

# 폐질환 조기 검출을 위한 결합 히스토그램 기반의 통계적 특징 인자에 대한 연구

## Study of Joint Histogram Based Statistical Features for Early Detection of Lung Disease

원철호\*

C. H. Won

### 요 약

본 논문에서는 폐질환 조기 검출을 위하여 Broncho vascular, Emphysema, Ground Glass Reticular, Ground Glass, Honeycomb, Normal의 6가지 폐조직에 대한 새로운 분류기법을 제안하였다. 단순 베이스 분류기와 아다부스트 학습 기법을 도입하여 459개의 결합 히스토그램 특징인자로부터 유효한 특징인자를 선별함으로써 폐조직을 분류하였다. 다중 해상도 해석, 체적 LBP 및 CT 휘도를 기반으로 하는 결합 히스토그램 특징인자는 정확도, 민감도, 특이도 결과에서 기존의 3D AMFM보다 우수한 결과를 보였다. 제안한 특징인자와 3D AMFM 특징인자의 정확도는 각각 90.1%과 85.3%로서 제안한 특징인자의 우수한 분류 성능을 확인하였다.

### ABSTRACT

In this paper, new method was proposed to classify lung tissues such as Broncho vascular, Emphysema, Ground Glass Reticular, Ground Glass, Honeycomb, Normal for early lung disease detection. 459 Statistical features was extraced from joint histogram matrix based on multi resolution analysis, volumetric LBP, and CT intensity, then dominant features was selected by using adaboost learning. Accuracy of proposed features and 3D AMFM was 90.1% and 85.3%, respectively. Proposed joint histogram based features shows better classification result than 3D AMFM in terms of accuracy, sensitivity, and specificity.

**Keyword** : Lung Tissue, Adaboost, Computer Aided Diagnosis, Volumetric LBP

### 1. 서론

폐암은 암중에서 가장 큰 사망의 원인이 되고 있으며 2008년에 전세계적으로 137만명을 사망에 이르게 하였다. 폐암의 5년 생존율은 15.9%인데 이러한 통계는 조기진단으로 폐암의 초기 단계에서는 52%까지 올라간다. 조기 진단으로 생존율을 높이는 것으로 인해, 흡연자와 과거 흡연자 같은 고위험군의 증상이 없는 개인에 대한 검사를 하게 되었다. 컴퓨터 단층 촬영(CT)에 의한 검사는 효과적인 기술로 밝혀졌다. 국립 폐 심사 시험(National Lung

Screening Trial)은 CT를 사용하여 검사하는 것은 투사 X선을 사용하여 것에 비해 20% 폐암 감소 폐암 사망률을 보이는 것을 발견했다.

최근에는 의료 영상에서 물체의 질감 특징을 추출해 분석하는 일이 큰 주목을 받고 있고 다양한 질환 분야에 대하여 연구가 진행되고 있다. 이러한 질감 특징 추출의 기본적인 개념은 관심 영역에서의 원하는 물체를 탐지하는 것에 기본적인 바탕을 두고 있다. 의료 영상에서 관심 영역중 하나인 폐 영역에서 질환을 하나의 물체로 보고 질환의 질감 특징을 추출해 환자의 상태를 진단하는 연구가 진행되고 있다. 컴퓨터 조력 진단(Computer Aided Diagnosis)시스템[1-3]은 입력된 환자의 CT 영상을 이용해 폐 영역에서 질환을 찾고 진단을 하게 된다.

이때 질환의 질감 특징을 찾는 것은 매우 중요한

접 수 일 : 2016.11.07

심사완료일 : 2016.11.17

게재확정일 : 2016.11.25

\* 원철호 : 경일대학교 의용공학과 교수

chulho@kiu.ac.kr (주저자)

과정이다. 폐 영역에서 발생할 수 있는 질환은 매우 다양한데 이들 각각의 질환은 서로 다른 특징들을 가지고 있기 때문이다. 질환의 특징을 추출하는 방법은 다양하다.

본 논문에서는 아다부스트 학습 방법 사용하여 반경 변화에 의한 다해상도 LBP 분석과 LBP-TOP 연산에 의한 체적 LBP의 유효성을 평가하였다. 다중 클래스 분류는 Broncho vascular, Emphysema, Ground Glass Reticular, Ground Glass, Honeycomb, Normal의 6가지 폐조직에 대하여 수행하였다. 3D AMFM과 제안한 특징인자, 단순 베이즈 분류기(Naive Bayes classifier)와 아다부스트 학습 분류기를 이용하여 6 클래스 분류 결과를 비교하였다.

