

흉부 X-ray 영상에서 유전자 알고리즘을 이용한 폐 결절 자동 추출

류지연⁰, 이경일, 장정란, 오명진, 이배호
전남대학교 컴퓨터공학과
multiji@hanmail.net⁰

Automated Detection of Pulmonary Nodules in Chest X-ray Radiography Using Genetic Algorithm

Ji-Youn Ryu⁰, Kyung-Il Lee, Jeong-Ran Jang, Myoung-Jin Oh, Bae-Ho Lee
Dept. of Computer Engineering, Chonnam National University

요 약

컴퓨터지원진단(Computer Aided Diagnosis: CAD) 시스템은 방사선 의사들이 흉부 X-ray 영상에서 결절을 탐지하는데 있어 실제적으로 발생할 수 있는 오진율을 줄이고, 폐 결절이 존재하는 폐야에서 결절의 존재 유무를 판단하여 검출을 표시함으로써 진단율을 개선시킬 수 있도록 하였다. 본 논문은 흉부 X-ray 영상에서의 폐 결절을 추출하는데 유전자 알고리즘(Genetic Algorithm)을 이용한 템플릿 매칭(Template Matching) 방법을 제안한다. 제안한 방법은 흉부 X-ray 영상에 존재하는 결절과 레퍼런스 이미지를 매칭시켜 적합도를 계산한 후, 그 값을 통하여 수치가 낮은 개체를 선택하여 높은 개체와 교차시킨다. 그리고 레퍼런스 이미지는 결절이 존재하는 환자 X-ray 영상에서 샘플 노들을 추출한 후 가우시안 분포를 갖는 512개의 레퍼런스 이미지를 생성하였다. 본 논문에서 사용된 영상은 결절 50개, 비결절 30개와 흉부 X-ray 영상에서 육안으로 판별이 가능한 결절 영상을 20개를 포함하여 총 100개 영상을 사용하였다. 실험 결과 83%의 결절을 자동 추출 하였으며, 가장 적절한 레퍼런스 이미지를 발견하고 이를 흉부영상에 매칭시켜 정확한 결절의 위치를 확인하였다.

1. 서 론

현재 우리나라에서 폐암 사망률은 암 사망률 중 1위를 차지하고 있으며, 그 비율은 점차 증가하고 있다. 흉부 X-ray 영상은 폐암 진단을 위한 집단검진검사에서 널리 사용하는 방법으로, X-ray 영상에서 의사의 폐암 발견 오진율은 30%정도이다.[2] 폐암의 조기발견은 폐암환자의 생존율을 높일 수 있는 효과적인 방법으로, 의사의 오진율이 개선된다면 폐암 환자의 생존율이 높아질 것으로 기대된다.

본 논문은 흉부 X-ray 영상에서 폐 결절이 존재하는 폐야에서 결절이 존재하는 부분을 컴퓨터가 자동으로 탐지해줌으로써 의사가 폐 결절을 간과하지 않도록 하여 진단율을 개선시킬 수 있도록 하였다. 본 논문에서는 흉부 X-ray 영상에 유전자 알고리즘(Genetic Algorithm)을 이용한 템플릿 매칭(Template Matching)을 적용하여 자동으로 결절을 추출하는 방법을 제안하고자 한다.

먼저, 결절 검출을 좀더 용이하게 하기 위해 흉부 X-ray 영상에서 관심대상이 되는 폐 영역만을 따로 분리하였다. 폐 영역의 자동 분리(automatic segmentation)를 구현하기 위하여 임계치(thresholding)와 라플라시안 필터링, 모폴로지(morphological) 기법 등이 사용된다. 다음 단계로, 분리된 폐 영역에 유전자 알고리즘(Genetic Algorithm)을 이용한 템플릿 매칭(Template Matching)을 적용하여 결절로 의심되는 부분을 추출한다. 여기서 템플릿은

가우시안 분포를 갖는 512개의 레퍼런스 이미지를 생성하여 사용하였다. 마지막으로, 검출된 결절 후보들 중에서 False-Positive 수를 감소시키는 방법을 적용하였다. False-Positive는 결절로 잘못 검출된 결절 후보로, 이를 줄여주기 위하여 히스토그램과 명도대비 등을 사용하였다.[1]

본 논문에서 사용된 흉부 X-ray 영상은 일본방사선학회 의사들의 진단에 의해 만들어진 표준 데이터베이스 영상을 사용하였고, 각각의 영상은 1024*1024 크기를 가지며, 픽셀당 256개의 명암값을 가진다.

