

## 통계적 패턴 분류법과 패턴 매칭을 이용한 유방영상의 미세석회화 검출

양윤석·김덕원·김은경\*

= Abstract =

### Detection of Mammographic Microcalcifications by Statistical Pattern Classification & Pattern Matching

Y.S. Yang, D.W. Kim, E.K. Kim\*

The early detection of breast cancer is clearly a key ingredient for reducing breast cancer mortality. Microcalcification is the only visible feature of the DCIS's(ductal carcinoma in situ) which consist 15 ~20% of screening-detected breast cancer. Therefore, the analysis of the shapes and distributions of microcalcifications is very significant for the early detection.

The automatic detection procedures have been the concern of digital image processing for many years. We proposed here one efficient method which is essentially statistical pattern classification accelerated by one representative feature, correlation coefficient. We compared the results by this additional feature with results by a simple gray level thresholding. The average detection rate was increased from 48% by gray level feature only to 83% by the proposed method. The performances were evaluated with TP rates and FP counts, and also with Bayes errors.

**Key words** : Breast cancer, Microcalcification, Correlation matching, Statistical pattern classification

### 서 론

미세석회화[1]는 종괴내에서 발견되는 석회화된 세포를 말한다. 유방암의 종괴의 크기가 10mm에 이르기 전에 발견하는 경우를 '조기 발견'이라고 정의하는데[2], 최근 우리 나라에서도 이를 위한 유방 스크리닝 검사가 산발적으로 시행되고 있는 추세이다. 한 연구에 따르면 스크리닝 검사에 의해 발견되는 유방암의 15~20%가 DCIS이며 이 경우 미세석회화만이 유일한 소견이라고 한다[2, 3]. 그러나 방사선 전문의가 많은 수의 필름영상에서 이러한 미세석회화 소견을 일일이 검출하는 것은 매우 힘들고 많은 시간이 요구되는 작업이기 때문에 여러 가지 디지털 영상 처리 기술을 이용하여 이를 자동화하려는 시도가 계속되어 왔다[4-7]. 본 연구에서는 통계적 문턱치 설

정 방법(statistical gray level thresholding)에 하나의 유용한 특징(feature)인 상관계수(correlation coefficient)를 추가하여 향상된 검출율을 얻었다. 두 방법 모두 Bayes 결정 법칙(Bayes decision rule)을 사용하여 최소 클래스 분류 오차(minimum classification error)를 갖는 경계를 얻었다.

### 재료 및 방법

먼저 디지털 스캐너의 필름 유닛을 사용하여 20장의 생검 확인된 미세석회화 유방영상필름으로부터 디지털 영상을 얻어냈다. 디지털 영상의 해상도는 자동검출을 위한 최소한의 화질[8]을 유지하며 계산량을 가능한 한 줄이기 위해 600 dpi로 하였고 한 화소당 256단계(8-bit)의

연세대학교 의과대학 의용공학교실

Dept. of Biomedical Eng. College of Medicine, Yonsei University

\*연세대학교 의과대학 진단방사선과학교실

\*Dept. of Diagnostic Radiology, College of Medicine, Yonsei University

본 연구는 1996년도 연세대학교 학술연구비의 지원을 받았음.

통신저자 : 김덕원, (120-752) 서울시 서대문구 신촌동 134, Tel. (02)361-5402, FAX. (02)364-1572

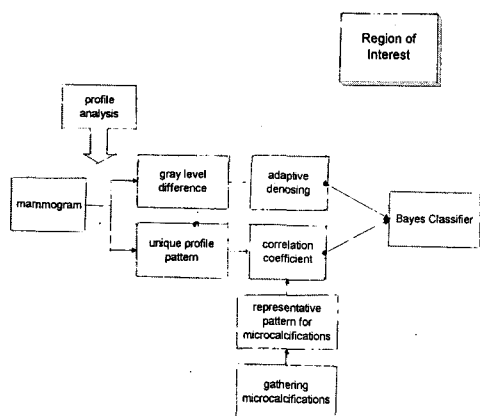


그림 1. 블록 다이어그램  
Fig. 1. Block diagram

gray level을 할당하였다. 모든 계산과 처리는 Matlab의 image processing tool을 사용하였으며 전체적인 처리 과정을 그림 1에 나타내었다.

