초음파 간영상의 특징벡터 분류 및 진단시스템 구현에 관한 연구

정정원°, 김동윤
연세대학교 보건과학대학 의용전자공학과

A Study on the Classification of Ultrasonic Liver Image Feature Vectors and the Design of Diagnosis System

Jeong-Won Jeong°, Dong-Youn Kim
Dept. of Biomed. Eng.,Yonsei Univ.

Abstract

Since one property(i.e. coarseness, orientation, regularity, granularity etc.) of ultrasound liver images was not sufficiently enough to classify the characteristics of livers, we used the multi-feature vectors from ultrasound images to diagnose the liver disease.

The proposed classifier, which uses the multi-feature vectors and Bayes decision rule, performed well for the classification of normal, fat and cirrhosis liver. In our simulation, we used the Battacharyya distance and Hotelling Trace Criterion to select the best multi-feature vectors for the classifier and obtained less classification errors than other methods using single feature vector.

1. 서 론

의료 영상 장비중 초음파 영상기기는 환자에 대한 안전성이 우수하고 다른 영상장비에 비해 가격이 상대적으로 저렴하므로 주로 간과 같은 복부 장기의 병리적 영상정보를 얻는데 널리 사용되고 있다. 특히 간질환중 흔히 발생하는 미만성 질환인 지방침윤간이나 간경화증과 같은 질환은 주로 초음파 단층 영상에 의해 간 실질 및 표면의 변성상태이나 간의 형태변형, 관상 구조물의 상태등을 관찰함으로서 진단을 내리는데 이중 간실질의 변성에 따른 초음파 반향의 세기와 분포 변화는 초음파 영상의 거칠기를 특징지우므로 초음파 실질영상의 거칠기는 간질환의 분류에 중요한 역할을 한다[1][2].

이러한 초음파 영상의 거칠기는 간질환별로 인체내 반사실질의 형태, 구성, 상태에 따라 크게 달라진다. 일반적으로 정상실질의 초음파영상은 전체적으로 균일하고 미세한 편이며, 미만성 실질(diffuse parenchyma)의 초음파 영상은 병변이나 그 진행정도에 따라서 미세한 흰 점들이 빽빽하게 분포하여 국부적으로 밝게 나타나거나, 그물모양의 굵은 점들이 불규칙하게 분포하여 미만성 간이 정상간에 비해 거칠게 보여지므로 정상간과 미만성 간의 분류에 있어서 거칠기는 중요한 파라메타가 된다.[1][2] 그러나 거친 정도가 유사한 지방간과 간경화 영상을 거칠기 파라메타만으로 분류할 때에는 많은 분류 오차가 발생하는데 이는 두 질환간 조직의 변성 역학과 정도가 상이하고 임의적이기 때문에 거칠기보다는 거칠기의 방향성이나 입자화 정도등과 같은 파라메타가 분류에 적합하기 때문이다. 따라서 본 논문에서는 정상간과 지방간, 간경화 초음파 영상의 분류 파라메타인 거칠기와 거칠기 입자의 균일성, 방향성, 입자화 정도등을 나타내는 특징벡터들을 미소 결(micro texture)의 분류에 주로 사용되던 기존 기법들과 본 논문에서 새로이 제안한 기법들을 사용하여 검출한뒤, 이러한 특징벡터의 질환 클래스간 분류 성능을 Hotelling Trace Criterion과 Battacharyya distance를 이용하여 통계적으로 평가를 하였다. 그리고 클래스간의 분리 정도가 우수한 특징벡터를 선정, 조합하여 새로운 거칠기 벡터계를 구성하고 이러한 거칠기 벡터계에서 정상간과 지방간, 간경화를 분류하는데 최소의 분류오차를 갖는 다중차원의 거칠기 벡터를 입출하는 기법을 제안하였다. 마지막으로 최적의 다중 거칠기 벡터 집합과 Bayes 분류기를 이용하여 질환별 실험모집단의 데이터 영상을 분류하였으며, 이때 구한 클라스간의 전이확률 (Transition Probability)로부터 평균 사후확률(*a posteriori probability*)를 구하여 분류에 이용하는 통계적 분류기를 설계하였다.

2. 본 론

2.1 특징벡터 검출

초음파 영상의 주요 특성인 밝기, 거칠기, 균일성, 방향성, 입자화 정도등을 검출하는데 사용한 기법들은 다음과 같다.

가. Amplitude 기법

a. 화소 밝기들의 평균과 분산

영상내 임의의 화소의 그레이 레벨 크기를 F(j, k)라고 할 때 관심 영역 w×w내 화소 밝기의 평균 M(j, k)와 밝기의 분산 V(j, k)는 아래와 같이 구한다[3][4].

$$M(j,k) = \frac{1}{W^2}\sum_{m=-w}^{w}\sum_{n=-w}^{w} F(j+m,k+n)$$

$$V(j,k) = \frac{1}{W^2}\left\{\sum_{m=-w}^{w}\sum_{n=-w}^{w}[F(j+m,k+n)-M(J+m,K+n)]^2\right\}$$

여기서 W=2w+1 로써 관심영역의 가로, 세로 길이를 나타낸다.

