변을 추출하는 Maximum Likelihood Segmentation 과정을 적용 할 수 있다[그림 3.16][10].

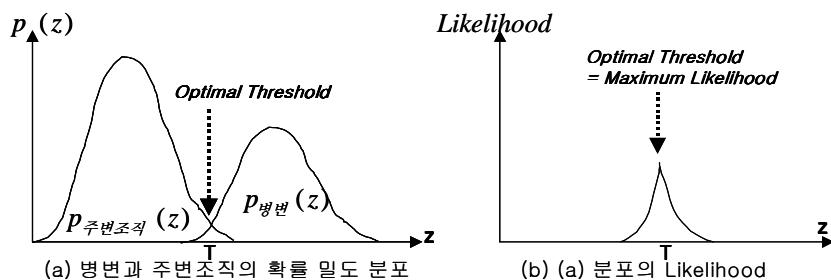


그림 3.16 : 이상적인 경우의 병변과 주변 조직 분류

그림 3.17은 병변을 포함한 영상과 그 영상의 히스토그램 분포를 나타낸 것이다. 그림 3.17의 (b)에서 보인 것처럼 히스토그램 분포로부터 사전에 병변 혹은 병변 외 조직의 확률 밀도 함수를 구할 수 없다.

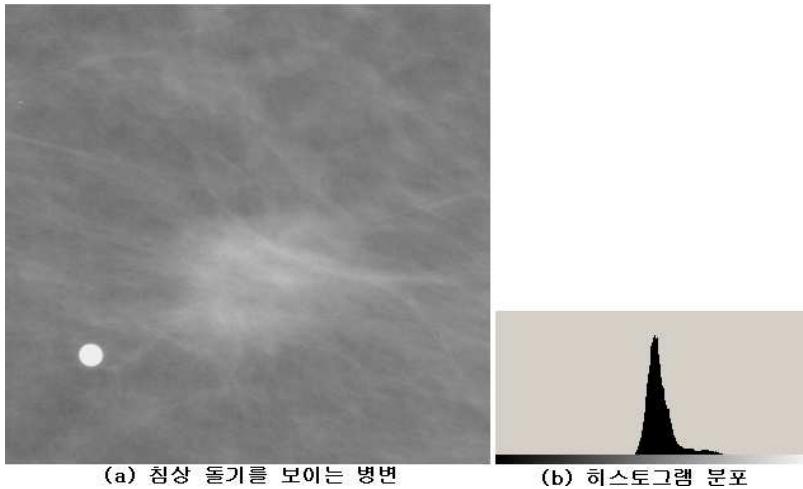


그림 3.17 : 병변과 주변 조직을 포함한 영상의 히스토그램 분포

그러므로 이들의 확률 밀도함수를 추정하기 위해 $[0 \sim L-1]$ 픽셀 값의 범위를 가지는 영상에서 동적으로 임계 값을 변화 시켜 병변과 주변조직을 분리하고 이들의 확률 밀도 함수를 각각 계산한다. 각각의 임계 값마다 추출되는 확률 밀도 함수를 n 개의 데이터 분류로 간주하여, Maximum Likelihood Estimation 방법에 따라 최상의 임계 값을 찾는 방법을 적용하였다.

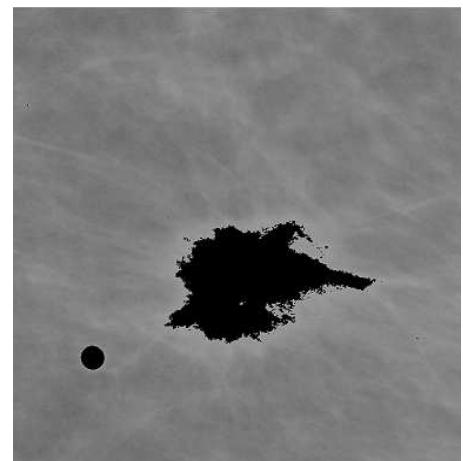
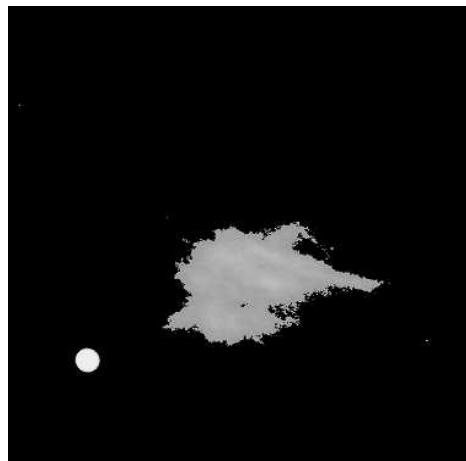
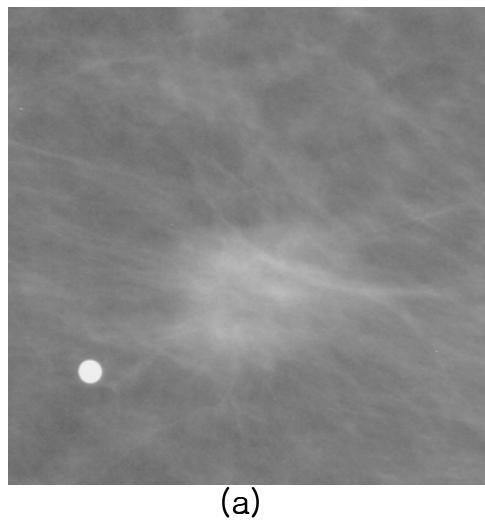
Maximum Likelihood Segmentation을 적용하기 위해 아래와 같은 단계를 따른다.

- 1) $[0 \sim L-1]$ 픽셀 값의 범위를 갖는 영상에서 $L-1$ 의 임계 값(t)을 시작으로 하여 영상에 (식 3-6) 과 (식 3-7)을 각각 적용 한다.

$$f(x,y) \begin{cases} 0 & : f(x,y) \text{ 가 임계 값 보다 작을 때} \\ f(x,y) & : f(x,y) \text{ 가 임계 값 보다 크거나 같을 때} \end{cases} \quad (3-6)$$

$$f(x,y) \begin{cases} f(x,y) & : f(x,y) \text{ 가 임계 값 보다 작을 때} \\ 0 & : f(x,y) \text{ 가 임계 값 보다 크거나 같을 때} \end{cases} \quad (3-7)$$

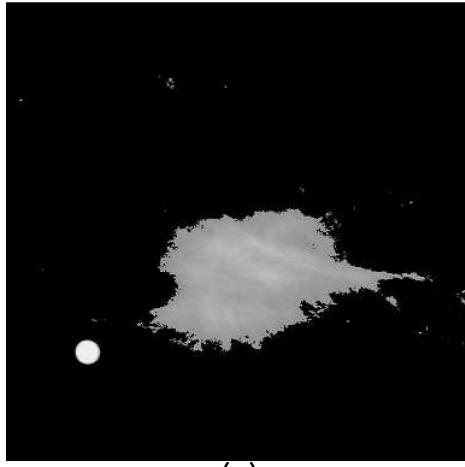
그림 3.18 의 (a) 는 영상 침상 돌기를 보이는 병변을 포함한 영상이며 (b)는 식 (3-6)을 적용한 것으로 임계 값이 2500 일 때 이 보다 큰 영역을 병변 영역으로 가정하고, (c) 의 영상은 식 (3-7)을 적용한 것으로 2500 보다 작은 영역을 병변 외 영역으로 가정하기 위한 것이다. 그림 3.19 은 같은 방법으로 임계 값이 2300 일 때 병변과 병변 외 영역의 예상 영상이다.



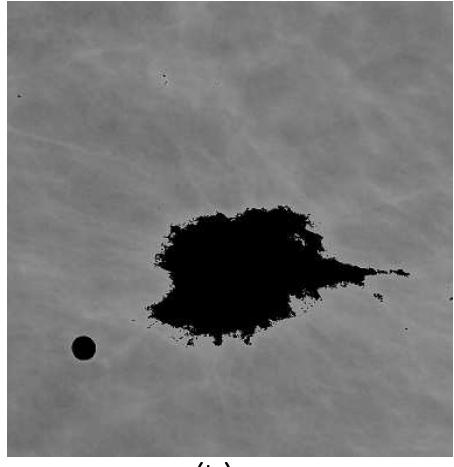
(a) 침상 돌기를 보이는 병변

(b) 병변 예상 영역 (c) 병변 외 영역의 예상 영역

그림 3.18 : 임계 값(2500)을 적용한 병변의 영상 예 I



(a)



(b)

(a) 병변 예상 영상

(b) 병변 외 영역의 예상 영상

그림 3.19 : 임계 값(2300)을 적용한 병변의 영상 예 II

2) 임계 값(t) 적용 시 생성된 병변 예상 영상의 확률 밀도 함수($p_{\text{병변}}(z)$) 와
병변 외 영역 예상 영상의 확률 밀도 함수($p_{\text{주변조직}}(z)$)를 구한다.