단면 분석(profile analysis) 과정을 통해 미세석회화로부터 두 개의 특징

1) 정상 조직과의 gray level 차이

2) 주변영상과 이루는 고유한 영상의 패턴

을 추출하였다. 영상에 포함된 noise를 줄이기 위해 디지털 영상을 noise가 혼합된 형태로 모델링한 후 국소적 적응 필터(local adaptive filter)를 사용하여 원 신호에 가깝게 복원한 후 사용하였다. 여기에 일반적으로 사용되는 통계적 문턱치 설정 방법을 적용하여 gray level만으로 미세석회화와 정상 조직을 구분하였다. 검출율을 보다 향상시키기 위해 본 연구에서 제안된 패턴 매칭(pattern matching) 방법을 적용하였다. 즉, 미리 선정된 미세석회화의 고유한 패턴과 영상의 각 위치에서의 유사정도를 표현하는 상관계수를 계산하였다. 이 때 미세석회화의 고유한 패턴 선정은 통계적인 분석에 의하였다. 두 개의 추출된 특징을 이용하여 유방영상의 각 화소를 2차원의 특징 벡터(feature vector)로 나타내면 feature공간상에 대략 두 클래스(class)를 이루며 분포하게 된다. 이 두 클래스 사이에 Bayes 결정 법칙에 의한 경계를 설정하여 최소의 클래스 구분 오차를 허용하며 두 개의 클래스를 구분하였다.

### 1. 단면 분석

미세석회화와 정상조직은 다른 gray level분포를 갖고 있기 때문에 이 둘의 중간값에 해당하는 문턱치를 이용하여 영상을 이진화하면 대략적인 영역추출을 할 수 있게 된다. 따라서 화소의 gray level을 미세석회화의 검출에 적합한 하나의 feature로 이용하였다. 또한 미세석회화와 그 주변의 화소들이 이루는 영상은 고유한 단면(profile)

패턴을 갖는데 이것을 두 영역을 구분하기 위한 또 하나의 feature로 추가하여 사용할 수 있다.

### 2. 전처리

유방 영상은 필름에 상이 맺힐 때의 X-선 산란 등 많은 잡음의 영향을 받게 되므로 적절한 잡음제거처리를 거쳐야 한다. 먼저 잡음의 영향이 포함된 유방영상을 식 (1)과 같이 순수한 영상신호, multiplicative noise와 additive noise가 혼합된 식으로 모델링하였다[8,9]. 식 (2)와 같이 영상에 적응 필터를 적용하여 필터의 출력신호가 처음의 순수한 영상신호에 가깝도록, 즉 M.S.E. (Mean Square Error)를 최소화하도록 식 (4)의 조건을 적용하면 적응 매개 변수  $b(k,l)$ 은 식 (3)과 같이 구해진다.

$$x(k, l) = u(k, l) + u(k, l)q(k, l) + r(k, l) \quad (1)$$

$x(k, l)$  : noise가 섞인 영상신호

$u(k, l)$  : noise가 섞이지 않은 순수한 영상신호

$q(k, l)$  : zero 평균을 갖는 multiplicative white noise

$r(k, l)$  : zero 평균을 갖는 additive white noise

$$y(k, l) = \bar{x}(k, l) + b(k, l)[x(k, l) - \bar{x}(k, l)] \quad (2)$$

$\bar{x}(k, l)$  : ensemble average

$$b(k, l) = \frac{\sigma_u^2(k, l)}{\sigma_x^2(k, l)} \quad (3)$$

$$\frac{\partial}{\partial b(k, l)} E[(u(k, l) - y(k, l))^2] = 0 \quad (4)$$

영상신호를 ergodic random signal로 가정하면 영상의 ensemble 통계치 대신 국소적(local) 통계치를 사용할 수 있게 된다. 따라서 계수  $b(k, l)$ 은 영상의 국소적인 특성을 나타내는 값임을 알 수 있다. 본 논문에서는 국소적 통계치를 얻기 위해  $5 \times 5$  크기의 윈도우내의 화소들을 사용하였다. 이렇게 하면 식 (2)의 적응 필터 처리는 영상의 외곽선 부분에 대해서만 원래 영상을 보존하고 그 외의 부분에 대해서는 국소적인 평균을 취하는 윤곽선을 보존하는 평균 필터(edge preserving mean filter)의 역할을 하게 된다[9].