M(j, k)와 밝기의 분산 V(j, k)는 아래와 같이 구한다[3][4].

$$M(j,k)=\frac{1}{W^2}\sum_{m=-w}^{w}\sum_{n=-w}^{w}F(j+m,k+n)$$

$$V(j,k)=\frac{1}{W^2}\left\{\sum_{m=-w}^{w}\sum_{n=-w}^{w}[F(j+m,k+n)-M(J+m,K+n)]^2\right\}$$

여기서 W=2w+1 로써 관심영역의 가로, 세로 길이를 나타낸다.

b. 1차 히스토그램의 에너지와 엔트로피[5][6]

관심 지역내 화소 밝기의 1차 확률 분포인 히스토그램을 이용하여 영상의 대조도를 영상의 거친 정도로 정량화 시킬 수 있는데, 그레이 레벨로 양자화된 영상의 히스토그램 P(b)는 아래와 같이 정의된다[4].

$$P(b)\approx\frac{N(b)}{M}$$

여기서, M은 관심지역내 화소들의 총개수이고 N(b)는 관심지역내에서 그레이 레벨 크기 b를 갖는 화소들의 총 개수이다.

이러한 히스토그램 P(b)의 분포형태가 영상의 특징으로 사용되는데 히스토그램의 형태는 아래와 같이 에너지 S_N과 엔트로피 S_E로써 정량화된다[5].

$$S_N=\sum_{b=0}^{L-1}[P(b)]^2$$

$$S_E=-\sum_{b=0}^{L-1}P(b)\log_2[P(b)]$$

나. 통계적 기법

a. coocurrence matrix[7][8][9]

임의의 영상내에서 일정거리 d와 방향 $\theta°$ 를 갖는 화소쌍들의 크기에 대한 2차 결합 조건부 확률 P(i, j ; d, $\theta°$)로써, 같은 거리와 방향을 갖는 화소쌍들의 크기의 분포를 2차원 히스토그램으로 나타낸 것이다. 이러한 P(i, j ; d, $\theta°$)의 각 성분값의 분포 정도를 정량화하는데에는 아래와 같은 수식을 이용한다[9].

$$SE(i,j;d,\theta°)=\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2P(m,n;d,\theta)$$

b. Glay level difference matrix[9]

임의 영상 I(x,y)에서 거리 $\delta(\Delta x,\Delta y)$를 갖는 화소쌍의 그레이 레벨 차는 아래와 같이 정의 할 수 있으며

$$I_\delta(x,y) = |I(x,y) - I(x+\Delta x,y+\Delta y)|$$

$I_\delta(x,y)$의 확률밀도 함수 $P_\delta(m)$는 아래와 같이 구할수 있다.

$$P_\delta(m) = \frac{I_\delta(x,y)가\ m인\ 화소쌍의\ 갯수}{거리\ \delta(\Delta x,\Delta y)를\ 갖는\ 화소쌍의\ 갯수}$$

이러한 $P_\delta(m)$값의 분포정도를 엔트로피로 정량화하여 거친 정도를 분류하는데 이용한다[8][13].

$$ENT = -\sum_i P_{\delta(m)}\log(P_\delta(m))$$

c. autocorrelation function[4]

Pratt등은 임의 영상 2T×2T F(m, n)에 대한 자기 상관 함수 A_F (m, n) 은 아래와 같이 정의하여 영상 인식 분야에 응용하였다[4].

$$A_F(m,n)=\sum_j\sum_k F(j,k)F(j-m,k-n)$$

여기서, $-T\le m,n\le T$ 아래 수식을 이용하여 영상의 균일성을 정량화한다.

$$S(u,v)=\sum_{m=0}^{T}\sum_{n=-T}^{T}(m,\eta_m)^u(n-\eta_n)^vA_F(m,n)$$

여기서, $\eta_m=\sum_{m=0}^{T}\sum_{n=-T}^{T}mA_F(m,n)$

$$\eta_n=\sum_{m=0}^{T}\sum_{n=-T}^{T}n\,A_F(m,n)$$

다. 구조적 기법

a. 윤곽선의 밀도

거칠기 입자의 크기와 분포의 지역적 차이는 영상내에서 지역간의 edge를 구성하며 이러한 edge의 밀도를 구함으로써 영상의 거친 정도를 나타낼 수 있는데 윤곽선 화소는 Robert, Rance 연산자를 이용하여 구한다[4][7].

두 연산자에서 구한 edge map에서 윤곽선 화소의 개수 R을 구하여 특징벡터로 사용한다.

b. 연속된 윤곽선 길이(Run Length)[7]

edge run length matrix P(i,j)를 아래와 같이 구성하여 edge의 크기와 분포정도를 정량화한다[9].

P(i,j) = gray level이 i이면서 임의 방향으로의 run길이가 j인 화소들의 확률밀도

이러한 edge run length matrix P(i,j)의 성분 분포를 나타내기 위해 아래와 같은 파라메타를 사용하여 특징벡터로 삼았다.