먼저 단순 베이즈 분류기를 이용하여 3D AMFM과 제안한 특징인자의 정확도를 비교하였으며 아다부스트 학습 방법 사용하여 분류 성능을 평가하였다. 이 실험에서 3D AMFM 특징인자 보다 제안한 특징인자의 우수성을 확인하였다. 6가지 폐 조직에 대한 전문가와 제안된 방법의 분류 성능을 평가하기 위해 오분류 행렬을 사용하였다. 또한, 3D AMFM 특징인자와 제안한 특징인자 (R=1), (R=1,2,3)에 대한 민감도 및 특이도를 계산하여 제안한 방법의 우수성을 평가하였다.

## 2. 방법

### 2.1 3D AMFM 특징인자

적응적 다중 피쳐 기법(Adaptive multiple feature method: AMFM)[4]는 질병의 병리를 고려할 수 있고, 동시적 다중 질병의 경과를 분류할 수 있는 텍스처 기반 방법으로 소개되었다. 3D AMFM 특징인자는 1차(first order) 함수와 2차(second order) 함수 특징인자, 히스토그램 특징인자 및 프랙탈 특징인자와 같은 통계적인 특징인자를 포함하고 있다. 1차 함수 특징인자는 VOI(volume of Interest)에서 각 복셀의 휘도 준위(gray level) 분포를 측정한다. 이러한 특징인자는 평균(mean), 분산(variance), 첨도(kurtosis), 및 엔트로피를 포함하고 있다. 2차 함수 특징인자는 런 력스(run-length) 동시발생(co-occurrence) 행렬 특징인자를 포함한다. 이는 short-run emphasis, long-run emphasis, gray-level non-uniformity, run length non-uniformity, run Percentage를 포함하고 있다. 동시 발생행렬(co-occurrence matrix)에서는 엔트로피, 각

2차 모멘트(angular second moment), 명암대조(contrast), 관성(inertia), 상관관계(correlation) 및 역차 모멘트(inverse difference moment)와 같은 second-order 특징인자가 계산된다. 네 가지 특징인자들은 VOI에 있는 복셀의 휘도 준위 분포의 히스토그램에 근거하여 계산된다.

또한, emphysema 영역 개수, normal 영역 개수, 모드(mode) 및 emphysema 비율이다. 평균, 분산, 비대칭(skewness), 왜도 그리고 휘도 준위 분포의 엔트로피와 같은 프랙탈 특징인자는 확률 프랙탈 영상(stochastic fractal images)에서 계산된다.

### 2.2 체적(volumetric LBP)에 근거한 통계적 특징인자의 제안

LBP는 원래 3×3 이웃 화소에서 국부 구조를 특징화하기 위한 휘도 스케일 불변 측정 인자로서 소개되었으며[5-6] 이후에 다중 해상도 분석(multi-resolution analysis)을 가능하게 하는 더 일반적인 공식이 발표되었다. LBP는 중앙 픽셀의 밝기와 관련된 국부 이웃화소의 임계치 설정에 의해 획득되며 다음 식에서 구해진다.

LBP 연산은 식 (1)과 같으며 현재 위치의 화소와 이웃 화소의 차이를 0과 1로 나타낸다. 여기서  $P, R, g_c, g_p$ 는 각각 인접 화소수와 원의 반경, 현재 화소 및 이웃화소를 의미한다.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{p-1} s(g_p - g_c)2^p, s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1)$$

다중 해상도 국부 질감(local texture) 분석은 반경 R의 원주 상에 존재하는 P 화소의 이웃 화소를 사용한다. 다중 해상도 분석은 공간상에 존재하는 다양한 크기의 미세 패턴의 형상을 특징짓는 데 유용하다.

