

## 의료이미지 데이터의 동적 분석을 위한 패턴 정형화 기술

고광만\*

### 요약

이 논문에서는 비정형 의료이미지 정형화 및 패턴 추출을 위해, 의료기기 또는 의료전문가로부터 생성 또는 감지되는 의료이미지 저장을 위한 데이터베이스 구축한다. 이러한 비정형 이미지의 특징을 정형화된 디지털 데이터로 변환한 후 정형화된 디지털 이미지 데이터로부터 의미있는 패턴 정보를 생성한다. 이러한 경험 기술 소개를 통해 많은 연구자들은 의료이미지 데이터베이스를 보다 쉽게 접근할 수 있고 다양한 분야에서 정형화된 의료이미지를 활용할 수 있다.

키워드 : 의료이미지, 정형화, 이미지 패턴 매칭

## Pattern Formalization Technique for Dynamic Analysis of the Medical Image Data

Kwang-man Ko\*

### Abstract

This paper suggested that medical image database construction technique that generated and recognized from variable medical device and professional medical experts for the formalization and pattern extraction from informal medical images. And then we transformed informal image characteristics to digital data, and generated the meaningful pattern matching informations. Through this experienced works, so many related researchers can easily access the medical images database and use this formalized image informations on the variable fields.

Keywords : Medical Images, Formalization, Image Pattern Matching

### 1. 서론

최근에는 의료정보 분야에서 의료기기, 의료전문가로부터 생성 또는 감지되는 비정형 의료이미지를 디지털 데이터로 추출하여 패턴 데이터로 변환한 후, 이를 의료정보 분석 플랫폼에 적

합하게 변환하여 진료, 처방, 예방 등에 유용한 정보를 생성하는 모델 구축에 관한 연구가 활성화되고 있다. 특히, CT, MRI 등의 의료기기에서 생성되는 비정형 이미지에 대해 윤곽분석, 이미지 향상(음->양), 이미지 잡음제거, 영상분할, 색처리(컬러->흑백)를 통해 의료이미지 분석에서 활용할 수 있는 정형화된 패턴 정보를 추출하는 것은 다양한 가공 및 분석을 위해 가장 기본적으로 선행되어야 하는 내용이다[1].

의료기기, IoT 기기, 의료전문가로부터 생성 또는 진단되는 대용량의 다양한 의료정보의 특징을 패턴화한 후 비교 및 분석할 수 있는 의료이미지 분석 인프라를 확보하고, 객관적 자료와 기술을 기반으로 의료기기에서 생성되는 수많은 데이터를 다양하고 정확하게 분석하여 이로부터 유용한 정보나 규칙을 추출하여 의료전문가들

※ Corresponding Author : Kwang-man Ko

Received : May 09, 2016

Revised : June 27, 2016

Accepted : June 30, 2016

\* Sangji University School of Computer and Information Engineering

Tel: +82-33-730-0486, Fax: +82-33-730-0480

email: [kkman@sangji.ac.kr](mailto:kkman@sangji.ac.kr)

▣ 이 논문은 2013년도 상지대학교 교내 연구비 지원에 의한 것임.

지원하는 것은 보다 신뢰성 높은 진단, 진료에 필요한 내용이며, 질병 의심자 또는 환자에게 불필요한 부정확하고 검증되지 않은 정보의 유통을 차단할 수 있는 방안이다. 또한 의료전문가, 질병 의심자, 환자에게 질병에 대한 진료, 진단을 위한 보다 신뢰성 높은 정보 제공과 더불어 오진단의 가능성을 줄여 양질의 의료서비스를 제공할 수 있다[2].

