

혼합형 질의 방법에 의한 온톨로지 기반 유물 검색 시스템

천 현 재*, 백 승 재**, 이 홍 철***

Ontology based Retrieval System for Cultural Assets Using Hybrid Text-Sketch Queries

Hyeon-Jae Cheon *, Seung-Jae Baek **, Hong-Chul Lee ***

요 약

최근 각종 정보가 증가함에 따라 효율적인 관리를 위해 정보 검색에 관한 연구가 더욱 활기를 띠고 있다. 현재 웹 환경에서 운영되고 있는 국내 유물 검색시스템의 경우 대부분이 키워드 기반의 텍스트 검색 방식을 채택하고 있다. 이러한 텍스트 검색 방식은 그 유물에 대한 정확한 이름이나 키워드를 질의자(user)가 미리 알고 있어야 한다. 하지만 검색대상에 관한 정보가 부족하여 키워드가 모호하거나 단순히 형상에 관한 기억만 있을 경우에는 검색이 쉽지 않았다. 이 논문에서는 기존 유물 검색 시스템의 문제점을 해결하기 위해 온톨로지 기반의 텍스트 질의와 사용자 스케치 이미지 질의를 사용하는 자바 기반의 혼합형 유물 검색시스템을 제안한다. 이 시스템은 국내 유물들을 대상으로 사용자가 기억하고 있는 유물에 관한 정보의 형태(텍스트, 형상 등)에 따라 다양한 질의방법을 제공하며, 검색결과 내에서 온톨로지 의미관계를 이용한 추가검색이 가능하다.

Abstract

With the rapidly growing information, the research on the efficient information retrieval is increasing. Most of the retrieval systems for domestic cultural assets on the web have adopted a keyword-based search method. Those systems have required users to know the exact information about cultural assets such as name, keyword, etc. However, it is not easy to search the cultural assets with little information or only a remembrance of the shape. In this paper, we propose the retrieval system for cultural assets using both ontology-based and sketch-based search method to solve the problems of existing systems. Our retrieval system allows users to use both text and sketch for a query regardless of the type of information about cultural assets and to search in results using the ontology.

▶ Keyword : 보검색(Information Retrieval), 온톨로지(Ontology), QBS(Query By Sketch), 검색엔진(Search Engine), 문화재(Cultural Assets)

• 제1저자 : 천현재

• 접수일 : 2005.08.25, 심사완료일 : 2005.10.20

* 고려대학교 산업시스템정보공학과 박사과정, ** 고려대학교 산업시스템정보공학과 석사과정

*** 고려대학교 산업시스템정보공학과 교수

1. 서론

최근 컴퓨터 및 네트워크 통신의 발달로 많은 양의 정보를 다양한 형태로 접하고 있으며, 이러한 방대한 양의 데이터를 효율적으로 관리하기 위해 정보 검색에 관련된 연구들이 진행되어 왔다. 현재 웹 환경에서 운영되고 있는 국내 유물 검색 시스템의 경우 대부분이 키워드 기반의 텍스트 검색 방식을 사용하고 있으며 이는 사용자의 편의성 면에서 볼 때, 다양한 질의 방법을 제시하지 못하고 있다. 이러한 텍스트 검색 방식은 사용자에게 검색 대상에 관한 유물의 정확한 이름이나 키워드를 요구하기 때문에 만약 사용자가 유물에 관한 정보가 부족하여 키워드가 모호하거나 단순히 형상에 관한 기억만 있을 경우에는 검색이 쉽지 않았다. 또한 텍스트 검색은 유물의 전체적인 의미를 몇 가지의 단어만으로 표현해야 하는 방식이라 대상에 대한 명백한 의미(semantics)를 표현하기가 쉽지 않다. 따라서 이 논문에서는 질의어의 의미와 관계를 명확히 파악할 수 있는 온톨로지 기반의 텍스트 검색과 스케치(sketch) 기반 이미지 검색을 동시에 사용하는 시스템을 제안하고자 한다.