### 3. 통계적 문턱치 설정에 의한 이진화(Statistical gray level thresholding)

미세석회화와 정상 조직의 gray level 차이에 의해 두 영역을 대략 구분할 수 있음을 앞에서 밝혔다. 이 때 사용되는 문턱치를 Bayes 결정 법칙을 적용하여 통계적인 오차를 최소화하는 값으로 결정할 수 있으나[10], 실제로 적용된 유방 영상에서 비교적 큰 오차를 포함하고 있다.

참고로 이 때 사용된 문턱치는 대부분의 영상에서 160~220 사이의 값으로 나타났다.

4. 상관계수를 이용한 패턴 매칭

미세석회화는 특별한 gray level 형태를 갖고 있으므로 유방영상에 패턴 매칭방법을 적용하였다. 먼저 매칭에 사용될 부분영상(subimage)의 선정이 필요하다. 유방영상에서 보여지는 미세석회화의 모양은 원형, 선형 등 매우 다양하지만 그 중 많은 종류가 원형, 또는 타원형을 갖고 있으므로 검출에 사용될 부분영상은 원형(circular shape)으로 선정하였다. 그 크기는 영상에 나타난 미세석회화들의 크기의 분포를 조사하여 분포도가 높은 11×11(=121) 화소로 선정하였다. 그리고 실제 부분영상은 9장의 유방영상에서 추출한 미세석회화 영상 27개를 평균하여 구하였다. 영상의 한 위치에서의 상관 계수는 그 점을 중심으로 하여 부분영상과 같은 크기만큼 화소들을 취하여 얻어진 국소영상(local image)과 앞서 선정된 부분영상에 대하여 식 (5)에 의해 계산되었다. 이 값은 -1과 1 사이에 분포하고 두 영상의 유사정도를 나타내므로 두 영상이 비슷할수록 큰 값을 갖게 된다.

$$r(s,t) = \frac{\sigma^2_{f(s,t)w}}{\sigma_{f(s,t)}\sigma_w} \tag{5}$$

$$= \frac{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)][w(x-s,y-t) - \bar{w}]}{\sqrt{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)]^2 \sum_x \sum_y [w(x-s,y-t) - \bar{w}]^2}}$$

$r(s,t)$  : 상관 계수

$f(x,y)$  : 국소 영상의 gray level

$w(x,y)$  : 부분 영상의 gray level

5. 이차원 패턴의 클래스 분류(Two-dimensional pattern classification)

이제 영상의 각 화소는 두 가지 특징, gray level(0~255)과 상관계수(-1~ 1)를 원소로 하는 2차원의 특징 벡터로 나타낼 수 있다. 이 벡터들을 2차원 특징 공간상에 나타내면 두 개의 구분 가능한 클래스를 형성하며 특징 공간상에 분포하게 된다. 여기에 마찬가지로 Bayes 결정법칙을 적용하여 최소의 오차를 갖도록 경계를 결정하면, 미세석회화와 정상조직으로 분류할 수 있다. 각 클래스에 속한 특징벡터의 확률적인 분포를 서로 다른 Gauss 분포함수로 가정하고 이에 따라 얻어진 2차 함수(quadratic function)인  $h(X)$ 를 매개적 분류함수(parametric classifier)로 하여 식 (6)과 같이 경계를 설정하였다.

$$h(X) = \frac{1}{2}(X - M_n)^T \Sigma_n^{-1} (X - M_n) - \frac{1}{2}(X - M_m)^T \Sigma_m^{-1} (X - M_m) + \frac{1}{2} \ln \frac{|\Sigma_n|}{|\Sigma_m|} \tag{6}$$

$$\omega_n < \ln \frac{P_n}{P_m} > \omega_m$$

$X$  : 특징 벡터

$M_n, M_m$  : 정상 조직과 미세석회화 특징 벡터들 각각의 평균

$\Sigma_n, \Sigma_m$  : 정상 조직과 미세석회화 특징 벡터들 각각의 covariance matrix

$\omega_n, \omega_m$  : 정상 조직과 미세석회화 각각의 클래스

결 과

여러 장의 유방영상에서 추출한 미세석회화의 평균으로 얻은 부분영상을 그림 2에 나타내었다.

