

MPEG-7 시각 기술자와 멀티 클래스 SVM을 이용한 의료 영상 분류와 검색

심정희*, 고병철*, 남재열*

*계명대학교 컴퓨터공학과

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

Medical Image Classification and Retrieval using MPEG-7 Visual Descriptors and Multi-Class SVM(Support Vector Machine)

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

*Dept of Computer Engineering, Kei-Myung University

요 약

본 논문은 의료 영상에 대한 효과적인 분류와 검색을 위한 알고리즘을 제안한다. 영상 분류와 검색을 위해서 MPEG-7 표준 기술자인 색 구조 기술자와 경계선 히스토그램 기술자를 사용해 영상들에 대한 특징 값을 추출한다. 이렇게 구해진 특징 값들을 의료 영상의 분류와 검색에 적용해 본 결과 비교적 낮은 성능을 보여줌을 확인하고 앞서 구해진 특징 값들을 교차 학습 방법인 SVM(Support Vector Machine)과 비교사 학습 방법인 FCM(Fuzzy C-means Clustering)에 적용시켰다. 기존 연구에서는 SVM과 FCM의 통합으로 의료 영상에 대한 분류와 검색을 시행하였지만 본 논문에서 실험한 결과 SVM과 MPEG-7 시각 기술자 중에 하나인 EHD(Edge Histogram Descriptor)를 가중치 선형 결합하여 실험한 결과가 더 정확한 분류와 높은 검색 성능을 나타냄을 확인하였다.

1. 서론

다양한 기술을 바탕으로 이루어진 PACS (Picture Archiving Communication System)[1]의 도입으로 최근 의료 진단에 필요한 방대한 양의 의료 영상이 디지털화 되어 데이터베이스를 구축하고 있다. 따라서 의료 영상에 대한 효과적인 저장과 검색이 중요 시 되기 시작하였고, 현재 이와 관련된 연구들이 활발히 진행 되고 있다. 의료 영상은 관찰자의 눈에 따라 다른 의미를 부여할 수 있으며, 영상 내에 중요한 의미를 가지고 있는 관심 영역과 단색의 배경으로 구성된 특징을 가지고 있기 때문에 일반적인 자연 영상과는 다른 특징 추출 기술자 사용 및 분류 방법이 요구된다. 의료 영상 분류에 관련된 기존연구에 대해 알아보면 Miller[2]등은 의료 영상 분류와 검색을 위해 사용된 시각적 특징 값으로 Canny, 소벨 기술자등을 통해 질감 정보를 획득하고 퓨리에 기술자를 사용해 형태 정보를 획득하였다. Greenspan[3]은 가우시안 혼합 모델을 통해 영상을 다수의 blob로 표현하고, KL(Kullback Leibler)을 사용하여 영상들 사이의 일치성(matching)을 판단하였다. Bhattacharya등[4]은 MPEG-7 표준 기술자인 색 계층 기술자, 경계선 히스토그램 기술자를 사용해 추출한 특징 값들을 교차 학습 분리기인 SVM(Support Vector Machine)와 비교사 학습 분리기인 FCM(Fuzzy C-means Clustering)을 통해 의료 영상에 대한 분류를 시행했다. 본 논문에서는 현재 개발 중에 있는 내용

기반 검색시스템인 MISS(Medical Information Searching System)의 구현을 위한 의료 영상 분류와 검색 알고리즘을 제안한다. 의료 영상의 분류, 검색을 위한 전단계로 영상 안에 들어있는 색, 질감, 형태 등의 고유 정보를 사용하여 특징 값을 추출하고 그 값들을 교차 학습 분리기인 SVM과 경계선 히스토그램 기술자(EHD)를 결합하여 적용함으로써 기존의 Bhattacharya등[4]의 방법의 문제점을 개선하고 의료 영상의 분류와 검색 성능을 높이는 방법을 제안한다.

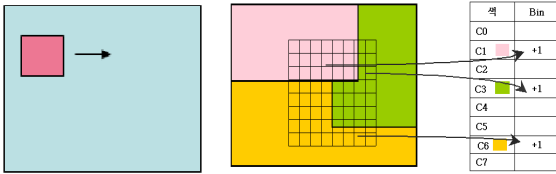
2. MPEG-7 기술자를 이용한 특징 추출

데이터베이스에 저장된 대용량의 의료 영상을 카테고리별로 보다 효과적으로 분류하고 진단하기 위해서 먼저 각 데이터베이스 안에 저장되어 있는 영상들에 대한 특징 값들을 추출해야 한다. 특징 추출에 사용되는 기술자로는 MPEG-7 표준에 정의되어 있는 색 구조 기술자(CSD)[5]로 색상 특징을 기술하며, 질감 특징은 경계선 히스토그램 기술자(EHD)[6]를 사용하여 추출한다.