SRE(short run emphasis)

$$=\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{p(i,j)}{j^2}\ /\ \sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j)$$

LRE(long run emphasis)

$$=\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}j^2p(i,j)\ /\ \sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j)$$

GLN(gray level nonuniformity)

$$=\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}p(i,j)\right)^2\ /\ \sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j)$$

RLN(run length nonuniformity)

$$=\sum_{i=1}^{N_r}\left(\sum_{j=1}^{N_g}p(i,j)\right)^2\ /\ \sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j)$$

여기서 N_g는 영상의 그레이 레벨 개수이며 N_r은 run length의 최대 길이 이다.

c. Template matching

영상의 거칠기는 유사한 형태의 거칠기 입자들의 구조적 반복에 의해 구성되므로 아래와 같은 template와의 상관정도 r(m,n)를 이용하여 거칠기의 입자화정도를 정량화 할 수 있다.

template 1 : fine texel	template 2 : coarse texel	template 3 : coarse texel
2 1 2 1 2	0 0 0 0 0	0 0 0 0 0
1 2 1 2 1	0 0 1 0 0	0 √2 √3 √2 0
2 1 2 1 2	0 1 2 1 0	0 √2 2 √3 0
1 2 1 2 1	0 0 1 0 0	0 √2 √3 √2 0
2 1 2 1 2	0 0 0 0 0	0 0 0 0 0

그림 1. Template 형태

$$r(m,n) = \sum_{i=1}^{N_r}\sum_{j=1}^{N_g}p(i,j)*f(i+m,j+n)$$

여기서, p(i,j)는 원영상이고 f(i,j)는 template, r(m,n)은 cross correlation 값이다. 본 논문에서는 $r(m,n) > R_{threshold}$ 를 만족하는 r(m,n)의 총 개수 R를 구함으로써 원영상내 texel의 형태와 크기정도

를 정량화하여 특징벡터로 삼았다.

라. 스펙트럼 기법

a. 2 차원 DFT[9][10]

N_1xN_1 영상 x(m,n)에 대한 DFT 계수 $X(k_1,k_2)$의 파워 스펙트럼 $\psi(k_1,k_2)$을 구하고, 효과적인 분류를 위해 2-D 인 계수 $\psi(k_1,k_2)$를 dc파워 $\psi(0,0)$로 규준화한 후 zig zag scanning으로 각 scan 선상의 계수들의 평균값을 구하여 1-D 파워스펙트럼으로 변환시켰다.

$$NFPS = [F(1),F(2),\cdots,F(r),\cdots,F(N_1)]$$

여기서

$F(r)= E[\psi 1(k_1,k_2) \mid \psi 1(k_1,k_2)$ on line r] 이고

$$\psi 1\ (k_1,k_2) = \frac{\psi(k_1,k_2)}{\psi(0,0)},\ 1 < r < N_1$$

NFPS 벡터의 저대역 계수 합 S_{FL} , 고대역 계수 합 S_{FH} 을 아래와 같이 정의하여 특징벡터로 사용한다. 여기서 L은 임계값이다.

$$S_{FL} = \sum_{m=1}^{L} F(m)$$

$$S_{FH} = \sum_{m=L+10}^{N} F(m)$$

b. 2 차원 DCT[3]

N_1xN_1 영상 x(m,n)에 대한 cosine 변환 행렬 C{c(k,n)}을 구한뒤 a.에서와 마찬가지 방법으로 규준화된 이산 cosine 변환 계수 벡터 NDCT를 구한고 NDCT벡터의 저대역 계수 합 S_{CL} , 고대역 계수합 S_{CH} 을 아래와 같이 정의하여 특징벡터로 사용한다. 여기서 L은 임계값이다.

$$S_{CL} = \sum_{m=1}^{L} \tilde{C}(m)$$

$$S_{CH} = \sum_{m=L+10}^{N} \tilde{C}(m)$$

마. 프랙탈 기법

a. 스케일별 밝기차 평균의 변화율

Chen등은 [11]에서

$$E[\Delta I_r] = K\Delta r^H, \qquad 0 < H \le 1$$

를 이용하여 거칠기 계수 H 를 구하였다. 여기서 연산자 E[·]는 평균을 나타내고, ΔI_r은 거리 r인 화소들간의 밝기차를 나타내며 K는 상수이다. 위 수식에서 거칠기 파라메타 H는 아래와 같이 구해진다.

$$H = \frac{\sum_{k=1}^{n} nfb(k)ndr(k)}{\sum_{k=1}^{n} ndr(k)^2}$$

여기서 nfb(k)는 거리 ndr(k)인 화소쌍들의 밝기차 평균을 규준화한 값이고 n은 nfb 곡선을 관찰하여 구한 단조로운 감소를 나타내는 거리 이다.

b. 스케일별 밝기차 분산의 변화율[10]

Mandelbrote[12]가 제안한 FBM 모델은 임의 표면의 거칠기를 random work B(x,y;H)의 end result로 간주하여 아래 식으로 나타낸다.

$$var[B(x_1,y_1;H)-B(x_2,y_2;H)] = V_H(\Delta r)^{2H}$$

여기서 var는 분산을 나타내며 V_H는 상수이다. 두 점간의 거리는 $\Delta r=|(x_1,y_1)-(x_2,y_2)|$이고 H 는 Hurst 계수라고 부르며 거친 정도를 나타내는 $0< H<1$ 인 비정수값이다.