LBP-TOP(Three Orthogonal Planes)은 일종의 체적 LBP 연산으로서 중앙 화소에서 교차하는 세 직교 평면을 사용한다. LBP-TOP는 세 개의 직각평면인 XY, XZ 및 YZ 평면을 연결하는 특징기술자이다. LBP-TOP 각각의 분리된 평면으로부터 특징인자 분포를 계산하고 이를 연결하여 하나의 특징인자 벡터로 구성한다. 의료용 3D 체적 영상은 Z축 상에 XY 평면의 스택으로 볼 수 있다. 세 가지 다른 평면의 LBP 히스토그램을 계산하고 이를 하나의 히스토그램으로 연결하여 LBP-TOP 특징인자 벡터를 획득할 수 있다.

### 2.3 제안한 특징인자

LBP 연산은 그 자체로 영상의 모노토닉한 휘도 스케일 변환에 강건한 특성이 있다. 하지만, 이 특성은 CT 영상을 처리할 때는 항상 바람직하지는 않다. CT영상이 생체조직의 물리적 성질을 측정하여 수치화하여 표현되기 때문이다.[12] 본 논문에서는 폐 조직의 밀도 정보를 포함하기 위해서 LBP와 휘도의 결합 히스토그램(joint histogram)을 사용한다.

폐조직의 분류를 위하여 본 논문에서는 다중 해상도 LBP 분석, 체적 LBP 그리고 LBP와 휘도 준위의 결합 히스토그램이라는 세 가지 주된 개념을 도입하였다. 다중 해상도 분석은 폐조직의 미세 패턴을 기술하기 사용되었는데 LBP 연산에서 반경 R을 변화시킴으로서 구할 수 있다. 3D 체적 LBP는 LBP-TOP의 형태로서 폐조직의 3차원 구조를 특징짓기 위하여 적용하였다. 그리고 결합 히스토그램은 조직의 휘도와 LBP 코드로서 폐조직을 기술하기 위하여 사용하였다. 구해진 결합 히스토그램 행렬에 표 1의 통계공식을 적용함으로써 폐조직을 구별하기 위한 최종 특징인자를 추출하였다.

표 1. 51개의 통계적 특징인자  
Table 1. 51 Statistical features.

Features
$f_1 = \sum_i \sum_j p(i,j)^2$
$f_2 = \sum_i \sum_j (i-j)^2 p(i,j)$
$f_3 = \sum_i \sum_j (i-j) p(i,j)$
$f_4 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
$f_5 = \sqrt{0.5(i-j)p(i,j)}$
$f_6 = \sum_i \sum_j p(i,j) \cdot \log(p(i,j))$
$f_7 = \max(p(i,j))$
$f_8 = \sum_i \sum_j \frac{p(i,j)}{j^2}$
$f_9 = \sum_i \sum_j p(i,j) \cdot j^2$
$f_{10} = \sum_i \sum_j \frac{p(i,j)}{i^2}$
$f_{11} = \sum_i \sum_j p(i,j) \cdot i^2$
$f_{12} = \sum_i \sum_j \frac{p(i,j)}{i^2 j^2}$
$f_{13} = \sum_i \sum_j \frac{p(i,j) \cdot i^2}{j^2}$

$$f_{14} = \sum_i \sum_j \frac{p(i,j) \cdot j^2}{i^2}$$

$$f_{15} = \sum_i \sum_j p(i,j) \cdot i^2 \cdot j^2$$

$$f_{16} = \sum_i (\sum_j p(i,j))^2$$

$$f_{17} = \sum_j (\sum_i p(i,j))^2$$

$$f_{18} = \sum_i \sum_j \frac{1}{p(i,j) \cdot j}$$

$$f_{19} = \sum_i \sum_j \frac{1}{p(i,j) \cdot i}$$

$$f_{20} = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i,j)$$

$$f_{21} = \sum_i \sum_j (j - \mu_y)^2 \cdot p(i,j)$$

$$f_{22} = \sum_i \sum_j (i - \mu_x) \cdot (j - \mu_y) \cdot p(i,j)$$

$$\mu_x = \sum_i \sum_j i \cdot p(i,j)$$

$$\mu_y = \sum_i \sum_j j \cdot p(i,j)$$

$$\sigma_x = \sqrt{\sum_i \sum_j (i - \mu_x)^2 \cdot p(i,j)}$$

$$\sigma_y = \sqrt{\sum_i \sum_j (j - \mu_y)^2 \cdot p(i,j)}$$

$$f_{23} = \sum_i \sum_j p(i,j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y}$$

