

유사객체 검색을 지원하는 협력 의료정보 시스템 설계

원정임⁰ 박형주 안상원 윤지희
한림대학교 컴퓨터공학과
(jiwon, hjpark, swan, jhyoon)@iris.ce.hallym.ac.kr

Design of a Cooperative Medical Information System which Supports Similarity-Based Object Retrieval

Jung-Im Won⁰ Hyung-Joo Park Sang-Won An Jee-Hee Yoon
Dept. of Computer Engineering, Hallym University

요 약

문자 정보 및 X-Ray, MRI, CT 등과 같은 의료영상 정보를 취급하는 의료정보 시스템에서 유사객체 검색을 지원하는 협력 의료정보 시스템의 설계에 대하여 논한다. 이를 위해 객체간 의미적 관련성을 기반으로 한 유사도 자동 추출 방식 및 지식베이스 구성 방식을 제안하고 이를 활용한 유사객체 검색에 대하여 논한다. 특히 의료영상을 객체 값으로 갖는 경우 객체간 유사도는 영상처리의 특징추출 방식에 의해 추출된 영상내에 출현하는 공간 객체의 위치, 면적, 둘레, 공간 객체간의 위상 관계 등의 공간 속성을 이용한다. 여기서 공간적 위치 에 근거한 유사도는 공간 위치를 대표하는 Hilbert 값의 분포와 빈도를 토대로 계산한다.

1. 서론

환자의 병력, 진단에 따른 처방 및 약제 정보, 진료비 계산업무 등과 같은 다양한 의료 정보의 증가로 인하여 날로 심각해져 가는 병원 업무의 효율을 향상시키기 위해 의료분야에서 최근 의료정보 시스템(Medical Information System)들이 개발되어 사용 중에 있다. 이들 시스템들은 크게 환자와 관련된 모든 정보를 관리하는 HIS(Hospital Information System), 방사선과에서 취급하는 의료정보를 관리하는 RIS(Radiology Information System), 각종 의료영상 정보를 수집, 저장, 전송하는 PACS(Picture Archiving and Communication System) 등으로 나눌 수 있다[1]. 특히, PACS는 X-Ray는 물론 MRI, CT, 초음파 등 최첨단 의료영상 진단 장비를 통해 얻은 영상 데이터를 디지털 신호로 바꾸어 컴퓨터에 저장하고 고속 네트워크를 통해 실시간으로 의사에게 전송하여 진단할 수 있게 해준다.

의료정보 시스템은 문자열이나 숫자와 같은 단순한 데이터 타입에 기반한 문자 정보 및 MRI, CT와 같은 의료영상 정보, 음성 정보 등의 다양한 형태의 의료정보를 취급하며, 질의 형태도 이들이 혼합된 복합적인 형태를 가진다. 또한 이들 시스템을 이용하는 사용자중도 일반 사용자, 병원관리자, 전문의 등 다양하며 전문의의 경우 전문의의 진단 결정을 보조할 수 있는 환자들의 병리학 적 통계자료나 시간적 흐름에 따른 질병의 전이상태를 예측하기 위한 환자의 유형별 분류 등 보다 전문적인 의료정보를 다룰 수 있는 질의유형을 제공받고자 할 것이다. 그러나 기존의 데이터베이스 시스템의 접근 기법, 질의 언어, 데이터 모델 등은 이를 제공하기 힘들다.

따라서 본 논문에서는 문자 정보를 기반으로 하는 HIS의 기능을 보완하면서 의료영상 정보를 체계적으로 저장, 관리, 운용하기 위한 유사객체 검색을 지원하는 협력 의료정보 시스템(Cooperative Medical Information System)의 설계에 대하여 논한다. 협력질의 응답(Cooperative Query Answering)[1,2,3]이란 사용자가 스키마 구조 및 문제 영역에 대한 정확한 기술을 하지 못한 경우나 질의 검색 결과 데이터베이스 상에 해당 해가 존재하지 않는 경우 사용자와의 협력과정을 거쳐 주어진 사용자의 질의 대상 영역을 확대 처리하여 질의에 대한 정확한 해 외에도 보다 일반화된 해의 집합, 관련 정보 등을 유연성 있게 제공할 수 있는 기능을 말한다.

