

유방암검출을 위한 컴퓨터 보조진단 시스템*

이철수, 김종국**, 박현욱
한국과학기술원 전기 및 전자공학과
** 삼성전자 소프트웨어센터

Computer-Aided Diagnosis System for the Detection of Breast Cancer

C. S. Lee, J. K. Kim** and H. W. Park

Department of Electrical Engineering, Korea Advanced Institute of Science and Technology

ABSTRACT

This paper presents a CAD (Computer-Aided Diagnosis) system for detection of breast cancer, which is composed of personal computer, X-ray film scanner, high resolution display and application softwares. There are three major algorithms implemented in the application software. The first algorithm is the adaptive enhancement of the digitized X-ray mammograms based on the first derivative and the local statistics. The second one is to detect the clustered microcalcifications by using the statistical texture analysis, and the third one is the classification of the clustered microcalcifications as malignant or benign by using the shape analysis. These algorithms were verified by real experiments.

서 론

유방암은 중년여성에 있어서 사망의 주원인중 하나이고, X-선 유방촬영술은 유방암의 조기진단을 위한 매우 효과적인 의료장비로 알려져있다[1]. 진단 방사선의사들은 촬영된 X선 맘모그램 필름을 통해서 유방암의 병소를 찾게된다. 하지만, 맘모그램영상의 대조도가 낮아서 진단에 어려움이 많다[2].

유방암의 주요 징후중의 하나인 석회화는 양성 및 악성으로 전이하는 과정에서 발생하는 부산물로서 유방조직에 형성되는 칼슘성분의 퇴적물로 알려져 있다. 직경이 0.7mm이하인 석회화를 미세석회화라고 부르고, 맘모그램 필름 해석을 통해 검출되는 유방암의 30-50%가 균집성 미세석회화의 형태를 나타내고 있으며, 조직검사를 통해 판명된 유방암의 60-80%가 균집성 미세석회화를 동반하고 있다고 보고되고 있다[3]. 따라서, 균집성 미세석회화의 존재는 유방암의 주요 징후라고 말할 수 있다.

본 논문에서는 유방암의 징후로 보여지는 균집성 미세석회화를 검출하는데 사용되는 보조 진단 시스템을 소개한다. 본 보조 진단 시스템에 구현된 응용 프로그램에서는 세가지 주된 알고리즘이 구현되어

있다. 첫번째는 낮은 대조도에서 관심영역의 가시성을 높이기 위한 적응적 영상 강조기법이고, 두번째는 맘모그램에서의 미세석회화를 찾는 알고리즘으로 영상에서의 텍스처를 기반으로 한 방법이다. 마지막 세번째는 검출된 미세석회화의 모양을 기반으로 하여 양성 및 악성으로의 판별을 할 수 있는 분류기에 대한 알고리즘이다. 여기에 사용된 알고리즘은 표본 영상에 대한 실험을 통하여 충분히 전문가들에 의해서 확인되었다. 본 시스템은 진단에 따른 보조장치로 유용하게 이용될 수 있다.

본 론

1. 진단 영역의 선택 및 필름 흡제거

진단을 위해서 원 유방영상으로부터 진단영역을 선택하여야 한다. 이는 이산화과정에서 영상이 픽셀당 0.1mm×0.1mm의 크기를 가지게 되고, 양자화과정에서는 픽셀당 12비트의 깊이를 가지므로 유방영상 이외의 배경영역을 제거하여 계산량을 줄일 수 있다. 진단영역으로부터 다시 부영역으로 세분화시켜서 연산을 취한다. 이 때 부영역의 크기는 128×128 픽셀이 되고, 이는 미세석회화를 판별하는데 있어서 1cm² 내에 미세석회화의 수가 일정수 이상으로 존재하면 악성으로 판정될 경우가 많으므로 이를 고려하여 부영역의 크기(1.28cm×1.28cm)를 정했다. 부영역으로 나눔에 있어서 각 미세석회화의 위치에 따라서 검출의 성능이 나빠질 경우를 대비하여 부영역을 64픽셀씩 중첩을 시켜서 설정하므로써 연산량은 늘어나지만, 미검출의 위험을 막을 수 있다.

