
디지털 흉부영상에서 주성분분석을 이용한 폐암인식

박형후* · 옥치상** · 강세식*** · 고성진*** · 최석윤****

Recognition for Lung Cancer using PCA in the Digital Chest Radiography

HyungHu Park* · ChiSang Ok** · SeSik Kang*** · SungJin Ko*** · SeokYoon Choi****

요 약

흉부의 폐질환으로 폐암발생은 꾸준히 증가하고 있다. 일차적인 폐암진단 방법에는 흉부X선영상이다. 흉부X선영상 이용하여 폐암진단을 하기 위해서는 임상경험이 풍부한 의사가 필요하다. 그러나 풍부한 경험을 가진 의사라도 오진이 발생할 수 있고 이는 폐암의 조기진단과 생존률을 낮게 한다. 본 논문에서는 주성분분석을 이용하여 학습영상의 데이터베이스와 질병이 있는 흉부영상을 진단함으로써 컴퓨터보조진단의 기반을 마련하고자 한다. 이를 의사가 진단하기 전의 예비판독의 단계로 이용한다면 오진으로 인한 환자의 조기진단률의 감소를 줄일 수가 있다. 실험은 정상흉부X선영상과 악성폐암인 기관지암(Bronchogenic Carcinoma)과 양성종양인 육아종(Granuloma)으로 실험하였다. 영상은 주성분분석 후 정상영상과 질환 영상의 고유영상을 추출하고 상호 비교한 뒤 인식효율을 비교하였다. 결과로는 정상영상과 질환영상간의 인식률은 높았으나 질환간의 인식효율은 정상에 비해 다소 떨어지는 것으로 나타났다. 흉부질환간의 인식효율을 높이기 위해서 관련 알고리즘에 관한 연구가 계속 이어진다면 컴퓨터보조진단에 좋은 연구기반이 되리라 생각한다.

ABSTRACT

Risk of lung cancer among lung-related diseases has gradually increased during last decades. The chest digital radiography is the primary diagnosis method for lung cancer. Diagnosing lung cancer using this method requires doctors of ripe experience. Despite their experience there are often wrong diagnoses, which decrease early diagnosis and survival rates of patients. The aim of this study was intended to establish the base on the Computer Aided Diagnosis (CAD) by analyzing Image Recognition Algorithm using Principle component Analysis (PCA) and diagnosing patient's chest X-ray image. The database obtained through this approach enables a doctor to significantly reduce misdiagnosis during the early diagnosis stage, if he or she utilizes it as the preliminary reading step. Case studies were carried out using normal organ, and organs suffering from bronchogenic carcinoma and granuloma. A normal image and unique disease images were extracted after PCA analysis, and their cross-recognition efficiency were compared each other. The result revealed that the recognition rate was much high between normal and disease images, but relatively low between two disease images. In order to increase the recognition efficiency among chest diseases the related algorithms have to be developed continuously in the future study, and such effort will establish the resolute base for CAD.

키워드

PCA, 영상인식, 폐암, 기관지암, 육아종

Key word

PCA, Image Recognition, Lung cancer, Bronchogenic Carcinoma, Granuloma

* 정희원 : 고신대학교대학원 (제1저자)

** 정희원 : 고신대학교대학원 보건과학과

*** 정희원 : 부산가톨릭대학교방사선과

**** 정희원 : 부산가톨릭대학교방사선과 (교신저자, marr35@nate.com)

접수일자 : 2011. 02. 17

심사완료일자 : 2011. 03. 23

I. 서 론

정보통신기술(Information Technology, IT)의 발전은 임상의료에 적용되어 인체의 내부를 시각화 하는데 큰 역할을 하고 있으며, 건강하게 오래 살고 싶은 인간의 욕망을 실현시키는데 큰 공헌을 하고 있다. 정보통신 기술은 인체의 영상분야에도 큰 성과를 나타내며, 1972년 컴퓨터 공학을 이용하여 X선을 투영한 사진을 3차원으로 볼 수 있게 한 컴퓨터단층촬영(Computed Tomography, CT)이 개발되었다. 2차원 영상을 3차원으로 발전시킨 이 연구를 기반으로 핵의학 분야에서 동위원소를 사용하여 인체의 신진대사를 연구하는 양전자 단층촬영(Positron Emission Tomography, PET)와 핵자기공명(Nuclear Magnetic Resonance, NMR)현상을 이용해 인체의 해부학적 변화를 보는 자기공명영상(Magnetic Resonance Imaging, MRI)이 개발되었다. 이 기기들을 이용하여 신체 내부를 정확히 파악할 수 있게 되어 암의 조기발견과 치료에 획기적인 발전이 있었다. 그리고 촬영된 영상이 의료영상저장전송시스템(Picture Archiving and Communication System, PACS)를 이용하여 디지털 이미지로 변환하는 동시에 대용량 기억장치에 저장시켜 영상의학과 전문의가 모니터를 통해 판독할 수 있도록 하고 있다. 최근에는 인공지능 컴퓨터를 이용하여 컴퓨터가 암으로 의심되는 부위를 찾아 의사의 진료업무를 지원하는 기술인 컴퓨터보조진단(Computer Aided Diagnosis, CAD)도 발전하고 있다. 이러한 기술의 발전으로 점점 정밀한 영상의학 기기가 개발되면서 의사가 없는 원격지에서의 진료와 진단이 가능하게 될 것으로 기대된다.