위 수식에서 거칠기 파라메타 H는 아래와 같이 구해진다.

$$H = \frac{\sum_{k=1}^{\Delta r_{max}} nfb(k)ndr(k)}{\sum_{k=1}^{\Delta r_{max}} ndr(k)^2}$$

여기서, nfb(k)는 거리 ndr(k)인 화소쌍들의 밝기차 분산을 규준화한 값이고 Δr_{max}를 nfb 곡선이 일정 범위 기울기내로 단조로운 감소를 나타내는 최초 거리이다.

c. 스케일별 표면면적의 변화율[13]

거리의 변화에 따른 표면면적 A(e)의 변화를 나타내는 log(A(e))-log(e)곡선의 세 점(e-1, e, e+1)간의 기울기 평균 S(e)를 계산하여 구해 영상의 분류에 사용했다.

본 논문에서는 원래의 영상 I(x,y)와 I(x,y)를 1회 subsampling한 영상 $I^1(x,y)$와 2회 subsampling한 영상 $I^2(x,y)$에 대해 a,b,c 방법을 적용하여 거칠기 특징벡터 H와 S(e)들을 각각 구하였다.

2.2 특징벡터 평가

검출된 특징 벡터들은 거칠기나 거칠기의 방향성, 균일성 등과 같은 영상의 물리적 성질을 나타내고 있다. 이러한 특징벡터들은 각각이 정규 가우시안 분포를 나타내고 있으며, 본 논문에서는 특징벡터의 클래스간 분리된 정도를 평가하기 위해 다음과 같은 두 방법을 사용하였다.

가. Battacharyya Distance(B-distance)[4][14]

class 1, 2가 $N_x(M_1,\Sigma_1), N_x(M_2,\Sigma_2)$ 인 정규 가우시안이라면 Cherrnoff error bound는 아래와 같이 된다.

$$\int p_1^s(x)p_2^{1-s}(x)dx = e^{-\mu(s)}$$

여기서,

$$\mu(s) = \frac{s(1-s)}{2}(M_2-M_1)^T[s\Sigma_1+(1-s)\Sigma_2]^{-1}(M_2-M_1) + \frac{1}{2}\ln\frac{|s\Sigma_1+(1-s)\Sigma_2|}{|\Sigma_1|^s|\Sigma_2|^{1-s}}$$

Chernoff error bound를 최소로 하는 s를 구하는 과정은 복잡하고 계산량이 많으므로 일반적으로 정규 가우시안일 경우 ε_u를 근사적으로 계산한다 [4][19].

$$\tilde{\varepsilon}_u = \sqrt{P_1P_2}\int\sqrt{p_1(x)p_2(x)}dx = \sqrt{P_1P_2}\,e^{-\mu\left(\frac{1}{2}\right)}$$

여기서,

$$\mu(1/2) = \frac{1}{8}(M_2-M_1)^T\left[\frac{\Sigma_1+\Sigma_2}{2}\right]^{-1}(M_2-M_1) + \frac{1}{2}\ln\frac{\left|\frac{\Sigma_1+\Sigma_2}{2}\right|}{\sqrt{|\Sigma_1||\Sigma_2|}}$$

이때 $\mu(1/2)$를 Battacharyya distance라고 부르고 정규 가우시안인 두 클래스의 분리정도를 계산하는 유용한 척도가 된다.

나. Hotelling Trace Criterion(HTC)[9]

HTC는 전체 class간의 분리정도를 전반적으로 평가하는데 간편하고 계산량이 적어 유용한 방법으로 알려져 있다[9]. HTC는 아래 수식과 같은 스칼라 양으로 되는데

$$J = \text{trace}\ (Sw^{-1}Sb)$$

여기서

$$S_w = \sum_{i=1}^{k} P(w_i) C_i$$

$$S_b = \sum_{i=1}^{k} P(w_i)(x^{(i)} - x^{(0)})(x^{(i)} - x^{(0)})^T$$

$P(w_i)$는 i번째 클라스의 *a priori probability*
$x^{(i)}$는 i번째 클래스의 기대값 벡터
$x^{(0)}$는 전체 클라스의 기대값 벡터 $\sum_{i=1}^{k} P(w_i) x^{(i)}$
k는 클라스 전체 개수

위에서 볼 수 있듯이 S_b는 각 클래스간 평균의 차이에 의존하고 S_w는 전체 클래스의 분산 정도에 의존하므로, 각 클래스간 분리정도는 HTC J의 값이 클수록 양호하다고 평가할 수 있다[9]. 따라서 본 논문에서는 HTC를 이용하여 각각의 기법을 통해 염출된 특징벡터들의 클래스간 분리정도를 전반적으로 평가하였으며 B-distance로는 두 클래스 쌍간의 분리정도를 세부적으로 평가하였다.