2.1 색 구조 기술자

MPEG-7에서 현재 정의되어 있는 색 기술자 중 하나인 색 구조 기술자는 영상의 유사한 색상 간의 히스토그램과 지역 공간 구조의 분포를 나타내는 기술자로서, 그

검색 성능이 다른 기술자에 비하여 상대적으로 뛰어나고 구현 또한 간단하다. 이 기술자에 사용되어지는 구조화 성분은 영상의 크기에 따라 유동적으로 그 크기와 서브샘플링 수가 정해지게 된다. 그림 1에서 보는 것과 같이 먼저 영상의 크기에 따라 정해진 크기 및 서브 샘플링 수의 구조화 성분을 결정한다.



(그림 1) 색 구조 기술자를 통한 특징추출과정

영상의 크기에 따른 구조화 성분의 크기 및 서브 샘플링 수는 식(1)을 통해 정해진다.

$$p = \max(0, \text{round}(0.5 \times (\text{height} \times \text{width}) - 8)) \quad (1)$$

$$K = 2^p, E = 8 \times K$$

위 식에서 *height*와 *width*는 각각 영상의 높이와 너비를 나타내며, $E \times E$ 는 구조화 성분의 크기를, K 는 서브 샘플링 수를 뜻한다. 위 식에 의해 결정된 구조화 성분을 영상 위에서 그 면적의 반씩 겹친 상태로 이동하면서 그 안에 있는 색의 종류 및 각 색의 상대적 히스토그램을 기록하게 된다. 예를 들어 그림1에서 8×8 의 구조화 성분에 존재하는 4개의 색과 8개의 색 구조 히스토그램이 생성되어 있다고 하면, 영상에서 구조화 성분이 위치된 부분이 미리 나누어 놓은 색 구조 히스토그램에 해당 될 때, 해당되는 빈의 개수가 1 증가된다. 이렇게 기록된 색 히스토그램을 이용하여 각각의 색의 히스토그램에 해당되는 구조화 성분의 개수 분포가 영상의 특징 값이 된다. 본 연구에서는 128개의 구조화 성분을 사용, 즉 총 128개의 특징 값을 사용하였다.

2.2 경계선 히스토그램 기술자

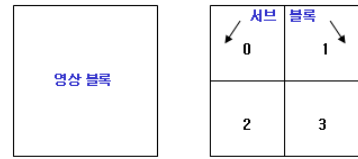
경계선 히스토그램 기술자[6]는 영상의 지역적 경계선의 분포를 나타낼 수 있는 기술자이다. 이 기술자는 하나의 영상에 대해 16개의 겹치지 않는 서브영상으로 분할하고 5종류의 경계선(수직, 수평, 45° 대각선, 135° 대각선, 비방향성)을 사용하여 영상의 지역적 영역의 공간적인 분포를 표현한다. 경계선 히스토그램 기술자는 먼저 입력된 영상들에서 동일한 크기로 나누어진 서브영상들에 대해 아래 식(2)를 적용하여 영상 블록의 크기를 결정 한다.

$$x = \sqrt{\frac{\text{width} \times \text{height}}{\text{desired Num block}}} \quad (2)$$

$$\text{blocksize} = \lfloor \frac{x}{2} \rfloor \times 2$$

위 식에서 *desired Num block* 은 영상 내에 있는 블록

들의 전체 개수를 뜻하는 것으로 실험에서는 MPEG-7에 정의된 기본 값인 1100을 사용하였다. 조건을 만족시키기 위해 최 외곽 픽셀은 생략 가능하다. 구해진 *blocksize* 값을 통해 나누어진 영상 블록은 또 다시 4개의 서브 블록들로 분할한다. 이는 각 서브 블록들에 대해 5종류의 경계선 추출 필터를 적용하기 위함이다. 그림 3은 4개의 서브 블록으로 나뉜 하나의 영상 블록을, 그림 4는 추출되어지는 경계선의 종류를 보여준다. 필터를 적용하여 특징 값을 추출하게 되면 하나의 서브 영상에 5개의 빈값 즉, 한 입력 영상에 대해 80개의 빈을 가진 히스토그램을 생성하게 된다.

