

# 의료 영상 검색 시스템의 설계 및 구현

문형석, 엄기현  
 동국대학교 컴퓨터공학과

## Design of Medical Image Retrieval System

Hyoungsuk Moon, Kyhyun Um  
 Dept. of Computer Engineering, Graduation School, Dongguk University  
 dabechi@dongguk.edu, khum@dgu.ac.kr

### 요약

대부분의 의료 영상 관리 시스템이 의료 영상의 저장, 전송 등의 기본적인 기능만 지원될 뿐 상위 응용 계층에서 내용기반 검색이 지원되지 않고 있다. 본 논문에서는 이를 위해 내용 기반 검색 기능을 지원하는 의료 영상 검색 시스템을 설계 및 구현한다. 의료 영상 검색 시스템은 질의 이미지의 내용기반 검색을 위해 색-공간, 질감, 모양 특징에 의한 유사 비교 기법을 사용하고 각각의 유사 비교 검색에 의해 생성된 결과 집합들을 통합하고 최종 결과 제시를 위해 복합 질의문 계획 생성 알고리즘을 제시한다.

### 1. 서론

대부분의 의료 영상 관리 시스템이 의료 영상 저장 전송 시스템(Picture Archiving and Communication Systems, PACS)을 활용하여 의료 영상의 저장, 전송 등의 시스템의 기본적인 기능에 충실하게 설계되었다. 하지만 이러한 시스템은 진료 활동에 필요한 상위 응용 계층에서 내용 기반(content-based) 검색 기능이 지원되지 않고 있다[1]. 이는 의료 영상은 용량이 크고 비정형화된 시각적인 데이터이기 때문이다. 따라서 영상 처리를 통해 추출된 내용 정보들을 구조화하여 데이터베이스에 저장하고 검색 정확도가 높은 내용 기반 검색 기법을 연구 개발하는 것이 필요하다.

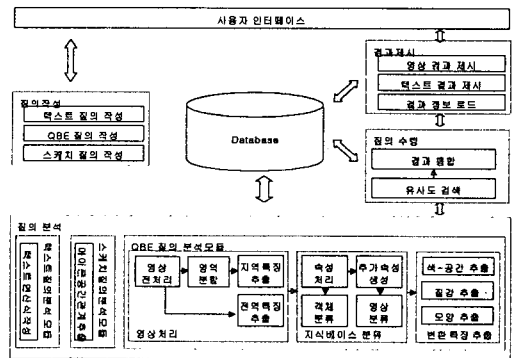
본 논문에서는 내용 기반 검색을 지원하는 의료 영상 검색 시스템을 소개한다. 의료 영상 검색 시스템에서는 영상 질의를 표현하는 질의어와 이를 처리하는 질의 처리 기술, 다양하고 복잡한 질의를 처리하여 검색 결과들을 통합하는 질의 통합 기법, 유사도 순으로 정렬된 질의 결과를 사용자에게 제시하는 질의 결과 제시 기법들이 요구된다. 이에 본 의료 영상 검색 시스템에서는 질의 유형들을 텍스트 키워드 질의, 예제에 의한 질의, 스케치 질의로 제안하고 질의 이미지의 내용기반 검색을 위해 색-공간, 질감, 모양 특징에 의한 유사 비교 기법을 사용하고 각각의 유사 비교 검색에 의해 생성된 결과 집합들을 통합하고 최종 결과 제시를 위해 복합 질의문 계획 생성 알고리즘을 이용하였다[2][3][4][5][6]. 의료 영상 검색 시스템에서 사용된 의료 영상은 MRI를 사용하였고, 인간의 뇌에 발생하는 질병을 대상으로 설계하였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 의료 영상 검색

시스템의 구조에 대해 설명하고 3장에서는 성능 실험과 결과 분석을 하고 마지막으로 4장에서는 결론을 맺는다.

### 2. 의료 영상 검색 시스템의 구조

의료 영상 검색 시스템은 질의 작성 모듈과 질의 분석 모듈, 질의 수행 모듈, 결과 제시 모듈로 구성된다. 다음 [그림 2.1]은 의료 영상 검색 시스템의 구조도이다.