이를 위해 본 논문에서는 문자 정보 및 의료영상 정보를 토대로 자동 추출되어 내부 지식베이스로 저장, 관리 되는 유사도 정보[4]를 활용한 유사객체 검색(Similarity-Based Object Retrieval)방식을 제안한다. 또한 사용자의 질의 응답 과정을 보조하여 편리하고 유연성 있는 질의응답 과정을 지원하기 위한 의료정보 시스템의 설계에 대하여 논한다.

2. 시스템 설계

문자 정보, 의료영상 정보, 음성 정보 등 다양한 형태의 의료정보를 통합, 관리, 운영하기 위해 제안된 본 의료정보 시스템은 의료영상 처리 모듈, 협력질의 처리 모듈, 지식 관리 모듈, 클래스 관리 모듈, 네트워크 관리 모듈 등으로 구성된다. 그림 1에 제안된 시스템의 전체적인 구성도를 보인다.

[1] 의료영상 처리 모듈

의료영상 처리 모듈에서는 의료영상 장비인 CT, MRI 등을 통해 얻어진 이미지를 의료 영상 표준안인 DICOM 포맷으로 획득하고, 획득된 의료영상의 특징(면적, 위치, 둘레, 위상관계 등)을 추출하여 데이터베이스에 저장하는 모듈이다. 또한 Zoom In/Out 기능, 기하학적 계산 기능, 메모 기능, 포맷 변환 기능 등을 제공한다.

[2] 협력질의 처리 모듈

주어진 사용자의 질의를 처리하는 모듈로 의료정보 시스템을 사용하는 다양한 사용자들의 편의성을 도모하기 위해 간편한 질의 입력 기능, 잘못 입력된 질의의 자동 수정 기능, 유사검색을 지원하는 질의 응답 기능 등을 제공한다.

[3] 지식 관리 모듈

데이터베이스 시스템에 저장된 문자 정보 및 의료영상 정보를 토대로 계산된 유사도 정보와 정의된 클래스의 스키마 정보를 지식베이스로 관리하는 모듈이다. 이들 자동 추출된 지식들은 협력질의 처리시 활용되며 전문가에 의해 편집되어 질 수 있다.

[4] 클래스 관리 모듈

사용자가 정의한 클래스의 서브클래스 관계성 및 속성, 도메인 정보 등을 포함한 스키마 정보를 그래프 출력하거나 이들에 대한 삽입, 삭제, 수정 작업등을 지원 하는 모듈이다.

[5] 네트워크 관리 모듈의 의료정보 시스템은 고속의 네트워크를 통해 Client/Server 시스템을 구축, 운영한다. 네트워크 관리 모듈들은 질의 수행 결과 검색된 문자 정보 및 의료영상 정보를 전송하는 모듈로 일반적으로 영상, 데이터, 음성 등 많은 양의 데이터를 빠르게 사용자에게 전송하기 위해 고속 네트워크를 사용한다.

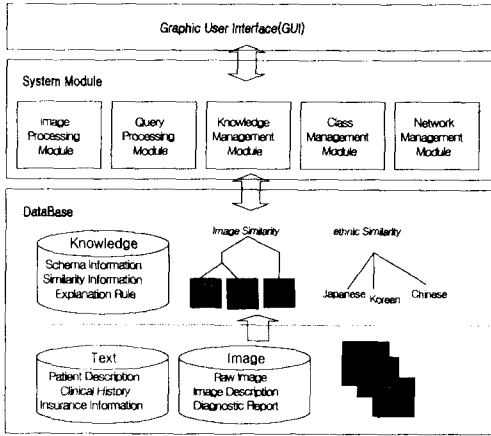


그림 1. 시스템 구성도

3. 협력질의

본 논문에서는 협력질의 응답을 지원하기 위한 지식 베이스로 객체 유사도 정보를 활용한다. 이를 위한 논리적 데이터 모델로 객체지향 데이터모델[5]을 이용하며 객체지향에서 객체는 클래스 객체, 인스턴스 객체, 또한 단순 객체, 복합객체 등으로 구분할 수 있다. 여기에서는 우선 이들 각 객체에 대하여 그 특성에 근거한 객체 유사도 계산 방식을 제안하고 이를 활용한 유사객체 검색에 대하여 논한다.