패턴인식(Pattern Recognition) 분야는 기본적인 영상 신호처리, 영상분할, 신경회로망 또는 통계적 패턴인식 기술 등을 종합하는 기술로서 선진 외국의 경우 기업은 물론 정부의 주도하에 1970년대부터 학제적 공동연구가 본격적으로 전개되기 시작하였으며[1] 현재에도 얼굴 인식의 영상처리에 대한 연구가 활발히 진행되고 있으며 의학적인 측면에서는 자동진단(Automatic Diagnosis)에도 응용되고 있다.

인식기술은 형태 확인의 응용에 널리 사용되며, 은행경비시스템, 개인적 신분확인, 국가안보 등에 적용되기도 한다. 객체인식(Object Recognition) 분야는 영

상의 목적물로부터 원하는 부분을 분리하는 과정, 목적물의 파라미터를 추출하는 과정, 표현된 특징 파라미터를 데이터베이스 상의 자료와 비교하여 인식하는 과정으로 나누어진다. 여기서 각각의 얼굴검출과정, 얼굴특징추출과정, 얼굴인식과정 이라는 세부 주제로 다시 구분하며 각각 독립적으로 연구되어 왔다[2]. 인식분야는 크게 Eigenface를 이용한 주성분분석(PCA, Principal Component Analysis) 방법론, 동적링크구조를 이용한 방법론, 신경회로망을 이용한 방법론 등으로 나누어져서 연구되어지고 있다[3]. 본 연구는 PCA를 이용한 객체인식의 방법으로 인식에 적용된 사례들의 의료영상에 적용해서 문제점과 활용가능성에 대하여 알아보려고 한다.

본 연구는 영상처리와 영상인식기술, 판독의사의 임상적인 경험을 바탕으로 컴퓨터 내에서 전문가적인 수준에 영상진단의 보조기능을 접목해서 진단하고자 하는 것이다. 본 연구의 목적은 인식 알고리즘에 많이 활용되어지고 있는 PCA를 이용해 흉부X선영상에서의 폐암 인식 효율을 알아보고, 이를 다른 흉부질환의 CAD에 활용할 수 있는 방안에 대하여 알아보려고 한다. 진단오류의 감소와 조기진단의 효율을 높이는 방안으로 PCA를 이용한 정상흉부영상과 질환영상의 구별에 관한 인식 평가를 시행하고 흉부 컴퓨터지원진단으로 활용 할 수 있는 방안에 관한 연구를 시행하고자 한다.

II. 이론적 배경

2.1. 컴퓨터보조진단(CAD)

W. R. Hendee는 논문에서 미래의 방사선학의 주안점으로 크게 4가지 분야를 거론하면서 자신의 정보관점으로 예측하였다. 방사선학의 미래분야로는 보다 나은 기술적, 영상적, 진단적 측면보다는 방사선 정보의 흐름에 초점이 맞추어져야 한다고 주장하였으며 이는 현재 기술의 시점에서 볼 때 의료영상저장전송시스템(PACS)의 예측과 일치 하였다[4]. 그러면 지금의 정보기반 사회의 시점에서 또 다시 미래의 방사선학 분야를 예측해 본다면 정보기반 이후의 지식기반 사회란 관점에서 CAD와 컴퓨터보조기술(Computer Aided Surgery, CAS)이라 판단된다.

의료영상에서 원하는 병변을 자동으로 검출하는 기능으로 제한된 시스템에서 컴퓨터보조검출(Computer Aided Detection)로도 기술되는데, 컴퓨터의 역할을 의심되는 병변(lesion)의 검출로 제한하고 그것에 대한 진단은 영상의학과 전문의에게 전임하는 것이 바람직하기 때문이다. 즉, 컴퓨터가 진단 및 수술을 도와주는 시대가 곧 도래 할 것이라 판단한다[5].

의료영상을 이용한 컴퓨터보조진단 시스템은 영상의 종류와 목적 병변에 따라 적용되는 알고리즘을 달리 하는데 적용범위와 중요도를 기준으로 그 종류를 크게 살펴보면 유방 초음파 영상에서 종양의 양성과 악성 감별진단, 유방 X선 영상에서의 석회화(Calcification)와 중앙검출의 시스템, 단순 X선 영상에서 병변의 검출시스템, 흉부CT에서의 폐 결절시스템에 대한 연구가 활발히 이루어지고 있다 [6]. 그리고 흉부 영상에서 폐암을 진단하기 위한 CAD연구로는 흉부X선영상과 CT를 이용한 연구가 이루어지고 있다.

CAD는 정량적으로 분석한 결과를 토대로 이상 부위를 표시함으로 영상을 판독하는 의사의 과도한 판독 업무를 줄여주고 질환으로 의심되는 영상에 대해서는 정량분석 결과를 토대로 의사가 최종적으로 진단을 내리게 하는 시스템을 말한다. 첨단의료장비들은 이전과는 비교 할 수 없을 만큼 고해상도의 고차원, 다중 영상 기법을 통해 영상들을 만들어 내어 고도의 기능으로 컴퓨터 처리의 도움이 없이는 일일이 해석하는 것과 진단에 어려움을 겪는 문제에 부딪히게 되었다. 이와 같이 CAD를 이용하여 반복적이거나 빠뜨리기 쉬운 부분을 검출, 계산 하는데 사용되어 진다면 의사의 진단 시간과 복잡한 업무를 대신 할 수 있고 진단 효율을 향상시키고 업무능률의 향상으로 효과적이다[7]. CAD 시스템이 개발되면, 현재 의사의 육안적 인식과 판별능력의 한계로 인한 진단의 불확실성을 상당히 개선할 수 있을 것으로 기대하고, 폐암과 유방암 등 다양한 질환에서의 조기에 암 진단을 위한 공익의료사업에 국가적 의료비용을 대폭 절감 할 수 있어서 향후 의료영상기술에 지대한 발전을 가져올 것으로 기대하며 이에 따른 큰 규모의 경제적 시장형성도 예측되어 진다[8].