3) 해당 임계 값(t)에서의 병변 예상 영상의 확률 밀도 함수($p_{\text{병변}}(z)$) 와 병변 외
영역 예상 영역의 확률 밀도 함수($p_{\text{주변조직}}(z)$)를 이용하여 Log likelihood
적용한다.

$$p_t = p_{\text{병변}}(z) \times p_{\text{주변조직}}(z)$$

$$\log(p_t) = \log(p_{\text{병변}}(z)) + \log(p_{\text{주변조직}}(z))$$

4) 임계 값(t)을 일정 간격으로 감소 시켜 1)~3)를 반복 하여 각 임계 값(t)
에서의 Log likelihood를 구한다.

5) 각 임계 값으로부터 생성된 Log likelihood를 더한다. 그림 3.19 는 계산된

Log likelihood를 보여준다.

$$\text{Log}(P_T) = \sum_{i=0}^n \log(p_i)$$

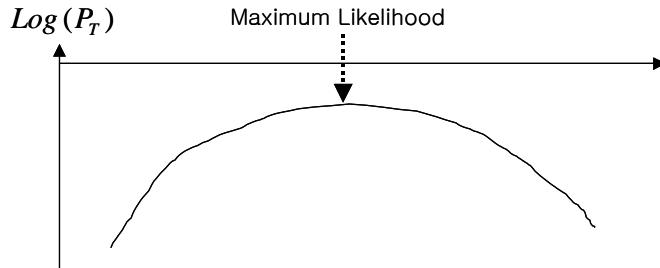


그림 3.20 : Log likelihood의 Maximum Likelihood

6) Maximum Likelihood 값을 선택한다.

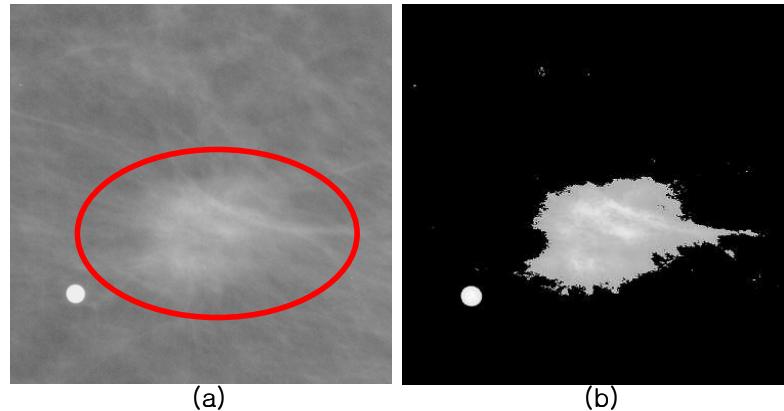
$$\text{Maximum Likelihood} = \text{argmax}(\text{Log}(P_T))$$

$$\text{Maximum Likelihood Change} = \text{argmax}\left(\frac{d(\text{Log}(P_T))}{dT}\right)$$

Babbs[9,11] 등이 침상돌기를 병변의 영상에 Maximum Likelihood

Segmentation을 적용하기 위한 최적의 임계 값으로 Log Likelihood의 최대 변화량을 선택하여 병변을 검출하였음을 보였다. 이는 병변과 병변 외 영역은 대조도에 있어 분명한 차이를 보이므로 병변과 병변 외 영역의 임계 값을 적용하려면 Log Likelihood의 변화가 가장 큰 값을 선택하는 것이 가장 의미 있는 값이 되기 때문이다.

그림 3.21 의 (a) 는 실제 병변 영역을 나타냈으며 (b) 영상은 Maximum Likelihood Segmentation을 적용하여 선택한 최적의 임계 값으로부터 병변이 있는 영상에서 병변 예상 영역을 최종적으로 검출한 영상이다.



(a) 실제 병변 영역

(b) 검출된 병변 예상 영역

그림 3.21 : Maximum Likelihood Segmentation을 적용하여 검출한 병변

제 4 장 침상 돌기를 보이는 병변의 특징 파라미터 추출 및 분류

X-선 유방 촬영 영상에 나타난 침상 돌기를 보이는 병변은 침상 돌기로 인해 방사형의 모양을 가지며 주변 조직과의 경계가 불분명하다. 또한 비 균질적인 조직으로 되어 있으며 대체적으로 주변 조직에 비해 높은 픽셀 값을 가지는 특징이 있다[8,9]. 침상 돌기를 보이는 병변의 특징 파라미터를 추출하기 위해 GCM(Gray level Co-Occurrence Matrix), ALOE(Analysis Local Oriented Edge), Law's texture, 분산, 평균, 엔트로피 등이 다양한 연구에서 다양하게 사용되고 있다. 침상 돌기를 보이는 병변의 가장 대표적인 특징을 보이기 위한 다양한 시도가 있으나, 현재까지 침상 돌기를 보이는 병변만의 절대적인 특징은 없는 것으로 알려져 있기 때문에 GCM, ALOE 등 여러 가지 방법을 사용해서 특징 파라미터를 추출하는 것이 일반적이다[7~11,27]. 병변을 대표하는 특징은 병변과 주변 조직을 분류하거나 혹은 악성 병변과 양성 병변을 분류하기 위한 기준으로 사용하므로 병변의 특징 파라미터를 정확하게 추출하는 것이 중요하다. 본 장에서는 병변 예상 지역으로부터 검출된 영역의 크기, 텍스쳐, ALOE, 통계 등 49 개의 특징 파라미터를 추출하는 방법을 설명하고 로지스틱 회귀분석을 통해 병변의 검출에 있어 유용한 특징 파라미터를 선별하는 로지스틱 회귀 모형 과정을 설명한다.

4.1 텍스쳐를 이용한 특징 추출

텍스처란 물체의 구성 요소나 표면의 특징을 표현하기 위한 것으로 영상에서는 단일 픽셀 보다는 물체를 구성하는 픽셀의 집합으로서 의미가 있는 값이다[21].

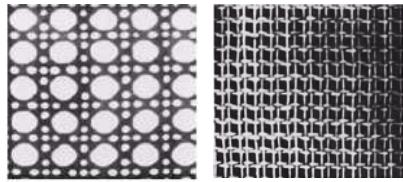


그림 4.1 : 다양한 텍스처

그림 4-1 은 규칙적인 표면의 특징을 가진 영상들로 두 영상의 특징이 다르기 때문에 각 영상의 특징을 텍스처를 이용하여 표현 할 수 있다. 여러 가지 연구에 따르면 침상 돌기를 보이는 병변은 비 균질적이며 엔트로피가 정상 조직에 비해 높은 특징을 가진다고 한다[7~11,27]. 이에 통계적인 패턴을 이용한 방법인 GCM(Graylevel Co-occurrence Matrix)을 이용하여 방향에 따른 엔트로피 등을 포함한 여러 가지 텍스처 특징을 측정하는 방법을 소개한다.

4.1.1 Graylevel Co-occurrence Matrix

GCM(Graylevel Co-occurrence Matrix)은 영상에서 규칙적으로 나타나는 통계적인 패턴을 0° , 45° , 90° , 135° 의 방향에서 측정하여 이의 정보로 매트릭스를 구성하고, 해당 매트릭스를 이용해서 에너지, 엔트로피, 대조도, 균질도 등을 측정하는 방법이다[23]. Co-occurrence Matrix 를 나타내는 $C_d[i, j]$ 는 영상 I를 구성하는 픽셀 값들 중 픽셀 값 i 와 픽셀 값 j 가 거리 d 에 인접하여 발생하는 빈도로 구성된다(식4-1).

$$C_d[i, j] = |\{[r, c] \mid I[r, c] = i \text{ and } I[r + dr, c + dc] = j\}| \quad (4-1)$$

그림 4.2 는 0° , 90° 방향으로 같은 픽셀 값이 인접하는 경우가 빈번한 특징을 가진 4×4 영상(a) 을 이용하여 거리 $d=1$ 일 때 0° , 90° , 45° 방향의 GCM을 구성하는 예를 보인 것이다. 0° , 90° 방향의 매트릭스가 45° 방향보다 큰 값을 가짐을 알 수 있다.