비정형 의료이미지(CT, MRI 등)는 의료전문가의 질병 진단 근거로 중요하게 활용하고 있지만, 이 과정에서 의료전문가는 이상징후 발견이나 판정에서 존재할 수 있는 오진단의 가능성을 최소화하기 위해 비정형 의료이미지에 포함된 다양한 데이터를 분석하는 패턴 모델링 및 분석 기술의 도입 시도는 의료전문가의 신뢰성과 적절한 의료판정을 통해 의료비를 절감하고 관련 산업시장을 활성화하는데 크게 기여할 수 있다. 비정형 의료 이미지에 대한 패턴 모델링을 통해 의료전문가의 진단 및 시술방법 결정시 데이터 분석기반의 객관적 자료를 근거로 한 합리적 진단을 위해, 디지털 의료이미지 특징을 추출하여 정형화하고 이를 저장할 수 있는 기술, 의료이미지를 통해 환자의 진단 및 치료계획에 필요한 데이터 모델링 기술, 정형 및 비정형 데이터의 통계적 분석 및 신뢰성 높은 정보 제공과 더불어 오진단의 가능성을 줄여 양질의 의료서비스 제공에 활용할 수 있다.

본 논문에서는 비정형 의료이미지 정형화 및 패턴 정보 추출을 위해, 의료기기 또는 의료전문가로부터 생성 또는 감지되는 의료이미지 저장을 위한 데이터베이스를 구축하며 이러한 비정형 이미지의 특징을 정형화된 디지털 데이터로 변환한 후 정형화된 디지털 이미지 데이터로부터 의미있는 패턴 정보를 생성한다[3,4].

본 논문의 구성은 제2장에서 현재 연구자를 위해 다양한 의료정보를 무료로 제공하고 있는 Physionet 데이터베이스에 이미지 분석에 활용한 OpenCV에 대해 소개한다. 제3장에서는 의료 이미지의 패턴 정형화를 위해 실험하였던 다양한 동작과 실험결과물에 대해 설명한다. 마지막으로 제4장에서 결론과 향후 연구방향에 대해 기술한다.

## 2. 영상 의료이미지 획득 및 데이터베이스 구축

### 2.1 Physionet

Physionet[5]은 생리학적인 신호에 대한 연구 자원을 지원하는 곳으로 생리학적인 신호가 기록된 컬렉션인 PhysioBank[6], 오픈소스 소프트웨어인 PhysioToolkit[7]을 지원한다. PhysioBank는 생체신호와 연관된 대규모의 데이터들을 생물학적인 연구 단체의 연구 목적을 위하여 디지털 기록으로 바꾸어 보관하고 있는 저장소로서 현재 건강한 사람들과 그리고 급성 심장사, 울혈성 심부전, 간질, 보행 장애, 불면증, 그리고 노화를 포함하는 주요 공중 보건 질환을 가지고 있는 다양한 조건의 환자들로부터 얻은 다중변수의 심폐, 신경, 그리고 다른 생물학적인 신호들에 대한 데이터베이스를 확보하여 무료로 다운로드할 수 있는 70개 이상의 데이터 보관하고 있다. PhysioBank ECG 데이터베이스는 현재까지 전세계적으로 MIT-BIH Arrhythmia Database대한 접근 및 호응이 가장 좋다.