3.1 객체 유사도

객체가 단순 문자열을 속성 값으로 갖는 인스턴스 객체인 경우 객체간 유사도는 객체가 갖는 속성 값에 근거한 객체의 의미적 관련성을 토대로 계산한다. 즉, M 이라는 클래스가 m 개의 속성을 갖는 경우 임의의 속성 값을 갖는 객체를 유사 분류하기 위하여 나머지 $m-1$ 개의 속성의 속성 값을 이용한 유사도 계산방식을 채택한다[3].

예를 들어 그림 2와 같은 예제 데이터베이스상에서, Patient 클래스의 patientID에 대한 객체간 유사도는 속성 patientID를 제외한 나머지 속성 name, age, address, room, doctor, sickness, images, bloodType, height, weight 등의 속성 값을 고려한 유사도 계산을 행한다. 이때 시스템 설계자는 유사도 계산시 특정 속성들만을 선택, 이용할 수 있으며 각 속성의 중요도에 따라 가중치를 부가할 수 있다. 또한 경우에 따라서는 유사도 계산시 자동추출 방식이 아닌 설계자의 수작업에 의한 편집방식을 사용하기도 한다. 예를 들어 문자열을 속성 값으로 갖는 속성 address와 sickness의 유사도는 이를 제외한 나머지 다른 속성들의 값에 근거한 계산 방식보다는 설계자의 전문 지식에 의존한 수작업 방식을 사용한다. 특히 속성 sickness의 경우는 환자들의 증상, 통증 부위, 의사의 소견 등의 정보를 토대로 자동 추출되어 정확한 임상진단을 위해 전문의와 시스템 설계자의 협의에 편집과정을 거치는 것이 일반적이다. 그림 3에 patientID에 대한 계산된 객체간 유사도를 보인다.

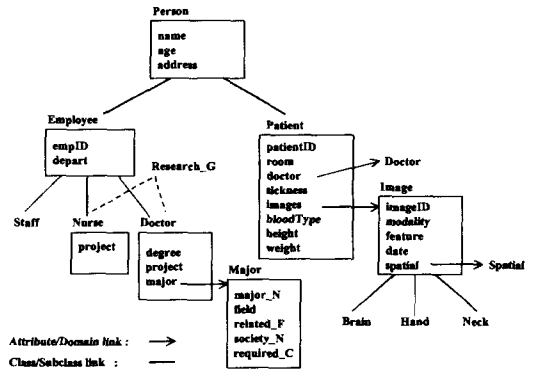


그림 2. 예제 데이터베이스

그러나 이 방식은 속성 값이 불연속적인 경우에는 적합하나 속성 age, height 등과 같이 연속적인 숫자 값을 갖는 객체의 경우에는 적합하지 못하다. 따라서 연속적인 값을 갖는 객체의 경우에는 객체의 분포와 빈도 등을 고려한 객체간 유사도를 계산하는 것이 효율적이다[1,3].

표 1. Patient 클래스의 인스턴스 예

OID	name	age	address	patientID	room	doctor	sickness
OID1	김소희	15	춘천	#1001	102호	OID200	뇌종양
OID2	이은수	10	서울	#1002	103호	OID203	안암
OID3	홍선주	10	서울	#1003	102호	OID200	뇌종양
OID4	최희철	20	대구	#1004	101호	OID202	뇌종양
OID5	장혁진	13	서울	#1005	103호	OID203	안암
OID6	박진영	10	춘천	#1006	102호	OID200	뇌종양
OID7	우희진	25	광주	#1007	105호	OID210	자궁암
OID8	방진숙	30	부산	#1008	105호	OID210	자궁암
OID9	김용준	16	대전	#1009	103호	OID203	안암
OID10	유성중	42	인천	#1010	101호	OID202	뇌종양

