

온-오프 타입 DNA 칩의 자동판독 시스템

유문호*, 김종대**

*전북대학교 생체정보공학부
*전북대학교 실버공학연구소
***한림대학교 정보통신공학부
e-mail : mhryu@chonbuk.ac.kr

Development of Automatic Reading System for On-Off Type DNA Chip

Munho Ryu*, Jong-Dae Kim**

*Div. of Bionics and Bioinformatics, Chonbuk National University
*Bioengineering Research Center for the Aged, Chonbuk National University
**Div. of Information Engineering & Telecommunications, Hallym University

요 약

본 연구는 진단용 DNA 칩의 자동판독 시스템을 제안하는 것을 목적으로 한다. 일반적인 자동판독 시스템의 사양을 정의하고 그 구현방법을 제안하였다. 응용 예로서 자궁경부암 진단용 DNA 칩을 대상으로 GenePix 스캐너 프로그램 환경에 적용하였다. 영상 획득은 GenePix 의 라이브러리를 사용하여 HTML 언어로 구현하였고, 영상의 판단과 보고서 생성은 Microsoft Visual C++ 6.0 를 사용하여 COM 형태로 구현하였다. 결과 보고서는 한글 2002 문서에 환자 정보와 결과 정보 등에 해당하는 곳에 미리 정의된 표지문자열들을 삽입하여 템플릿을 만들었다. 판독 시스템은 템플릿을 읽어 들여 처리 결과의 내용으로 표지문자열들을 치환하여 보고서를 생성하였다. 제안한 시스템을 통해서 스캐닝을 통한 영상 획득, 영상의 판독, 결과 보고서 생성으로 구성된 전체 판독과정이 사용자의 개입 없이 자동으로 처리될 수 있었다. 본 시스템은 기존에 수작업을 자동화하여 판독 시간을 단축하고 판독 기준을 정량화하여 진단용 DNA 칩이 대량검사 활용되는데 공헌할 것으로 기대된다.

1. 서론

DNA 칩을 이용한 진단은 DNA 샘플을 칩과 반응시키는 생화학 처리 과정과 칩을 스캐닝 하여 진단하는 판독 과정으로 구성되는데[1,2], 판독 과정은 DNA 칩을 스캐닝 하여 영상을 획득하는 과정, 영상으로부터 특정 유전형질의 존재유무를 판단하는 과정, 판단결과에 대한 보고서를 작성하는 과정으로 구성된다. 대량 검사를 가속화하기 위해서 생화학 처리과정의 단순화를 위한 랩온어칩(Lab on a chip)에 대한 연구가 활발히 진행되고 있는 한편, 판독 과정의 자동화 또한 절실히 요구된다[3].

DNA 칩 전용 스캐너의 소프트웨어들이 영상 획득과 함께 판단에 필요한 기능을 제공하는데는 하지만 탐색자의 위치를 찾고 교잡반응 여부를 판단하는 등의 판독 과정은 특정 칩의 설계와 밀접하게 관계되기 때

문에 그대로 사용할 수 없다. 또한 환자정보와 판독 영상을 포함한 판독 결과 등을 보고서에 포함 시켜야 하기 때문에 사용자의 많은 수작업을 필요로 하고, 보고서의 형태 등에도 한계가 있다[4,5]. 특히 보고서의 필수 정보를 제외한 나머지 서식들은 사용자가 임의로 구성하고 필요에 따라 추가, 변경할 수 있도록 보편적인 상용 문서 편집기로 만드는 것은 칩 보급의 중요한 요소가 된다. 이상과 같은 이유로 판독 과정을 자동화 하기 위하여 칩 판독 전용프로그램이 필요하다.

보고서 작성을 포함한 칩 판독 과정을 자동화하는 것은 대량 검사를 위하여 필수적이나, 자동 판독에 대한 많은 연구들은 판독과정 전체를 자동화하지는 못하고 있다[6,7].

