

시각 기술자들의 앙상블 결합을 이용한 의료 영상 분류와 검색

박기희*, 심정희*, 고병철*, 남재열*

*계명대학교 컴퓨터공학과

e-mail: khp5500@kmu.ac.kr, sjh0229@kmu.ac.kr, niceko@kmu.ac.kr, jynam@kmu.ac.kr

Medical Image Classification and Retrieval Using Ensemble Combination of Visual Descriptors

Ki-Hee Park*, Jeong-Hee Shim*, Byoung-Chul Ko*, Jae-Yeal Nam*

*Dept. of Computer Engineering, Kei-Myung University

요 약

본 논문은 의료 영상을 효과적으로 분류하고 검색 하기 위한 새로운 알고리즘을 제안한다. 의료 영상 중 X-Ray 영상은 어두운 배경에 반해 밝은 전경을 갖고 있기 때문에, 전경의 두드러진 부분에서만 시각 기술자로 추출한다. 우선, 색 구조 기술자(H-CSD)에서 헤리스 코너 검출기로 검출한 관심 포인들에서 색상 특징을 추출하고, 경계선 히스토그램 기술자에서 영상의 전역 및 지역적 질감 특징을 추출한다. 추출된 특징 벡터는 멀티클래스 SVM 에 적용되어 각 영상을 위한 멤버십 스코어를 얻는다. 이후, H-CSD 와 EHD 에 대한 SVM 의 멤버십 스코어를 앙상블 결합하여 하나의 특징 벡터로 생성하고, K-nearest Neighborhood 방법을 이용하여 상위-K 개의 영상을 검색을 하도록 하였다. imageCLEFmed2007 을 이용한 실험 결과에서 다른 전역적 속성 또는 분류 기반 검색 방법에 비교하여 보다 개선된 검색 성능을 나타낼 수 확인하였다.

Keywords: H-CSD, EHD, SVM, ensemble vector, MISS

1. 서론

최근 디지털화된 의료 영상의 증가와 다양한 기술을 바탕으로 이루어진 PACS(Picture Archiving Communication System)[1]와 같은 의료 지원 시스템의 도입으로 의료 영상들의 분류와 검색은 중요한 쟁점으로 떠오르게 되었다. 기존 의료 영상의 분류와 검색은 전문가의 주관과 관점에 따라서 달라질 수 있고, 많은 비용과 시간이 걸리는 단점이 있다. 따라서, 이 문제를 극복하기 위하여 다양한 유형의 분류와 검색 방법들이 지난 수십 년간 제시되었다[1]~[4].

기존의 영상 분류 알고리즘을 살펴보면, Mojsilovic[2]등은 영상으로부터 획득한 특징의 의미에 기반한 시각 특징 세트를 이용하여 영상을 분류하는 방법을 제안하였다. Greenspan[3]은 가우시안 혼합 모델을 통해 영상을 다수의 blob으로 표현하고, KL(Kullback-Leibler)을 사용하여 영상들 사이의 측정값을 판단하였다. Bhattacharya[4] 등은 MPEG-7 표준 기술자인 색 계층 기술자(CSD)와 히스토그램 기술자를 이용한 특징 벡터를 추출하여 특징 벡터들을 얻고 SVM과 FCM에 적용하여 분류하였다.

본 논문에서는, 의료 영상들의 분류와 검색을 위한 새로운 알고리즘을 제안한다. 우선, 의료 영상을 분류하기 위하여 헤리스 코너 검출기를 이용한 색 구조 기술자(H-CSD)를 제안하고, 질감 특징을 위한 경계선

히스토그램 기술자(EHD)를 사용하여 특징을 추출한다. 각각 추출된 특징벡터는 멀티 클래스 SVM에 적용하여 각 분류 범주에 대한 멤버십 스코어를 얻는다. 다시, H-CSD와 EHD의 멤버십 스코어를 하나의 앙상블 특징 벡터로 결합하고 영상 검색을 위하여 K-NI(K-Nearest Images)에 적용함으로써 기존의 관련 방법들의 문제를 개선하고 의료 영상의 분류와 검색 성능을 높이는 효과적인 방법을 제안한다.