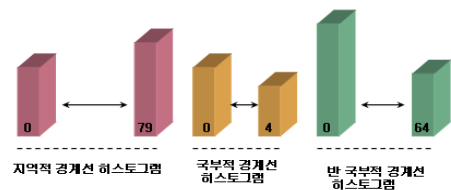


(그림 3) 하나의 영상 블록에 대해 4분할 된 서브 블록



(그림 4) 경계선 추출에 사용되는 5종류 경계선

의료 영상의 적절한 분류와 검색은 환자의 생명과도 직결된 사항이기 때문에 더욱더 높은 정확성이 요구되는 작업이다. 따라서 지역적 에지 히스토그램 값만으로는 불충분하다 판단되어 입력 영상 전체에 대한 5개의 국부적 경계선 히스토그램 값과 13개의 집단으로 그룹화한 Semi(반)-국부적 경계선 히스토그램 값을 추가하였다. 즉, 그림 5에서 보는 것과 같이 지역적 경계선 히스토그램 80개, 국부적 경계선 히스토그램 5개, Semi(반)-국부적 경계선 히스토그램 65개로 총 150개의 경계선 히스토그램을 생성하여 경계선 히스토그램 특징 값으로 사용하였다.



(그림 5) 경계선 히스토그램 기술자에 사용된 특징벡터

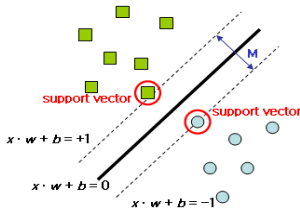
3. 의료 영상 분류와 검색

앞서 MPEG-7 특징 기술자인 색 구조 기술자와 경계선 히스토그램 기술자를 통해 추출한 특징 값들만을 사용하여 내용기반 검색 시스템인 MISS에 적용시킨 결과 의료 영상 분류를 하기엔 정보가 불충분하다고 판단되었다. 따라서 Bhattacharya등[4]의 방법을 토대로 교차 학습 방법인 SVM(Support Vector Machine)과 FCM(Fuzzy C-means

Clustering)이라는 비교사 학습 방법을 적용하여 시행해보고 보다 정확하고 세밀한 분류와 검색을 위한 알고리즘을 제안하고자 한다.

3.1 교사 학습(Supervised learning) 방법인 SVM (Support Vector Machine)

지도 학습의 일종인 SVM(Support Vector Machine)은 선형 분리 가능한 영상들에 대해 차원을 확대시켜 분류시키는 방법이다. 먼저 각 카테고리 별로 일정개수의 의료 영상들을 학습시키는 과정이 필요하다. 본 연구에서는 20개의 카테고리에 대해 각각 일정 개수의 의료 영상을 학습 시킨다. 적은 개수의 의료 영상을 가진 카테고리들인 발가락 48장, 허벅지 47장, 팔꿈치 21장, 발 38장의 의료 영상을 학습 시킨 것을 제외하고 나머지 카테고리에 대해서는 각각 100장의 영상을 학습, 총 1,754장의 영상을 학습시켰다. SVM에서 가장 중요한 것은 정확한 분류를 위한 최적의 초평면을 찾는 것이다. 그림 6 에서 보는 것과 같이 $x \cdot w + b = +1$, $x \cdot w + b = -1$ 인 두 개의 클래스가 있다고 할 때, $x \cdot w + b = 0$ 의 w 와 b 의 값들에 대해 구해진 최적의 값이 초평면이 된다. 이 초평면을 결정 경계라 하며 그 결정 경계에 가장 가까운 값들을 서포트 벡터(Support Vector)라 부른다. 여기서 초평면은 서포트 벡터와의 거리가 최대가 될 때 최적화 된다.



(그림 6) SVM에서의 서포트 벡터와 초평면

아래의 제약식(3)을 통해 목적식(4)이 계산되어지는데 식(4)에서 w 값이 최소 값을 가질 때 초평면과 서포트 벡터의 거리가 최대가 된다. 이때 구해지는 값이 최적의 초평면을 위한 w 와 b 의 값이 된다.

$$y_i((W^T\Phi(X_i)+b) \geq 1-\xi_i \quad : \text{제약식 (3)}$$

$$\frac{1}{2}w^T w + C \sum_{i=0}^N \xi_i \quad : \text{목적식 (4)}$$

식(3),(4)에서 w 는 가중치, b 는 바이어스, N 은 영상의 총 개수, ξ_i 는 라그랑제 승수를 나타낸다. 결정된 값들은 SVM 분류 함수인 식(5)에 적용되어 결과를 계산하며 입력되어지는 영상들에 대한 분류를 하게 된다.