$$f_{24} = -\sum_i \sum_j p(i,j) \log(p_x(i) p_y(j))$$

$$f_{25} = -\sum_i \sum_j p_x(i) p_y(j) \log(p_x(i) p_y(j))$$

$$f_{26} = \sum_k k p_{x+y}(k)$$

$$p_x(i) = \sum_j p(i,j)$$

$$p_y(j) = \sum_i p(i,j)$$

$$p_{x+y}(k) = \sum_i \sum_j p(i,j),$$

$$k = i + j = \{0, 1, 2, \dots, (N_p + N_y - 1)\}$$

$$p_{x-y}(k) = \sum_i \sum_j p(i,j),$$

$$k = i - j = \{0, 1, 2, \dots, (|N_p - N_y| - 1)\}$$

$$f_{27} = \sum_k p_{x+y}(k) \log(p_{x+y}(k))$$

$$f_{28} = \sum_k (k - f_{26})^2 p_{x+y}(k)$$

$$f_{29} = \sum_k k p_{x-y}(k)$$

$$f_{30} = \sum_k p_{x-y}(k) \log(p_{x-y}(k))$$

$$f_{31} = \sum_k (k - f_{29})^2 p_{x-y}(k)$$

$$HX = -\sum_i p_x(i) \log(p_x(i))$$

$$HY = -\sum_j p_y(j) \log(p_y(j))$$

$$HXY = -\sum_i \sum_j p(i,j) \cdot \log(p(i,j))$$

$$HXY1 = -\sum_i \sum_j p(i,j) \cdot \log(p_x(i) p_y(j))$$

$$\begin{aligned}
 HXY2 &= - \sum_i \sum_j p_x(i)p_y(j) \cdot \log(p_x(i)p_y(j)) \\
 f_{32} &= \frac{HXY - HXY1}{\max(HX, HY)} \\
 f_{33} &= (1 - \exp[-2(HXY2 - HXY)])^{1/2} \\
 f_{34} = \mu_x &= \frac{1}{N_g} \sum_i i p_x(i) \\
 p_x(i) &= \sum_j p(i, j) \\
 f_{35} &= \frac{1}{N_g^2} \sum_i (p_x(i) - \mu_x)^2 \\
 f_{36} &= \frac{1}{N_g^3} \sum_i (p_x(i) - \mu_x)^3 \\
 f_{37} &= \sum_i p_x(i) \log(p_x(i)) \\
 f_{38} = \mu_y &= \frac{1}{N_p} \sum_j j p_y(j) \\
 p_y(j) &= \sum_i p(i, j) \\
 f_{39} &= \frac{1}{N_p^2} \sum_j (p_y(j) - \mu_y)^2 \\
 f_{40} &= \frac{1}{N_p^3} \sum_j (p_y(j) - \mu_y)^3 \\
 f_{41} &= \sum_j p_y(j) \log(p_y(j)) \\
 f_{42} \sim f_{48} &= \sum_{i=g_n}^{g_n + N_g/10} p_x(i), \\
 g_0 &= 0, g_{n+1} = g_n + N_g/10, n = 0 \sim 6 \\
 f_{49} = \mu &= \frac{1}{N_p N_g} \sum_i \sum_j p(i, j) \\
 f_{50} &= \frac{1}{(N_p N_g)^2} \sum_i \sum_j (p(i, j) - \mu)^2 \\
 f_{51} &= \frac{1}{(N_p N_g)^3} \sum_i \sum_j (p(i, j) - \mu)^3
 \end{aligned}$$

LBP 기반 특징인자를 구현함에 있어 59개의 ULBP 코드 개수에 정합하기 위하여 CT 값을 64 휘도 준위로 양자화를 수행하였으며, 64(CT 휘도)×59(ULBP 코드) 차원의 결합 히스토그램을 얻었다.  $LBP_{8,1}$ ,  $LBP_{8,2}$  및  $LBP_{8,3}$  연산이 다해상도 분석을 위하여 사용되었고 체적 분석을 위하여 XY, YZ 및 XZ 평면상의 LBP-TOP는 각각의 VOI에서 계산되면 총 9(3(다해상도 LBP)×3(체적 LBP))개의 결합 히스토그램이 얻어진다.