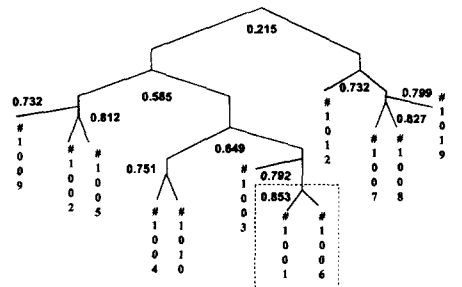


그림 3. patientID의 객체 유사도

객체지향 데이터 모델에서 객체는 문자열, 숫자 등 다양한 데이터 타입을 갖는 복수 개의 속성들의 집합으로 이루어진다. 즉, 표 1에 제시된 바와 같이 Patient 클래스의 인스턴스 객체는 단순 문자열을 값으로 갖는 속성 name, address, sickness 등과 숫자 값을 갖는 속성 age, weight, height 그리고 OID를 속성 값으로 갖는 속성

doctor가 복합된 형태를 갖는다. 따라서 본 논문에서는 객체지향에서 객체를 대변하는 OID에 대한 유사도는 이 둘을 구성하는 각각의 속성이 갖는 속성 값에 따라 서로 다른 방식에 의해 자동 추출된 변수 개의 유사도 정보를 혼합한 방식을 이용한다. 즉, 클래스 Patient의 OID에 대한 유사도는 문자열 값에 의한 patientID의 유사도와 숫자 값에 의한 age, weight, height의 유사도 그리고 시스템 계층에 주속업한 address와 sickness의 유사도 등을 혼합, 적용하여 계산한다.

특히, 객체가 MRI, CT와 같은 의료영상을 객체 값으로 갖는 경우 의료영상 객체간 유사도는 영상의 특징을 결정하는 속성 즉, 영상내의 공간 객체의 위치, 모양, 크기, 공간 객체간의 상호 위상 관계 등의 공간 속성을 활용할 수 있다. 이들 공간 속성은 대부분 획득된 의료영상에서 영상의 특징을 추출하기 위한 기하학적 계산 방식에 의해 산출된 일련의 숫자들을 속성 값으로 한다. 따라서 본 논문에서는 의료영상 객체간 유사도 계산시 숫자를 객체의 값으로 갖는 유사도 계산 방식을 확대, 적용한다. 즉, 의료영상내에 출현하는 공간 객체의 분포와 빈도 등을 고려한 유사도 계산 방식을 채택한다.

표 2에 그림 2의 예제 데이터베이스에서 Brain클래스의 공간 속성 Spatial의 도메인 클래스의 구체적인 인스턴스 예를 보인다. 여기서 객체는 Brain영상에서 특징 추출된 종양객체로 종양의 위치, 면적, 둘레, 모양 등의 공간 속성을 갖으며 각각의 속성은 속성이 갖는 숫자 값

표 2. Spatial클래스의 인스턴스 예

OID	hvalue	x	y	area	perimeter	circularity
OID100	460	12	22	12	13	1.32
OID101	846	21	15	24	19	1.4
OID102	513	17	16	12	13	1.33
OID103	877	18	9	11	11	1.52
OID104	298	7	23	24	19	1.87
OID105	115	11	6	17	16	1.34
OID106	155	14	11	22	18	1.44
OID107	283	6	19	25	20	1.46
OID108	857	16	13	19	19	1.47
OID109	503	13	18	15	14	1.29

의 분포와 빈도를 근거로 계산된 유사도를 갖는다. 즉 종양의 위치에 근거한 유사도, 면적에 근거한 유사도, 둘레에 근거한 유사도 등이다. 또한 이들 계산된 각각의 유사도를 혼합, 적용하여 공간 객체의 OID에 대한 유사도를 계산할 수 있다.