2. 시각 기술자를 이용한 특징 추출

대용량의 의료 영상들을 효과적으로 분류하기 위하여 먼저 데이터베이스에 저장된 영상들에서 특징 벡터들을 추출한다. 본 논문에서는, MPEG-7 표준에 정의되어 있는 색 구조 기술자 중 하나인 CSD (Color Structure Descriptor)로 색상 특징을 사용하고, 경계선 히스토그램 기술자인 EHD (Edge Histogram Descriptor)로 질감 특징을 사용하며, 특히, CSD는 헤리스 코너 검출기를 이용하여 전경 영역들에서만 특징 추출하기 위하여 변형을 하였고, H-CSD라고 명명하였다.

2.1. 헤리스 코너를 이용한 색 구조 기술자(H-CSD)

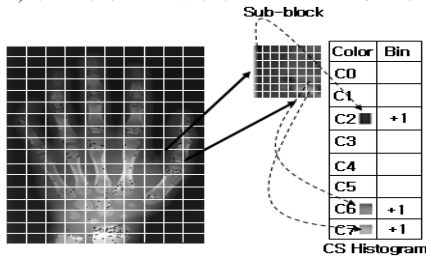
색상은 보는 각도, 변환 및 영상의 회전에 강건하기 때문에 영상 검색에서 널리 이용되는 시각적인 특징

중의 하나이다. 본 논문에서는, 작은 윈도우를 이용하여 로컬화된 색상 분포들을 식별하도록 CSD를 이용하여 색상 특징 벡터를 추출하였다. CSD는 영상의 색 히스토그램과 색상의 지역 부분 구조로 영상을 표현하는 특징을 가졌다 [5].

CSD를 적용하기 위해 X-ray 영상은 어두운 배경과 밝은 진경을 가지는 특성을 이용하여, 영상을 128 명암도 색상으로 양자화시키고, NxN 개의 서브블록들로 분리한다. 본 논문에서는 서브블록의 크기를 MPEG-7 표준에 정의된 구조 성분의 스케일과 같은 크기의 8x8 픽셀로 설정하였다. (그림 1)에서 보는 것처럼, X-ray 영상들은 사용하지 않는 배경 영역을 포함하기 때문에, 색 구조 히스토그램을 생성하기 전에 제거할 필요가 있다. 이를 위해, 먼저 양자화된 영상에서 헤리스 코너[6] 를 검출하고, 하나 또는 다수의 헤리스 코너 포인트들을 포함하는 서브블록들만을 선택한다. 헤리스 코너 검출기는 시점의 변화, 조명 방향 및 스케일과 잡음에 강한 불변성과 안정성 때문에 유명한 포인트 검출기이다. 선택된 서브블록들에서 128-빈 색 구조 히스토그램은 128개의 양자화된 명암도 영상에서 추출된다. CSD는 m 비트로 양자화된 값들의 1차원 배열이다.

$$CSD = \bar{h}_s(m), \quad m \in \{1, \dots, M\} \quad (1)$$

위 식에서 M 은 {256, 128, 64, 32} 세트에서 선택되고 s 는 연관되는 구조 성분(서브블록)의 스케일이다. 각 서브블록들의 위치에서, 색 구조 히스토그램은 서브블록 내의 색상을 표현하기 위하여 갱신된다. 예를 들어, 8 x 8 서브-블록에서 네 개 또는 여덟 개의 색 구조 히스토그램이 생성된다면, 영상에서 CSD의 서브블록의 위치가 이전 분할된 색 구조 히스토그램과 일치될 때 적절한 빈 수는 1씩 증가한다. 이 방법에서, 영상의 특징 값들은 기록된 색상 히스토그램에 의하여 CSD의 서브블록 수의 분포와 각 색상 히스토그램은 일치한다. 마지막으로, 추출된 128-빈 CSD는 SVM(Support Vector Machine)의 트레이닝을 위하여 0~1 범위로 정규화된다.