[그림 2.1] 의료 영상 검색 시스템의 구조도

사용자 질의 작성 모듈은 웹 상에서 사용자가 질의를 작성하는 화면으로 간단한 텍스트 키워드 질의와 예제에 의한 질의(QBE, Query By Example), 사용자가 직접 질의 영상을 스케치할 수 있는 스케치 질의를 작성할 수 있는 모듈이고, 질의 분석 모듈은 사용자가 작성한 질의를 질의 유형에 따라 분석하고 그에 따른 과정을 수행한다. 질의 수행 모듈은 작성된 내부

본 논문은 한국과학재단 특정기초연구(과제 번호: R01-1999-000-00234-0) 내용의 일부임.

질의를 이용해서 데이터베이스내의 데이터를 검색한다. 의료 영상 검색 시스템에서 텍스트 질의에 대해서는 일반적인 완전 일치 검색 기법을 사용하고, QBE질의와 스케치 질의에 대해서는 유사도 검색 방안을 따라 검색을 수행한다. 검색된 각 질의 결과들은 하나의 결과 집합으로 통합된 후 유사도 순서로 정렬된다.

### 2.1. 질의 유형

의료 영상 검색 시스템에 지원되는 질의의 종류는 텍스트 키워드 질의와 사용자가 미리 정의된 질명 아이콘과 템플릿 영상을 이용하여 질의를 작성할 수 있도록 하는 스케치 질의, 그리고 QBE 질의가 있다.

텍스트 키워드 질의는 일반적인 검색 시스템에서 사용하는 키워드 중심의 질의의 방법이며, 환자의 뇌 MRI 이미지와 환자 정보를 데이터베이스로 구축한 상태에서 이용한다.

스케치 질의는 사용자가 정의하고자 하는 질의 영상의 객체 모양, 색, 화면 레이아웃 등을 직접 조작 방식으로 그려서 질의하는 방법이다. 하지만 뇌 MRI는 이미지를 구성하는 객체들의 속성 값이 미세하고 객체들 사이의 관계가 복잡하기 때문에 이미지 편집 도구들을 사용하여 질의 이미지를 작성하기가 매우 어려운 단점이 있다.

QBE 질의는 사용자가 시스템에 예제 영상을 제시하여 가장 유사한 영상을 검색하는 방법이다. QBE 질의 방법은 사용자가 영상의 특징을 알지 못해도 질의할 수 있고, 질의 작성에 개입할 필요가 없다는 장점이 있다. 특히 의료 영상이 담고 있는 정보를 그대로 표현하여 질의할 수 있어 스케치 질의 보다 정확한 질의를 할 수 있다. 본 논문에서는 색-공간, 질감, 모양 특징에 기반한 QBE 질의 처리 방법을 사용하였으며, 텍스트 질의와 함께 QBE 질의를 사용할 수 있는 복합 질의 형태를 지원한다.

### 2.2. 질의 처리 과정

텍스트 키워드 질의를 제외한 QBE 질의와 스케치 질의는 유사한 영상을 검색하는 내용 기반 검색 기법을 사용한다. 따라서 기존의 완전 일치 검색이 아닌 유사 검색을 사용한다. 즉, 의료 영상 검색 시스템의 질의 처리 기법은, 완전 일치 검색을 지원하는 관계형 데이터베이스 관리시스템의 질의 처리 기법과는 다르게 된다[7][8]. QBE 질의는 질의 정보에 대한 검증과정을 거친 다음 질의 이미지에 대한 전처리 및 기본 특징 추출, 지식베이스 분류, 변환 특징 추출 과정이 거쳐 검색에 사용될 비교 특징을 추출한다. 특징 추출 과정이 끝나면 유사 비교 과정을 수행한다. 본 절에서는 QBE 질의에 대한 복합 질의를 수행하는 질의처리 과정을 소개한다.

#### 1) 복합 질의를 수행하기 위한 전처리 과정

의료 영상 검색 시스템은 QBE 질의에 대한 복합 질의문을 수행하기 위해 영상 처리와 이미지 클래스를 분류하는 과정을 거친다. 영상 처리 과정에서는 질의 이미지의 기본 특징 정보들을 추출하고, 이상 객체 영역의 MBR(Minimum Boundary Rectangle)정보를 추출한다. 이미지 클래스 분류하는 과정에서는 의료 영상 지식베이스 모듈에 의해 질의 이미지가 속하는 클래스 정보를 추출한다.