본 연구는 앞서 설명한 판독과정을 자동화하는 시스

템을 제안하고, 실질적인 응용 예로 자궁경부암 진단용 DNA 칩인 HPV DNAChip ((주)바이오메드랩)을 대상으로 보편적으로 많이 사용되는 스캐너 프로그램 중의 하나인 GenePix (Axon Instruments, Union City, CA) 환경에 자동진단 시스템을 구현한다. 제안한 시스템은 칩을 스캐너에 장착한 후부터 보고서가 생성될 때까지의 전 과정을 자동화하고, 보고서 양식으로는 가장 일반적인 한글 워드프로세서를 사용한다.

2. 자동판독 시스템 요구사항 및 설계

2.1 시스템 요구사항 및 설계방법

자동판독 시스템의 요구 사양은 다음과 같다.

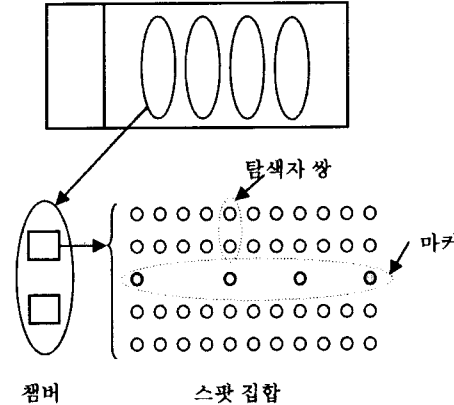
- 1) DNA 칩을 스캐닝 하여 영상을 획득하는 영상획득 과정, 영상으로부터 특정 유전형질의 존재유무를 판단하는 영상처리과정, 판단결과에 대한 보고서를 작성하는 보고서 생성과정을 사용자의 개입 없이 처리되도록 자동화한다.
- 2) 영상처리과정에서 탐색자들의 위치를 찾고 탐색자 교잡반응 되었는지의 여부를 판정하기 위한 진단용 DNA 칩의 특성에 적합한 척도를 제공한다.
- 3) 보고서 생성에 필요한 환자정보를 입력하는 기능을 제공한다.
- 4) 결과 보고서에는 칩의 이미지, 유전형질의 존재유무 판독 결과, 환자정보 등을 포함한다.
- 5) 결과 보고서의 양식은 사용자가 서식의 구성과 내용의 수정이 용이하고, 구하기 쉬운 상용 문서편집기를 사용한다.

첫 번째 사항은 영상 획득을 위한 스캐너 프로그램의 조작에서부터, 보고서 생성까지의 전 과정을 일련의 배치작업으로 단순화시켜 구현한다. 특히, 영상 획득을 위한 스캐너의 파라미터와 이미지 영역 설정 등이 DNA 칩의 조건과 형식에 맞추어 자동으로 이루어지도록 해야 한다. 두 번째 사항은 획득된 영상을 읽어 들여, 기준 위치를 결정하고 결정된 위치를 기준으로 정해진 위치의 탐색자 스팟들을 자동으로 판독하는 알고리즘을 적용하여 구현한다. 이때 진단용 DNA 칩의 종류에 따라 진단 성능을 최대화할 특정한 판단 척도가 필요하다. 세 번째 사항은 진단과는 직접적인 관계는 없지만, 결과 보고서 생성에 추가될 환자의 정보를 적절한 시기에 입력할 수 있도록 제공한다. 마지막으로 네 번째와 다섯 번째 사항들은, 일반적으로 사용되는 상용 문서편집기 형식으로 보고서 템플릿을 준비하여 이를 기반으로 보고서를 생성하도록 구현할 수 있다. 템플릿의 규칙을 단순화하여 사용자가 쉽게 보고서 서식이나 구성의 변경을 용이하게 할 수 있도록 한다.

2.2 자궁경부암 진단용 HPV DNAChip

HPV DNAChip ((주)바이오메드랩)은 인유두종 바이러스 (human papillomavirus, HPV)의 감염을 진단하는 DNA 칩으로서 그림 1 과 같은 구조를 갖는다[8,9]. 칩 하나에 4 개의 챔버가 있어 각각 한 명의 환자 샘플을 검사하고, 진단 신뢰도를 높이기 위해 각 챔버에