### 2.2 OpenCV

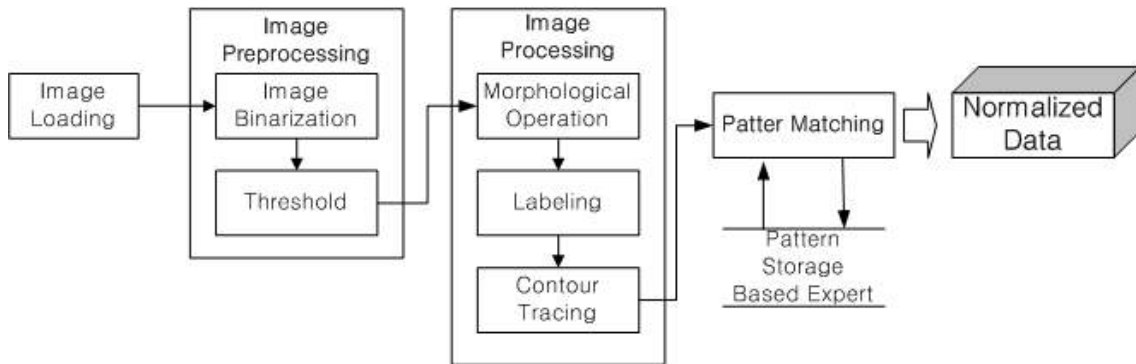
OpenCV[8]는 오픈 소스로 만들어지고 있는 컴퓨터 비전 라이브러리이다. OpenCV 라이브러리는 C와 C++ 언어를 이용하여 작성되었고, 리눅스, 윈도우, 맥 OS X 등에서 동작하며 주요목적 중 하나는 사용하기 쉬운 컴퓨터 비전 기반 구조를 제공함으로써 정교한 컴퓨터 비전 응용 프로그램을 쉽고 빠르게 만들 수 있도록 도와주는 것이다. OpenCV 라이브러리는 500개가 넘는 함수로 구성되어 있으며 이들 라이브러리 함수는 의료 영상, 보안, 사용자 인터페이스, 카메라 보정, 스테레오 비전, 로봇 등 다양한 컴퓨터 비전 분야에 활용되고 있다.

## 3. 의료이미지의 패턴 정형화 기술

### 3.1 전체 시스템 구성

OpenCV를 이용하여 의료이미지를 처리하는 과정은 (그림 1)과 같이 전처리 과정인 Image Binarization, Threshold 과정을 진행하고, 생성된

(그림 1) 비정형 이미지 변환 및 패턴 추출기



(Figure 1) Informal Image Translation and Pattern Extraction

데이터를 기반으로 Morphological Operation, Labeling, Contour Tracing 과 같은 작업을 거쳐 패턴 매칭하기 위한 데이터를 생성하여 데이터 베이스로 구축한다. 패턴매칭 과정에서 필요한

(그림 1) 비정형 이미지 변환 및 패턴 추출기

PSBE(Pattern Storage Based Expert)는 기존 영상 데이터를 기반으로 의료 전문가의 자문을 통해 생성한 Database로 정규화된 데이터를 생성하는데 사용된다.

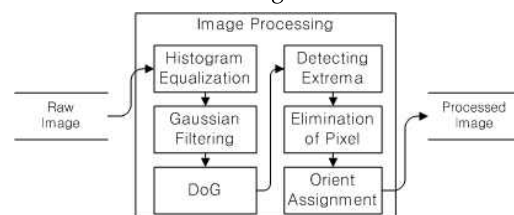
패턴 생성 추출을 위한 이미지 처리를 위해 Image Binarization 단계에서는 배경과 특정 부분의 분리를 위해 사진의 각 픽셀에 접근을 해서 픽셀 값이 임계값 보다 크면 흰색 임계값 보다 작으면 검은색으로 바꿔주는 작업을 진행하였다. Threshold 단계에서는 이미지의 일정 픽셀 값을 잘라내 특정한 영상 부분을 찾아내기 위해 사용한다. Morphological Operation 단계중 침식(erosion) 과정에서는 주변보다 밝고 작은 크기의 얼룩 잡음들을 제거하며, 팽창(dilation) 과정에서는 연결된 유사한 색상 또는 밝기 값을 가진 큰 영역의 구성요소를 찾는다. Labeling 단계에서는 이진화된 0부터 255(1-channel)의 값을 갖을 경우 인접한 영역끼리 그룹을 짓는 작업을 진행하며, Contour Tracing 단계에서는 Labeling 객체의 엣지를 검출하기 위해 엣지 Tracing이나 Boundary Flowing 알고리즘을 수행하며 OpenCV의 cvFindcontours() 함수를 이용하여 외곽선을 추출하였다. 마지막으로, 패턴매칭 단계에서는 이미지 처리에 의해 결과 이미지들에서 정형화된 데이터를 추출하기 위해 PSBE 데이터

를 이용하며, Pattern Matching의 정확도를 높이기 위해 많은 영상데이터를 준비하여 전문가의 자문을 기반으로 다량의 데이터를 학습하고, 이를 통해 정규화된 Matching data set인 PSBE를 생성하였다. 패턴매칭은 이러한 PSBE를 기반으로 이미지의 유사도를 검사하고 그 결과로 의료 이미지의 분석을 위한 정형화된 데이터를 추출하였다[9].