2.3 다중 거칠기 벡터 구성

각 특징벡터들은 초음파 영상의 한가지 특성만을 나타내므로 정상간과 미만성간의 분류에는 적합하나 미만성 간질환의 분류에는 부적합한 것으로 알려져 있다. 따라서 초음파 간 영상에 다양한 기법을 적용하여 특징벡터를 검출한뒤 클래스간 분리정도가 우수한 특징벡터들을 조합하여 새로운 특징벡터계를 구성하는 방법을 제안하고자 한다. 먼저 각각의 특징벡터에 대해 HTC와 B-distance계산을 통해 클래스간 분리정도를 평가한 뒤 분리정도가 양호한 특징벡터를 선별, 조합하여 새로운 특징벡터계 S를 아래와 같이 구성하고 S의 부분집합중 원소의 개수가 2인 X_m를 다중 거칠기 벡터로 사용한다.

$$S = \{ x_1, x_2, \cdots x_n \}$$
$$X_m = [x_j, x_k] \quad , \quad j \neq k$$

여기서 X_m는 다중 거칠기 벡터를 나타내며 x_j, x_k는 특징벡터를 나타낸다.

2.4 분류기

가. Bayes 분류기[9][14]

입력벡터 X_m에 대해 각 클래스의 discriminant function값을 계산하여 가장 큰 값의 클래스로 입력벡터를 decision하는 Bayes분류기를 설계하였다.

$$g_i(X_m) = \log p_i(X_m) + \log P(w_i)$$

여기서,
$g_i(X_m)$는 클래스 i의 discriminant function
$p_i(X_m)$ 는 클래스 i의 conditional density function
. $P(w_i)$ 는 클래스 i의 *a priori probability*
입력 벡터 X_m은 일반적으로 다중 가우시안 분포를 나타내므로, 각 클래스의 *a priori probability*가 같다고 가정하면, 각 클래스의 discriminant function $g_i(X_m)$은 아래와 같이 정리된다.

$$g_i(X_m) = -\frac{1}{2}(X_m - \mu_i)^T C_i^{-1}(X_m - \mu_i) - \frac{1}{2}\log|C_i|$$

여기서 μ_i는 클래스 i의 평균 벡터이고 C_i는 클래스 i의 covariance 벡터이다.

나. 통계적 분류기

본 논문에서는 Bayes 분류오차를 최소화하고 분류에 대한 신뢰도를 향상시키기 위하여 다수의 거칠기 벡터와 동수의 Bayes 분류기를 사용하여 실험 모집단에서 구한 각 클래스간의 전이확률을 이용하는 통계적 분류기를 설계하였다. 그림 2에서 X_1, X_2, X_3는 2.3절에서 언급한 다중 거칠기 벡터 집합 S중에서 B-distance와 HTC가 가장 크고 실험 모집단의 데이터 영상에 대한 분류율이 우수한 2차원의 다중 거칠기 벡터이고 N은 정상간 클래스, F는 지방간 클래스, C는 간경화 클래스를 나타낸다. 그리고 각 Bayes 분류기의 출력 D_i는 N_i, F_i, C_i(i=1,2,3) 가운데 하나로 선택된다. 입력영상이 랜덤하다고 가정하여 사전확률(*a priori probability*) P(N)=P(F)=P(C)=1/3로 정하였다. 그리고 입력영상의 질병(N,F,C)과 각 Bayes 분류기의 분류결과 D_i에 대한 전이확률 $P(N|D_i)$, $P(F|D_i)$, $P(C|D_i)$는 실험 모집단 데이터 영상에 대해 X_1, X_2, X_3를 사용하여 각 Bayes 분류기로 분류한 클래스 분류확률로 구하였다. 즉, 3개의 다중 거칠기 벡터를 사용할 경우 27(=3*9)개의 전이확률을 사용하게 된다. 기본 분류 알고리듬은 한 개의 입력영상으로 부터 X_1, X_2, X_3 다중 거칠기 벡터를 구성하고 X_1, X_2, X_3 각각에 대한 Bayes 분류기의 decision $D_1(N_1,F_1,C_1)$, $D_2(N_2,F_2,C_2)$, $D_3(N_3,F_3,C_3)$를 구한 뒤 D_1, D_2, D_3에 대한 N,F,C 클래스의 평균 사후확률(*a posteriori probability*) S_N, S_F, S_C를 구하여 그 크기가 최대인 사후확률에 대해 최종적인 decison을 내리는 것이다. 여기서 S_N, S_F, S_C는 그림 2의 식에서 볼 수 있듯이 임의의 입력영상이 각 Bayes 분류기에서 D_1, D_2, D_3로 decision되었을때 각 decision이 N, F, C 클래스일 평균 사후확률을 말한다. 예를 들어 임의의 입력 영상에서 얻어진 X_1, X_2, X_3에 대한 Bayes 분류기들의 decision이 D_1=정상(N_1), D_2=정상(N_2), D_3=지방(F_3)이라면 입력 영상이 정상일 평균 사후확률 S_N, 입력영상이 지방간일 평균 사후확률 S_F를 계산하여 이중 가장 큰 값으로 최종 분류하게 된다.