는 동일한 스팟 집합이 2 개씩 있다. 스팟 집합에는 4 개의 양성 기준 마커와 22 쌍의 HPV 유형 특이적 올리고뉴클레오티드 탐색자가 정착되어 있다. 인간의 베타 글로빈에 해당하는 양성 기준 마커는 교잡반응을 확인하고 스팟 집합의 기준 좌표를 설정하는데 사용된다. 표적 DNA 는 임상 샘플에서 추출된 후 PCR 증폭되어 칩에 교잡반응된다. PCR 증폭 과정에서 Cy5 형광물질이 불규칙하게 붙어, 스캐닝 후에 교잡반응의 위치를 시각화한다. 마커 위치를 찾은 다음, 이 위치를 기준으로 각 탐색자 스팟의 존재 유무를 판단하여 해당 HPV 유형의 존재를 온-오프로 판단한다.



(그림 1) HPV DNAChip 의 구조

2.3 판단용 척도

한편, 저자들의 선행 연구에서 템플릿 정합법이 이와 같은 온-오프 형태의 DNA 칩을 판독하는데 우수한 성능을 내는 것을 보였다 [10-12]. 스팟의 모양과 배열에 관한 사전정보를 융합하여 마커 찾기와 탐색자 구분에 성공적으로 적용하였다. 그리고, 템플릿 영역 내의 밝기를 밝은 화소와 어두운 화소로 양분하여 스팟 영역 내의 밝은 화소 개수와 배경 영역내의 밝은 화소 개수의 차이를 취하는 것으로 정의되는 비선형 정합도가 정규상관도와 비교하여 우수하였다.

$$Counting(i, j) = \#\{(k, l) | I(i+k, j+l) > Th, (k, l) \in O\} - \#\{(k, l) | I(i+k, j+l) > Th, (k, l) \in B\} \quad (1)$$

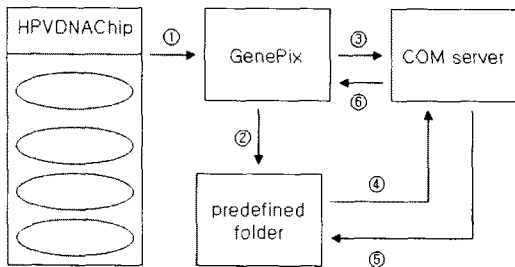
여기서, $I(i, j)$ 와 $I(i, j)$ 는 화소 위치 (i, j) 에서의 템플릿과 영상의 강도이다. $\#\{ \cdot \}$, O , B , Th 는 집합 원소 개수, 스팟 영역, 배경, 문턱 값이다.

3. 자동판독 시스템 구현

실질적인 응용 예로 자궁경부암 진단용 DNA 칩인 HPV DNAChip (그림 1)을 대상으로 보편적으로 많이 사용되는 스캐너 프로그램 중의 하나인 GenePix (Axon Instruments, Union City, CA) 환경에서 자동판독 시스템을 구현하였다. GenePix 의 사용자 인터페이스는 HTML 언어로 작성되어 개방성이 우수하여, 환자 정보를 입력하고 자동판독을 시작하는 메뉴를 쉽게

구현할 수 있었다. 영상획득은 GenePix 에서 제공하는 라이브러리를 사용하였고, 영상의 판단과 보고서 생성은 자동진단 COM (component object model)을 별도로 작성하여 GenePix 에서 호출하였다. 자동진단 COM 은 Microsoft Visual C++ 6.0 를 사용하여, 영상을 판단하여 결과보고서를 생성하도록 하였다. 결과 보고서는 한글 2002 문서로 템플릿을 준비하고, 템플릿 내에 필요한 정보마다 미리 정의된 표지문자열을 삽입하였다. 자동진단 COM 은 템플릿 문서를 읽어서 표지문자열 들을 해당 정보로 치환하여 생성되도록 하였다.