2.2. 폐암(Lung Cancer)

암(Cancer)은 생활수준의 향상과 의료 수준의 발달로 전체적인 사망자 수는 줄고 있지만 폐암으로 사망하는

사람의 수는 줄어들지 않고 오히려 증가하고 있다[9]. 통계청이 발표한 2000년도 사망원인 통계결과에 따르면 그동안 사망 원인 1위를 차지하던 위암을 제치고 폐암이 암 가운데 가장 높은 사망률을 기록했다고 보고하고 있다. 일반적으로 단순 흉부X선 검사로 간단하고 쉽게 폐에 종양 유무를 알 수가 있지만 종양의 크기와 형태, 정확한 위치를 파악 하려면 추가적으로 CT나 MRI를 추가로 검사해야 한다[10].

폐암의 분류는 90~95%가 기관지의 상피에서 기원하는 기관지암(bronchogenic carcinoma, BC)이며, 5~10%가 기관지유암(bronchial carcinoid), 세기관지폐포암(bronchioloalveolar carcinoma)과 기타의 암으로 구성되며, 폐 결절(SPN)과 육아종(Granuloma) 등은 암으로 판단되기 전까지는 모두 초기 암으로 진단되어 진다[11]. 폐암은 대부분이 병이 진행된 상태에서 발견되기 때문에 근시적 수술, 항암요법, 방사선치료 등을 포함하여 적극적인 치료에도 불구하고 그 예후가 좋지 않으므로 [12] 폐암의 조기 진단과 치료의 중요성은 매우 강조된다. 폐암의 조기 발견은 폐암 환자의 생존율을 높일 수 있는 효과적인 방법으로 폐암의 검진에 기본적으로 사용되는 단순 흉부 X선 검사와 객담검사를 실시한다.

2.3. 주성분분석(PCA)

PCA 알고리즘은 평균과 분산의 통계적 특징을 이용한 2차적 통계적 기법으로 입력데이터에 대하여 최대 공분산의 각 방향을 가리키는 직교와 정규화된 일련의 축들의 집합을 찾는다. PCA의 중요한 목적은 데이터 무리에서 존재하는 변화를 가능한 많이 보유 하면서 서로 연관된 많은 변수들이 있는 데이터 무리의 차원(dimension)을 줄이는 것이다. 이것은 서로 상관관계가 없고 첫 몇 성분이 본래 변수 모두에 존재하는 대부분의 변화를 간직하는 변수의 새 집합, 즉 주성분으로 변환해서 나타낼 수 있다. PCA는 직접적으로 많은 얼굴 영상들 간의 상관관계에 대한 해석이 어려울 때 서로 상관되어 있는 변수들을 적절히 선형 변환시켜서 소수의 서로 독립적인 주성분들을 추출 및 해석하는데 목적이 있다. 차원축소의 결과로 얻어지는 주성분 점의 수들은 통계분석들을 위한 입력 자료로 이용되어 일련의 분석과정에서 하나의 중간단계의 역할을 한다. PCA를 이용한 영상 인식(image recognition) 기술은 학습영역과 테스트영역의 두 영역으로 구분한다. 먼저 학습영역은 고유영상

(Eigenvector)을 생성시키고 모든 학습영상을 이 영상 공간에 투사시켜 몇 개의 성분 값을 저장한다. 그 후 각각의 영상마다 저장된 성분들의 평균은 대표 값으로 저장해 놓는다. 그림 1은 단순 흉부 X선 영상을 1차원의 벡터(vector)로 변환하는 과정을 나타내고 그림 2는 기본적인 PCA를 이용한 흉부영상에서의 폐질환 인식기술의 흐름도(diagram)를 나타낸다.

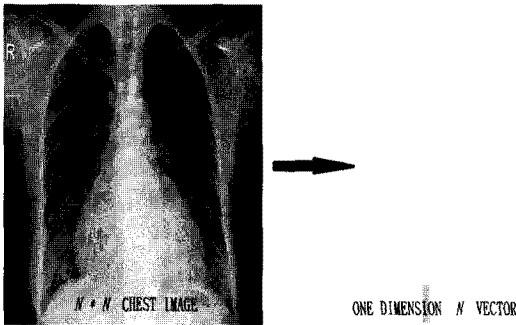


그림 1. 1차원의 벡터로 변환
Fig. 1. Transformation of one dimension vector

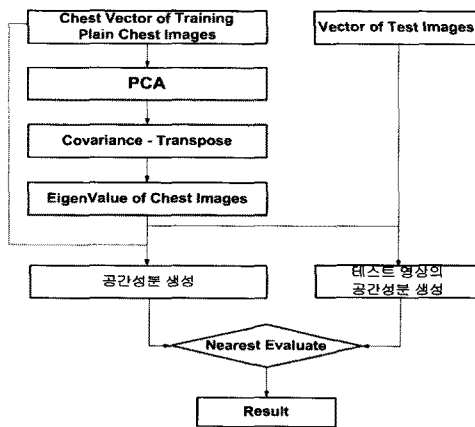


그림 2. PCA 이용한 폐질환인식 모식도
Fig. 2. Diagram of Lung lesion recognition using PCA

2.4. 고유영상

고유영상은 흉부영상에서의 질환의 인식을 위한 영상의 특징점을 이용하는 것으로 고유값과 고유벡터를 나타내며, 각 학습영상(training set)에 대해 특정한 고유값을 가지며, 그 고유값을 이용해 유사한 영상을 찾기 위한 방법이다. 행렬을 사용하는 이유는 2차원 데이터를