(그림 1) H-CSD의 특징 추출 처리과정

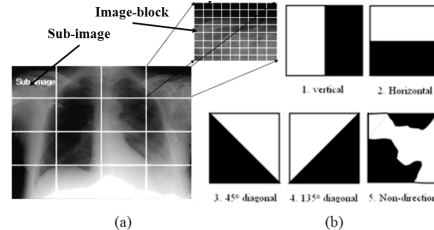
2.2. 경계선 히스토그램 기술자(EHD)

EHD[7]는 영상의 영역 경계선의 분포를 표현하는 기술자이다. 특히, 4x4 겹쳐지지 않은 서브영상들에서 영상 공간을 분리하고 각 서브영상은 (그림 2a)에서 보는 것처럼 겹쳐지지 않은 사각형 영상 블록들로 분리된다. 각 서브영상을 위하여 로컬 경계선 분포는 히스토그램으로 표현할 수 있고 서브영상들에서

경계선들은 다섯 가지 유형으로 범주화된다. (그림 2b)에서 보는 것과 같이 수직, 수평, 45 대각선, 135 대각선과 무방향 경계선 영상 블록의 크기는 입력영상을 동일한 크기를 가진 서브영상들로 분리하는 식2 를 사용하여 결정된다.

$$x = \sqrt{\frac{\text{width} \times \text{height}}{\text{desired Num block}}}, \quad \text{blocksize} = \left\lfloor \frac{x}{2} \right\rfloor \times 2 \quad (2)$$

위 식에서 desired Num block 은 영상 내 블록들의 전체 수이며, 실험을 통하여 1100으로 임의값을 정하였다. 각각 영상 블록들은 다섯 개의 경계선 범주 중 하나 또는 경계선이 없는 블록으로 분류된다. 적용된 각 필터로 특징값들이 추출된다면, 경계선 검출기는 최대 경계선 값으로 식별하고, 경계선 값이 주어진 임계값(threshold)을 넘으면, 일치하는 경계선 오리엔테이션은 영상 블록으로 조합된다. 여기에서 사용된 것은 16개의 서브영상들이므로, 총 5x16=80 히스토그램 빈이 생성된다.



(그림 2) (a)서브 영상과 영상 블록 (b) 다섯 가지 경계선 유형

의료 영상 중 특히, X-ray 영상에서 경계선은 정확도 개선에 결정적인 역할을 하는 특징이다. 따라서, 서브영상을 이용한 경계선 특징 이외에, 5개의 빈을 가진 글로벌 경계선 히스토그램을 추출하였다. 또한, 세미-글로벌 경계선 히스토그램을 위해, 13개의 다른 서브세트들을 그룹화하고 다섯 가지 경계선 유형을 위한 경계선 분포들을 생성하였다. 마지막으로, 80개의 영역 경계선 히스토그램, 5개의 로컬 경계선 히스토그램과 65(5x13)개의 세미-글로벌 경계선 히스토그램으로 총 150개의 경계선 히스토그램 특징 값을 사용한다.

3. 의료 영상 분류와 검색

특징 추출 후, 영상들은 미리 정의한 클래스들 중의 하나로 분류된다. 이를 위해, 멀티클래스 SVM과 두 개의 특징 벡터인 H-CSD와 EHD를 사용한다. 트레이닝 결과에 의하여, 각 영상은 전체 20개의 범주로 멤버십 스코어를 가지며, H-CSD 와 EHD를 하나의 특징 벡터로 앙상블 하는 것처럼 각각 결합하여 평가된다. 앙상블 특징 벡터는 마지막에 의료 영상 검색 시스템인 MISS(Medical Information Searching System)에 사용된다.

3.1. H-CSD와 EHD 특징 벡터를 이용한 SVM 분류

SVM은 패턴 분류 문제를 위하여 좋은 일반화 성능을 제공할 수 있으며, 영상 분류 결정을 위해 실험적인 특징 매개변수들을 요구하지 않는다.