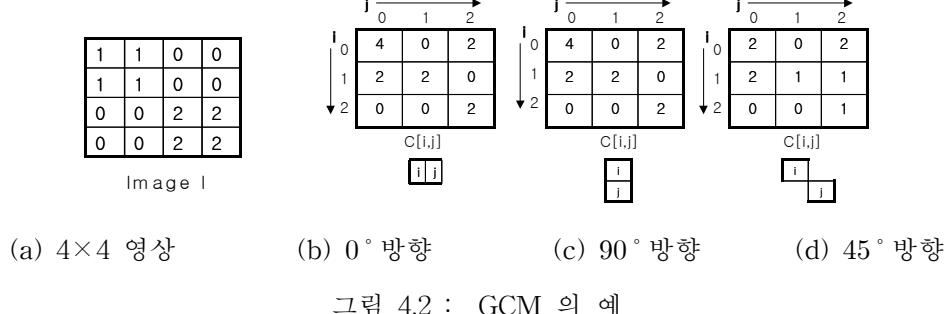


그림 4.2 : GCM 의 예

(식 4-1)에 의해 생성된 GCM은 (식 4-2)와 같이 정규화 하여 각 방향에서 7개의 텍스쳐 특징과 모든 방향의 평균으로 구성된 7개의 특징을 얻을 수 있다(식 4-3 ~ 4-8).

$$N_d[i, j] = \frac{C_d[i, j]}{\sum_i \sum_j C_d[i, j]} \quad (4-2)$$

$$Energy = \sum_i \sum_j N_d^2[i, j] \quad (4-3)$$

$$Entropy = - \sum_i \sum_j N_d[i, j] \log_2 N_d[i, j] \quad (4-4)$$

$$Contrast = \sum_i \sum_j (i - j)^2 N_d[i, j] \quad (4-5)$$

$$Homogeneity = \sum_j \sum_j \frac{N_d[i, j]}{1 + |i - j|} \quad (4-6)$$

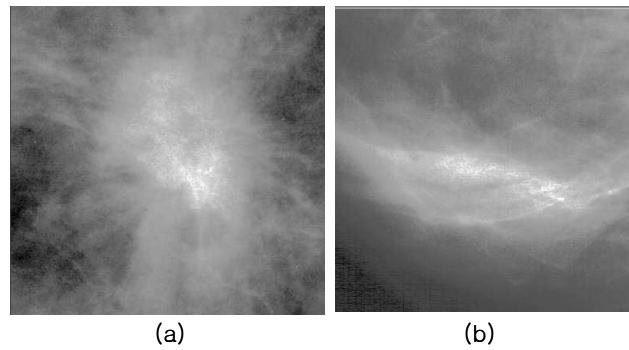
$$ASM = \sum_i \sum_j N_d[i, j] \quad (4-7)$$

$$Mean_i = \sum_i \sum_j i N_d[i, j] \quad (4-8)$$

$$Variance_i = \sum_i \sum_j N_d[i, j] (i - Mean_i)^2 \quad (4-9)$$

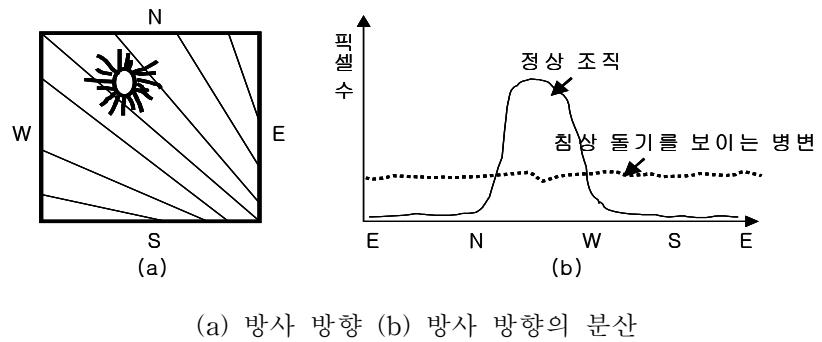
4.2 Analysis Local Oriented Edge를 이용한 특징 추출

침상돌기를 보이는 병변과 정상 조직의 가장 특징적인 구분은 병변의 가장 자리에 발생하는 돌기(spicule)다. 일반적으로 침상 돌기를 보이는 병변의 돌기는 병변의 모든 방향으로 방사하는 형태를 보이는 반면 정상 조직은 유두(nipple)을 향해 일정한 방향으로 방사한다[7,9].



(a) 침상 돌기를 보이는 병변 (b) 정상 조직

그림 4.3 : 유방 촬영 영상에서 조직의 방사 형태의 예



(a) 방사 방향 (b) 방사 방향의 분산

그림 4.4 : 병변과 정상 조직의 방사 방향

그림 4.3 은 X-선 유방 촬영 영상에 나타난 침상 돌기를 보이는 병변과 정상 조

직의 방사 방향을 보여주는 것으로 병변 영상인 (a)는 모든 방향으로 방사하는 형태임에 반해 정상 조직은 한쪽 방향으로 향하는 것을 볼 수 있다. 그림 4.4 의 (a)는 이러한 방사 방향을 일반화 시켜 보여주고, (b)는 방사 방향을 측정한 후 각 방향에 따른 분산을 보여준 것으로 병변의 분산은 일정함에 비해 정상 조직의 분산은 일정 방향에서 두드러짐을 확인 할 수 있다[7,9]. 추출된 병변 예상 영역의 방사 방향에 따른 분산을 추출하는 ALOE(Analysis Local Oriented Edge)는 (식 4-9) 와 같이 계산 한다.

$$\sigma_{hist}(i, j) = \sqrt{\frac{1}{L-1} \sum_0^{L-1} [hist_{ij}(n) - \overline{hist}(i, j)]^2} : \text{픽셀 범위 } [0 \sim L-1]$$

$$\theta(i, j) = \tan^{-1} \left[\frac{D_y(x, y)}{D_x(x, y)} \right] : D_y(x, y), D_x(x, y) \text{ 는 한 점에서 } x, y \text{ 방향으로의 변화 값}$$

$hist_{ij}$: 후보 영역에서의 θ 히스토그램

$$hist_{ij}(n) : [-\frac{\pi}{2} + \frac{n\pi}{L}, -\frac{\pi}{2} + \frac{(n+1)\pi}{L}] \text{ 범위, } \theta = n \text{ 의 픽셀 수}$$

$$\overline{hist}(i, j) = \frac{1}{L} \sum_{n=0}^{L-1} hist_{ij}(n) \quad (\text{식 4-9})$$

4.3 기타 특징 추출

X-선 유방 촬영 영상에 나타난 침상 돌기를 보이는 병변 중 일부 병변은 주변 조직에 비해 밝은 값을 가지는 픽셀들이 모여, 병변 지역의 평균은 높고 분산은 낮은 경향을 보인다[13]. 그리고 양성 유방암의 종괴는 둥근 모양이나 악성 병변에 해당하는 종괴는 모양이 불규칙한 특성을 보인다[그림 2.3][1]. 이에 침상 돌기를 보이는 병변의 특징을 좀더 알아보기 위해 통계적인 특징과 병변의 형태에 따른 특징을 추출하였다.

4.3.1 통계를 이용한 특징

병변을 포함한 영상의 평균과, 분산 그리고 병변 예상 영역에 대한 평균과 분산을 각각 구하였다(식 4-10)(식 4-11).

$$m_G = \frac{1}{N_G} \sum_x \sum_y f(x, y) : N_G 은 전체 데이터 수, f(x, y) 는 픽셀 값$$

$$\sigma_G = \sqrt{\frac{1}{N_G} \sum_x \sum_y (f(x, y) - m_G)^2} \quad (\text{식4-10})$$

$$m_R = \frac{1}{N_R} \sum_x \sum_y f(x, y) : N_R 은 병변 예상 지역 데이터 수, f(x, y) 는 픽셀 값$$

$$\sigma_R = \sqrt{\frac{1}{N_R} \sum_x \sum_y (f(x, y) - m_R)^2} \quad (\text{식4-11})$$

히스토그램 분포의 형태에 따른 특징의 유무를 알기 위해 병변을 포함한 전체 영상과 병변 예상 영역의 Skewness 와 Kurtosis를 각각 추출하였다. Skewness 는 분포의 비 대칭도를 측정하며, Kurtosis 는 분포의 첨도를 측정하는 방법이다 (식 4-12)(식 4-13).

$$Skewness_G = \frac{1}{N_G} \frac{\sum_x \sum_y (f(x, y) - m_G)^3}{\sqrt{\sum_x \sum_y (f(x, y) - m_G)^3}}$$

$$Kurtosis_G = \frac{1}{N_G} \frac{\sum_x \sum_y (f(x, y) - m_G)^4}{\sqrt{\sum_x \sum_y (f(x, y) - m_G)^4}} \quad (\text{식4-12})$$

$$Skewness_R = \frac{1}{N_R} \frac{\sum_x \sum_y (f(x, y) - m_R)^3}{\sqrt{\sum_x \sum_y (f(x, y) - m_R)^3}}$$

$$Kurtosis_R = \frac{1}{N_R} \frac{\sum_x \sum_y (f(x,y) - m_R)^4}{\sqrt{\sum_x \sum_y (f(x,y) - m_R)^4}} \quad (\text{식 4-13})$$