이 부분영상을 사용하여 유방영상의 각 점에서 상관계수를 계산하였다. 그림 3과 그림 4는 표 1, 표 2의 실험 6의 경우에 대한 유방 영상과 계산된 상관 계수를 각각 나타내고 있다.

이 두 개의 값으로 구성된 2차원의 특징 벡터는 그림 5에 나타난 것과 같이 특징 공간상에 대략 2개의 클래스를 중심으로 분포하였다. 미세석회화의 특징 벡터들은 그림과 같이 오른쪽 상단에, 정상조직의 특징 벡터들은 왼쪽 하단에 주로 분포하였다. Bayes 결정법칙에 의해 이들을 그림 6과 같이 두 개의 클래스로 분류하였다.

유방영상을 클래스 분류에 따라 두 개의 클래스로 나누는 것은 이진영상의 결과가 얻어짐을 뜻한다. 그림 7은 (a) 처음 영상, (b) 방사선전문의를 의해 표시된 미세석회화 영상, (c) gray level 문턱치 설정에 의해 이진화된

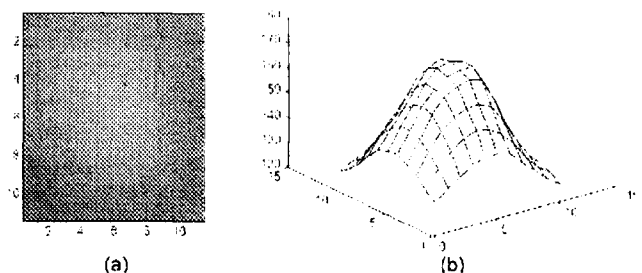


그림 2. 매칭에 사용된 부분영상. (a) 영상, (b) contour Fig. 2. Matching subimage. (a) image, (b) contour

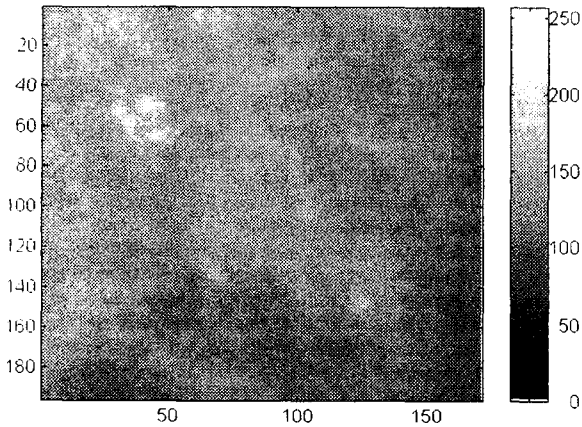


그림 3. 석회화 유방 영상 : 10개의 미세석회화를 포함. 오른쪽의 graybar는 영상의 밝기에 대한 실제 gray값을 나타낸다  
 Fig. 3. Calcified mammogram : it contains 10 microcalcifications. The reference graybar at righthand indicates the real values

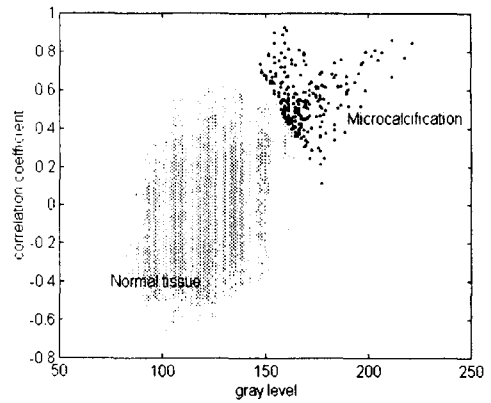


그림 6. Bayes 결정 법칙에 따른 경계에 의해 클래스 구분된 결과  
 Fig. 6. Classified result by Bayes decision boundary