아래 수식은 쌍선형 분류 함수 일반식으로, 주어진 트레이닝 데이터 $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ 는 $\mathbf{x}_i \in \mathbb{R}^d$ 공간에서 벡터들이고, $\mathbf{y}_i \in \{+1, -1\}^N$ 에서 라벨 $(\mathbf{y}_1, \dots, \mathbf{y}_N)$ 이며, 수식의 결과는 분리되는 초평면과 일치된다.

$$\mathbf{g}(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \quad (3)$$

수식 (3)에서 \mathbf{x} 는 입력 벡터, \mathbf{w} 는 가중치, 그리고 b 는 바이어스이다. SVM 분류의 주된 목적은 초평면에서 클래스간의 거리를 최대로 하는 동안 정확하게 데이터 포인트들의 상당한 일부분들을 떼어놓는 최적의 초평면을 위한 매개변수 \mathbf{w} 와 b 를 찾는다.

SVM 분류 함수는 다음 식으로 정의된다[8]:

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^n v_i \cdot k(\mathbf{x}, \mathbf{x}_i) + b\right) \quad (4)$$

위 식에서 $k(\cdot, \cdot)$ 는 커널 함수, v_i 는 각 커널 출력들의 가중치, b 는 바이어스 이고 $f(\mathbf{x})$ 의 사인은 +1 클래스와 -1 클래스처럼 \mathbf{x} 의 클래스 멤버십을 결정한다. 초평면에서 결정 함수 $f(\mathbf{x})$ 는 영상이 속하는 한 범주(+1)가 다른 범주들(-1)로부터 얼마나 다른지 측정 가능한 소프트 벡터로 결정된다. 직관적으로, 초평면에서 더 멀리 떨어진 포인트, 즉 상당한 양의 $f(\mathbf{x})$ 일수록, 더욱 신뢰할 수 있는 분류 결과이다.

선형 SVM에서의 커널 함수는 입력 공간에서 단지 단순한 점이다. 그러나 비선형 SVM에서의 커널 함수는 높은 차원 F의 특징 공간에 샘플을 효과적으로 추정하고 F에서 초평면을 만든다[8]. SVM 트레이닝 알고리즘은 가장 큰 마진을 사용하는 두 클래스들 내에 F에서 데이터를 분리하는 초평면으로 판단한다.

본 논문에서는, 의료 영상들이 신체의 각 부분들에 따라 분류되는 20개의 클래스들이기 때문에 바이너리 SVM 대신에 RBF (Radial-Basis Function) 커널로 멀티클래스 SVM을 사용하는데, 이는 일대다, 일대일, 비순환 그래프와 같이 일반적으로 사용된 방법들이다 [9]. 본 논문에서는 n개의 분류기를 만드는 일대다 방법을 적용하는 하였다. 이 방법에서는 원래 클래스와 일치하는 i번째 특징벡터들은 +1을 부여하고, 나머지 클래스의 특징벡터들은 -1을 부여하여 학습 하는 방법을 사용 한다. 학습 후에 테스트 샘플 \mathbf{x} 는 n개의 SVM 분류기들에 적용되고 가장 높은 출력 스코어(dj)가 최종 클래스로 선택된다.

$$d_j(\mathbf{x}) = \max_{i=1, \dots, n} d_i(\mathbf{x}) \quad (5)$$

위 식에서 d_i 는 입력 \mathbf{x} 영상들을 위한 i번째 클래스에 대한 출력 스코어이다.

본 논문에서는 트레이닝 실행부분에서, 1,754 영상들을 20개의 영상 범주에서 임의로 선택하고, Image CLEF Med2007에서 사용된 IRMA(Image Retrieval in Medical Applications)의 X-ray영상을 사용하였다 [10].

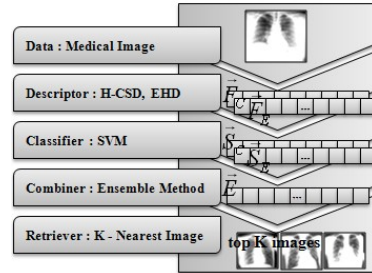
본 논문에서는, 학습과정에서 두 특징 벡터들을 각각 독립적으로 사용하기 때문에, $2n$ SVM 분류기가 생성된다.