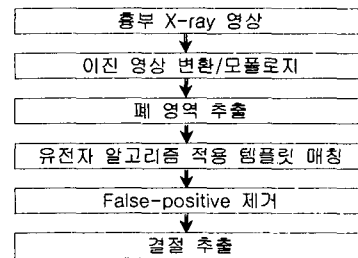


그림 1. 폐 결절 자동 추출 과정

2. 폐 영역 분리

2.1 임계치(Threshold) 계산

본 논문에서는 흉부 X-ray 영상의 각 행의 Peak 값의 평균값과 전체 영상에서 가장 낮은 값 사이의 2/3지점을 임계치값으로 지정하고 정확한 폐 영역을 분리(Segmentation)하기 위해 사용하였다.[5]

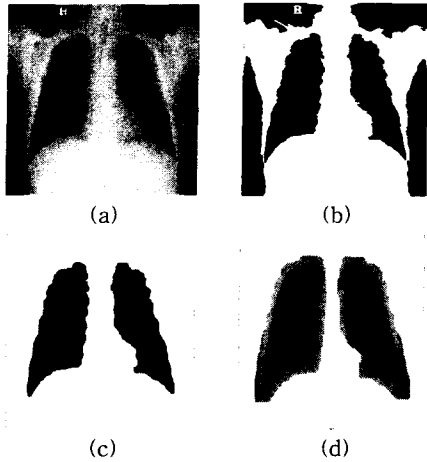


그림 2. (a) 원 영상, (b) 모폴로지 형태학(Morphology)기법 적용 결과, (c) 에지 추출 결과, (d) 폐 영역 분리 결과

2.2 이진영상 변환 및 폐 영역 추출

그림 2.(a)는 원영상이고 그림 2.(b)는 임계치를 이용해 폐영역을 이진화시킨 후 모폴로지 기법을 적용한 것이다. 그림 2.(c)는 이진영상으로부터 얻어진 에지 외곽선 정보이다. 폐 영역의 정확도를 측정하기 위한 방법으로 라플라시안 필터를 사용하였다. 라플라시안 필터 적용 후 영상은 원 픽셀로 폐 경계를 구성하게 된다. 그림 2.(d)는 앞에서 구한 에지 정보를 이용하여 폐 영역을 분리한 결과이다. 그림 2.(d)를 이용하여 결절 추출하는 과정을 진행할 것이다.[5]

3. 폐 결절 추출

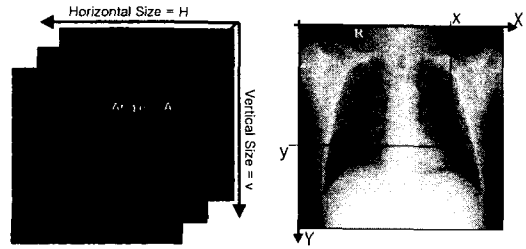
3.1 유전자 알고리즘 적용 템플릿 매칭

결절을 자동 추출하기 위한 방법으로 유전자 알고리즘을 이용한 템플릿 매칭기술을 사용하였다. 유전자 알고리즘은 염색체(Chromosome)를 선택(selection), 교차(crossover), 변이(mutation)로 구성된 세대(generation)을 반복함으로써 최적의 해에 도달하는 알고리즘이다. 본 논문에서는 레퍼런스 이미지를 템플릿으로 하여 흉부 X-ray에 유전자 알고리즘을 사용함으로써 최적의 해인 결절을 찾는다.[4]

3.1.1 염색체(Chromosome)