맘모그램에 있어서 필름의 흠은 미세석회화의 자동감지 및 감별에 있어서 위양성(false-positive)을 증가시키는 요인이 된다. 필름의 흠은 미세석회화와 같이 밝게 나타나고 크기 또한 매우 작기 때문에 미세석회화로 보이는 경우가 있다. 하지만, 필름의 흠은 밝기의 변화가 극심하여 이러한 성질을 이용한 필름의 흠 제거 필터를 구현하였다.

현재 픽셀 (x,y)를 중심으로 3×3 윈도우 내의 픽셀에 대한 구배(gradient)를 구하고, 이렇게 얻어진 9개의 구배값의 최대값이 특정 임계치를 초과할 경우 현재 픽셀은 필름의 흠이라 간주되어 5×5 윈도우내의

* 이 연구는 의공학 선도기술과제의 일환으로 수행되었음

중간값으로 대체시킨다. 여기서 사용한 임계치는 미세석회화는 보존하고 필름의 흠은 제거할 수 있는 값으로 실험을 통해서 구하였다. 그리고, 구배값은 이산 라플라시안 연산자를 사용하였다. 이는 필름의 흠이 3×3보다 작다는 가정 하에서 연산을 취하는 것이므로 가정보다 큰 필름의 흠은 효과적으로 제거될 수 없다.

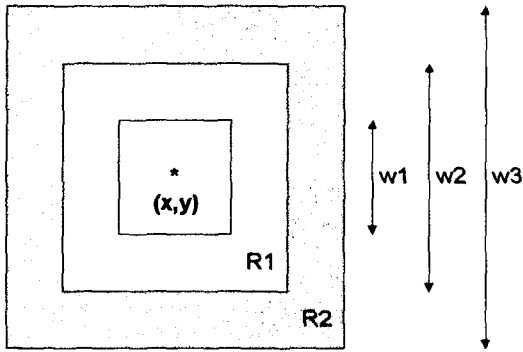


그림 1 현재 픽셀(x,y)에 대한 인접영역의 구성도

2. 영역의 영상강조

개개의 미세석회화는 모양이 이질적이며, 크기가 직경 0.5mm 이내의 경우가 대부분이고, 정상적인 유조직과 같이 있을 경우 위치의 파악이나, 악성 및 양성 감별에 있어서도 판별이 쉽지 않다. 이에 이미 제안된 일차 미분과 국부 통계를 이용한 적응적 영상강조 기법[4]을 적용하여 미세석회화의 위치 및 형태에 대한 가시성을 높일 수 있다.

3. 미세석회화의 위치검출[5]

진단영역에서 나뉘어진 영역에 대해서 Surrounding Region Dependence Method (SRDM)을 적용한다. SRDM은 군집성 미세석회화를 검출하기 위해 이미 제안된 통계적인 텍스처 분석의 한 방법이다[5].

그림 1에서 보는 바와 같이 현재 픽셀을 중심으로 한 두 가지의 인접 영역을 설정하고 각각의 영역을 R1, R2라고 정의하고, 각각의 변을 w1, w2 및 w3로 정한다. 본 시스템에서는 각각의 변을 3, 5 그리고 7로 선택하였다. 이는 X선 mamogram의 스캐닝 해상도와 미세석회화의 병변의 크기를 고려하여 선택되었다. 필름의 흠을 제거한 후 SRDM 행렬을 아래와 같이 산출한다.