4.3.2 병변의 형태를 이용한 특징

병변 예상 영역의 둑근 정도를 측정하는 Roundness 와 밀집 도를 측정하는 Compactness , 그리고 가로 비와 세로 비를 알 수 있는 Aspect Ratio 를 측정하여 형태의 특징을 추출하였다(식 4.13 ~ 4.15).

$$Roundness = \frac{4 \times Area}{\pi \times Max \ Diameter^2} \quad (\text{식 4-13})$$

$$Compactness = \frac{\sqrt{(\frac{4}{\pi})Area}}{Max \ Diameter} : \quad Area \text{ 는 병변 예상 지역의 면적 } \quad (\text{식 4-14})$$

$$Aspect \ Ratio = \frac{Max \ Diameter}{Min \ Diameter} \quad (\text{식 4-15})$$

4.4 로지스틱 회귀 분석을 이용한 병변 분류

보건, 의학 분야에서는 로지스틱 회귀 분석은 빈번히 사용된다. 보건, 의학 분야에서 다변량 분석으로 로지스틱 회귀분석이 가장 많이 쓰이는 이유는 질병의 발생과 가장 연관이 깊은 위험 인자(Risk Factor)가 무엇인지 추정하고자 하기 때문이다[30] 질병의 위험 인자를 평가하는 방법은 환자 대조군 연구나 코호트 연구를 통하여 교차비와 비교위험도를 측정하는 단변량 분석 방법이 있으나, 질병의 위험 인자가 단일 인자일 경우는 거의 없고 대부분이 복합적이므로 다 요인적인 인자를 변수로 하여 질병의 유무 등에 미치는 영향을 파악하는 로지스틱 회귀 분석을 의학 연구 자료의 많은 부분에서 사용하고 있다[30]. 이에 본 연구에서는 훈련 데이터의 병변 예상 영역으로부터 추출한 다양한 특징 파라미터들 중 병변과 관련 깊은 특징 파라미터를 로지스틱 회귀분석을 통하여 선별하고, 로지스틱 회귀

모형을 이용하여 병변 예상 영역을 병변과 병변 외 영역으로 분류 하였다. 이를 위한 과정으로 SPSS의 데이터 마이닝 프로세서를 이용하였다.

4.4.1 로지스틱 회귀 모형

성공, 실패등과 같이 반응범주의 개수가 두개(binary)인 자료들을 연속형의 특징들을 이용하여 설명하고자 하는 경우에 사용되는 모형이 로지스틱회귀모형이다. 특정 파라미터 x_1, x_2, \dots, x_p 가 주어졌을 때 반응변수가 병변일 경우의 확률을 $P(Y=1)$ 로 병변 외 영역일 경우의 확률을 $P(Y=0)$ 로 나타내면 로지스틱회귀 모형은 (식 4-20)와 같이 표현할 수 있다.

$$\log \frac{P(Y=1)}{P(Y=0)} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (\text{식 4-20})$$

따라서 특정 파라미터 x_1, x_2, \dots, x_p 가 주어졌을 때 검출된 병변 예상 영역이 병변 일 경우의 사후 확률은 (식 4-21)로 표현할 수 있다.

$$P(Y=1) = \frac{1}{1 + e^{-Z}}, \quad (\text{단 } Z = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p) \quad (\text{식 4-21})$$

4.4.2 병변을 대표하는 특징 파라미터 선별

여러 가지 추출 방법으로 추출한 특징 파라미터들이 실제로 모두 병변을 대표하는 특징 파라미터인가는 로지스틱 회귀 모형 시 판별 변수의 선택을 통해 결정한다. 판별 변수의 선택과정은 제공되는 특징 파라미터들이 병변과 병변 외 지역을 분류하기 위한 가치가 있는지 혹은 다른 특징 파라미터들과 결합되어 사용될 때 의미 있는지를 알아보고, 중복되거나 병변을 대표하는 특징 파라미터가 아니라고 판단되는 특징 파라미터들을 제거하는 과정이다. 특징 파라미터를 선택하기 위한 방법은 의미 있는 특징 파라미터의 선택과 제거를 반복하면서 변수를 찾아내는 방법을 사용한다. 이 방법은 중요한 특징 파라미터를 하나씩 추가로 선택하면서 이미 선택된 특징 파라미터들이 제거될 수 있는지 매 단계마다 검토하는 방법이다. 이 방법은 로지스틱 회귀 모형 시 선택되는 특징 파라미터들 사이의 최적 결합을 찾는 데에 매우 논리적이고 효율적이어서 많이 이용되고 있다[31,34].

제 5 장 데이터 및 실험

본 장에서는 침상 돌기를 보이는 병변의 검출 방법 적용에 사용한 X-선 유방 촬영 영상들과 병변의 검출 결과에 대해 설명한다.

5.1 데이터 준비

실험에 사용된 영상은 X-선 유방 촬영 영상 분석 연구를 위해 사용할 수 있도록 한 Digital Database for Screening Mammography(DDSM)의 영상을 이용하였다. DDSM은 Massachusetts General Hospital, University of South Florida, Sandia National Laboratories, Washington University School of Medicine에 의해 설립, 유지되는 증례 DB로, 유방암과 관련된 질병의 검출과, 진단, 교육 또는 훈련을 보조하는 컴퓨터 알고리즘 개발의 연구를 육성하는 것을 목적으로 한다. 약 2,500 개의 증례를 포함하고 각 증례는 환자 정보, ACR(American College of Radiology)에 근거한 유방 밀도와 병적 평가, 병변의 위치 및 소견을 제공한다. 주로 백인 여성을 대상으로 한 영상이며 각 영상에 대해서는 악성 병변은 최소한 한번 이상의 조직학적 검사를 통해 확진되었고 양성 병변은 어떤 수단으로도 악성이 아님이 감별되었다. 각 영상은 필름 형태의 유방 촬영 영상을 HOWTEK 스캐너를 이용하여 픽셀 당 공간 해상도 $43.5\mu m$ 로 디지털화 된 영상이다. 실험에 사용된 모든 영상은 12 bit이며 Lossless JPEG으로 압축된 Big Endian Encoding이고, 해상도는 1000×2000 에서 5000×7000 로 다양하다[33]. 본 연구를 위해 DDSM에서 침상 돌기를 보이는 병변의 연구를 위해 제공하는 훈련 데이터와 테스트 데이터 볼륨을 이용하였으며, 이를 데이터 볼륨의 총 68 케이스로부터 160개의 X-선 유방 촬영 영상을 선택하였다.

5.1.1 데이터 습득

먼저 DDSM 으로부터 제공되는 Decompressed 모듈을 이용하여 Lossless JPEG 압축 영상을 RAW 영상으로 변환하고 영상의 Big Endian Encoding을 Little Endian Encoding 으로 바꾸었다. X-선 유방 촬영 영상의 해상도가 약 5000×7000 의 대용량 영상도 다수 존재하기 때문에 병변의 검출을 위한 연구에서 빈번하게 사용되는 ROI(Region of Interest) 영상을 이용하여 실험하였다. 표 5.1은 훈련 데이터 와 테스트 데이터의 데이터 구성은 보여주는 것으로 130 개의 침상 돌기를 보이는 병변을 포함한 영상과 30 개의 정상 영상으로부터 130개의 병변을 포함한 ROI 영상과 321 개의 정상 혹은 병변 외 영역의 영상을 임의로 추출하였다. ROI 영상의 해상도는 768×768 이며 병변의 사이즈가 클 경우 1024×1024 로 ROI 영상의 크기를 조절하였다. 실험에 사용한 영상의 침상 돌기를 보이는 병변의 크기는 약 4 mm 에서 약 5 cm 까지 다양하다.

	훈련 데이터	테스트 데이터	Total
병변 영상	95	35	130
정상 영상	246	75	321
Total	341	110	451

표 5.1 데이터 구성

5.2 실험 결과 분석

실험은 훈련 데이터를 사용하여 침상 돌기를 보이는 병변 예상 지역을 검출하는 과정과 검출된 예상 지역의 특징 파라미터를 추출하여 이를 병변과 병변 외 지역으로 분류하는 과정으로 나누었다. 그리고 테스트 데이터를 사용하여 실험 과정을 검증하고, 실험 결과와 검증 결과를 FROC(free response operating characteristic) 를 이용하여 각각의 TPF(True Positivie Fraction)과 영상 당 평균 FP(False Positive)를 분석하였다. 그림 5.1 은 본 연구에서 침상 돌기를 보이는 병변을 검출하기 위한 단계를 나타낸 것이다.