#### 2) 변환 특징을 이용한 유사 비교 검색

QBE 질의는 복합 질의이기 때문에 질의 이미지에 대한 기본적인 특징만 가지고 복합 질의문을 수행하는데 어려움이 따른다. 이에 전처리 과정에서 추출된 기본 특징을 색-공간, 질감, 모양 특징으로 변환한 변환 특징을 이용하여 유사 비교 검색을 수행한다.

#### ① 색-공간 특징을 이용한 유사 비교 기법

이미지에 대한 색-공간 특징으로서 추출한 DCC(Dominant Color Composition)와 DCD(Dominant Color Distribution) 시그너처 특징을 이용한 유사 영상 검색은 우선 DCC 시그너처를 사용하여 유사한 대표 색 구성을 가지는 후보 유사 영상들을 추출한다. 그 후, 실제 색에 대한 유사도와 DCD 시그너처를 이용하여 색-공간 유사도를 구한 다음 유사도 점수가 높은 후보 유사 영상들 순서로 미리 정의한 임계값  $t$ 를 만족하는 영상들을 최종 결과로 출력한다[3].

#### ② 질감 특징을 이용한 유사 비교 기법

의료 영상의 질감 특징을 이용한 검색 과정은 3단계 웨이블릿 변환을 통해, 원본 영상으로부터 방향성을 가진 특징 집합을 얻어서 검색을 하게 된다. 먼저 저주파 성분을 가진 영상에 대하여 미리 계산한 각 객체의 저주파 영상의 특징과 유사도를 계산한다. 이후에 유사성이 높은 객체의 나머지 특징들에 대하여 대각선 성분, 수평 성분, 수직 성분에 대한 특징치에 따른 객체를 검색하고 검색의 유사도 계산에 따라 순서대로 영상을 출력한다[5].

#### ③ 모양 특징을 이용한 유사 비교 기법

모양 특징을 이용한 유사 비교 검색을 위해 제귀적 윤곽선 근사 알고리즘에서 얻은 우세 점들을 무게 중심과 우세 점들과의 거리  $r$ 과 각도  $\theta$ 로 정규화하여 모양의 지역 특징을 추출한 Shape\_Similarity 유사성 검색 알고리즘을 사용하였다[4].

#### ④ 결과 통합

의료 영상 검색 시스템에서 QBE와 같은 복합 질의를 처리하기 위해 퍼지 집합 이론을 이용한 복합 질의문 계획 생성 알고리즘을 이용하였다. 질의 최적화를 거친 복합 질의문은 각각의 논리합과 논리 곱 복합 질의문으로 논리적으로 동등하면서 같은 결과를 가지는 질의로 대체되어 수행된다. 논리곱 복합 질의이면 SSCC(Similarity Search for Conjunction Combination query) 알고리즘을 이용하여 검색하고 논리합 복합 질의이면 SSDC(Similarity Search for Disjunction Combination query) 알고리즘을 이용하여 검색한다[6]. 따라서 색-공간, 질감, 모양과 같은 변환 특징에 대한 단순 질의가 포함된 QBE 질의는 복합 질의문 생성 계획 알고리즘에 따라 수행되어 검색된 결과 집합은 각각의 특징에 대한 유사 비교 검색 수행 후 각각의 결과 집합들을 통합한 최종결과 집합과 논리적으로 동등한 결과 집합을 생성하게 된다.

## 3. 성능 실험과 결과 분석

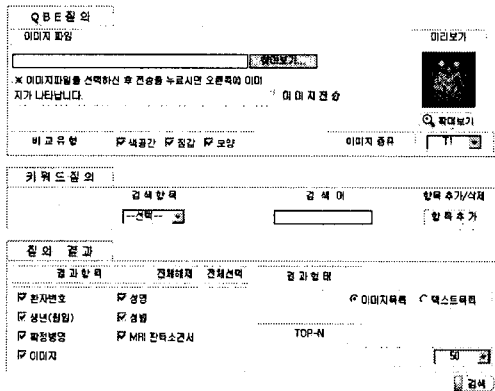
의료 영상 검색 시스템은 기본적으로 인터넷 환경에서 웹브라우저를 통해 사용자가 접속하여 텍스트 키워드 질의, QBE 질의, 스케치질의를 사용할 수 있도록 설계되었다. 의료 영상 데이터베이스 질의를 위한 사용자 인터페이스를 HTML, ASP, Java Script, ActiveX 및 Java Applet으로 구현하였다. 웹 서버는 윈도우즈 2000 서버 운영체제에서 IIS 5.0을 기준으로 구현하였다.