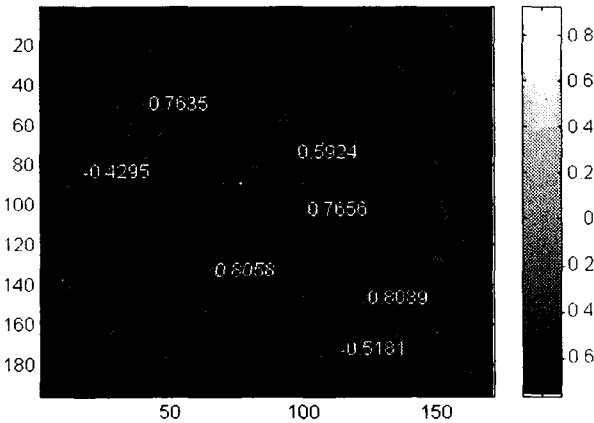


그림 4. 그림 3에 대하여 계산된 상관계수. 오른쪽의 graybar는 영상의 밝기에 대한 실제 상관계수의 값을 나타낸다  
 Fig. 4. Computed correlation coefficients. The reference graybar at righthand indicates the real values

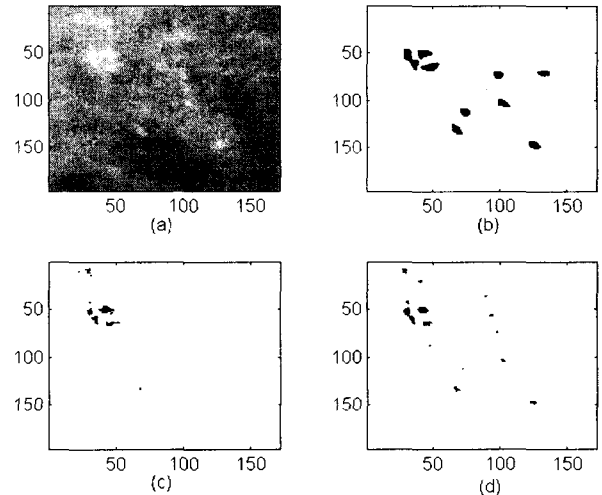


그림 7. 유방영상에 적용된 클래스 구분결과 : (a) 유방 영상. (b) 방사선 전문의에 의해 표시된 미세석회화 영상. (c) gray level 문턱치 설정에 의해 구해진 이진 영상. (d) 본 연구에서 제안된 방법에 의해 구해진 이진 영상  
 Fig. 7. Classifications applied to a mammogram : (a) original image : (a) microcalcifications are marked by radiologist to be used as truth image. (c) binary image acquired by the gray level thresholding. (d) binary image acquired by the proposed method

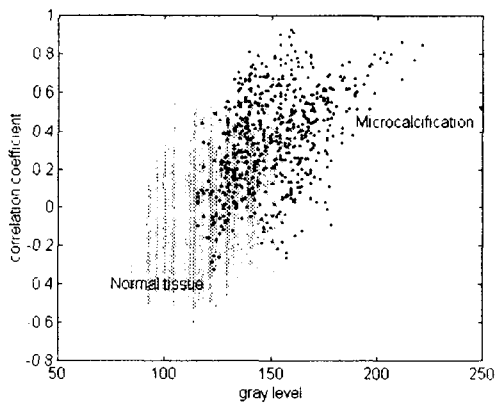


그림 5. 각 2차원 특징 벡터들은 특징 공간상에 두 개의 클래스를 이루며 분포한다  
 Fig. 5. Individual 2-D feature vectors are distributed in feature space, composing two distinguishable classes

영상, (d) 본 연구에서 제안된 방법으로 이진화된 영상을 차례로 나타낸 것이다. 본 연구에서는 이상의 방법을 13장의 미세석회화 영상에 적용하여 클래스 분류 오차와 검출율을 얻었다. 일반적인 gray level 문턱치 설정 방법과 본 연구에서 제안된 방법에 의한 결과를 비교하기 위해 각각 표 1, 표 2에 나타내었다. 통계적 gray level 문턱치 설정 방법에 의한 평균 검출율은 48.0%이었으나 본 연구에서 제안된 상관 계수를 추가한 2차원 클래스 분류법에