결합 히스토그램 행렬에 51개의 통계적인 특징인자를 적용하였다. 통계적 특징인자는 1차 및 2차 합수 특징인자, 히스토그램 특징인자, 런랜스, 동시발생 행렬 등의 특징인자를 포함한다. 각 VOI로부터 6가지 서로 다른 폐조직을 정량화하기 위해서 총 459(51(특징인자) × 9(결합 히스토그램))개의 폐조직 특징인자를 생성하였다.

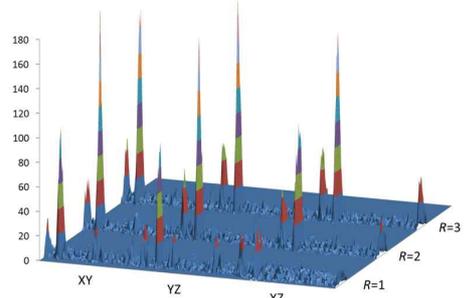


그림 1. 다중 해상도 분석과 체적 LBP분석을 위한 결합 히스토그램

Fig. 1. Joint histogram for multi-resolution analysis and volumetric LBP

### 2.4 아다부스트 학습에 의한 특징인자 선정 및 분류

각 VOI로부터 얻어진 459개의 특징인자에 대하여 의미있는 특징인자를 폐조직 분석에 포함하기 위해서는 중요 특징인자의 선별이 필요하다. 주요한 특징인자가 선별하기 위하여 아다부스트 학습 방법을 사용하였다. 특징인자 선별방법의 검증과 아다부스트 학습의 기본 분류기로서 나이브 베이즈 분류기를 사용하였다. 아다부스트 학습 방법은 이진 분류 문제를 해결하기 위해 제안되었으나 다중 분류 문제로 확장될 수 있다. 다중 클래스 아다부스트 방법은 Shpire[7]에 의해 제안되었으며 다중 분류 문제를 k개의 직교하는 이진 분류 문제로 분해하는 방법으로서 본 논문에서 이 방법을 6개 다중 폐조직 분류에 적용시켰다.

---

```

Given:  $(x_1, Y_1), \dots, (x_m, Y_m) : x_i \in \mathcal{X}, Y_i \subseteq \mathcal{Y}$ 
Initialize  $D_1(i, l) = 1/(mk)$ 
For  $t = 1, \dots, T$ 
  Train weak learner using distribution  $D_t$ 
  Get weak hypothesis  $h_t: \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ 
  Choose  $\alpha_t \in \mathbb{R}$ 
  Update  $D_{t+1}(i, l) = \frac{D_t(i, l) \exp(-\alpha_t Y_i(l) h_t(x_i, l))}{Z_t}$ 
Final hypothesis
 $H(x, l) = \text{sign} \left( \sum_{t=1}^T \alpha_t h_t(x, l) \right)$ 

```

---

그림 2. 다중 클래스 아다부스트 학습  
Fig. 2. Multi-class adaboost learning.

### 2.5 실험방법

필립스 MX8000 MDCT(multi detector CT) 스캐너로부터 120 kV, 40/100 mAs, 피치 1.5, 1.3mm 콜리메이션(collimation), 0.6mm 슬라이스 증가, FOV

290 - 500mm, B 재구성 필터 등의 파라미터를 설정하여 영상을 획득하였다. 34명의 지원자는 9명의 흡연자는 심각한 COPD, 10명의 흡연자는 가벼운 COPD, 8명의 비흡연자는 정상 폐기능, 7명의 흡연자는 정상 폐기능을 가진 인원으로 구성되었다. 폐기능 검사로부터 각 피실험자들의 vital capacity (VC), forced vital capacity (FVC), forced expiratory volume을 얻었다.

제안된 방법을 성능을 평가하기 위해서 민감도 (sensitivity), 특이도 (specificity) 및 정확도 (accuracy)의 지표가 사용되었다. 민감도는 바르게 분류된 표본의 수를 전체 양 (positive)의 표본의 수로 나눈 것이다. 특이도는 바르게 분류된 표본의 수를 전체 음 (negative)의 표본의 수로 나눈 것이다. 정확도는 바르게 분류된 표본의 수를 전체 표본의 수로 나눈 것이다. 성능 평가를 위하여 기본적으로 단일관측치제거 (leave-one-out) 교차검증법을 사용하였다.