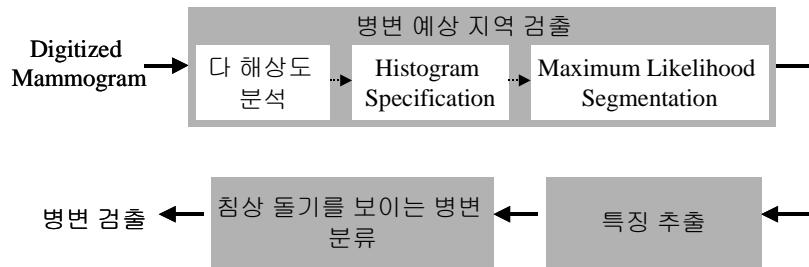


그림 5.1 : 침상 돌기를 보이는 병변의 검출 과정

5.2.1 병변 예상 영역 검출 결과

병변 예상 영역 검출 과정에서는 다 해상도 분석을 하기 위한 적절한 분석 레벨을 찾고 Maximum Likelihood Segmentation 과정을 통해 병변 예상 영역을 검출하고자 하였다. 병변을 포함한 영상을 다 해상도 분석을 적용하기 위해서 고려해야 할 사항은 병변의 크기이다. 표 5.2 은 병변 의 크기가 5mm 일 때 다 해상도 분석 시 병변의 크기 변화와 영상에서 나타나는 픽셀 수를 나타낸 것이다. J 레벨에서 228 픽셀 수를 가지는 병변의 영상이 J-3 레벨에서는 28 픽셀로 감소하므로, 5mm 와 같은 작은 크기의 병변에 대해서도 검출이 가능한 분석 레벨을 찾아야 한다. 이에 본 실험에서는 분석 레벨을 J, J-1, J-2 그리고 J-3 레벨로 나누어 병변 예상 영역 검출 방법을 적용하여 보았다. 실험에서 J 레벨의 영상은 $N \times N$ 의 해상도를 가지며 J-1 레벨은 $\frac{N}{2} \times \frac{N}{2}$, J-2 레벨은 $\frac{N}{4} \times \frac{N}{4}$ 그리고 J-3 레벨은 $\frac{N}{8} \times \frac{N}{8}$ 의 해상도를 갖는다.

다 해상도 분석 레벨	병변의 지름	영상의 지름(픽셀 수)
J 레벨	5 mm	228
J-1 레벨	2.5 mm	114
J-2 레벨	1.25 mm	57
J-3 레벨	0.625 mm	28

표 5.2 다 해상도 분석 상의 병변의 크기

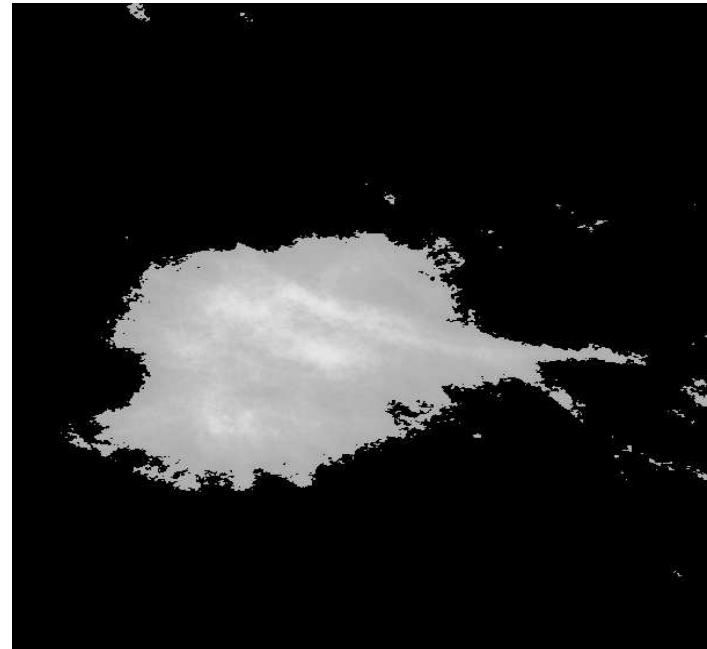
다 해상도 분석 레벨	J 레벨	J-1 레벨	J-2 레벨	J-3 레벨
병변 영역	108	108	108	89
병변 외 영역	648	654	650	583
Total	756	762	758	672

표 5.3 훈련 데이터를 이용한 병변 예상 영역 검출 결과

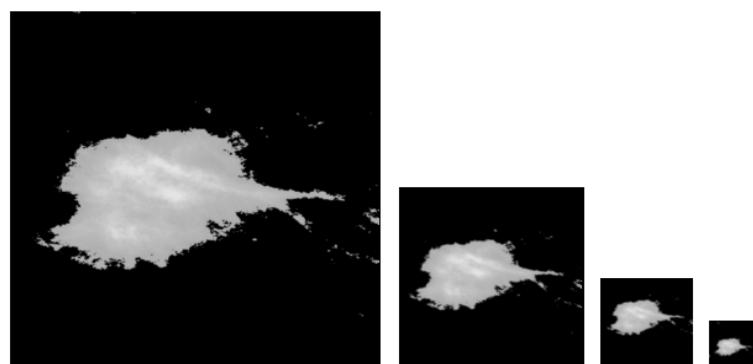
표 5.3 는 훈련 데이터에 병변 예상 영역 검출 방법을 적용하였을 때 검출된 데이터 결과를 나타낸 것으로 검출된 병변 예상 영역이 실제 병변과 일치된 것은 병변 영역이라 하고, 실제 병변이 아닌 영역을 병변 예상 영역으로 검출한 결과는 병변 외 영역이라 구분하여 나타낸 표이다. J 레벨, J-1 레벨과 J-2 레벨에서 모든 병변 영역이 병변 예상 영역으로 검출 되었으나 J-3 레벨에서는 검출이 감소함을 보여준다. 이는 레벨이 변화하면서 병변의 크기가 감소하였기 때문으로 다 해상도 분석 시 병변의 크기를 고려하여 레벨을 선택해야 함을 알 수 있다. 훈련 데이터의 병변 영역은 95개이나, J 레벨과 J-1, J-2 레벨에서 95개 이상의 병변 영역이 검출 된 이유는 실제 병변이 균일하지 못한 특성을 가지고 있어 한 개의 병변이 여러 개의 병변으로 분리되어 검출기 때문이다.

다 해상도 분석과 관련해서 두 번째로 고려해야 할 점은 다 해상도 분석을 적용하여 생성된 가우시안 피라미드 영상에서 침상 돌기를 보이는 병변의 특징을 유지하는 분석 레벨을 찾는 것이다. 그림 5.2 은 Maximum Likelihood Segmentation 만을 적용하여 침상 돌기를 보이는 병변 영역을 검출한 뒤 다 해상도 분석 레벨을 변화시켜 나타나는 병변의 형태를 보인 것으로, 병변의 중요한 특징인 돌기의 형태를 다 해상도 상에서 관찰하기 위한 것이다. (a)의 영상은 J 레벨의 해상도 분석이며 (b)는 J-1 레벨, (c) 의 영상은 J-2 레벨, (d) 는 J-3 레벨, (e) 는 J-4 레벨을 보인 것으로 병변의 특징인 돌기의 형태가 (c), (d), (e) 영상에서 점점 사라지는 것을 볼 수 있다. 이는 병변 예상 지역을 검출하기 위해 사용한 영상이 다 해상도 분석 시 생성되는 저 주파수 성분의 가우시안 피라미드의 영상으로 돌기의 형태와 같은 고 주파수 성분이 레벨이 변화함에 따라 제거되기 때문이다. 또한 다

해상도 분석이 진행됨에 따라 가우시안 피라미드에 나타나는 영상은 스무드(smooth)하게 나타나기 특징이 있기 때문에 엔트로피가 감소하고 Roundness가 증가하는 등 병변의 특징이 점점 감소함을 알 수 있다[표 5.4].



(a) $N \times N$



(b) $\frac{N}{4} \times \frac{N}{4}$ (c) $\frac{N}{8} \times \frac{N}{8}$ (d) $\frac{N}{16} \times \frac{N}{16}$ (e) $\frac{N}{32} \times \frac{N}{32}$

그림 5.2 : 다 해상도 상에 나타난 침상돌기를 보이는 병변의 돌기 변화

	135° 엔트로피	Roundness
J-1 레벨	10.58352	0.47961165
J-2 레벨	10.47788	0.48737864

표 5.4 그림 5.2 영상의 특징들

이에 본 연구에서는 병변 예상 영역의 검출을 위해 J 레벨, J-1 레벨과 J-2 레벨을 선택하였다. Liu[9] 등이 다 해상도 분석의 J-1 의 레벨을 이용하였고, Mudigonda[28] 등은 J-3 레벨을 선택한 연구 결과와 비교해 보았을 때 다 해상도 분석을 이용한 병변의 검출 레벨이 기존의 연구와 크게 다르지 않음을 알 수 있다.