$$f(X) = \text{sign}(\sum_{i=1}^N \alpha_i y_i K(X_i, X) + b) \quad (5)$$

식(5)에서 X 는 입력 영상을, K 는 커널함수 값을 나타

낸다. 본 연구에서는 의료 영상을 그 부위에 따라 20개의 클래스로 분류해야 하므로 Binary SVM이 아닌 멀티 클래스 SVM을 사용한다. 멀티 클래스 SVM에 대한 계산 방법으로는 One to others 방법과 Pair-wise 방법이 있는데 우리는 식(6)을 이용하여 One to others 방법을 사용하였다.

$$d_j(X) = \max_{i=1, \dots, q} d_i(x) \quad (6)$$

위 식(1)에서 i 는 클래스의 수를 의미하고 d_i 는 입력된 X 영상이 각 클래스에 처리된 결과 값을 의미한다. 즉 X 영상은 가장 값이 큰 클래스로 분류된다.

3.2 비교사 학습(Unsupervised learning) 방법인 FCM(Fuzzy C-means Clustering)

비교사 학습 방법인 FCM은 위에서 언급한 SVM 방법과 달리 교사 되지 않은 영상들에 대해 임의의 C개의 군집 개수를 정하여 분류 하는 방법이다. FCM 알고리즘은 식(7), (8)을 반복적으로 계산함으로써 구현되어진다.

- 1단계 : 군집 개수 C와 랜덤 한 초기 중심 값 $V_j(0)$ 을 지정하고 식(7)을 통해 집단 함수 $U(0)$ 값을 초기화 한다.

$$\mu_{ji} = \frac{\text{dist}(X_i, V_j)^{-2/(m-1)}}{\sum_{k=1}^c \text{dist}(X_i, V_k)^{-2/(m-1)}} \quad (7)$$

- 2단계 : 식(8)을 이용하여 새로운 군집에 대한 중심 값을 갱신한다.

$$V_j(t) = \frac{\sum_{i=1}^N \mu_{ji}^m X_i}{\sum_{i=1}^N \mu_{ji}^m} \quad (8)$$

- 3단계 : 식(8)에 의해 구해진 새로운 중심 값을 1단계에 식(7)에 적용하여 $U(t)$ 값을 계산한다.

- 4단계 : $|V_j(t) - V_j(t-1)| < \epsilon$ 즉, 중심점이 이동하지 않을 때까지 2, 3 단계를 반복 수행한다.

4. 실험 및 결론

제안된 시스템의 성능평가를 위해 본 논문에서는 ImageCLEFmed2007[7]에 사용된 IRMA(Image Retrieval in Medical Applications) 영상 중 963장(20개 카테고리)을 테스트에 적용하였다. MPEG-7 시각적 기술자 중 특징 값 추출을 위해 사용된 색 공간 기술자와 경계선 히스토그램 기술자는 각각 128, 150개의 특징 값을 추출하여 실험에 사용하였다. SVM 적용을 위해 우선적으로 이루어진 학습 과정은 20개 카테고리에 대해 각각 일정 개수의 의료 영상을 학습시키도록 하였다. 표 1은 20개의 카테고리 별 명칭과 학습된 영상들의 개수를 보여준다. 영상 자체의 개수가 적은 발가락, 허벅지, 팔꿈치를 제외한 나머지

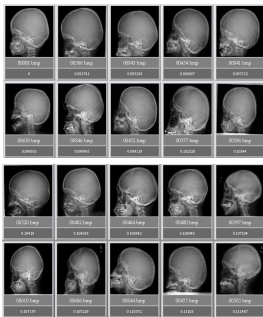
<표 1> SVM 학습을 위해 분류한 카테고리 명칭과 학습된 영상의 개수

번호	명칭	학습 개수	번호	명칭	학습 개수
1	가슴	100	11	손가락	100
2	팔반	100	12	손목	100
3	머리정면	100	13	슬개골	100
4	머리측면	100	14	어깨	100
5	목	100	15	척추	100
6	무릎	100	16	흉부정면	100
7	발가락	48	17	흉부측면	100
8	발목정면	100	18	허벅지	47
9	발목측면	100	19	팔꿈치	21
10	손	100	20	발	38