3.2. 앙상블 특징 벡터 결합과 유사성 측정

SVM 트레이닝 후, 테스트 영상으로부터 H-CSD와 EHD의 특징 벡터가 추출되고 각각의 특징값을 위한 SVM 분류기에 적용되어 각 범주에 대한 멤버십 스코어를 갖게 된다. (그림 3)에서, 추출된 특징 벡터 F_c 와 F_E (F_c : H-CSD 특징 벡터, F_E : EHD 특징 벡터)는 SVM 분류기에 각각 제공되고, SVM 분류기는 각 특징 벡터를 위해 20개의 멤버십 스코어인 \vec{S}_c 와 \vec{S}_E 를 출력한다. 마지막으로, 두 개의 멤버십 스코어 벡터는 하나의 앙상블 벡터 ($\vec{E} = [s_{c1}, s_{c2}, \dots, s_{c20}, s_{e1}, s_{e2}, \dots, s_{e20}]$)로 합쳐진다. 이 앙상블 벡터는 주어진 영상으로 최상위 k를 예측하기 위해 최종적으로 K-Nearest Neighbor 클러스터에 적용 된다. 가장 유사한 범주를 검색하기 위하여, 거리는 식 (6)으로 판단하고 그 결과에 따라 최상위 근접 영상들(the top nearest images)을 오름차순으로 보여준다.

$$S(q, t) = \sum_{i=1}^{40} |s_i^q - s_i^t| \quad (6)$$

위 식에서 q와 t는 쿼리와 목표 영상을 나타낸다.



(그림 3) 영상 분류 및 검색 처리과정

4. 실험 결과

제안하는 시스템은 오프라인 트레이닝을 위해 Visual C++ 6.0 언어를 사용하여 개발되었고, 테스트 시스템은 C# 언어를 사용하여 ASP.NET 2.0을 기반으로 개발되었다. 테스트를 위해서, IRMA(Image Retrieval in Medical Applications)[10]에서 1,000개의 영상들(20 범주)을 사용했다. 웹사이트 (<http://cvpr.kmu.ac.kr>)에서 MISS 시스템을 시연해볼 수 있다.

<표 1> 각 클래스별 이름과 테스트 영상 수

범주	신체부분	테스트 데이터 수	범주	신체부분	테스트 데이터 수
1	가슴	60	11	손가락	50
2	골반	60	12	손목	50
3	머리정면	50	13	슬개골	60
4	머리측면	50	14	어깨	50
5	목	50	15	척추	60
6	무릎	50	16	흉부정면	60
7	발가락	45	17	흉부측면	50
8	발목정면	50	18	허벅지	40
9	발목측면	50	19	팔꿈치	20
10	손	60	20	발	35

<표 2> 각 기술자 실험 결과

	Top = 5	Top = 10	Top = 20	Average Precision
CSD	0.52	0.426	0.32	0.422
H-CSD	0.744	0.693	0.658	0.698
EHD	0.81	0.77	0.7415	0.773
Ensemble	0.976	0.968	0.952	0.965

<표 3> 결합 기술자 실험 결과

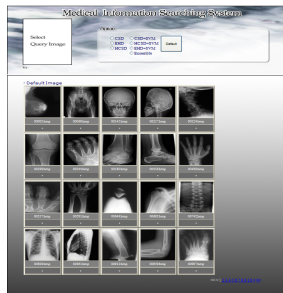
	Top = 5	Top = 10	Top = 20	Average Precision
Bhattacharya's method [SVM+FCM]	0.63	0.605	0.53	0.588
Ensemble vector	0.976	0.968	0.952	0.965

쿼리완료를 위해 사용자는 '디폴트' 버튼을 누르고 일곱 가지 방법 중에 한 검색 방법을 선택한다. 그 후, 사용자는 검색하기 원하는 다른 영상을 클릭하면 (그림 4)와 같이 가장 유사한 20개의 영상들이 나타난다.

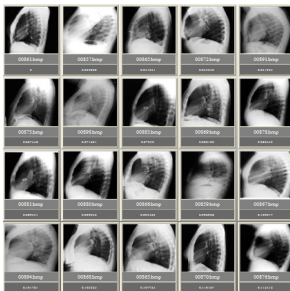
본 논문에서는, 세가지 특징 벡터와 제안하는 앙상블 특징 벡터를 각각 사용하여 유사성을 측정하고, 유사성 측정은 precision 방법을 사용한다.