염색체는 생물체에서 가장 중요한 정보를 담고 있는 부분으로써, 유전자 알고리즘 구현을 위해 우리는 염색

체를 Binary digit로 나타내었다. 이 Binary digit는 흉부 X-ray에서 결절의 위치(x,y)와 적절한 레퍼런스 이미지의 선택을 위한 정보(H, V, A)를 가지고 있다. 우리는 흉부 X-ray 영상에서 결절로 검출된 형태를 바탕으로 가우시안 분포를 갖는 512개의 레퍼런스 이미지를 생성하였다. 염색체의 각 필드 x, y, H, V, A는 10bit, 10bit, 4bit, 4bit, 3bit로 표현한다.[4]



(a) 레퍼런스 이미지 (b) 흉부 X-ray 영상
그림 3. 유전자 알고리즘을 이용한 템플릿 매칭

3.1.2 적합도(Fitness)

적합도는 흉부 X-ray 영상에 추출된 샘플 레퍼런스 이미지와 환자 흉부 X-ray 영상과의 픽셀값의 차이를 나타내는 정도이다. 즉, 선택된 레퍼런스 이미지에서 배경부분을 제외한 영역을 선택함으로써 적합도 계산이 이루어지게 된다.[3,4]

두 이미지의 선택된 영역의 평균값 계산은 다음과 같다.

$$m_a = \frac{1}{n} \sum_{i=0}^{n-1} a_i, \quad m_b = \frac{1}{n} \sum_{i=0}^{n-1} b_i, \quad (1)$$

m_a, m_b : pixel average value
 a : observed image, b : reference image

그리고 cross-correlation을 이용한 적합도 계산식은 다음과 같다.

$$Similarity_{a, b} = \frac{\sum_{i=0}^{n-1} (a_i - m_a) (a_i - m_b)}{\sqrt{\sum_{i=0}^{n-1} (a_i - m_a)^2} \sqrt{\sum_{i=0}^{n-1} (b_i - m_b)^2}} \quad (2)$$

3.1.3 GA(Genetic Algorithm) 처리과정

GA 처리과정은 초기 모집단 생성, 적합도(fitness) 계산, 선택과 교차, 변이로 이루어져있다. 먼저, 120개 염색체의 각 bit값을 랜덤하게 생성함으로써 초기 모집단을 구성한 후, 앞서 소개한 적합도 계산식을 이용하여 각 개체(염색체)의 적합도를 계산한다. 적합도가 높은 절반의 개체들을 선택하고, 남은 절반의 개체들은 적합도가 높은 개체들과 일점교차(one point crossover)를 시켜줌으로써 적합도를 높여준다. 각 개체들의 bit마다 난수를 발생시켜 발생된 난수가 변이 발생 확률(P_m)보다 작으면, 그 개체는 변이를 수행한다. 확률(P_m)은 세대가 거듭

될수록 0.25에서 0.01로 줄어들게 된다. 적합도를 높이기 위해 가장 높은 적합도를 갖는 10%이내의 개체들은 변이에서 제외시킨다. 적합도 계산, 선택, 교차, 변이가 한 세대를 이루며, 위 과정이 끝나면 세대를 반복해 줌으로써 적합도를 높여준다.[6] 세대 수는 200으로 하였고, GA 처리과정이 끝난 후 선택된 위치가 결절이 있을 것으로 의심되는 결절 후보들이 된다.

3.2 False-positive 제거

False-positive는 결절이라고 오인되어 검출된 결절 후보를 뜻한다. False-positive는 폐 경계부분, 혈관, 뼈와 같은 신체 조직에 많이 나타난다. 우리는 선택적으로 False-positive를 제거하는 알고리즘으로, 첫 번째는 결절 후보들의 히스토그램을 해석하여 낮은 픽셀 값을 갖는 영역을 제거하고, 두 번째는 결절 후보들의 벡터 이미지 히스토그램 해석에 기반하여 뼈를 제거하였다. 마지막 단계는 결절 후보의 주의를 둘러싼 주변 픽셀값의 차이를 계산하여 그 대비가 높은 값을 혈관으로 검출하고 이 부분을 제거하였다.[1]