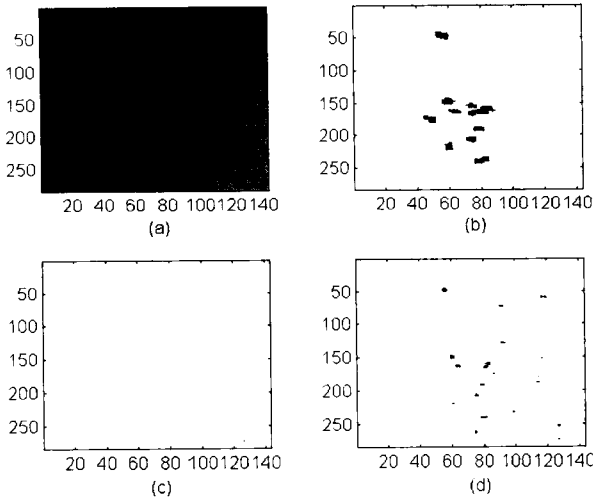


그림 8. Case 6의 결과 : (a) 유방 영상. (b) 방사선 전문의에 의해 표시된 미세석회화 영상. (c) gray level 문턱치 설정에 의해 구해진 이진 영상. (d) 제안된 방법에 의해 구해진 이진 영상

Fig. 8. Results for case 6 : (a) original image. (b) truth image by radiologist. (c) binary image by the gray level thresholding. (d) binary image by the proposed method

의해서는 82.6%로 증가하였다. 따라서 상관계수, 즉 패턴 매칭 계수는 미세석회화를 검출하는 데 유용한 또 하나의 특징으로 사용될 수 있음을 확인하였다.

특히 2차원 클래스 구분법은 상관계수를 사용함으로써

전체적으로 어둡고 미세석회화가 뚜렷이 보이지 않는 희미한 영상, 즉 gray level의 차이가 비교적 크지 않은 영상에 대하여 gray level 문턱치 설정 방법보다 훨씬 높은 검출율을 보였다. 표 1, 표 2의 실험 6의 0.0%에서 75.0%로 향상된 검출율이 이 사실을 잘 나타내고 있다. 두 경우를 비교한 영상을 그림 8에 나타내었다.

## 고 찰

한 가지 주목할 만한 것은 13개의 케이스 중에서 세 케이스(7, 10, 13)가 total error에 있어 gray level 문턱치 설정 방법에 의한 결과보다 더 높은 수치를 보였지만 결과적인 검출율에 있어서는 감소를 보이지 않았다. 그 이유는 하나의 미세석회화는 영상에서 많은 화소들로 이루어져 있어서 그 중 몇 개의 화소들만 정확히 분류되어도 그 미세석회화는 검출되기 때문이다. 또한 gray level과 상관계수 모두 미세석회화의 중심에서 큰 값을 가지므로 검출된 위치도 비교적 정확하게 된다. 이와 같은 이유들로 실제 검출율은 화소단위의 클래스 분류 성능보다 좋을 수 있다.

대개의 경우 TP(True Positive)의 비율이 높으면 이와 더불어 FP(False Positive)의 개수도 다소 증가하게 된다. 마찬가지로 FP의 개수를 줄이기 위해 결정 경계를 조정하면 TP의 비율도 동시에 줄어들게 된다. 따라서 높은 TP 비율을 얻을 것인가 또는 낮은 FP의 개수를 유지

표 1. Gray level 문턱치 설정에 의한 정확도

Table 1. Accuracy of gray level thresholding with Bayes decision rule

실험 번호	Bayes error (%)			True Positive rate (%)	False Positive counts
	error 1(%)*	error 2(%)'	tctal error(%)		
1	3.1	1.0	4.1	87.0	1
2	4.3	0.0	4.3	8.6	0
3	6.2	5.9	12.0	61.4	1
4	0.2	1.4	1.5	50.0	1
5	0.9	0.0	0.9	36.4	0
6	1.7	0.0	1.7	0.0	0
7	1.0	0.5	1.5	100.0	1
8	4.5	0.5	5.0	52.4	2
9	4.1	0.2	4.3	31.4	1
10	3.0	0.1	3.1	20.9	1
11	3.8	0.1	3.9	19.6	2
12	2.3	0.6	2.9	60.0	1
13	3.0	0.6	3.6	96.2	5
average	2.9	0.8	3.8	48.0	1.2

\*error 1 : 정상조직으로 잘못 분류된 미세석회화 화소 수의 전체 화소 수에 대한 비율

'error 2 : 미세석회화로 잘못 분류된 정상 조직 화소 수의 전체 화소 수에 대한 비율

표 2. Gray level과 상관계수를 사용한 2차원 클래스 분류의 정확도

Table 2. Accuracy of 2-dimensional(gray level and correlation coefficient) classification with Bayes decision rule