쉽게 표현하기 위한 방법이며 연산을 사용하기 위해 사용되며 영상은 2차원으로 구성되어 있다. 즉, 영상의 크기가 256×256이면, 가로 256개 픽셀과 세로 256개 픽셀로 이루어져 있으며 각 픽셀은 흑백의 Gray 영상 경우는 0~255값을 가지게 된다. 0이면 검은색, 255는 흰색이며 픽셀값이 커질수록 흰색에 가까운 색을 띠게 된다. 이 기본요소는 영상 이미지의 변화를 나타내는 일련의 특징을 의미하는 것으로 수학적으로 표현하면 하나의 영상 이미지를 나타내는 고변행렬의 고유벡터를 나타낸다. 영상은 주변 픽셀과 매우 상관관계가 높아 공간영역에서 영상을 기술하는 것은 많은 중복성을 갖게 한다. 이렇게 상관관계가 높은 데이터의 경우 PCA를 적용하면 데이터의 중복을 제거할 수 있다. 따라서 상관관계가 높은 영상의 경우 PCA를 통해 새로운 영역에서 효율적으로 데이터를 표현할 수 있다.

2.5. 고유값 생성[13]

수집할 영상은 유사한 형태의 영상을 선택해야 된다. 특정영상의 고유한 값을 가지기 위해서 평균값을 이용하므로 평균값이 너무 넓게 분포되어 있으면 고유값을 나타내기 어려우므로 데이터베이스 내에서 특정영상을 찾기가 힘들다.

$$I_1, I_2, I_3 \dots I_N \tag{1}$$

2차원 영상을 1차원 Vector로 표현하는 과정이다.

$$I_i \text{ as a vector } \Gamma_i \tag{2}$$

2차원 행렬의 계산값이나 1차원 벡터로 표현하나 값은 변화가 없다. 수집영상의 평균영상을 구한다.

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \tag{3}$$

평균영상에서 특정영상을 빼면(drift) 그 영상의 특징값이 되나 그 값을 이용하기에는 너무 적기 때문에 기저값(Eigenvalue) 및 기저벡터(Eigenvector)를 구한다.

$$\Phi_i = \Gamma_i - \Psi \tag{4}$$

공분산(covariance)은 두 변수의 관계를 나타내는 양 또는 두개 이상 변수들 간에 공유하는 변량이며 두 변수 간의 선형 연관성을 나타내는 도구이다. 즉, 하고자 하는 작업은 특정영상을 찾는 것이므로 특정영상은 평균 영상에서 얼마의 특징값을 가지는지를 계산하는, 공통적으로 분포된 값이라 할 수 있다. 따라서 그 특징 값은 가장 간단하게 평균영상에서 얼마나 떨어져 있는나를 계산하는 것으로 예를 들면, 얼굴이 큰 사람은 평균적 사람보다 2차원 영상으로 표현하고, 위의 과정들을 수행하면 평균얼굴보다 얼마나 차이가 나는나의 값은 얼마나 얼굴이 큰지로 표현할 수 있다. 그러므로 얼굴이 큰 사람과 작은 사람, 얼마나 크고 얼마나 작은지를 표현할 수 있으며, 더 나아가 특정 값을 가지지 않더라도 그룹화 시켜서 그 그룹에 속하면 그룹단위로 어떤 처리가 가능하다.

공분산 C 를 수식으로 표현하면 아래와 같다.

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T \quad (N^2 \times N^2) \quad (5)$$

where $A = [\Phi_1 \Phi_2 \dots \Phi_M] \quad (N^2 \times M)$

2.6. 정규화

정규화(Normalization)는 명암값은 분포가 한쪽으로 치우치거나 균일치 못한 영상을 정규화에 의해 명암값의 분포가 균일하게 되어 영상이 향상되는 과정이다. 즉 단순흉부X선영상의 학습영상에서 정규화는 질환형태가 일관성을 가지도록 질환영상을 일정한 영상을 가지도록 재생성하는 과정으로 변환하는 과정이다. 그러므로 고유벡터를 정규화 함으로써 정규화한 영상을 출력한다.

2.7. 평균영상

평균영상을 구하는 방법에서는 미디언필터를 이용하는 방법과 최대값, 최소값 필터를 이용하는 방법 등이 있는데 본 논문에서는 영상전체의 특징점을 이용하는 PCA 방법에 적합한 필터를 사용하였다. 식은 다음과 같다.

$$\text{평균영상} = \frac{(\text{학습영상벡터} - \text{평균벡터})}{(S.D. \times \text{size})} \times S.D. \quad (6)$$

2.8. 고유영상추출

각각의 흉부의 질환영상은 고유영상의 선형결합으로 완벽하게 복원될 수 있다. 학습영상으로부터 구할 수 있는 고유영상의 개수는 학습영상의 수와 같다. 고유치 및 고유벡터를 순서별로 정렬한 후 평균을 뺀 질환영상을 입력하여 투영(Projection) 시키면 고유흉부영상의 질환영상 인식에 사용되는 학습영상에 대한 특징벡터를 구할 수 있다. 또한 똑같은 방법으로 테스트영상에 대한 특징벡터를 추출한다. 이것은 고유의 흉부질환 영상의 각각에 영상차이를 나타내는 특징점이다.