$$M(q) = [\alpha(i, j)], \quad 0 \leq i \leq m, 0 \leq j \leq n \quad (2)$$

여기서 m, n은 R1, R2영역내의 픽셀의 갯수를 나타내고 q는 미세석회화와 주변 영역의 밝기 차이에 대한 임계치가 된다. 그리고, 각 행렬원소 $\alpha(i, j)$ 는 아래와 같이 정의된다.

$$\alpha(i, j) = \#\{(x, y) | c_{R1}(x, y) = i \text{ and } c_{R2}(x, y) = j, (x, y) \in L_x \times L_y\} \quad (3)$$

여기서 #은 집합원의 갯수를 나타내며 $L_x \times L_y$ 는 2차

원 공간이다. c_{R1} 과 c_{R2} 는 아래의 식으로 정의 된다.

$$c_{R1}(x, y) = \#\{(k, l) | (k, l) \in R_1 \text{ and } [S(x, y) - S(k, l)] > q\} \quad (4)$$

$$c_{R2}(x, y) = \#\{(k, l) | (k, l) \in R_2 \text{ and } [S(x, y) - S(k, l)] > q\} \quad (5)$$

여기서 R1과 R2는 그림1에서 정의된 주변영역을 나타낸다. 일반적으로 임계치 q값이 클수록 더 많은 미세석회화를 검출하지 못하고, q값이 작을수록 잡음이 미세석회화로 검출될 가능성이 높게 된다. 이렇게 구하여진 행렬 $M[\alpha(i, j)]$ 로부터 네가지의 특징값(Horizontal Weighted Sum, Vertical Weighted Sum, Diagonal Weighted Sum, Grid Weighted Sum)을 추출하여 신경회로망을 통하여 학습을 시키고, 학습된 신경회로망을 이용하여 검출시스템을 구현하였다. 네가지 특징값의 수식은 참고문헌에 나와있다.

4. 검출된 미세석회화의 판별

관심영역으로부터 미세석회화가 검출된 부영역에 대해서만 판별작업을 하게 된다. 위치 검출에서는 텍스처기반으로 구현이 된 반면에 판별은 각각의 미세석회화의 모양에 기반한 특징들을 이용하여 악성 및 양성으로의 판별을 한다.

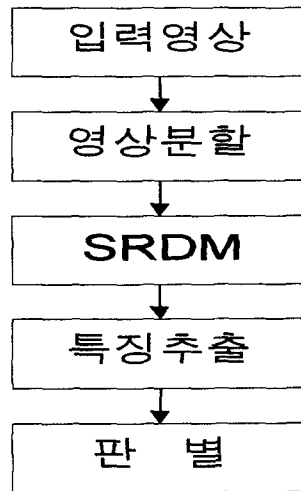


그림 2 검출된 영역의 판별

검출된 미세석회화의 판별은 그림 2에 나와있는 순서로 진행된다. 2절에서 선택된 관심영역을 64픽셀씩 움직여가며 각 부영역이 미세석회화가 검출된 경우에만 판별의 작업을 취한다. 128×128 픽셀크기의 부영역에서 식(4)에서 c_{R1} 이 8보다 크고, 식(5)에서 c_{R2} 가 12보다 큰 픽셀은 미세석회화에 포함된 픽셀로 정의한다. 일차적으로 정의된 이진영상으로부터 세 픽셀 이하의 개체는 잡음으로 여기고 무시한다. 잡음을 제거한 이진영상으로부터 각각의 객체를 표현하기 위한 특징값들을 다음과 같이 정의한다. 이들에 대한 자세한 수식은 참고문헌을 통해서 알 수 있다 [6].

- 가. 크기
- 나. 밝기의 평균

다. 최적 타원의 장축과 단축의 비
라. 이심률

이 네가지는 하나의 객체에 대해서 나타나는 값들이다. 하지만, 하나의 부영역에서 검출된 다수의 미세석회화를 표현하기 위해서는 각 특징들을 잘 표현할 수 있는 방법이 요구되고 이에 제안된 방법을 통해서 각각의 특징들을 나타낼 수 있다. 하나의 부영역에 나타난 미세석회화의 각각의 특징을 아래와 같은 방법으로 나타낸다.