그림 2 는 구현된 자동판독 시스템의 구성요소들 간의 관계도를 나타낸다. 동작 순서는 다음과 같다. ① 스캐너에 DNA 칩을 넣고, GenePix 에서 환자 정보를 입력한 다음 보고서 생성 버튼을 누른다. ② 이때 GenePix 는 DNA 칩 영역 중 4 개의 챔버에서 8 개의 스팟 집합을 충분히 포함하도록 미리 정해진 슬라이드 영역을 스캐닝 하여 영상 파일을 저장한다. 영상 파일은 미리 정의된 위치에 저장되며, 유일한 이름을 갖도록 환자 이름과 환자 정보 중에 환자마다 유일한 키값 (예를 들면 등록번호)을 사용하였다. ③ 그 다음, 자동진단 COM 개체를 호출하여 판독을 시작한다. 이때 영상 파일의 이름을 넘겨준다. ④ 자동진단 COM 개체는 저장된 한 챔버의 두 스팟 집합에 해당하는 영상을 두 개 읽어서 마커 위치를 찾고 탐색자를 분류한다. ⑤ 환자정보, 탐색자 분류 결과, 마커를 기준으로 하여 잘라낸 두 개의 영상 등을 보고서에 전술한 치환방법을 사용하여 삽입한다. 4 개의 챔버에 대해서 ④, ⑤ 과정을 반복한다. ⑥ 결과를 GenePix 에 알려 표시한다.



(그림 2) 구성 요소 관계도

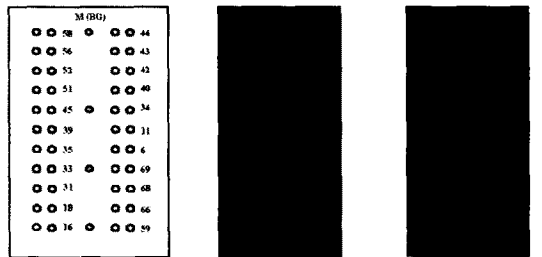
그림 3 에 자동진단을 시작하는 GenePix 화면의 일부를 표시하였다. 테이블의 좌측 열부터 검사 접수번호, 환자정보 (이름, 주민등록번호, 성별, 나이), 기타 정보의 순으로 입력한다. 4 개의 행은 각각 DNA 칩의 챔버에 해당한다. 필요한 정보를 입력하고 '보고서만들기' 버튼을 누르면 앞서 설명한 자동판독을 실행하고 결과를 표시한다.

Ch	접수번호	이름	주민등록번호	성별	나이	의뢰기관명	의뢰기관번호	Chart No	검체채취일	검사
1										
2										
3										
4										

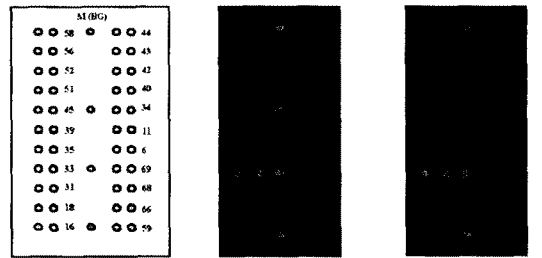
보고서만들기 자취기

(그림 3) 자동진단 시작 화면 일부

그림 4 와 5 에 각각 보고서 템플릿과 실제로 생성된 결과 보고서의 예를 일부 보였다. 그림 4 에서 좌측에 HPV DNA 칩의 형식을 참고용으로 표시하고, 우측에는 결과 영상으로 대체될 비트맵 영상을 위치한다. 하단에는 판단된 유전 형질의 존재 유무를 표시하는 표가 존재한다. 비트맵 이미지의 시작 영역과 유전 형질 판단 부분에는 각각 표지문자열이 존재하여 보고서 생성시 대체된다. 그림 5 에서 우측의 2 개의 영상은 템플릿에 있는 비트맵 영상이 결과 영상으로 대체된 것임을 알 수 있다. 좌측의 형식을 참고해서 알 수 있듯이 HPV 33 번에 해당하는 스팟으로부터 해당 바이러스의 존재가 확인된다. 그리고 하단의 테이블에서 보듯이 HPV 33 번 만이 존재하는 것으로 체크되는 것을 확인할 수 있다.



(그림 4) 보고서 템플릿 일부



(그림 5) 자동 생성된 보고서 예

4. 결론 및 토의

본 연구는 진단용 DNA 칩의 자동판독 시스템을 제안하고, 응용 예로서 자가경부암 진단용 DNA 칩을 대상으로 GenePix 스캐너 프로그램 환경에 적용하였다. 본 시스템을 통해서 스캐닝을 통한 영상획득, 영상의 판독, 결과 보고서 생성으로 구성된 전체 판독과정이

사용자의 개입 없이 자동으로 처리될 수 있었다.