4. 실험 결과

본 실험에서는 앞에서 제안한 방법을 통해 잘못 선택된 False-positive 수와 결절을 자동 추출하기 위해 히스토그램 균등화, 이진영상변환, 모폴로지, 폐 영역 분리 및 GA, 템플릿 매칭등의 작업을 수행하였다. 기존 논문에서 또한 어떻게 False-positive를 줄 일수 있는지가 문제시 되고 있다. 본 실험에서는 히스토그램과 명도 대비를 통해 False-positive로 의심되는 수를 줄이고, 이를 통해 좀더 정확한 결절의 수를 자동 추출 하였다.

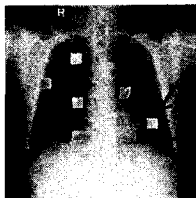


그림 5. 결절로 추출된 영상

그림 5는 결절이 추출된 X-ray 영상을 보여준다. 실험 결과 X-ray 영상의 오른쪽 아랫부분에 화살표로 표시한 부분이 실제 결절이 위치한 곳이고, 나머지 5개의 사각형 모양은 비결절로 선택 되었다.

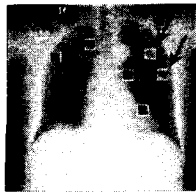


그림 6. 결절로 추출된 영상

그림 6은 오른쪽 윗부분이 화살표로 표시한 두 부분이 실제 결절이 위치한 곳이며, 나머지 4개의 사각형 모양은 비결절로 나타났다.(그림 5, 6)는 방사선 의사의 확인을 통해 화살표로 나타낸 곳이 결절이 존재하는 것으로 나타났으며, 자동 추출을 통하여 결절은 좀더 효율적으로 찾았다. 그리고 False-positive로 나타난 노들을 육안으로 구별하는데 더 효과적 있었음을 확인할 수 가 있었다.

5. 결론

본 시스템에서는 X-ray 영상에 팽창(Dilation)연산과 침식(Erosion)연산을 수행하여 결절 추출이 어려운 폐 경계부분을 좀더 확장함으로써 False-positive 수를 감소시켰고, 이를 통해 정확한 관독이 어려운 폐 경계부분의 결절을 효과적으로 탐지 할 수 있었다. 아울러 유전자 알고리즘을 이용한 템플릿 매칭으로 흉부 X-ray 영상에서 찾고자 하는 폐 결절을 좀 더 효율적으로 찾을 수 있었다.

본 논문에서는 결절 50개, 비결절 30개와 흉부 X-ray 영상에서 육안으로 관별이 가능한 결절 X-ray 영상 20개를 포함하여 총 100개 영상을 사용하였다. 본 시스템에서 실험한 결과 83%의 결절을 효과적으로 자동 추출 할 수 있었다.

참고문헌

- [1] T. Hara, H. Fujita and J. Xu, "Development of Automated Detection System for Lung Nodules in Chest Radiograms," in Proc. *IASTED Intelligent Information System(IEEE computer Society)*, pp. 71-74, 1997.
- [2] K. Doi, H. MacMahon, M. L. Giger and K. R. Hoffmann, "Computer-Aided Diagnosis in Medical Imaging," Elsevier Science B.V., pp. 57-59, 1999.
- [3] Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Nodule Detection on Chest Helical CT Scans by Using A Genetic Algorithm," in Proc. *IASTED Intelligent Information System (IEEE computer Society)*, pp. 67-70, 1997.
- [4] Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," *IEEE Trans. Med. Image.*, vol. 20, no.7, pp. 595-599, 2001.
- [5] F. Hoare and G. D. Jager, "Model based Segmentation of Medical X-ray Images," Department of Electrical Engineering, University of Cape Town, South Africa.
- [6] Y. Lee, T. Hara, and H. Fujita, "Evaluation of GA template-matching method by simulation using chest helical X-ray CT images," *Med. Imag. Inform. Sci.*, vol. 17, no. 3, pp. 118-129, 2000.