### 3. 결과

제안한 결합 히스토그램 특징인자의 정확성을 세 가지 비유사도 측정도구 (dissimilarity measures)를 이용하여 분류 성능을 비교하였다. 59 ULBP 코드와 64 휘도 준위의 2D 결합 히스토그램을 이용하고, 3가지 반경에 대하여 XY, YZ 및 XZ 평면에서의 LBP를 계산하였다. 총 9개의 결합 히스토그램 특징인자는 각각의 293개의 VOI (31×31×31, 31×31×30 화소)에서 수집하였고, 단일관측치제거 교차 검증 (cross-validation) 방법이 제안한 특징인자의 성능을 평가하기 위해 사용되었다. 2D 결합 히스토그램 특징인자에서 79.9%의 높은 정확도가 로그 유사도 측정도구 (log likelihood measure)에 의해 얻어졌다. 카이 제곱 (Chi-square)을 사용한 분류 성능은 77.8%의 정확성을 나타내었고, 히스토그램 교차에 의해서는 75.1%가 얻어졌다. 이 세 가지 측정도구를 이용한 결과에서는 만족할 만한 결과를 얻을 수 없었다.

다중 해상도와 체적 LBP 연산의 효과를 평가하기 위하여 다양한 분류 실험을 통해 정밀도를 비교하였다. 다양한 반경 (R=1, 2, 3)과 평면 (XY, YZ, XZ)의 변화에 따라서 9개의 결합 히스토그램으로부터 최대 459가지의 특징인자를 얻을 수 있었다. 처음 LBP-TOP에 대한 세 개의 실험에서 반경은 1에서 3까지 변화하면서 정확도를 계산하였고, 다음 세 실험에서 정확도는 세 반경 (R=1, 2, 3)에 대하여 각 평면에서 계산하였다. 이 6 가지 실험에서 153개의 특징인자가 생성되었다. 반경 R=1의 XY 평면 실험과 세 가지 반경의 LBP-TOP에서는 각각 51 및 459가지의 특징인자가 각각 생성되었다.

표 2. 비유사도 측정도구에 의한 정확도 비교

Table 2. Accuracy comparison by dissimilarity.

Classification	Chi-square	Histogram intersection	Log likelihood
#. of TP	222	220	234/293
Accuracy(%)	75.8	75.1	79.9

세 개의 평면과 세 반경의 실험이 가장 높은 정확도 89.8% (263/293)의 결과를 예상했으나 가장 높은 정확도 90.1% (264/293)는 반경 R=1인 LBP-TOP에서 나타났다. 반경 크기는 세 평면 (XY, YZ, XZ) LBP 연산의 정확성에 반비례하였다. XY 평면의 정확도는 87.7% (257/293)로서 세 평면 중에서 가장 높았다.

표 3. 제안한 특징인자의 정확도 비교

Table 3. Accuracy comparison of proposed features

	R=1	R=2	R=3	XY	YZ	XZ	XY	LBP-TOP
	LBP-TOP			R=1,2,3			R=1	R=1,2,3
No. of TP	264	256	248	257	255	252	252	261
Accuracy (%)	90.1	87.4	84.6	87.7	87.0	86.0	86.0	89.1

3D AMFM과 제안한 특징인자의 정확성은 특징인자에 대한 선택이 없는 경우 단순 베이스 분류기에 의해서 각각 74.1%, 74.4%이었다. 아다부스트 학습을 사용한 제안한 특징인자의 정확도는 아다부스트 학습을 사용한 3D AMFM 특징인자의 85.3%보다 우수한 90.1%였다. 실험에서는 25개의 3D AMFM 특징인자 중에서 15개의 특징인자와 153개의 특징인자 (R=1) 중에서 23개의 특징인자가 아다부스트 학습에 의해 선정되었다. 중요한 특징인자는 불필요한 특징인자를 제외하여 적절하게 선택함으로써 주요 특징인자가 선별됨에 따라 정확도는 증가되었다.

표 4. 3D AMFM과 제안한 특징인자의 정확도 비교

Table 4. Accuracy comparison of 3D AMFM and proposed features.