마커 찾기에 실패하여 판독이 중단되는 사례가 있었다. 그 이유는 두 가지 경우로 나누어 진다. 첫째, 베타글로빈 마커에도 반응하지 않은 음성 컨트롤 샘플을 검사에 포함한 경우로서, 이는 검사 환경의 오염을 확인하기 위한 것으로서 판독자에 의해 무시되기 때문에 문제가 되지 않는다. 둘째, 반응된 DNA 칩의 영상이 열악하여 자동으로 마커를 찾을 수 없는 경우도 있었다. 이를 고려하여 마커 찾기에 실패한 경우 GenePix 에서 판독자로 하여금 수동으로 마커를 찾고 다음 작업을 계속하는 기능을 제공하였다. 하지만, 근본적으로는 DNA 칩의 생산 품질이나 칩 사용자의 숙련도에서 기인하는 문제이기 때문에, 원인에 따라 해당 측에 피드백을 보내어 개선하는 과정이 필요했다.

본 시스템은 기존에 수작업을 자동화하여 판독 시간을 단축하고 판독 기준을 정량화하여 진단용 DNA 칩이 대량검사 활용되는데 공헌할 것으로 기대된다. 비록, 특정한 DNA 칩과 스캐너 프로그램을 대상으로 구현하여 검증하였지만, 본 연구에서 제안하는 시스템은 일반적인 DNA 칩의 자동판독 시스템에도 충분히 적용가능 할 것이라 예상된다.

후기

이 논문은 2005 년도 전북대학교 지원 연구비에 의하여 연구되었음.

참고문헌

- [1] Robert S. H. Istepanian, "Microarray image processing: current status and future directions", IEEE Trans. Nanobioscience 2(4), 173-175 (2003)
- [2] Yihua Zhou and Jingdong Liu, "AVA: visual analysis of gene expression microarray data", Bioinformatics. 19(2), 293-294 (2003)
- [3] Lloyd J Ridley, "Guide to the radiology report", Australasian Radiology 46(4), 366-369 (2002)
- [4] *GenePix 4000A User's Guide*, Axon Instruments, Inc., 1999
- [5] *Operators' Manual, GSI Lumonics*, 1999
- [6] M.R. Fielden, R. G. Halgren, E. Dere, T.R. Zacharewski, "GP3: GenePix post-processing program for automated analysis of raw microarray data", Bioinformatics. 18(5) 771-773 (2002)
- [7] Daniel Bozinov, "Autonomous system for Web-based microarray image analysis", IEEE Trans. Nanobioscience 2(4), 215-220 (2003)
- [8] Chan Joo Kim, Jeongmi Kim Jeong, Misun Park, Tae Shin Park, Tae Chul Park, Sung Eun Namkoong, and Jong Sup Park, "HPV oligonucleotide microarray-based detection of HPV genotypes in cervical neoplastic lesions," Gynecol. Oncol. 89, 210-217 (2003).
- [9] Byeong-Sun Choi, Okjin Kim, Mi Sun Park, Ki Soo Kim, Jeongmi Kim Jeong, and Joo-Shil Lee, "Genital Human Papillomavirus Genotyping by HPV Oligonucleotide Microarray in Korea Commercial Sex Workers", J. Med. Virology 71, 440-445 (2003).

- [10] J. D. Kim, S. K. Kim, J. S. Cho, and J. Kim, "Knowledge-based image processing for on-off type DNA microarray", Proc. SPIE 4623, 38-46 (2002)
- [11] Munho Ryu, Jong Dae Kim, and Byoung Goo Min, "Robust template-matching measurements for variations of signal amplitude", Opt. Eng. 43(2), 482-488 (2004)
- [12] Munho Ryu, Jong Dae Kim, Byoung Goo Min, Myung-Geol Pang, and Jongwon Kim, "Nonlinear matching measure for the analysis of on-off type DNA microarray images", J. Biomed. Opt. 9(3), 482-438 (2004)