2.9. 유클리디언 거리

질환영상의 데이터베이스에 저장되어 있는 고유치를 학습영상에 곱하고 입력된 영상에 고유치를 곱한 영상 사이의 거리를 구한다. 이렇게 구한 유클리디언 거리를 가장 작은 값을 갖게 하는 패턴이 최종 결과로 출력되며, 이 최소값을 영상의 인식을 비교를 통해 결과를 도출하게 된다. 예를 들어 흉부 질환영상 A의 특징벡터를 $X = X_1, X_2, X_3, \dots, X_N$ 이라고하고, 흉부질환영상 B의 특징벡터를 $Y = Y_1, Y_2, Y_3, \dots, Y_N$ 이라고 하고 D를 유클리디언 거리라고 한다면

$$D^2 = (X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_N - Y_N)^2 + (X_N - Y_N)^2 = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (7)$$

즉, D^2 값이 가장 작을 때의 결과가 인식 하고자 하는 입력 흉부질환영상 이라고 단정할 수 있다. 실험에서는 Matlab 실험을 위한 프로그램 소스를 코딩하고 프로그램을 실행하여 디버깅 하였다. 그림을 유클리디언 거리를 구하고 작은 순서에서 큰순서로 정렬한 후 가장 가까운 이미지를 찾는다. 즉 최소치 이미지를 가지고 비교 분석하였다.

III. 재료 및 방법

3.1. 실험영상

실험에서 사용된 영상은 영상의학과 전문의의 영상 진단과 진단검사의학과의 진단을 토대로 정상으로 진단된 단순흉부X선영상 50중례와 BC로 진단된 흉부영

상 38증례, Granuloma로 진단된 흉부 영상 32증례의 영상으로 질환부위의 영상을 획득하였고, 정상흉부영상 획득은 폐를 구분하는 5개 구획구간인 좌폐상엽 10증례, 좌폐하엽 10증례, 우폐상엽 10증례, 우폐중엽 10증례, 우폐하엽 10증례를 고르게 획득하였다. 실험영상 획득원칙은 실험영상 중 두개 이상의 질병이 있어 영상에 중복되어 나타난 것은 실험영상에 포함시키지 않았다.

환자의 자세는 선 자세에서 후진방향의 자세(posterior to anterior) 검사를 원칙으로 하고 누운자세(supine) 및 기타 자세의 영상은 실험에서 제외 하였다. 실험에 사용된 흉부X선영상은 DCM파일 형식으로 크기가 15.4MB로 용량이 커서 분석이 적합하지 않아, 15.8KB의 BMP파일 형식으로 변환하여 실험하였다. 영상변환은 Adobe Photo Shop-CS3 를 이용해 100*100 Pixel의 크기로 실험영상을 획득하였다. 그림 3의 영상들은 실험에서 사용될 학습영상이다. 실험영상은 크게 두 가지로 학습영상과 테스트영상으로 구분한다. 고유영상을 생성시키기 위한 학습영상(training image)은 정상흉부영상과 BC 및 Granuloma가 있고, 랜덤영상(random image)의 인식을 알아보기 위한 테스트영상으로 정상흉부영상, BC 및 Granuloma영상으로 실험하였다.

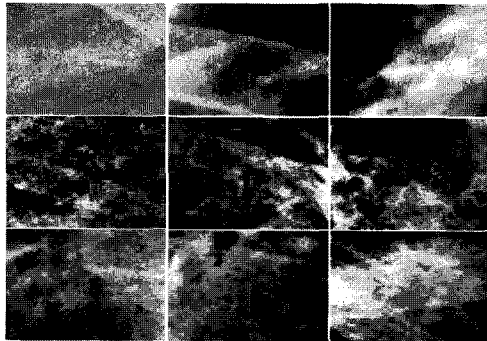


그림 3. 학습영상
Fig. 3. Training image

3.2. 실험방법

정상흉부X선영상 50증례, 질환영상인 BC 38증례, Granuloma 32증례를 실험영상으로 사용하였다. 정상 흉부 영상은 폐의 5개 구역에서 10개씩 균등히 획득한다. 이때 정상흉부영상의 크기는 100*100 pixel의 크기로 cut off 하여 실험영상을 획득 하였다(그림 4).

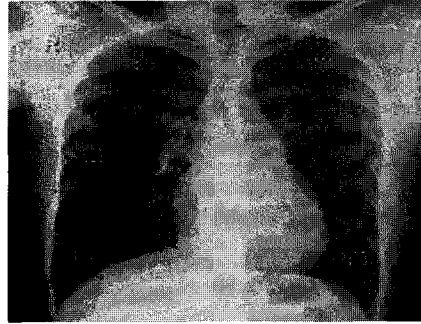


그림 4. 정상흉부 영상
Fig. 4. Normal image on Chest Digital image

질병영상인 BC와 Granuloma는 질환부위를 정확히 획득하기 위해서 그 환자 영상의 CT 단층영상(axial image)과 관상면영상(coronal image)을 참고 하였다. 흉부영상의 질환 부위를 100*100 pixel 크기로 cut off 하여 실험 영상을 생성하였다(그림 5, 그림 6).