Classification	Naive Bayes		Adaboost	
	3D AMFM	Proposed features (R=1,2,3)	3D AMFM	Proposed features (R=1)
No. features	26	459	15	23
No. of TP	217	218	250	264
Accuracy(%)	74.1	74.4	85.3	90.1

전문가 및 제안된 방법 간의 분류 결과를 비교하기 위해 분류 결과를 위한 오차행렬(confusion matrix)을 사용하였다. 테스트를 위해 단일관측치 제거 교차검증법을 사용하였다.

표 5. 제안한 방법을 이용한 6가지 폐 조직패턴에 대한 분류결과

Table 5. Classification Result Using Proposed Method.

	True Class					
	Broncho	Emphysema	GG_Reticular	Green Glass	Honeycomb	Normal
Broncho	66.7	0.0	12.1	3.03	18.2	0.0
Emphysema	0.0	92.0	0.0	0.0	4.0	4.0
GG_Reticular	4.5	0.0	90.9	4.5	0.0	0
GGlass	0.0	0.0	1.5	97.1	0.0	1.5
Honeycomb	4.7	4.7	0.0	0.0	90.6	0.0
Normal	0.0	0.0	0.0	6.8	0.0	93.2

6개의 폐질환영역의 분류 성능을 평가하기 위하여 3D AMFM, 제안한 특징인자, 그리고 제안한 특징인자(R=1,2,3)의 세 가지 실험에 대한 민감도(sensitivity)와 특이도(specificity)를 계산하였다. 아다부스트 학습을 사용한 제안한 특징인자 (R=1)에 의해 가장 높은 정밀도를 얻었다. 그러나, 제안한 특징인자 (R=1) 와 (R=1,2,3)의 정확성에는 의미있는 차이가 없었다, 따라서 민감도와 특이도의 비교에서 제안한 특징인자 (R=1,2,3)도 포함하였다.

표 6. 3D AMFM과 제안한 특징인자(R=1), (R=1,2,3)에 의한 방법에 의한 6가지 폐조직에 대한 민감도 비교

Table 6. Sensitivity comparison of six tissues classification using 3D AMFM, proposed (R=1), and proposed (R=1,2,3) features.

	Broncho	Emphysema	Green Glass Reticular	Green Glass	Honeycomb	Normal
3D AMFM	63.6	84.0	86.3	94.1	81.3	91.5
Proposed R=1	66.7	96.0	93.1	95.6	87.5	93.2
Proposed R=1,2,3	66.7	92.0	90.9	97.1	90.6	93.2

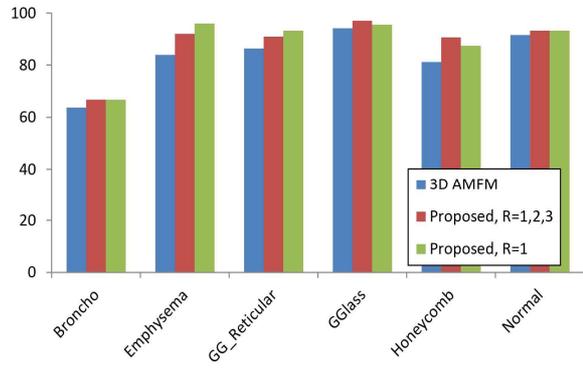


그림 3. 3D AMFM과 제안한 특징인자(R=1), (R=1,2,3)에 의한 방법에 의한 6가지 폐조직에 대한 민감도 비교 그래프

Fig. 3. Sensitivity comparison graph of six-class classification of 3D AMFM and proposed features.

표 7. 3D AMFM과 제안한 특징인자(R=1), (R=1,2,3)에 의한 방법에 의한 6가지 폐조직에 대한 특이도 비교

Table 7. Specificity comparison graph of six-class classification of 3D AMFM and proposed features.

	Broncho	Emphysema	GG_Reticular	Green Glass	Honeycomb	Normal
3D AMFM	98.1	98.7	96.8	95.6	95.2	97.9
Proposed R=1	97.7	99.3	97.2	97.3	97.4	98.7
Proposed R=1,2,3	98.1	98.8	98.0	96.9	97.0	99.1

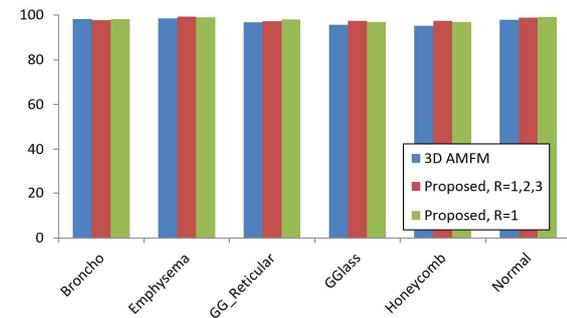


그림 4. 3D AMFM과 제안한 특징인자(R=1), (R=1,2,3)에 의한 방법에 의한 6가지 폐조직에 대한 특이도 비교 그래프

Fig. 4. Specificity comparison graph of six-class classification of 3D AMFM and proposed features.