실험 번호	Bayes error (%)			True Positive rate (%)	False Positive counts
	error 1(%)	error 2(%)	total error(%)		
1	3.4	0.6	4.0	78.3	1
2	4.0	0.1	4.1	51.4	11
3	7.8	4.2	12.0	79.6	1
4	1.3	0.2	1.5	100.0	3
5	0.9	0.0	0.9	63.6	3
6	1.4	0.1	1.5	75.0	2
7	1.3	0.4	1.7	100.0	0
8	3.3	1.6	4.9	100.0	6
9	4.0	0.3	4.3	90.2	23
10	2.8	0.5	3.3	86.1	20
11	3.8	0.1	3.8	53.6	5
12	1.9	0.6	2.5	100.0	2
13	3.3	0.5	3.8	96.2	7
average	3.0	0.7	3.7	82.6	6.5

할 것인가에 관해 적절한 사전 계획(strategy)에 의한 조정이 필요하다. 본 연구에서는 미세석회화의 가능성이 있는 영역의 검출이 보다 중요하므로 FP 개수가 다소 증가하는 것을 허용하여 TP의 비율을 높이는 데 중점을 두었다.

참 고 문 헌

1. B.S. Monsees, "Evaluation of Breast Microcalcification", Radiologic Clinics of North America, vol. 33, pp. 1109-1121, 1995.
2. 오기근, 유방 영상학. 서울, 고려의학, 71-84, 1996.
3. F.M. Hall, "Mammography in the diagnosis of in situ breast carcinoma", Radiology, vol. 168, pp. 279-280, 1988.
4. H.P. Chan and B.L. Shin-Chung, "Computer-aided detection of microcalcifications: Pattern recognition with an artificial neural network", Medical Physics, pp. 1555-1567, 1995.
5. L. Shen and R.M. Rangayyan, "Application of Shape Analysis to Mammographic Calcifications", IEEE Transactions on Medical Imaging, vol. 13, pp. 263-274, 1994.
6. C.J. D'Orsi and M.D. Debor, "Communication Issues in Breast Imaging", Radiologic Clinics of North America, vol. 33, pp. 1231-1245, 1995.
7. S.A. Feig and M.J. Yaffe, "Digital Mammography, Computer-aided diagnosis, and Telemammography", Radiologic Clinics of North America, vol. 33, pp. 1205-1230, 1995.
8. W. Qian and L.P. Clarke, "Computer Assisted Diagnosis for Digital Mammography", IEEE Engineering in Medicine and Biology Magazine, vol. 14, pp. 561-569, 1995.
9. R. Bernstein, "Adaptive Nonlinear Filters for Simultaneous Removal of Different Kinds of Noise in Images", IEEE Transactions on Circuits and Systems, vol. 34, pp. 1275-1291, 1987.
10. 이원진, 박광석, "최적 역치 방법을 이용한 정자의 영역분할 방법에 관한 연구", 대한 PACS 학회지, 1권, 59-64, 1995.
11. K. Fukunaga, Introduction to Statistical Pattern Recognition, Boston, Academic Press, pp. 51-55, 1990.

**=국문초록=**

유방암은 그 조기 발견이 암환자의 사망률을 줄이는 데 있어서 가장 중요한 요소임이 알려져 있다. 스크리닝 검사에 의해 발견되는 유방암의 20% 정도를 차지하는 DCIS(ductal carcinoma in situ)의 경우 미세석회화만이 필름 상에서 볼 수 있는 유일한 소견이다. 따라서 미세석회화를 발견하고 그 형태와 분포의 분석을 통한 진단이 암의 조기 발견에 매우 중요하다.

이 검출과정을 자동화하려는 시도가 디지털 영상처리 기술의 관심이 되어 왔다. 본 연구에서는 상관계수를 특징(feature)으로 사용하여 성능을 향상시킨 통계적 패턴 분류법을 제안하였다. 결과적인 검출율은 통계적 문턱치 설정에 의한 이진화 방법과 비교하여 48%에서 83%로 향상되었다. 성능은 TP와 FP로 평가되었으며 클래스 구분시의 오차도 함께 나타내었다.