그림 5. 흉부CT 영상(단층)
Fig. 5. Chest CT image(axial)



그림 6. 흉부CT 영상(관상면)
Fig. 6. Chest CT Image(coronal)

흉부정상영상을 학습영상으로 데이터베이스 내에서 고유영상을 생성하고 테스트영상으로 정상흉부 영상과 Granuloma, BC로 PCA 알고리즘을 적용하였다. BC를 학습 영상으로 데이터베이스 내에서 고유영상을 생성하고 테스트 영상으로 BC영상, 정상흉부영상, Granuloma로 PCA 알고리즘을 적용하였다. Granuloma를 학습영상으로 데이터베이스 내에서 고유영상을 생성하고 테스트영상으로 정상 흉부영상, BC영상, Granuloma로 PCA 알고리즘을 적용하였다. 분석과정은 PCA분석이론에 따라 Matlab 프로그램을 이용하여 정상영상과 질병영상 간의 인식률을 알아보았다. 그림 7은 흉부영상의 PCA 과정이고 표 1은 학습영상의 실험영상과 테스트영상의 관심영역을 나타낸다.

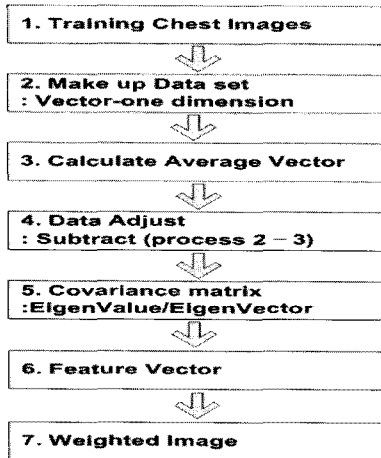


그림 7. 흉부영상의 PCA분석 과정
Fig. 7. Process of PCA on Chest image

표 1. 실험영상의 관심영역
Table 1. Cut-off of ROI on experiment image

Training image	Nor	BC	Gr
Input images (Random images)	Nor	Nor	Nor
	BC	BC	BC
	Gr	Gr	Gr

Nor=Normal, BC=Bronchogenic Carcinoma, Gr=Granuloma .

IV. 실험 결과

학습영상의 데이터베이스 내에서 랜덤의 테스트영상으로 입력영상의 유클리디언 거리를 나타낸다. 기본적인 데이터베이스 내 영상은 최종적으로 고유치인 Eigenvalue

에 대해서 입력영상의 고유치와 비교하여 거리를 계산하고 가까운 순서대로 이미지를 정렬하여 입력영상의 인식을 최종적으로 나타낸다. 데이터베이스 내에 공분산에 대하여 구한 고유치를 영상에 곱하여 가중이미지(weighted image)를 생성한다. 그림 8의 결과는 Matlab 프로그램에서의 가중이미지와 유클리디언 결과를 나타낸다.

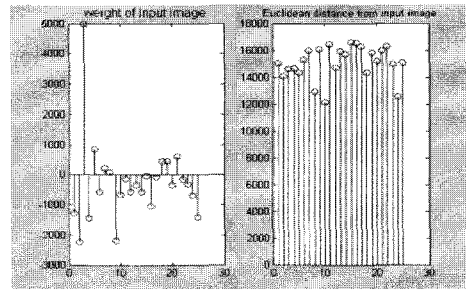


그림 8. Matlab 분석결과(공분산에서 구한 고유치의 가중이미지)
Fig. 8. The result of Matlab analysis

수치 중 유클리디언 거리에 가장 가까운 최소치를 가지고 흉부질환의 폐암 인식률을 비교하였다. 유클리디언 거리를 바탕으로 정상흉부영상, BC와 Granuloma의 인식 효율을 그래프로 도식화 하였다. 그리고 질병영상인 BC, Granuloma에서는 정상영상과 질병들(BC, Granuloma)의 질환별로 구분하여 입력하고 질병의 인식 효율을 그래프로 도식화 하였다. 표 2는 유클리디언 거리 분석결과로 유클리디언 거리의 최대값과 최소값을 나타낸다.

표 2. Matlab program의 유클리디언 거리
Table 2. Euclidean distance of Matlab program

NAME	Value	Min	Max	Std
A	25×10000 double	0	255	70.7034
B	1×25 double	6232...	2827...	5593763
DiffWeigh	25×1 double	-110...	8023...	3073.557
Difference	10000×1 double	-226...	371...	87.9987
InImWeigh	25×1 double	-224...	4990...	1350.281
InImage	10000×1 double	0	255	35.5006
InputImage	100×100 unit 8	0	255	
L	25×25 double	7416...	1615...	1657889..
M	25	25	25	0
MaxValue	16628.4217	1662...	1662...	0
MinValue	12134.8204	1213...	1213...	0
NorImage	10000×1 double	-126...	448...	80
ReshapedImage	100×100 double	-8.26...	365...	57.0315

그림 9는 정규화 영상으로 입력영상이 균일하지 못하므로 PCA에 적용하기 위해서는 일정한 영상으로 재생성한 영상을 나타낸다.

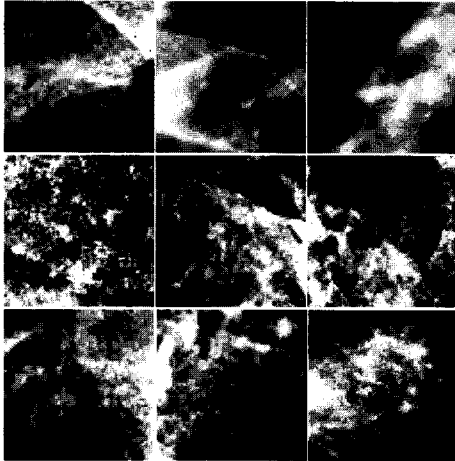


그림 9. 정규화 영상
Fig. 9. Normalization image

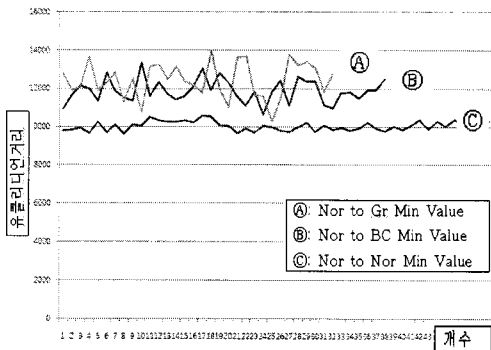


그림 10. 학습영상과 입력영상간의 비교
(학습영상: 정상-정상, 입력: 정상-BC & 정상-Gr)
Fig. 10. Comparison between training image and test image (Training image: N-N, Test image: N-BC & N-Gr)