#### 4. 결론

제안한 특징인자와 3D AMFM 특징인자의 정확도는 각각 90.1%과 85.3%로서 제안한 특징인자가 3D AMFM 특징인자보다 분류 성능이 우수하다는 것을 보여주고 있다. 제안한 방법에 의한 성능 개선은 Emphysema, GG\_Reticular, Honeycomb 표본의 민감도에서 현저하게 나타난다. 반경 R=1과 R=1,2,3에서 제안한 특징인자를 사용하였을 때 Emphysema의 민감도는 각각 96.0%와 92.0%였지만, 3D AMFM은 84%였다. 특이도의 개선은 두드러지지 않지만 전반적인 분류 성능은 제안한 방법이 3D AMFM보다 우수하였다. Emphysema의 특이도는 제안한 특징인자 R=1과 R=1,2,3에서 각각 99.3%와 98.7%였고, 3D AMFM은 98.6%였다. 반경 R=1에서 제안한 특징인자의 정확도는 R=1,2,3 보다 높았으나 그 차이는 유의미한 수준은 아니었다. 민감도 비교에서 GGlass과 Honeycomb에 대하여 반경 R=1,2,3의 제안한 특징인자가 R=1보다 우수한 성능을 보여주었다. 특이도 비교에서는 제안한 특징인자 (R=1,2,3)이 Broncho, GG\_Reticular 및 normal에서 더 나은 분류 성능을 보여주었다. 제안한 특징인자 (R=1)이 Emphysema에서 제안한 특징인자 (R=1,2,3) 보다 높은 민감도와 특이성을 보여주고 있다. 민감도 비교에서 가장 작은 반경 1인 LBP연산이 R=1,2,3 보다 Emphysema와 GG\_Reticular를 분류함에 더 나은 성능을 보였으며 이는 Emphysema, GG\_Reticular 폐조직은 작은 LBP 연산에서 더 잘 식별될 수 있음을 알 수 있다.

#### REFERENCE

[1] J. W. Gurney, "Determining the Likelihood of Malignancy in Solitary Pulmonary Nodules with Bayesian Analysis: Part I. Theory", *Radiology*, vol. 186, pp. 405-413, 1993.

[2] B. Sahiner, H.-P. Chan, N. Petrick, M. A. Helvie, and M. M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis", *Medical Physics*, vol. 25, pp. 516-526, 1998.

[3] K. I. Laws, "Rapid texture identification", in *Image Processing for Missile Guidance*. vol. 238, T. F. Wiener, Ed., ed Bellingham, Washington:

The Society of Photo-Optical Instrumentation Engineers, pp. 376-280, 1980.

[4] X. Ye, "Computer-aided Classification of Interstitial Lung Diseases Via MDCT: 3D Adaptive Multiple Feature Method (3D AMFM)", *Acad Radiol.*, vol. 13, no. 8, pp. 969-78. 2006 .

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multi-resolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[6] C. Shan, S. Gong, and P.W. McOwan, "Facial Expression Recognition based on "Local Binary Patterns: A Comprehensive Study", *Journal of Image and Vision Computing*, vol. 27, no. 6, pp. 803-816, 2009.

[7] Y. Freund and R.E. Schapire. "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, vol. 14, no. 5, pp. 771-780, 1999.



원 철 호(Won Chul-ho)

1998년 8월 경북대학교 대학원 전자공학과 박사  
 1998년 2월 - 1999년 11월 계명대학교 동산의료원 의공학교실 연구강사  
 1999년 11월 - 2002년 2월 Univ of Iowa, Dept. of Radiology, Post Doc. Fellow  
 2002년 2월 - 현재 경일대학교 의용공학과 교수

Interest: Bio-medical signal processing