다 해상도 분석을 적용 한 후 각 분석 레벨에서 병변 후보 영역의 검출을 위해 Maximum Likelihood Segmentation를 적용하면 그림 5.3(a) 과 같은 Log Likelihood 를 얻고 병변의 검출을 위한 Maximum Likelihood Change 값을 찾을 수 있다. 그림 5.3(b) 는 Maximum Likelihood Change 부분을 확대한 것이다.

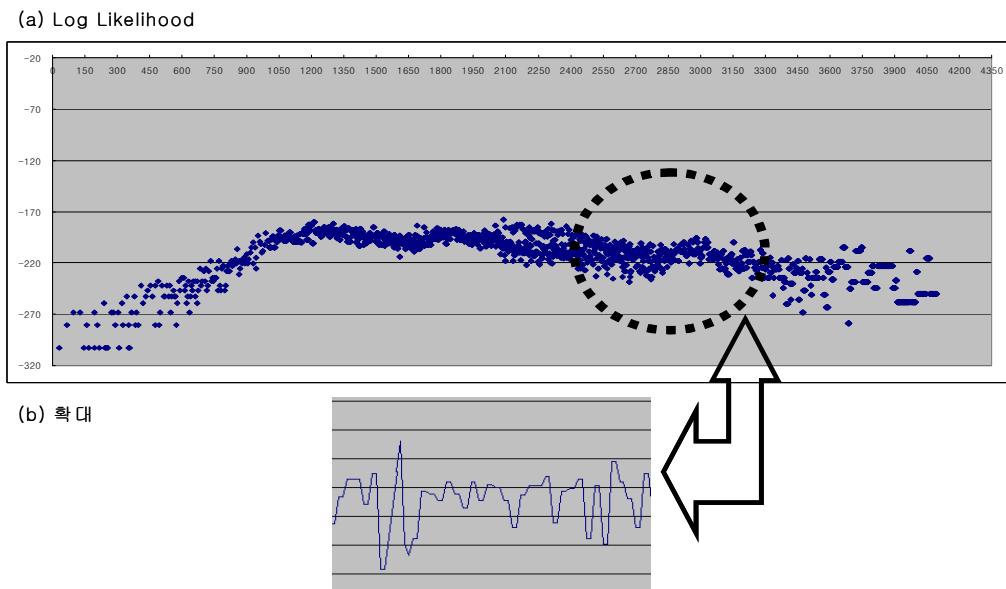


그림 5.3 : 침상 돌기를 포함한 영상의 Log Maximum Likelihood

Babbs[9,11] 등은 Maximum Likelihood Segmentation을 적용한 연구에서 병변 외 조직의 영상에서의 Maximum Likelihood Change 가 병변을 포함한 영상의 Maximum Likelihood Change 보다 훨씬 큰 연구 결과가 관찰되었다고 한다. 그러나 본 연구에서는 Maximum Likelihood Change 에 있어 병변 영상과 병변 외 영역의 영상이 뚜렷한 차이를 보이지 않았다. 이는 Maximum Likelihood Segmentation을 적용하기에 앞서 Histogram Specification 과정을 거쳐 병변을 포함한 영상과 정상 조직의 영상의 대조도를 동일하게 하였기 때문이다.

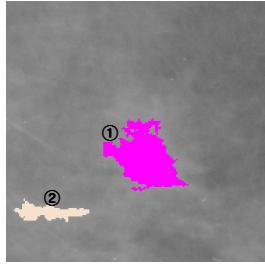


그림 5.4 : J-2 레벨에서 검출된 병변 예상 영역

그림 5.4 는 J-2 레벨에서 검출된 병변 예상 영역을 나타낸 것으로 그림의 ① 은 실제 병변이 검출된 것이며, ② 는 병변 외 영역이 검출 된 것을 보여준다.

본 연구에서 제안한 병변 예상 영역의 검출 방법에 의해 검출된 영역이 DDSM 으로부터 확인한 전문가의 소견에 의한 병변 영역에 포함되거나 일치하면 실제 병변 영역을 검출한 것으로 분류하였다.

5.2.2 침상 돌기를 보이는 병변 검출 및 FROC 분석

검출된 병변 예상 영역으로부터 특징 파라미터를 추출하고 이의 로지스틱 회귀 모형을 이용하여 병변과 병변 외 영역으로 분류함으로써 실제 병변을 검출하고자 하였다. 표 5.5는 로지스틱 회귀 모형 추정 시 각 해상도에서 선별된 특징 파라미터를 나타낸 것이다. Naga[27] 등은 침상 돌기를 보이는 병변은 불균일한 특징을 가져 균일도가 낮고 엔트로피가 높음을 보였으며 Kegelmeyer[7] 등과 Liu[9] 등은

병변의 돌기를 측정하는 ALOE를 주요 특징 파라미터로 이용하였다. 본 연구의 J 레벨에서는 침상 돌기를 보이는 병변의 특징인 Homogeneity 와 엔트로피, 그리고 ALOE 와 형태 특징인 Aspect Ratio 가 병변을 대표하는 특징 파라미터로 선별되었다. J-1 레벨과 J-2 레벨은 GCM 135° ASM을 제외하고 공통으로 Roundness, GCM 0° Entropy, GCM 135° Entropy, ALOE가 선택되었는데 J 레벨에 비해 병변의 특징인 돌기 부분이 두드러지면서 형태 특징인 Roundness 가 중요한 특징 파라미터로 선택되었음을 알 수 있다.

다 해상도 분석 레벨	J 레벨	J-1 레벨	J-2 레벨
선별된 특징	<ul style="list-style-type: none"> ● GCM 135° Entropy ● GCM 평균 Entropy ● GCM 135° Homogeneity ● Aspect Ratio ● ALOE 	<ul style="list-style-type: none"> ● Roundness ● GCM 0° Entropy ● GCM 135° Entropy ● GCM 135° ASM ● ALOE 	<ul style="list-style-type: none"> ● Roundness ● GCM 0° Entropy ● GCM 135° Entropy ● ALOE

표 5.5 훈련 데이터의 각 해상도에서 선별된 특징 파라미터

각 해상도 레벨에서 선별된 특징 파라미터들을 사용하여 추정한 로지스틱 회귀 모형은 (식 5-1), (식 5-2) 그리고 (식 5-3)과 같다. (식 5-1)은 J 레벨의 로지스틱 회귀 모형이며 (식 5-2)는 J-1 레벨, (식 5-3)은 J-2 레벨의 로지스틱 회귀 모형이다.

$$\log \frac{P(Y=1)}{P(Y=0)} = -20.545 + 205.714 GCM135^\circ Homogeneity + 11.539 GCM135^\circ Entropy \quad (\text{식 5-1}) \\ - 10.001 GCM Total Entropy - 1.362 AspectRatio - 0.185 ALOE$$

$$\log \frac{P(Y=1)}{P(Y=0)} = -17.99 - 3.780 GCM0^\circ Entropy + 211.49 GCM135^\circ ASM \quad (\text{식 5-2}) \\ + 5.09 GCM135^\circ Entropy + 2.87 Roundness - 2.44 ALOE$$

$$\log \frac{P(Y=1)}{P(Y=0)} = -4.54 - 1.5 GCM0^\circ Entropy + 1.82 GCM135^\circ Entropy \quad (\text{식 5-3}) \\ + 3.39 Roundness + 0.12 ALOE$$

훈련 데이터의 각 해상도 레벨에서 얻어진 로지스틱 회귀 모형을 이용하여 테스트 데이터의 병변 예상 영역에 적용하여 사후 확률, 즉 병변 가능치를 계산하였다. 병변 가능치는 [0~1]의 범위를 갖는 값으로써 특정 임계 값을 기준으로 임계 값보다 크면 병변으로, 임계 값보다 작으면 병변 외 영역으로 구분하기 위한 수치이다. 유방 촬영 영상에서의 병변 검출에 대한 연구 결과를 평가하기 위해서는 주로 FROC(free response operating characteristic)에 나타난 TPF와 영상 당 평균 FP를 이용하여 성능을 보인다. FROC은 병변의 분류 기준이 되는 임계 값 [0~1] 사이의 모든 값에 대하여 TPF와 영상 당 평균 FP를 계산해서 만들어 내는 곡선이다[32,36].

$$TPF(\text{True Positive Fraction}) = \frac{TP}{TP+FN}$$

$$FP \text{ per image} = \frac{FP}{\text{총 영상 수}} = \frac{\text{실제 : 정상, 분류 : 병변}}{(\text{실제 : 병변, 분류 : 병변}) + (\text{실제 : 병변, 분류 : 정상})}$$

TPF는 실제 병변의 검출율을 보인 것이며 영상 당 평균 FP는 실험에 사용한 영상 당 발생하는 평균 오진율을 나타낸 것으로, TPF는 1에 가까울수록 영상 당 평균 FP는 0에 가까울수록 성능이 좋은 검출 모델이라 할 수 있다.