4.2. 폐암의 인식도

그림 10의 실험결과에서 총 50개의 정상흉부영상 중에서, 정상영상(N으로 표기) 25증례를 학습영상으로 하고 나머지 25개 정상영상 25증례를 테스트영상으로 해서 PCA 분석결과의 최소치와 학습된 정상흉부영상 25증례와 테스트영상인 BC영상 25증례를 PCA 분석결과의 최소치와 비교한 그래프로 25개증례 모두 구별이 가

능하여 인식률이 100%를 나타냈다.

정상영상 25증례를 학습영상으로 하고 또 다른 정상 흉부영상 25증례를 테스트 영상으로 해서 PCA 분석결과의 최소치와 정상흉부 25증례를 학습영상으로 하고 Granuloma 25증례를 테스트영상으로 한 비교에서도 25증례 모두 구별이 가능해서 인식률이 100%를 나타냈다. 그림 10은 학습영상으로 흉부영상의 데이터베이스 내에 기준영상으로 흉부정상영상으로 하고 질환영상(BC, Gr)을 입력영상으로 결과를 나타낸다.

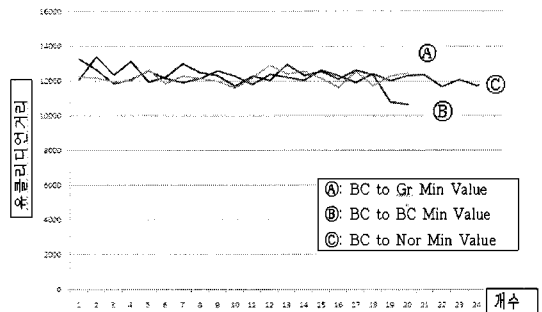


그림 11. 학습영상과 입력영상간의 비교
(학습영상:BC-정상, 입력영상:BC-BC & BC-Gr)
Fig. 11. Comparison between training image and test images (Training image: BC-Nor, Test image: BC-BC & BC-Gr)

그림 11의 실험결과에서 BC 20증례를 학습영상으로 정상흉부영상 20증례를 테스트영상으로 PCA 분석 후 결과의 최소치와 BC 20증례를 학습영상으로 Granuloma 20증례를 테스트영상으로 PCA 분석결과의 최소치를 비교 하였을 때의 그래프로 구별 가능한 증례가 9개로 구별확률이 45%를 나타냈다.

BC 20증례를 학습영상으로 정상영상 20증례를 테스트영상으로 PCA 분석결과의 최소치와 BC 20증례를 학습영상으로 또 다른 BC 20증례를 테스트영상으로 PCA 분석결과의 최소치를 비교한 그래프에서 보면 8증례가 구별 가능하여 인식률이 40%를 나타냈다. 그림 11은 흉부영상의 데이터베이스 내에 기준영상을 BC영상으로 하고 질환영상인(BC, Gr)과 정상영상을 입력영상으로 실험한 결과를 나타낸다.

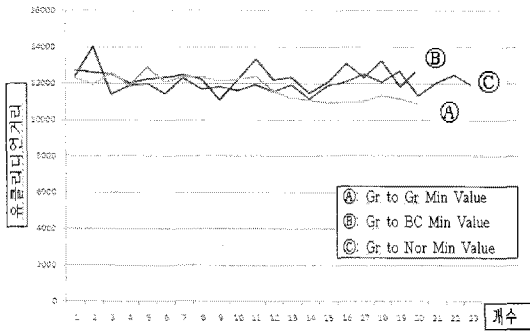


그림 12. 학습영상과 입력영상간의 비교
(학습영상:Gr-정상, 입력영상: Gr-Gr & Gr-BC)

Fig. 12. Comparison between training image and test image(Training image: Gr-Nor, Test image: Gr-Gr & Gr-BC)

그림 12의 실험결과에서 학습된 Granuloma 20증례와 테스트영상으로 정상영상 20증례를 PCA 분석 후에 결과의 최소치와 Granuloma 20증례를 학습영상으로 BC 20증례를 테스트영상으로 PCA 분석결과의 최소치를 비교 하였을 때의 그래프의 인식률은 8개의 증례가 구별 가능하여 40%를 나타냈다. Granuloma 20증례를 학습영상으로 정상영상 20증례를 테스트영상으로 PCA 분석 결과의 최소치와 Granuloma 20증례를 학습영상으로 Granuloma 20증례를 테스트영상으로 PCA 분석결과의 최소치를 비교 하였을 때의 결과 그래프의 인식률은 11개의 증례가 구별 가능하여 55%를 나타냈다.

그림 12는 흉부영상의 데이터베이스 내에 기준 영상으로 Granuloma 영상으로 하고 질환영상 (BC, G)과 정상 영상을 입력영상으로 실험한 결과를 나타낸다.