### 3.2 비정형 이미지 처리

비정형 이미지 처리는 (그림 2)와 같은 세부 과정을 진행하며, SIFT 알고리즘 기반 기존의 노이즈 제거에 활용되어온 알고리즘을 개선하여 각 이미지의 Feature를 보다 정확하고 세밀하게 특징화할 수 있는 기술을 제공한다.

(Figure 2) Processing Procedure of Informal Images

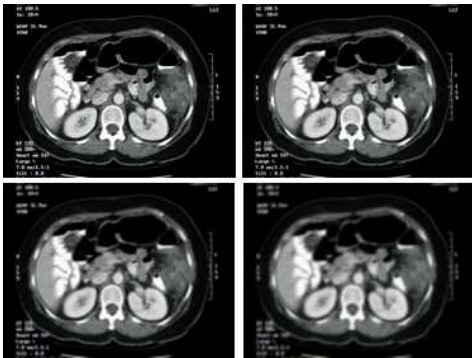


(그림 2) 비정형 이미지 처리 과정

Histogram Equalization 과정에서는 너무 어둡거나 밝은 영상을 보정하여 밝거나 어두운 쪽으로 편중된 픽셀 값의 분포를 그레이 스케일 전체로 분포하도록 한다. 가우시안 필터링과

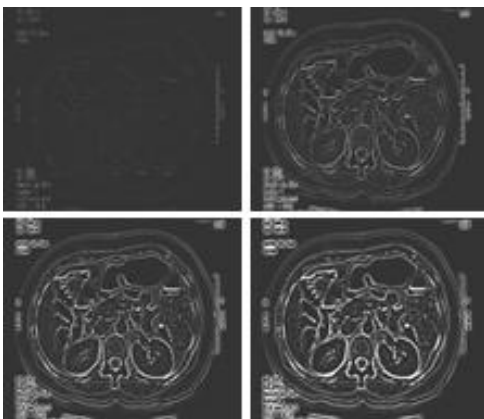
DoG(Difference of Gaussian)는 SIFT 특징점 추출 알고리즘에 적용되는 4단계의 가우시안 이미지를 생성하기 위해 (그림 3)과 같이 각 4단계별로 분산값을 생성하고 적용하여 이미지를 필터링 한다. (그림 4)와 같은 DoG 처리에서는 가우시안의 특성상 다른 분산값을 이용하여 가우시안 영상을 생성하면 영상의 엣지부분에서 흐릿해 지는 양이 틀리기 때문에 이러한 영상차이를 이용하여 엣지 기반의 이미지로 변환한다.

(그림 3) 가우시안 필터링 4단계 과정



(Figure 3) Gaussian Filtering 4 Steps

(그림 4) DoG 처리 과정



(Figure 4) DoG Processing Procedures

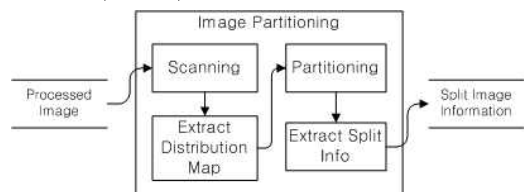
Detecting Extrema 과정에서는 생성된 3개의 DoG 영상을 이용하여 현재 픽셀을 기준으로 주위 26개의 픽셀의 극값이 임계치를 만족하게 하여 특징점 후보를 추출한다. 픽셀 제거 단계에서는 주변 픽셀값과 차이가 많이 나지 않는 점들

은 Taylor Series를 통하여 제거하고, 엣지 부근에서의 점들은 Hessian Matrix를 사용하여 제거하였으며 최종 단계인 Orient Assignment 과정에서 특징점 주위에 픽셀값의 경사도를 계산하여 특징점의 경사도를 36개 방향으로 특징점 주변의 경사도를 제거하여 비정형 이미지를 처리하였다.