테스트는 각 범주에서 5개의 쿼리 영상으로 실행되었고, 성능은 평균 검색 precision을 사용하여 평가되었다. 표 3에서 보는 것과 같이, 검색 성능은 각각 50%, 69.8%, 77.3%, 그리고 96.5%로서 다른 세 개의 방법들을 능가한다. 특히, 제안하는 H-CSD의 검색 성능은 기존 CSD보다 19.8% 개선된 것을 볼 수 있다.

또한 Bhattacharya 등[4]의 알고리즘과 검색 성능을 비교결과 제안하는 방법이 37.7%의 개선된 검색 성능을 보여줌을 알 수 있었다. (그림 5)는 MISS 시스템의 검색 결과이다.



(그림 4) MISS 검색 인터페이스



(그림 5) '흉부측면' 검색 결과

5. 결론

본 논문에서는, 의료 영상의 효과적인 분류와 검색을 위한 새로운 알고리즘을 제안한다. 의료 영상들을 분류하기 위해, 먼저 제안하는 해리스 코너 검출기 기반 색 구조 기술자 (H-CSD)로 색상 특징을 추출하고 경계선 히스토그램 기술자는 영상들의 글로벌과 로컬 부분에서 질감 특징을 추출한다. 추출된 특징벡터는 멀티클래스 SVM에 적용하여 각 영상을 위해 멤버십 스코어를 얻는다. H-CSD 와 EHD의 멤버십 벡터로부터, 하나의 앙상블 벡터를 생성하고 MISS 시스템에 적용하여 그 성능을 비교하여 보았다. IRMA 영상을 사용한 실험 결과는 제안하는 시스템이 다른 글로벌 속성 기반 또는 다른 분류 기반 검색 방법들과 비교하여 실제로 검색 성능이 개선되는지를 보여준다.

앞으로의 연구에서 범주 분류와 영상 분류 기반 자동 주석을 위한 개선된 알고리즘이 요구된다. 특히, 목 대 척추와 손가락 대 발가락처럼 유사한 범주들에서 분류 성능을 개선하기 위해 새로운 특징 개발이 요구된다.

감사의글

본 연구는 산업자원부의 지방기술혁신사업(RT104-01-01)지원으로 수행되었습니다.

참고문헌

[1] Qi H., Snyder W.E.: "Content-based image retrieval in PACS", Journal of Digital Imaging, Vol. 2, (1999), 81-83.
 [2] Mojsilovc A. and Gomes J.: "Semantic based categorization, browsing and retrieval in medical image databases", Int.Conf. on Image Processing, Vol. 3, (2002) 145-148.
 [3] Greenspan H.: "Medical Image Categorization and Retrieval for PACS Using the GMM-KL Framework", IEEE Transactions on Information Technology in BioMedicine, Vol. 11 (2007) 190-202
 [4] Bhattacharya P. and Rahman M. M., "Image Representation and Retrieval Using Support Vector Machine and Fuzzy C-means Clustering Based Semantical Spaces", International Conference on Pattern Recognition, Vol. 2 (2006) 1162-1168.
 [5] Manjunath B. S., Salembier P. and Sikora T., "Introduction to MPEG-7", John Willy & Sons, LTD, (2002).
 [6] Harris C. and Stephens M. J.: "A combined corner and edge detector", Alvey Vision Conference (1998) 147-152.
 [7] Won C. S. and Park D. K.: "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, Vol. 24, (2002), 23-30.
 [8] Vapnik V., : "The Nature of Statistical Learning Theory". Springer-Verlag. (1999).
 [9] Chen S-C. and Murphy R.F.: "A graphical model approach to automated classification of protein subcellular location patterns in multi-cell images", BMC Bioinformatics, Vol 7, (2006) 1-13.
 [10] Deselaers T.: "The CLEF 2005 Automatic Medical Image Annotation Task", International Journal of Computer Vision Vol.74, (2007) 55-58.