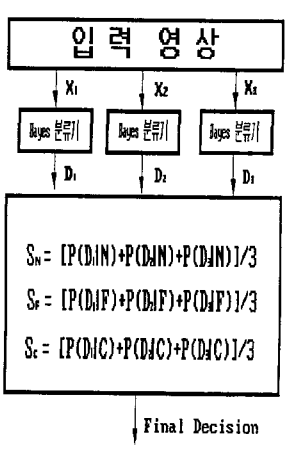

그림 2. 통계적 분류기

초음파 간영상의 거칠기 특성을 나타내는 특징벡터 x_i로부터 다중 거칠기 벡터 X_m을 구하고 통계적 분류기를 이용하여 분류하는 알고리듬을 정리하면 다음과 같다.

1. 각 기법들로부터 특징벡터 x_n를 구한다.

2. 각각의 x_n에 대해 클래스간 분리정도를 나타내는 HTC와 B-distance를 구한다.
3. x_n의 HTC와 B-distance를 비교하여 큰 값을 갖는 x_n를 선택한다.
4. 선택된 x_n를 조합하여 다중 거칠기 벡터 $\mathbf{X}_m$ 집합을 구한다.
5. 다중 거칠기 벡터 $\mathbf{X}_m$의 HTC와 B-distance를 2와 마찬가지 방법으로 구한다.
6. 큰 값의 HTC와 B-distance를 갖는 다중 거칠기 벡터 $\mathbf{X}_m$를 선택한다.
7. 실험 모집단의 각 데이터 영상으로 부터 다중 거칠기 벡터 $\mathbf{X}_m$의 성분값을 구하고 이를 Bayes분류기로데이터 영상을 분류하여 각 다중 거칠기 벡터 $\mathbf{X}_m$에 대한 클래스별 전이확률을 구한다.
8. 각 D_i에 대한 입력영상이 N, F, C의 평균 사후 확률 S_N, S_F, S_C를 구한다.
9. S_N, S_F, S_C의 크기를 비교하여 최대값인 클래스로 최종 분류한다.

3. 실험 및 고찰

3.1 실험 데이터 획득 및 실험 집단 구성

실험에 사용한 초음파 탐침자 주사방식은 환자에 따라 sagital , subcostal, transverse 방식을 사용하였으며 성별이나 연령에 관계없이 질환에 따라 전문의가 육안으로 진단을 내린 질환별 영상을 동일한 기기에서 동일한 이득(2 dB)으로 비디오 테잎에 녹화한 뒤 PC의 프레임 그레버를 이용하여 디지탈화된 그레이 영상을 640x480 , 128 그레이 레벨로 획득했다. 획득된 총영상의 수는 정상간 50장, 지방간 50장, 간경화 50장씩 총 150장이다. 이렇게 획득된 질환별 영상으로부터 진단에 근거가 될 수 있는 깊이 5-10 ㎝의 간 실질부위를 32x32 크기, 128 그레이레벨로 샘플링하여 실험 데이터로 사용하였는데 전체 영상 1개당 중복되지 않게 3개의 실험 데이타를 추출하였다. 질환별 150개의 표본 영상중에서 100개는 특징벡터의 분류성능을 측정하고 분류에 최적인 다중 거칠기 벡터를 검출하는 실험 모집단으로 구성하였으며 나머지 50개는 실험 집단으로 구성하여 모집단에서 구한 최적의 다중 거칠기 벡터와 통계적 분류기의 분류 성능을 평가하는데 사용하였다. 모든 프로그램은 Boland C 3.0으로 제작하여 실험하였다.

3.2 클래스별 분류성능이 우수한 특징벡터

표 1은 각 기법들의 특징벡터들 중에서 각 클래스 쌍간의 B-distance가 0.3이상이고 해당 클래스로 정확하게 분류하는 비율이 0.8 이상인 특징벡터들이다.

3.3 2차원 거칠기 벡터 구성 및 분류 결과

표 1.의 특징벡터들을 조합하여 거칠기 벡터계 S를 구성한 뒤 S의 부분집합중에서 원소의 개수가 2인 부분집합을 다중 벡터 $\mathbf{X}_i$를 구성하였다.

$$\mathbf{X}_i = [\, x_k, x_j \,] \qquad k \neq j$$

여기서 x_k, x_j는 표 1.에 나와 있는 특징벡터들이다.

이렇게 구성한 다중 결 벡터 $\mathbf{X}_i$ 중에서 비교적 큰 HTC와 B-distance를 갖는 $\mathbf{X}_i$를 표 2에 정리하였다.