### 3.3 의료이미지 분할 및 패턴 매칭

분석이 필요한 이미지가 포함하고 있는 특징들을 포함하고 있는 영역을 추출하여 데이터 분석 속도를 향상한다. 이는 분할된 이미지를 이용하여 병렬 처리가 가능하기 때문에 분석 속도 뿐 아니라 정확도 측면의 품질이 향상된다. 이미지 분할을 위해서는 (그림 5)와 같은 세부적인 4단계 과정을 거친다. 스캐닝 단계에서는 SURF[10]로 추출된 특징점 및 대응점의 분포를 스캔한다. 특징점 추출은 Fast-Hessian Detector[11]를 사용하여 추출한다. 특징점 추출 기준을 변경하여 최상의 결과를 도출할 수 있도록 할 예정이다. 표현자는 필터 크기가 2인 “Haar Wavelet Response[12]“에 기반하여 표현자를 생성(64차원)하지만, 48차원 불변하는 컬러 정보를 추가하여 112차원의 특징으로 표현자를 구하고, 표현자들의 거리로 대응점을 추출한다. Extract Distribution Map 단계에서는 정점의 거리를 측정하여 특징점들이 분포되어 있는 분포도를 추출한다. Partitioning 단계에서는 군집되어 있는 특징점을 기준으로 영역을 분할하며, Extract Split Info 과정에서 분할된 영역으로 이미지를 스플릿하여 영역 및 위치에 대한 정보를 생성하여 패턴 매칭을 위한 정보를 생성한다.

(그림 5) 의료이미지 분할 과정



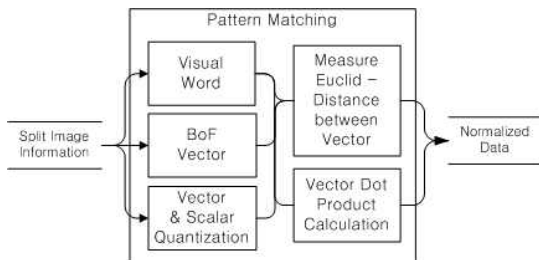
(Figure 5) Medical Image Partitioning Procedures

기존 매칭 알고리즘을 테스트하여 분석 결과

를 확인하고 개선된 알고리즘을 개발하고 구현한다. 결과 품질의 향상을 위해서 다양한 형태의 데이터를 적용하여 분석 실험하고, 이미지의 Feature에 대한 분석 비교 알고리즘을 활용하여 매칭 Flow를 확립하여 이미지 Feature들에 대한 매칭 속도와 정확도를 높인다.

패턴 매칭 알고리즘은 벡터 또는 스칼라 양자화를 인수로 활용하여 정확도를 향상시키고, 벡터 내적계산과 유클리드 거리 계산으로 매칭되는 데이터를 분석하였다. 분석 정확도와 속도 향상을 위해 특징점의 적용 방법에 대한 다양한 변수 적용으로 분석 결과를 분석하여 정확도를 향상시켰다.

(그림 6) 이미지 패턴 매칭 과정



(Figure 6) Image Pattern Matching Procedures

#### 4. 결론 및 향후연구

비정형 의료이미지(CT, MRI 등)는 의료전문가의 질병 진단 근거로 중요하게 활용하고 있지만, 이 과정에서 의료전문가는 이상징후 발견이나 판정에서 존재할 수 있는 오진단의 가능성을 최소화하기 위해 비정형 의료이미지에 포함된 다양한 데이터를 분석하는 패턴 모델링 및 분석 기술의 도입 시도는 의료전문가의 신뢰성과 적절한 의료판정을 통해 의료비를 절감하고 관련 산업시장을 활성화하는데 크게 기여할 수 있다.