#### 1) 예제에 의한 질의(QBE) 화면

QBE 질의 화면의 구성은 이미지 선택 기능, 특징정보 선택(색상, 질감, 모양, 공간), 키워드 입력, 결과 제시 유형/항목 선택, 페이지당 결과 수, 유사도 정렬 기준으로 구성되어 있고 화면은 [그림 3.1]과 같다.

텍스트 키워드 질의와 QBE 질의를 동시에 질의할 수도 있고 텍스트 키워드 질의만 수행을 하거나 QBE 질의만 수행할 수 있게 지원한다. QBE 질의에 대한 검색 과정으로 색-공간, 질감, 모양 특징 추출을 위해 전처리 과정, 영역 분할 과정, 지

역 특징 추출 과정, 객체 매핑 과정, 클래스 분류 과정을 거친 후, 이 해당 특징들을 이용하여 데이터베이스내 후보 집합을 검색한다. 이 때, 클래스에 속한 이미지 검색, 검색 결과 집합에 대해 키워드 검색을 수행한다. 다음으로 후보 집합들 중에서 각 특징 비교(유사도 계산)를 진행함으로써 최종 집합 검색을 수행한다. 여기서 이미지 불러오기 기능과 의료 이미지 특징 추출 기능은 VC++로 작성하였고 웹상에서 지원할 수 있도록 ActiveX DLL로 변환하여 ASP내에 ActiveX를 포함하여 구현하였다.



[그림 3.1] QBE 질의와 텍스트 질의 화면

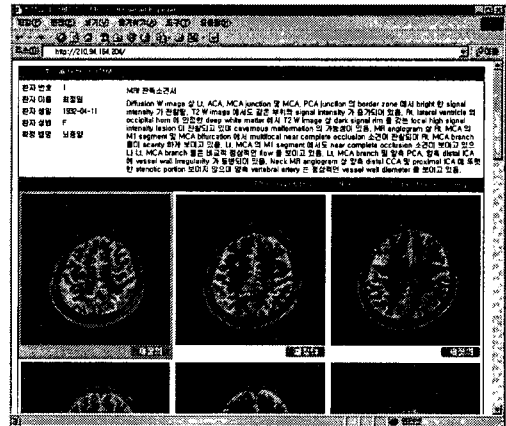
2) 결과 제시 화면

텍스트 키워드 질의, QBE 질의, 스케치 질의는 각 질의 분석 모듈과 질의 수행 모듈을 통해 처리되며, 검색 결과는 결과 제시 화면을 통해 유사도 순서대로 정렬되어 디스플레이 된다. 웹 상에서의 결과 제시 화면은 [그림 3.2]와 같다.

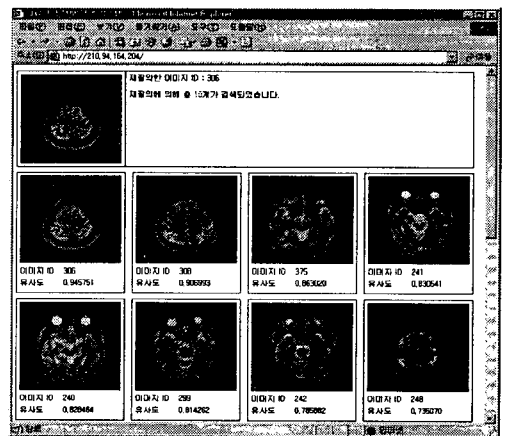


[그림 3.2] 결과 제시 화면

결과 제시 화면에 나타난 결과들에 대한 자세한 정보는 결과 그림 밑의 상세보기 버튼[그림 3.3]을 누르면 볼 수 있다. 또한 [그림 3.4]와 같이 제시된 결과를 이용해서 특정 결과 이미지와 유사한 이미지를 재질의할 수 있는 재질의 기능도 제공된다.