실험결과, 하나의 특징벡터로부터 구한 B-distance에 비하여 2차원의 다중 거칠기 벡터로 구한 B-distance가 전반적으로 크게 나타났다. 특히 미만성 질환인 지방간과 간경화 클래스사이의 B-distance가 크게 증가하여 두 질환을 효과적으로 분류할 수 있게 되었다. 대체적으로 전반적인 클래스의 분리정도를 나타내는 HTC의 값이 클수록 각 B-distance의 값이 증가하여 HTC도 클래스의 분리정도를 측정하는데 유용한 측정치임을 알 수 있었으며 이는 Wu등[9]의 결과와도 일치하는 것이다.

표 1. 클래스별 분류성능이 우수한 특징벡터

정상간	지방간	간경화
V(j.k)	M(j,k)	S_N
S_E	Tempate 3	S_E
$S_E(6,45^\circ)$	SRE	$S_E(6,45^\circ)$
ENT	LRE	ENT
Template 1		GLN
Template 2		$S_1(e)$
S_{FH}		$H_{\Omega\cdot 40}$
Robert		$H_{n\cdot 16}$
Operator		
Range		
Operator		
$H_{\Omega\cdot 40}$		
$H_{\Omega\cdot 10}$		
$H_{n\cdot 16}$		

표 2. 2차원 거칠기 벡터의 B-distance

다중 벡터	x_i	x_j	HTC	Battacharrya Distance 정상간 ↕ 지방간	지방간 ↕ 간경화	간경화 ↕ 정상간
X_1	M(j,k)	ENT	1.671098	0.733485	1.198985	1.394064
X_2	V(j,k)	Template 3	1.799048	0.932120	1.390045	1.288752
X_3	V(j,k)	LRE	1.545782	0.678407	1.277697	1.299659
X_4	S_E	LRE	1.589507	0.725092	1.220534	1.284374
X_5	Robert	$SE(6,45^\circ)$	1.533724	1.113976	1.221407	1.190657
X_6	Tempalate 3	ENT	1.966480	1.003647	1.372197	1.394218
X_7	LRE	ENT	1.673554	0.751952	1.249230	1.400254
X_8	LRE	S_E	1.589507	0.725092	1.220534	1.284374
X_9	$S_1(e)$	ENT	1.461488	0.642798	1.020456	1.541142

표 3. 2차원 거칠기 벡터로 분류한 결과

다중 벡터	실험 모 집단 정상간 N	F	C	지방간 N	F	C	간경화 N	F	C
X_1	88	10	2	8	91	1	3	4	93
X_2	87	11	2	8	91	1	3	8	89
X_3	90	8	2	9	90	1	3	4	93
X_4	89	9	2	9	90	1	3	4	93
X_5	90	7	3	4	94	2	7	7	86
X_6	88	11	1	7	92	1	3	9	88
X_7	95	4	1	8	90	2	3	4	93
X_8	89	9	2	9	90	1	3	4	93
X_9	87	12	1	8	87	5	3	4	93

표 3.은 실험 모집단의 데이터 영상에서 각 X_i를 추출한 뒤 이를 Bayes분류기로 분류한 결과이다. 특징벡터만으로 분류한 분류율에 비하여 각 클래스

간의 분류율이 현저히 증가하였다. 특히 거친 정도가 유사한 지방간과 간경화 클래스의 경우 90% 정도의 높은 분류율을 얻을 수 있었는데 이는 거친 정도를 나타내는 특징벡터와 거칠기의 방향성이나 입자화정도를 나타내는 특징벡터를 혼합하여 함께 분류함으로써 거칠기의 크기만으로는 분류가 안되었던 미만성 질환이 효과적으로 분류될 수 있었던 것 같다. 또한 정상간 클래스의 데이터 영상을 간경화로 잘못 분류한 개수보다 지방간으로 잘못 분류한 개수가 훨씬 많고 지방간 클래스의 데이터 영상을 간경화로 잘못 분류한 개수보다 정상간으로 잘못 분류한 개수가 많은데 이는 표 2.에서 볼 수 있듯이 정상간↔지방간의 B-distance 크기가 비교적 적은데서 비롯된 것으로 지방간이나 정상간 데이터 영상중 일부는 초기 병변을 갖은 영상들이여서 거칠기의 차이가 미묘하기 때문이다. 그리고 각 X_i별로 잘못 분류한 영상 데이터를 검사해본 결과 대체적으로 일치하였다. 표 4, 5는 통계적 분류기로 분류한 결과이다. 클래스별간의 전이확률은 표 3.에 나와 있는 분류결과를 100으로 나누어 사용하였다.

표 4. 통계적 분류기에 의한 실험 모집단 분류

다중 벡터	실험 모 집단								
	정상간			지방간			간경화		
	N	F	C	N	F	C	N	F	C
X_3	90	8	2	9	90	1	3	4	93
X_7	95	4	1	8	90	2	3	4	93
X_8	89	9	2	9	90	1	3	4	93
X	90	9	1	9	90	1	3	4	93

표 5. 통계적 분류기에 의한 실험 집단 분류

다중 벡터	실험 집단								
	정상간			지방간			간경화		
	N	F	C	N	F	C	N	F	C
X_3	45	2	3	5	39	6	3	3	44
X_7	45	2	3	5	42	3	3	3	44
X_8	45	2	3	5	42	3	3	4	43
X	45	2	5	5	42	3	3	3	44

표 4, 5에서 볼 수 있듯이 통계적 분류기를 사용한 결과 X_3, 와 X_8을 사용하여 얻은 분류결과보다는 양호한 결과를 얻을 수 있었으나 X_7만을 사용했을 때보다는 미흡한 결과를 얻었다. 그러나 총 4개의 특징벡터를 사용하고 확률공간으로부터 얻은 클래스간 전이확률의 합을 비교하여 분류하였으므로 모든 클래스에 대해 90% 정도의 양호한 결과를 얻을 수 있었고 초음파 간영상의 질환별 임의성등을 고려해 본다면 분류에 적합한 분류기라고 말할 수 있다.