본 논문에서는 비정형 의료이미지 정형화 및 패턴 정보 추출을 위해, 의료기기 또는 의료전문가로부터 생성 또는 감지되는 의료이미지 저장을 위한 데이터베이스를 구축하며 이러한 비정형 이미지의 특징을 정형화된 디지털 데이터로 변환한 후 정형화된 디지털 이미지 데이터로부터 의미있는 패턴 정보를 생성하였다. 패턴 매칭 알고리즘은 벡터 또는 스칼라 양자화를 인수로

활용하여 정확도를 향상시키고, 벡터 내적계산과 유클리드 거리 계산으로 매칭되는 데이터를 분석하였으며, 패턴 매칭 속도 향상을 위해 메모리 내장 기법을 적용하였다.

현재까지의 연구 내용 및 결과를 기반으로 향후에는 보다 정밀한 의료이미지의 패턴 매칭 정보를 추출하기 위한 새로운 알고리즘을 개발할 예정이며, 보다 다양한 의료이미지에 적용하여 패턴 매칭의 정확성과 속도 향상에 대한 우수한 방법 및 결과를 제시할 예정이다.

#### References

- [1] J. Henriques, T. Rocha, S. Paredes, R. Cabiddu, D. Mendes, R. Couceiro, P. Carvalho, "ECG analysis tool for heart failure management and cardiovascular risk assessment," Int'l Conf. Health Informatics and Medical Systems(HIMS2015), pp.195-200, July 2015.
- [2] A. Jonathan Garza, B. Sishir Subedi, C. Yuntian Zhang and D. Hong Lin, "A Web-Based System for EEG Data Visualization and Analysis," Int'l Conf. Health Informatics and Medical Systems(HIMS2015), pp.119-124, July 2015.
- [3] J. Cruz-Mota, I. Bogdanova, B. Paquier, M. Bierlaire, J.-P. Thiran, "Scale invariant feature transform on the sphere: Theory and applications", IJCV, Vol. 98 (2), pp. 217-241, 2012.
- [4] Ana C. Murillo, Gautam Sinth, Jana Kosecka, J. J. Guerrero, "Localization in Urban Environments Using a Panoramic Gist Descriptor", IEEE Transactions on Robotics, Vol. 29, No. 1, pp. 146-160, 2013.
- [5] Physionet, <http://www.physionet.org>
- [6] Physiobank, <http://www.physionet.org/physiobank>
- [7] PhysioToolkit, <http://www.physionet.org/physiotool>
- [8] OpenCV, <http://opencv.org>
- [9] Lindeberg T, "Scale-space theory in computer visio

n”, Springer Science & Business Media, Vol. 256, 2013.

[10] Bay, H., Tuytelaars, T., & Van Gool, L., “Surf: Speeded up robust features”, In Computer vision - ECCV Springer Berlin Heidelberg, pp. 404-417, 2006.

[11] Akshay Bhatia, “Hessian-Laplace Feature Detector and Haar Descriptor for Image Matching“, University of Ottawa, 2007.

[12] P. Raviraj and M.Y. Sanavullah, “The Modified 2D-Haar Wavelet Transformation in Image Compression”, Middle-East Journal of Scientific Research Vol. 2, pp. 73-78, 2007



### 고 광 만

1991년: 원광대학교 컴퓨터공학과 (공학사)

1993년: 동국대학교 일반대학원 컴퓨터공학과(공학석사)

1998년: 동국대학교 일반대학원 컴퓨터공학과(공학박사)

1998년~2001년: 광주여자대학교 컴퓨터학과 교수

2002년:한국연구재단지원 QUT(호주) 방문연구원

2008년:한국연구재단지원 퀘백대학교(캐나다) 방문연구원

2009년~2010년: NRF 대학교수해외방문연구사업 UC at Irvien(미국) 교환교수

2001년~현재: 상지대학교 컴퓨터정보공학부 교수  
관심분야 : 프로그래밍언어론, 모바일 클라우드 컴퓨팅(MCC), U-헬스케어