- 가. 평균
- 나. 표준편차
- 다. 최대값
- 라. 최대값 - 평균값
- 마. 평균값 - 최소값

따라서, 각각의 모양 특징별로 다섯 가지의 값이 나오고, 부영역에 나타나는 미세석회화의 개수를 또 하나의 특징으로 설정한다. 이렇게 구하여진 21개의 값으로부터 검출과정에서와 마찬가지로 신경회로망을 이용하여 학습과 실험을 하였다. 신경회로망의 구조는 입력 뉴런층, 은닉 뉴런층, 출력 뉴런층의 구조를 가지는 3계층 신경회로망을 사용하였고, 학습 알고리즘은 EBP(Error Back Propagation) 알고리즘을 사용하였다. 학습의 과정에서 위에서 구한 특징값들의 범위가 클 경우는 학습에 오랜 시간이 요구되므로 정규화 과정을 통해 특징값의 범위를 줄인 후 학습하였고 이 때 사용한 정규화 인자를 저장하여 응용 프로그램에서도 같은 인자를 사용하였다.

5. 실험결과

이상에서 기술한 관심영역의 설정, 필름의 흠 제거 필터, 적응적 영상강조, 미세석회화의 검출과 양성 및 악성의 감별등을 포함하는 기능을 가진 응용 프로그램을 구현하였다. 이 시스템은 그림 3에 나타난 것처럼 PC, 고해상도 모니터 및 X선 맘모그램 스캐너로 구성되어 있다.

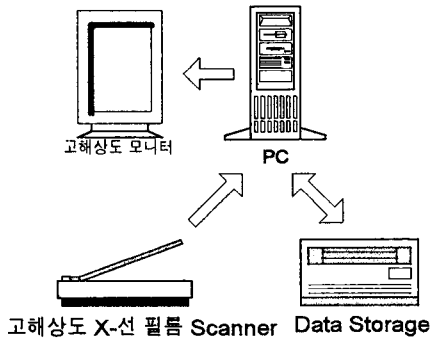


그림 3. 보조 진단 시스템의 블록도

X선 맘모그램을 고해상도 X선 필름 스캐너로 이산화하고, PC를 통해서 모니터에 원 영상 및 처리된 영상을 디스플레이할 수 있다. 스캐닝한 데이터는 대용량 저장장치에 저장할 수 있다.

이상에서 제시된 알고리즘들은 C++을 이용하여 원

도우 응용프로그램으로 구현되었다. 프로그램에 구현된 내용들은 일반적인 영상처리 알고리즘과, 본 논문에서 소개한 맘모그램의 진단을 위한 알고리즘으로 구성되어있다.

사용자가 이산화된 맘모그램 영상을 선택하여, 관심영역을 선택하고 진단에 필요한 과정들을 순차적으로 행할 수 있게끔 구현되었다.

서울 중앙병원으로부터 받은 맘모그램 영상 111매를 이용하여 본 응용 프로그램을 시험하였다. 이들의 영상은 전문의에 의한 육안 또는 조직검사를 통해 미세석회화의 유무 및 악성 및 양성인 양성인 것들이다.

관심영역은 전체 영상에서 유방조직을 전부 포함하는 방법으로 선택되었다. 이러한 과정을 거치지 않고 영상의 모든 부분을 처리할 경우 약 9Mbyte의 메모리가 소요되는 반면, 유방영역만을 선택하여 처리할 경우는 약 3Mbyte에서 5Mbyte 가량의 메모리만을 소요하게 된다. 이는 메모리 용량 뿐만 아니라, 연산시간에 있어서도 많은 이점이 있다.