[그림 3.3] 상세 결과 보기 화면



[그림 3.4] 재질의 결과 화면

3) 의료 영상 검색 시스템의 성능 평가

의료 영상 검색 시스템의 성능을 평가하기 위해 QBE 질의를 구성하는 각 단순 질의인 색-공간, 질감, 모양 질의에 대한 정확도 실험을 각각 수행하였다. 이들 결과를 통합하는 결과 통합 모듈에 대한 정확도와 재현율, 수행 시간을 평가하였다.

DCC와 DCD 시그니처를 이용한 이미지의 색-공간 특징 표현 방법의 타당성을 실험하기 위해 이미지를 32×32의 부영역으로 나누어 추출한 DCC와 ZonedDCD 시그니처 조합에 대해 QBE 질의와 스케치 질의를 수행하였다. QBE 질의의 결과에 대해 평균 재현율과 평균 정확도는 0.91로 스케치 질의에 대한 평균 재현율은 0.89, 평균 정확도는 0.56으로 측정되었다. 스케치 질의의 특성상 정확도가 낮게 측정되는 것을 보였다.

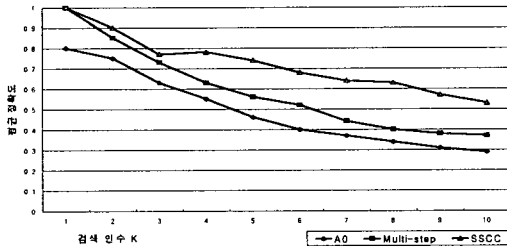
질감 질의의 성능은 구축한 이미지 데이터베이스에서 임의의 이미지 50개를 선택하여 QBE 질의에 대한 정확도를 측정하여 [표 1]과 같은 결과를 얻었다.

검색 방법	검색 정확도
선별된 전역 질감 특징을 이용한 검색	0.61
전역 질감 특징을 이용한 검색	0.61
선별된 지역 질감 특징을 이용한 검색	0.63
지역 질감 특징을 이용한 검색	0.59
전역-지역 질감 특징을 이용한 검색	0.65
선별된 전역-지역 질감 특징을 이용한 검색	0.71

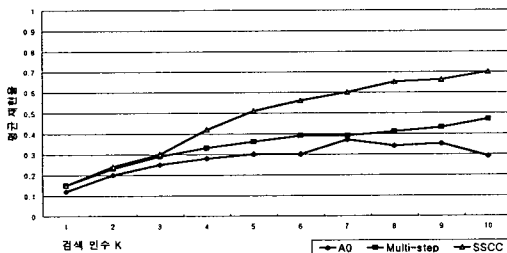
[표 1] 질감 특징을 이용한 영상 검색 결과

[표 1]에 의하면 전역 질감특징을 이용하는 것은 전체 10개의 특징을 모두 사용한 것과 10개중 2개를 사용한 것 모두 0.61의 검색 정확도를 보여주고 있다. 이는 의료 영상의 뇌 MRI 이미지의 특성상 전체 영역으로 특징을 추출했을 경우 검은 바탕에 회색과 흰색의 뇌 영상이 모두 비슷하게 보이기 때문이다.

본 논문에서 사용된 복합 질의문 계획 생성 알고리즘에 대한 성능 평가를 위해 Fagin의  $A_0$  알고리즘, Suya의 Multi-step 알고리즘과 비교, 분석하였다[9][10]. QBE 질의를 복합 질의로 사용하였으며 복합 질의를 이루는 단순 질의는 색-공간 질의, 질감 질의, 모양 질의를 사용하였다. 그리고 복합 질의문 계획 생성 알고리즘의 SSCC 알고리즘은 논리합 질의로만 구성된 복합 질의를  $A_0$  알고리즘을 사용하였을 때와 논리적으로 같기 때문에 비교 분석에서 제외하였다. 각 알고리즘에 대해 검색 결과의 수,  $k$ 를 증가시키면서 정확도와 재현율을 비교하였으며 이미지 질의 수를 증가시키면서 10개의 이미지 질의에 대한 정확도, 재현율을 비교하였다. [그림 3.5]는 10개의 예제 이미지 질의에 대해 검색인수  $k$ 의 수를 증가시키면서 측정한 세 알고리즘에 대한 평균 정확도를 나타내며 [그림 3.6]은 평균 재현율을 나타내고 있다.