V. 결론

실험결과를 데이터베이스 내의 학습영상 기준에 따라 인식률이 달랐으며, 흉부 폐암인식률은 질환흉부영상보다 정상흉부영상에서 인식률이 100%로 높게 나타났다. 질환이 있는 데이터베이스에서 영상의 폐암인식률은 정상에 비해 다소 미흡하나, 단순 흉부영상에서 일반적인 영상처리 알고리즘인 주성분분석을 적용하는 실험의 임상증례로 충분하게 의의가 있다고 사료된다. 그리고 단순흉부영상에서 검사 시 환자의 호흡과 자세,

관전압(KVP)과 관전류(mAs) 조건에 따라 영상의 표현이 달랐고 또한 같은 질환에서도 환자의 검사 조건에 따라 달리 표현되는 흉부영상으로 실험하였기 때문에 동일 PCA 알고리즘을 이용해 높은 인식률을 나타내지는 못하였으나 정규화과정을 통해 개선 방향을 찾을 수가 있었고 단순흉부영상의 컴퓨터보조진단에 적용하기 위해서는 활용도가 높은 개선된 알고리즘의 개발이 필요로 하며, 정확한 정량적인 분석을 통하여 여러 가지 질환이 포함된 복합질환의 흉부영상을 구별해 내는 연구도 같이 병행되어야 할 것이다.

본 논문 결과를 토대로 흉부컴퓨터보조진단에 적용되는 프로그램을 개발한다면 영상의학과 전문의의 흉부 영상 판독의 보조수단으로서 판독 오진 감소와 판독 시간을 단축할 수 있으며, 조기폐암 진단에 매우 중요한 역할을 하리라 기대 한다. 또한 다른 질병에도 컴퓨터보조진단의 임상적용 실험이 이루어진다면 실제로 상용화 되고 있는 유방영상과 같이 여러 분야에 임상적 증례를 실험하는 계기가 될 것이며, 방사선 영상의학에서 다양한 분야에 질환별 적용도 가능하리라 사료된다.

참고문헌

- [1] 정병수, "PCA와 k-Nearest Neighbor 방법을 이용한 모델 기반형 물체 인식", 전남대학교 대학원, pp.2-3, 2008.
- [2] R. Chellappa, Charles. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Face" A Survey.;Proc. IEEE Vol.83 No 5, pp 704-740, May, 1995.
- [3] 이승철, "동적 링크 구조상에서의 얼굴 인식에 관한 연구", 연세대학교 대학원, pp.1-4, 1999.
- [4] W R .Hendee, et. al, "The information perspective of radiology", AJR pp. 1387-1389, 1992.
- [5] 김진철, "흉부영상에서 컴퓨터지원진단을 위한 간질성 폐질환의 특징분석과 검출", 전남대학교 대학원, pp. 4-7. 2003.
- [6] 이정원, "흉부 CT에서의 폐 분할 알고리즘과 고립 특성을 이용한 폐 결절 자동검출 알고리즘 개발", 서울대학교 대학원, pp. 12-13. 2002.
- [7] 류지원, "단순흉부영상의 유전자 알고리즘을 이용

한 폐 결절 자동 추출”, 전남대학교 대학원, pp. 5-6, 2003.

- [8] K. Doi, H. MacMahon, M. L. Giger, and K. R. hoffman, "Computer Aided Diagnosis in Medical imaging", Elsevier Science B. V, pp. 57-59, 1999.
- [9] R. T. Greenlee, T. Nurray, S. Bolden, P. A. Wingo, "Cancer Statistics 2000", CA Cancer J clinic, 50 pp. 7-33, 2000.
- [10] C. V. Zwirowich, S. Veda, R. Miller, "Solitary Pulmonary Nodule High-Resolution CT and Radiological Pathologic Correlation", Radiology, 179 pp. 469-476, 1991.
- [11] D. C. Chan, W. R. Wu, "Image contrast enhancement based on a histogram transformation of local standard deviation", IEEE transactions on Medical imaging, (68);vol.17. no.4 pp. 518-531, 1998.
- [12] 임정기 등, "흉부방사선과학", 일조각, pp. 202-283. 2006.
- [13] 김정희, "A study on real-time Face Recognition using PCA", 동신대학교 대학원, pp. 7-32, 2006.

저자소개



박형후(Hyung-Hu Park)

2009년 부산가톨릭대학교
방사선학과 졸업(이학석사)
2010년 고신대학교 보건과학과
박사과정

1997 ~ 현재 봉생병원 영상의학과 재직중
※관심분야: 의료영상신호처리, CAD.



옥치상(Chi-Sang Ok)

연세대학교 대학원 보건학과
(보건학박사)
현재 고신대학교 보건환경학부
교수

※관심분야: 보건환경, 폐기물관리 처리기술, 수질
관리 및 처리, 자연건강학, 대체의학



강세식(Se-Sik Kang)

1991년 원광대학교 농화학과
(농학박사)
1997년 ~ 현재 부산가톨릭대학교
방사선학과 교수

※관심분야: 방사선 치료, 방사선 기기



고성진(Seong-Jin Ko)

1997년 경성대학교 생물학과
(이학박사)
1992년 ~ 현재 부산가톨릭대학교
방사선학과 교수

※관심분야: 방사선 생물, 방사선 계측



최석윤(Seok-Yoon Choi)

2009년 고려대학교 의공학협동
(공학박사수료)
2010년 ~ 현재 부산가톨릭대학교
방사선학과 조교수

※관심분야: 컴퓨터 비전, Large scale problem, Partial
Differential Equations(PDE)