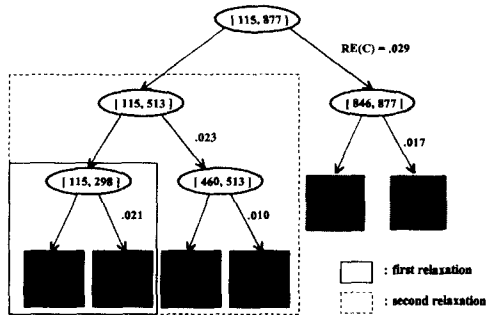


그림 4. hvalue의 객체 유사도

그림 4는 표 2에서 종양의 위치를 근거로 계산된 유사도이다. 본 논문에서는 이때 공간 객체의 위치를 대변하는 주색인의 탐색키인 Hilbert값을 이용한다[4]. Hilbert 곡선은 다차원 공간상에서의 객체간 근접성을 유지시켜 주는 공간 순서화 곡선의 일종으로 이를 이용하여 2차원

공간상의 종양의 위치를 나타내는 속성 x, y의 값을 1차원의 Hilbert값을 갖는 속성 hvalue로 매핑한 후 hvalue의 빈도와 분포에 근거한 유사도 계산 방식을 이용한다.

3.2 유사 객체 검색

유사객체 검색은 주어진 사용자의 질의에 대하여 해가 존재하지 않는 경우나 사용자가 의도적으로 유사 객체를 검색하기 위해 명시한 유사 연산자(@)를 근거로 사용자와 시스템간의 협력과정을 통하여 주어진 질의 대상영역을 확대 처리함으로써 이루어진다[4]. 다음 (예제 1), (예제 2), (예제 3)에 유사객체 검색에 대한 예를 보인다.

```
(예제 1) SELECT *
FROM Patient
WHERE patientID=@ '#1001' AND age = 10
AND sickness = '뇌종양';
```

주어진 질의는 환자번호가 '#1001'인 10살된 '뇌종양' 환자를 검색하기 위한 질의이다. 명시된 유사연산자를 근거로 patientID '#1001'은 '#1006'으로 확대 처리될 수 있다. 이때 그림 3의 patientID의 객체 유사도를 활용한다.

```
(예제 2) SELECT Y.Image
FROM Brain X, Y
WHERE X.spatial.hvalue @= Y.spatial.hvalue
AND Y.feature = '종양';
```

(예제 2)는 스크린 상에서 선택한 부위 근처에 종양이 있는 환자의 의료영상을 검색하기 위한 질의로 그림 4에서 제시된 유사도 구조를 활용하여 질의 대상 영역을 확대 처리한다. 만약 현재 스크린 상에서 선택한 위치의 hvalue가 130이라면 이를 그림 4의 유사도 구조를 근거로 115~298로 질의 대상 영역을 확대 처리할 수 있다.

```
(예제 3) SELECT *
FROM Doctor@
WHERE major.major_N = '산부인과'
AND project = '자궁경부암';
```

(예제 3)는 유사 연산자@ 클래스명에 명시된 경우로 이때 클래스 객체간 유사도는 클래스의 스키마 정보에 근거하거나 시스템에서 내부적으로 유사객체 클래스를 임의로 선정할 수 있다. 예제는 질의대상 범위가 Doctor 클래스에서 시스템이 선정한 Research_G를 통해 Nurse 까지 확대 처리된다.

4. 결론

본 시스템은 현재 개발 중으로 Windows NT 상에서 의료영상 처리기, 협력 질의 처리기, 네트워크 관리기 등 협력 질의 용달 기능을 수행할 모듈별 프로그램을 Visual C++를 사용하여 개발 중에 있다.

5. 참고 문헌

[1] Chu, W.W., Cardenas, A.F., Taira, R.K., "KMED : A Knowledge-Based Multimedia Medical Distributed Database System", Information Systems, Vol.19, No.4, pp.33-55, 1994.
 [2] Motro, A., "FLEX: A Tolerant and Cooperative User Interface to Databases", IEEE Trans. on Knowledge and Data Engineering, Vol. 2, No. 2, pp. 231-246, 1990.
 [3] Chu, W. W., Chiang, K., Hsu, C. C., Yau, H., "An Error-based Conceptual Clustering Method for Providing Approximate Query Answers", Communication of ACM, 1996.
 [4] 원정업, 홍동완, 손영일, 윤지희, "근사질의응답을 지원하는 협력지리정보시스템의 설계 및 구현", '98동계제이더베이스 학술대회 논문집, pp. 1-6, 1998.
 [5] Kim, W., Kifer, M., "Querying Object-Oriented Databases", ACM SIGMOD Conf. on Management of Data, pp. 393-402, 1992.