[그림 3.5] k의 변화에 따른 10개의 이미지 질의에 대한 평균 정확도



[그림 3.6] k의 변화에 따른 10개의 이미지 질의에 대한 평균 재현율

정확도 비교에서 보면 복합 질의문 계획 생성 알고리즘의 SSCC 알고리즘을 사용하였을 때 평균 정확도가 0.46이고  $A_0$  알고리즘의 평균 정확도가 0.32, Multi-step 알고리즘의 평균 정확도가 0.38로 복합 질의문 계획 생성 알고리즘의 SSCC 알고리즘이 가장 높음을 알 수 있다. 재현율에 있어서도 SSCC 알고리즘이 평균 재현율 0.59, Multi-step 알고리즘이 0.5,  $A_0$  알고리즘이 0.41로 SSCC 알고리즘이 가장 높음을 알 수 있다.

#### 4. 결론

본 논문에서는, 의료 영상을 질의 이미지로 사용하여 질의하는 지식베이스를 활용한 내용 기반 검색이 지원되는 의료 영상 검색 시스템의 구현 내용을 소개하였다. 의료 영상 검색 시스템은 질의 이미지의 내용기반 검색을 위해 색-공간, 질감, 모양 특징에 의한 유사 비교 기법을 복합 질의문 계획 생성 알고리즘을 통하여 수행하도록 설계 및 구현하였다. 성능 실험 및 결과 분석에서 시스템을 구성하는 각 단순 질의에 대해 성능평가와 결과 통합 모듈의 성능을 평가함으로써 의료 영상 검색 시스템의 성능을 살펴보았다.

본 의료 영상 검색 시스템은 진료할 때 필요한 정보를 검색하거나 질병에 대한 객관적인 영상 데이터를 요구하는 시스템에 적용할 수 있다. 본 의료 영상 검색 시스템은 인간의 뇌에 발생하는 질병에 대해서만 국한하여 다루었기 때문에 더 많은 질병에 대해 적용될 수 있는 연구 과제가 남아있다.

#### [참고문헌]

- [1] 동국대학교, 보건복지부 한국보건사회연구원, "PACS 조사 기술서", 1999
- [2] 이낙훈, 엄기현, "의료 영상 검색을 위한 아이콘 기반의 스케치 질의 작성 방안", 한국정보과학회, 가을 학술발표논문집, 27권2호, pp 122~124, 2000.
- [3] 조성택, 엄기현, "색-공간 이미지 질의를 위한 시그너처 기반 인덱싱", KDBC 학술발표논문집, pp.7-16, 2000
- [4] 김영태, 엄기현, "이미지 객체의 모양 특징에 기반한 검색 방안", 한국멀티미디어학회 춘계 학술발표 논문집, 제 4권 제 1호, pp.53~56, 2001.
- [5] 구혜영, 엄기현, "웨이블릿 변환을 이용한 적응적 뇌영상 검색방안", 멀티미디어학회 추계 학술발표논문집, pp.447-452, 2001.
- [6] 박미화, 엄기현, "내용 기반 영상 검색을 위한 복합 질의문 계획 생성 기법", 한국정보과학회논문지:데이터베이스, 27권 4호, pp 562~571, 2000.
- [7] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, W. Equiz, "Efficient and Effective Querying by Image Content", Journal of Intelligent Information System(JIIS), 3(3), pp.231-262, July 1994.
- [8] 이석호, 송병호, 김범수, "멀티미디어 데이터베이스 관리 시스템에서의 내용기반 검색 기법에 관한 연구", 한국정보과학회 데이터베이스연구회지, 11권 4호. pp.102-119. 1995.
- [9] Ronald Fagin, "Combining fuzzy information form multiple systems", in 15th ACM Symposium on Principles of Database systems, pp.216-226, June 1996.
- [10] Surya Nepal and M. V. Ramakrishna, "Query Processing Issues in Image(multimedia) Databases", in Proceedings of Data Engineering, March 23-26, Sydney, Australia, pp.22-29, 1999.