그림 4에서는 유조직 근처에 존재하는 미세석회화의 원 영상과 영상강조된 영상이다. 보이는 바와 같이 미세석회화가 유조직내에 존재할 경우, 원영상과 강조된 영상을 비교하면, 강조된 영상이 위치파악이나, 형태등의 식별에 용이하다.

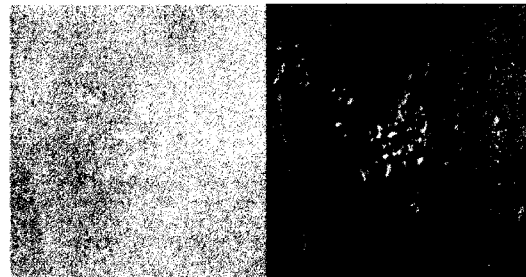


그림 4. (가) 원영상 (나) 강조된 영상

미세석회화의 검출에 대한 실험은 보유한 맘모그램 데이터 베이스를 이용하여 시행하였다. 이 데이터 베이스는 미세석회화를 포함한 영상 69매와 정상적인 영상 42매를 사용하였다. 미세석회화를 포함한 영상은 69매중 49매가 검출에 성공하여 진양성이(True Positive) 71%를 나타내었고, 정상적인 영상에서는 42매중 28매가 정상적인 영상으로 판별되고 나머지 14매가 부분적으로 미세석회화가 있는 것으로 판별이 되어 진음성이(True Negative) 67%를 나타내었다. 정상조직이 양성으로 판별되는 영역은 대부분 밝기가 급격히 변하는 부분, 또는 조직이 밀하게 구성되거나 제거되지 않은 흠들이 양성으로 판별되는 경향이 발견되었다.

위 과정에서 선별된 미세석회화를 포함한 49매의 영상을 이용하여 분류 시험을 시행하였다. 이 49매의 영상에서 미세석회화를 포함하는 영역은 52부분이 있었다. 이 영역은 양성 및 악성의 미세석회화를 포함하는 영역들이고, 이 영역을 정확하게 판별한 영역은 39부분으로 75%의 진양성을 나타내었다.

결 론

본 연구를 통해서 유방암의 초기 상태에서 많이 나타나는 미세석회화의 검출 및 판별의 기능을 가지는 컴퓨터 진단시스템을 개발하고 그에 대한 성능을 시험하였다. 이 응용프로그램의 개발로 컴퓨터를 이용한 미세석회화의 보조진단이 가능하게 되었다. 앞으로 진양성을 증가시키는 방향으로 본 시스템이 사용한 각종 매개변수를 조절시켜 나갈 것이다.

감사의 글

이 연구를 하는데 여러 가지 자료 및 의료 자문을 해주신 서울 중앙병원 진단 방사선과 송준식 박사님과 박정미 박사님께 감사를 드립니다.

참 고 문 헌

- [1] P. C. Johns and M. J. Yaffe, "X-ray characterization of normal and neoplastic breast tissues," *Phys. in Med. and Biol.*, vol. 32, pp 675 - 695, 1987
- [2] D. B. Kopans, "Breast Imaging", Philadelphia, J. B. Lippincott Company, 1989
- [3] W. A. Murphy and K. D. Kecskemeti, "Isolated clustered calcifications in the breast : Radiologic - Pathologic correlation", *Radiology*, vol. 127, pp. 335 - 341, 1978
- [4] J. K. Kim, J. M. Park, K. S. Song and H. W. Park, "Adaptive mammographic image enhancement using first derivative and local statistics", *IEEE Trans. on Medical Imaging*, Oct. 1997
- [5] J. K. Kim, J. M. Park, K. S. Song and H. W. Park, "Texture analysis and artificial neural network for detection of clustered microcalcifications on mammograms", *Proceeding of the 1997 IEEE Workshop, NNSP VII*, pp. 199-206
- [6] A. K. Jain, *Fundamentals of digital image processing*, Englewood Cliffs, NJ, Prentice-Hall, 1989.