카테고리에 대해서는 각각 100개의 영상을 학습시켜 총 1,754장의 학습 데이터를 생성하였다. FCM 적용을 위해서 군집 클래스의 개수는 20개로 설정, 초기 중심 값은 랜덤한 값을 할당하였고, 무한 반복을 피하기 위해 최대 반복횟수는 500번, ϵ 은 0.001의 값을 할당하였다. 성능 평가는 MPEG-7 기술자인 색 구조 기술자, 경계선 히스토그램 기술자, SVM, FCM을 독립적 혹은 결합하여 의료 영상 분류와 검색을 실시하였다. 기존에 Bhattacharya등[4]이 제안한 알고리즘은 SVM과 FCM의 결합을 통한 영상 검색이 가장 우수한 성능을 보인다 했지만 본 연구의 실험을 통해 식(9)와 같이 SVM과 MPEG-7 시각적 기술자인 EHD를 결합했을 때 보다 뛰어난 성능을 나타냄을 확인하였다.

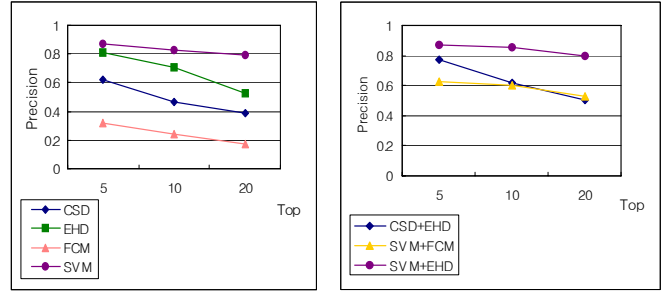
$$S(query, database) = w_1SVM + w_2EHD \quad (9)$$

식(9)에서 w_1 과 w_2 는 각각 SVM과 EHD를 위한 가중치를 의미하고 본 실험에서는 각각 0.5, 0.5의 가중치를 부여하였다. 그림 7은 MISS에 적용한 실험 결과를 보여준다.



(그림 7) 머리측면 카테고리에 대한 검색 결과(Top 20)

본 논문에서는 Precision방법을 이용하여 성능을 비교하였고 실험 결과는 그림 8과 같다. Bhattacharya등[4]이 제안한 SVM과 FCM의 결합은 평균 58%의 성능을 보임에 반해 제안하는 방법은 SVM과 EHD의 결합은 평균 84%의 성능으로 기존 방법보다 우수한 성능을 나타냈다. 실험을 통해 비슷한 형태의 특징을 가진 척추, 목 카테고리나 손, 발 카테고리의 유사성 검색은 비교적 낮은 성능



(그림 8) 유사성 검색 성능 평가 결과

을 보였다. 향후 연구로는 보다 정확한 유사성 검색을 위한 알고리즘 개발과 각 카테고리별 자동적인 주석 처리 기능에 대한 연구를 할 계획이다.

감사의 글

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

실험 데이터는 독일 RWTH Aachen 대학의 ImageCLEFmed2007에 사용된 IRMA(Image Retrieval in Medical Applications)영상을 사용하였습니다.

참고문헌

- [1] H. Qi, W. E. Snyder "Content-based image retrieval in PACS", Journal of Digital Imaging 12(2) pp. 81-83, 1999
- [2] H. Muller, N. Michoux "A Review of Content - Based Image Retrieval Systems in Medical Applications-Clinical Benefits and Future Directions", International Journal of Medical Informatics, Vol. 73, Issue 1, pp.1-23
- [3] H. Greenspan "Medical Image Categorization and Retrieval For PACS Using the GMM-KL Framework", IEEE Transactions on Information Technology in BioMedicine, pp. 190-202, 2007
- [4] P. Bhattacharya, Md. M. Rahman, "Image Representation and Retrieval Using Support Vector Machine and Fuzzy C-means Clustering Based Semantical Spaces", Pattern Recognition, 2006. ICPR 2006. 18th International Conference, pp. 1162-1168, 2006
- [5] A. Buturovic, "MPEG-7 Color Structure Descriptor for visual information retrieval project VizIR", <http://vizir.ima.tuwien.ac.at>
- [6] Chee Sun Won, Dong Kwon Park, "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, Vol. 24, Number 1, 2002
- [7] T. Deselaers, "The CLEF 2005 Automatic Medical Image Annotation Task", International Journal of Computer Vision 74(1), pp.55-58, 2007