그림 5.5는 본 검출 방법을 훈련 데이터에 적용시켰을 때의 FROC이다. J 레벨, J-1 레벨과 J-2 레벨에서 병변 분류를 위한 최적의 로지스틱 판별분석을 추출하고, 임계 값을 변화 시켜 병변과 병변 외 영역으로 분류한 뒤 각 임계 값을 해당하는 TPF와 영상 당 평균 FP를 계산하였다. J 레벨에서 TPF가 90% 일 때 영상 당 평균 FP는 약 0.9이며 TPF가 100% 일 때 영상 당 평균 FP는 약 1.5 다. 또한 J-1 레벨에서 TPF가 90% 일 때 영상 당 평균 FP는 약 1, TPF가 100% 일 때 영상 당 평균 FP는 1.6이며 J-2 레벨에서 TPF가 90% 일 때 영상 당 평균 FP는 약 1.9, TPF가 100% 일 때 영상 당 평균 FP는 3.2의 결과를 보였다[표 5.6].

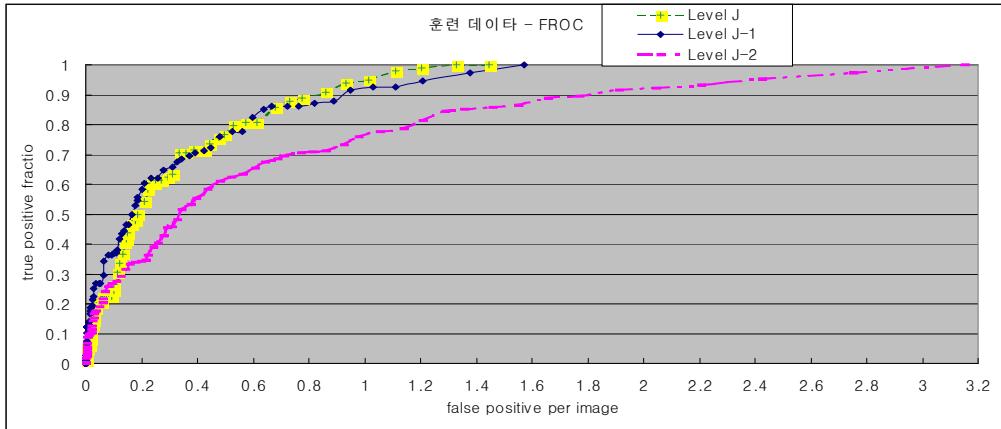


그림 5.5 : 훈련 데이터의 FROC

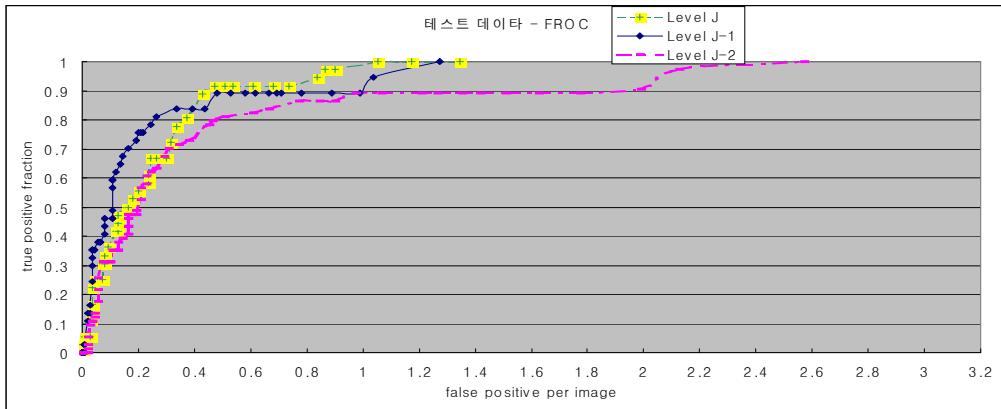


그림 5.6 : 테스트 데이터의 FROC

훈련 데이터에 적용한 병변 검출 결과를 분석한 FROC를 통해 본 연구의 병변 검출 모델의 성능을 확인한 후 테스트 데이터에 병변 검출 모델을 적용하고 해상도 분석을 위해 J, J-1 레벨과 J-2 레벨 각각에서 TPF와 영상 당 평균 FP를 계산하였다. 그림 5.6 은 테스트 데이터에 적용하였을 때의 FROC를 나타낸 것으로 J 레벨의 해상도에서 TPF가 90% 일 때 영상 당 평균 FP는 약 0.8, TPF가 100% 일 때 영상 당 평균 FP는 1.3 이다. 그리고 J-1 레벨의 해상도에서 TPF가 90% 일 때 영상 당 평균 FP는 1, TPF가 100% 일 때 영상 당 평균 FP는 1.3 이며 J-2 레벨의 해상도에서는 TPF가 90% 일 때 영상 당 평균 FP 2, TPF가

100% 일 때 영상 당 평균 FP는 2.6를 보였다. Liu[9] 등이 J-1 레벨의 다 해상도 분석과 ALOE, 이진 결정 트리를 이용하여 TPF가 90% 일 때 영상 당 평균 FP 가 약 1.8, TPF가 100% 일 때 영상 당 평균 FP가 약 2.2 의 검출 결과를 보인 것과 비교해 볼 때, 본 연구에서 병변의 검출을 위한 특징을 다양하게 추출하여 각 해상도 레벨에 맞는 최적화 된 특징 파라미터를 사용함으로써 J 레벨과 J-1 레 벨에서 TPF 와 FP per image 율을 높였음을 알 수 있다. 또한 Cheng[36] 등이 J 레벨에서 GCM 과 펴지 신경회로망을 이용하여 병변을 검출한 연구에서 TPF가 90% 일 때 영상 당 평균 FP는 약 1.4 , TPF가 100% 영상 당 평균 FP 는 약 2.1 의 연구 결과를 보인 것과 비교해 볼 때 , 펴지 신경회로망과 같은 복잡한 분류 방법을 사용하지 않고 병변의 특징을 다양하게 추출함으로써 TPF 와 영상 당 평균 FP를 높였음을 확인 할 수 있다[표 5.6]. 본 연구의 J-2 레벨의 결과 또한 TPF가 90% 일 때 영상 당 평균 FP 가 2 로 병변 검출을 위한 가능성성이 있는 분 석 레벨임을 제시하였다.

해상도 레벨	FP per image	
	TPF 90%	TPF 100%
J	0.8	1.3
J-1	1	1.3
J-2	2	2.6

표 5.6 훈련 데이터의 FROC 분석 결과

	해상도 레벨	FP per image	
		TPF 90%	TPF 100%
Liu등의 연구	J-1	1.8	2.2
Cheng등의 연구	J	1.4	2.1
본 연구	J	0.8	1.3
	J-1	1	1.3
	J-2	2	2.6

표 5.7 테스트 데이터의 FROC 분석 결과

제 6 장 결론

본 논문에서는 DDSM에서 제공되는 X-선 유방 촬영 영상을 사용하여 침상 돌기 를 보이는 병변을 검출 하였다. 본 논문에서 제안한 침상 돌기를 보이는 병변의 검출 방법은 다 해상도 상에서 병변의 특징을 분석하고 병변의 검출이 가능한 분석 레벨을 찾음으로써 검출의 정확도를 높이고 검출 시간은 줄이고자 하였다. 또한 기존의 연구 방법이 병변 예상 영역 검출 시 실험 값을 이용하였던 것과 달리 병변, 병변 외 영역의 확률 밀도를 이용한 Maximum Likelihood Segmentation을 사용하여 보다 일반적인 방법을 사용하였다. 병변 예상 영역의 검출 후 병변과 병변 외 영역으로의 분류를 위해 Analysis Local Oriented Edge, 통계, 텍스쳐, 형태적 특징 등 다양한 특징 파라미터를 추출하였다. 특징 파라미터를 선별하는 과정으로 다변량 분석 방법인 로지스틱 회귀 분석을 이용하였으며, 검출된 병변 예상 영역을 추정된 로지스틱 회귀 모형을 사용하여 병변과 병변 외 영역으로 분류하고 FROC로 분류 결과를 분석 하였다. FROC 분석 결과 J 레벨과 J-1 레벨에서 TPF가 90% 일 때 영상 당 평균 FP는 0.8 과 1이며 TPF가 100% 일 때 영상 당 평균 FP는 각각 1.3을 보였으며 이는 Cheng등의 타 연구에 비해 TPF는 높이 되 영상 당 평균 FP는 낮추는 방법임을 보였다.