4. 결 론

초음파 간영상의 진단은 간의 형태변형이나 간실질 및 표면의 변성상태, 관상 구조물의 상태등을 관찰함으로서 이루어지는데 이중 간실질의 변성에 따른 초음파 반향의 세기와 분포 변화에 의해 특징지워지는 초음파 영상의 거칠기는 정상간과 미만성 질환의 분류에 주로 사용되어 왔다.

이러한 초음파 간영상의 거칠기는 외부변수나 관찰자의 주관에 따라 달라지는데 거칠기의 특성(거친 정도, 방향성, 균일성, 입자화 정도)을 객관화하고 정량화하는 여러 가지 특징벡터들이 발표되어 왔다. 그러나 이러한 특징벡터들은 거칠기의 특성중 한 특성만을 나타내므로 지방간과 간경화와 같은 미만성 질환간의 분류에는 적합하지 않다. 따라서 본 논문에서는 여러 가지 기법을 사용하여 검출한 특징벡터들을 HTC와 B-distance로 분류성능을 평가한 뒤 우수한 특징벡터들을 선별하여 거칠기의 특성을 복합적으로 반영하는 다중 거칠기 벡터를 구성하였다. 또한 분류의 신뢰도를 향상시키기위해 이러한 다중 거칠기 벡터와 확률공간에서 Bayes 분류기로 얻은 정상간, 지방간, 간경화 클래스간 전이확률을 이용하여 각 클래스의 평균 사후확률을 구하는 통계적 분류기로 간영상을 분류하였다.

각 클래스의 실험 모집단과 실험 집단의 데이터 영상에 대한 실험 결과 하나의 특징벡터만으로 분류한 결과보다 다중 거칠기 벡터로 분류한 결과가 훨씬 우수하였으며 85%-90%의 높은 분류율을 얻을 수 있었다.

참 고 문 헌

[1] Francis S. Weill, " Ultrasound Diagnosis of Digestive Diseases ", 3th, Springer-Verlag, 1989

[2] Robert A. Wilkins & Heather B. Nunnerley, " Imaging of the Liver,Pancreas and Spleen ". Blackwell Scientific Publications, Oxford London, 1990

[3] A. K. Jain, " Fundamentals of Digital Image Processing, Englewood Cliffs, NY; Prentice-Hall. Inc, 1989.

[4] William K. Pratt, " Digital Image Processing ", 2nd, Willey Interscience, 1991

[5] Rafael C. Gonzalez, Richard E. Woods, " Digital Image Processing",Addison Wesley, 1992

[6] Robert M., Haralick, " Textural Features for Image Classification ", IEEE Trans. SMC , VOL smc-3 , NO. 6, NOVEMBER 1973

[7] Robert M., Haralick, " Statistical and Structural Approaches to Texture ", Processings of the IEEE, VOL 67, NO. 5, MAY 1979

[8] 홍정우, 김선일, 이두구, "통계적 특성과 신경망을 이용한 초음파 화상진단", 대한생체공학회 1992년 추계학술대회 논문집 제 14 권 제 2 호, pp 26-29, 1992

[9] Chung-Ming Wu, Yung-Chang Chen, " Texture Feature for Ultrasonic Liver Images ", IEEE Trans. Med. Imaging, vol. 11, NO 2, pp. 141-152, 1992.

[10] 정정원, 김동윤, "프랙탈차원과 퓨리에 파워 스펙트럼을 이용한 간조직 분류", 대한생체공학회 1995년 춘계학술대회 논문집 제 17권 제 1 호, pp 37-41, 1995

[11] C.C. Chen, J.S. Daponte, and M.D. Fox, " Fractal feature analysis and classification in medicalimaging", IEEE Trans. Med. Imaging, vol. 8, pp. 133-142, 1989.

[12] Torbjörn Lundahl, William J, Ohley, Steven M. Kay, Robert Siffert, " Fractional Brownian Motion; A Maximum Likelihood Estimator and Its Application to Image Texture ", IEEE Trans. Med. Imaging, vol. MI-5, NO 3, pp. 152-161,1986.

[13] Shamuel Peleg, " Multi Resolution Texture Analysis and Classification ", IEEE Trans. PAMI, Vol. PAMI-6, No 4, JULY 1984.

[14] Keinosuke Fukunaga, " Introduction to Statistical Patturn Recognition ", 2nd, Academic Press, 1990