침상 돌기를 보이는 병변은 악성 유방암의 대표적인 병변으로 컴퓨터 보조 진단 시 이들 병변의 정확한 검출이 핵심이다. 본 연구의 병변 검출 결과 병변의 핵 부분에 대한 검출은 성공적이나 돌기에 대한 검출은 문제점을 보였다. 그러므로 돌기의 정확한 검출에 대한 연구와 보다 많은 임상 데이터의 적용이 필요하고 ROI 영상이 아닌 전체 영상에 적용하는 연구도 필요하다. 로지스틱 회귀 분석 이외에 페지 신경회로망 등의 분류방법에 대한 연구도 계속되어져야 한다고 생각한다.

본 연구에서 제시한 다 해상도 분석을 통한 침상 돌기의 검출 방법은 병변 검출의 정확도를 높이고 검출 시간을 줄이는 모델을 제시하여 컴퓨터 보조 진단의 연구 시 도움이 되리라 생각한다.

참고 문헌

- [1] 연세대학교 의과대학 진단방사선과학 교실 편, “진단 방사선학”, pp. 512-520, 고려의학, 2003
- [2] 원철호 외5명, “영역 성장법을 기반으로 자동적인 임계치 설정을 이용한 미세 석회화 추출”, 의공학회지, vol 25, No 4, August 2004
- [3] 진료영상학연구회 편저, “진료영상학(1)-방사선영상학 I”, pp. 319-323, 대학서 립, 2001
- [4] 한국 유방암 협회, <http://www.kbcs.or.kr>
- [5] 한만청 편, “진단방사선과학”, pp364, 일조각, 2003
- [6] M. L.Giger, “COMPUTER-AIDED DIAGNOSIS OF BREAST LESIONS IN MEDICAL IMAGES”, IEEE, October 2000
- [7] W. P Kegelmeyer, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs and M. L. Nipper, “Computer-aided Mammographic Screening for Spiculated Lesions”, Radiology 191, pp. 331-337, 1994
- [8] N. Youssary, E. Z. Fatma, Abou-Chadi and Alas'M. El-Sayed, “Early Detectikon of Masses In Digitized Mammograms Using Texture Features and Neuro-Fuzzy Model”, IEEE CNF, March 2003
- [9] S. Liu, C. F. Babbs, E. J. Delp “Multiresolution Detection of Spiculated Lesions in Digital Mammograms”, IEEE Trans. Imaging, vol 10, no 6, June 2001
- [10] L. Kinnard, S. C. B. Lo, P. Wang, M. Freedman, M. Chouikha “Separation of Malignant and Benign Masses using Maximum-Likelihood Modeling and Neural Networks”, SPIE, Med. Imaging, vol 4684, 2002
- [11] M. A. Kupinski and M. L. Giger, “Automated Seeded Lesion Segmentation on Digital Mammograms”, IEEE Trans. Med. Imaging, vol 17, no 4, August 1998
- [12] N. Karssemeijer and G. M. Brake, “Detection of stellate distortions in mammograms”, IEEE Trans. Med. Imaging, vol 15, pp. 611-619, October 1996
- [13] Méndez AJ, Tahoces PG, Lado MJ, Souto M, Vidal JJ. “Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms”, pp. 149-163, Med. Inform., vol26, No 3, 2001
- [14] L. Tabar and P. B. Dean, “Teaching Atlas of Mammography”, Georg Thieme Verlag, 1985

- [15] F. Winsberg, M. Elkin, J. Macy, V. Bordaz and W. Weymouth, "Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis", *Radiology* 89, pp. 211–215, 1967
- [16] F. F. Yin, M. L. Giger, K. Doi, C. E. Metz, C. J. Vyborny and R. A. Schmidt, "Computerized detection of masses in digital mammograms: Analysis of bilateral-subtraction images", *Med. Phys.*, 18, pp. 955–963, 1991
- [17] S. M. Lai, X. Li, and W. Bischof, "On techniques for detection circumscribed masses in mammograms", *IEEE Trans. Med. Imaging* 8, pp. 377–386, 1989
- [18] D. Brazakovic, X. M. Luo and P. Brzakovic "An approach to automated detection of tumors in mammograms", *IEEE Trans. Med. Imaging* 9, pp. 233–241, 1990
- [19] P. J. BURT, E. H. ADELSON "The Laplacian Pyramid as a Compact Image Code", *IEEE Trans. vol. com-31*, no 4, April 1983
- [20] R. C. Gonzalez, R. E. Woods "Digital Image Processing", Prentice Hall, Second Edition
- [21] M. Sonka, V. Hlavac, R. Boyle "Image Processing, Analysis, and Machine Vision", PWS Publishing, Second Edition, pp. 123–188
- [22] R. O. Duda, P. E. Hart, D. G. Stork "Pattern Classification", WILEY-INTERSCIENCE, Second Edition, pp. 84–90
- [23] L. G. Shapiro, G. C. Stockman "COMPUTER VISION", Prentice Hall, pp. 217–224, 2001
- [24] J. C. Russ "The IMAGE PROCESSING Handbook", Second Edition, IEEE PRESSI, pp. 520–530, 1995
- [25] P. J. BURT, "Fast filter transforms for image processing", *Comput. Graphics, Image Processing*, vol 16, pp. 20–51, 1981
- [26] T. K. Lee, M. S. Atkins "A new approach to measure border irregularity for melanocytic lesions", *SPIE. Img. Processing*, vol. 3979, February 2000
- [27] N. R. Mudigonda, R. M. Rangayyan, J. E. L. Desautels "Gradient and Texture Analysis for the Classification of Mammographic Masses", *IEEE. Trans. Med. Imaging*, vol. 19, no. 10, October 2000
- [28] N. R. Mudigonda, R. M. Rangayyan, J. E. L. Desautels "Segmentation and Classification of Mammographic masses", *SPIE. Img. Processing*, vol. 3979, February 2000

- [29] H. Qi, W. E. Snyder "Lesion detection and characterization in digital mammography by Bezier histograms", IEEE. Trans, 1998
- [30] 안재억, 유근영 공저 "의학, 보건학 통계 분석", SPSS 아카데미, pp. 531-554, 2003
- [31] 강현철, 한상태, 최종후, 김은석, 김미경 "데이터 마이닝", 자유 아카데미, pp. 47-60, 2001
- [32] 이재훈, "컴퓨터보조분석에 의한 유방촬영상 종괴와 밀도의 방사선학적 특징 추정 및 위험도 평가", 연세대학교 대학원 석사 학위논문, 2003
- [33] Digital Database for Screening Mammography, <http://marathon.csee.usf.edu/Mammography/Database.html>
- [34] 김기영, 전명식 "다면량 통계자료 분석", 자유 아카데미, pp. 226-231, 1997
- [35] Interactive Mammography Analysis Web Tutorial, <http://sprojects.mmi.mcgill.ca/mammography/index.htm>
- [36] H. D. Cheng, M. Cui "Mass Lesion Detection with a Fuzzy Neural Network", IEEE. Trans, 2003

ABSTRACT

A Study on the Detection of Lesions with Spiculated Border in Digital Mammograms using Multiresolution Analysis

Jo, Jeong-Jin

Dept. of Biomedical Engineering

The Graduate School

Yonsei University

X-ray mammography is currently the best technique in the screening and diagnosis of breast cancer and it is proved to be useful for early detection of breast cancer. Masses are one of the signs that has to be detected in X-ray mammograms. But masses can be hard to detect because, due to the projection, they are often partially covered by glandular tissue. Recent work has shown that many of the tumors that are missed by radiologists can be detected by a system that automatically detects masses. A computer aided diagnosis(CAD) system that prompts suspicious regions can draw the attention of the radiologist to a tumor they might otherwise overlook. In this study, we implemented a method to detect lesions with spiculated border using multiresolution analysis in mammograms. At first step, a

candidate for lesion area is segmented from the training group by multiresolution analysis and maximum likelihood segmentation. And 49 feature parameters extracted from segmented candidate for lesion areas, after that, important features which are concerned with a lesion with spiculated border are selected by logistic regression analysis. Finally, for the test group, we separate lesion with spiculated border and normal tissue using logistic discrimination made of selected features.

The performance of the classification for each analysis level was evaluated by the curve of a free-response receiver operating characteristic (FROC). In the training group, the number of FP(false positive) per image is 0.9 and 1 for level J and J-1 when TPF(true positive fraction) is 90%. The number of FP per image is 1.5 and 1.6 for level J and J-1 when TPF is 100%. For the test group, when TPF is 90%, the achieved number of FP per image is 0.8 and 1 for leve J and J-1. When TPF is 100%, that of FP per image is 1.3 for level J and J-1. This result is superior to that of Cheng who achieved the number of FP is 1.4 when TPF is 90% and that of FP is 2.1 when TPF is 100% at level J for the test group.

Key words : X-ray mammography, Breast cancer, Lesions with spiculated border, Multiresolution Analysis, Logistic regression analysis, FROC