온톨로지(Ontology)란 W3C(World Wide Web Consortium)에서 정한 시맨틱 웹(Semantic Web)의 핵심기술로서 웹상의 정보에 잘 정의된 의미를 부여함으로써 사람뿐만 아니라 컴퓨터도 쉽게 문서의 의미를 이해할 수 있도록 하여 정보의 검색, 해석, 공유 등의 자동화를 하기 위한 목적으로 제안되었다. 온톨로지 기술은 XML(Extensible Markup Language) 기반의 마크업(markup) 언어로 표현되며 그 종류에는 RDF(Resource Description Framework)/RDFS(schema)와 DAML+ OIL(DARPA Agent Markup Language + Ontology Inference Layer), OWL(Web Ontology Language)이 있으며 현재 W3C에서 OWL을 온톨로지 언어의 표준으로 정하고 있다[1].

온톨로지 기술은 문서정보의 메타데이터를 마크업 언어 형태로 정의하여 스크립트에 기술함으로써 컴퓨터가 그 의미를 알 수 있게 한다. 이는 정보의 검색, 해석, 공유 등의 자동화를 가능하게 한다는 점에서 큰 장점이 있다. 온톨로지를 이용한 문서 정보의 표현은 기본적으로 트리플(triple) 형태의 표현 방법이 있으며 트리플 형태의 메커니즘을 확장

시켜 클래스 자원의 기술과 클래스의 다양한 관계를 기술함으로써 메타정보에 대해서 명확하게 기술할 수 있다. 이러한 온톨로지 기술을 활용하면 Web portal, Multimedia collections, Design documentation, Intelligent agents, Ubiquitous computing 등 활용범위가 매우 넓다[2].

이 연구에서 제안하는 온톨로지 기반 검색방법은 검색대상의 메타데이터를 온톨로지 언어에 기술함으로써 질의자의 질의목적에 가깝게 결과를 추론하는 것이다. 이는 기존의 단순한 키워드 매칭을 통한 데이터베이스 질의방법과 달리 온톨로지 지식베이스를 통해 질의정보의 의미를 파악하여 추론하기 때문에 질의자의 질의목적에 더 가까운 검색결과를 가져온다. 온톨로지 기반의 검색시스템을 구축하기 위해서는 검색대상의 메타 데이터에 대한 추론을 규정하고 규칙을 표현할 수 있는 온톨로지의 설계가 우선적으로 이루어져야 하며 구축된 온톨로지 지식베이스와 함께 온톨로지 질의를 가능하게 하는 프레임워크의 구축이 필요하다. 이 연구의 온톨로지 검색 부분에서는 OWL을 이용하여 온톨로지를 구축하였으며 검색, 질의를 하기 위해 Jena[3]를 사용하여 시맨틱 프레임워크를 구축하였다.

한편 스케치 기반 검색은 사용자가 기억하고 있는 이미지의 형상(shape), 색상(color), 질감(texture) 등의 정보를 바탕으로 컴퓨터에 이미지를 직접 그리는 방식으로 질의하는 것을 의미한다[4]. 이러한 스케치 기반 검색은 전형적인 이미지 검색과 달리 사용자가 질의 영상을 가지고 있을 필요가 없다는 장점이 있다. 하지만 이러한 스케치 기반 검색 시스템은 사용자의 스케치 능력이나 상세묘사의 정도에 따라 그 검색 성능에 많은 영향을 미치게 된다[5]. 따라서 이 논문에서는 사용자가 직접 그리게 될 스케치는 대부분 단순한 윤곽선 위주의 형상만 표현 가능하다는 가정 하에 형상 정보만을 다루기로 하였으며, 이러한 스케치 기반 검색은 앞서 소개한 온톨로지 기반 텍스트 검색과 함께 상호 보완될 것이다. 따라서 이 논문에서는 사용자가 기억하고 있는 유물에 관한 정보의 형태(텍스트, 형상 등)에 상관없이 다양한 질의 방법을 제공하는 혼합형 질의 방법에 의한 유물 검색 시스템을 제안한다.

이 논문의 구성은 다음과 같다. 2장에서는 관련 연구들을 살펴보고, 3장과 4장에서는 제안한 온톨로지 기반 검색과 스케치 기반 이미지 검색을 각각 소개하게 된다. 5장에서는 3, 4장에서 제안된 두 가지 검색방법을 통합한 전체 시스템의 구조에 대해 설명한다. 6장은 제안한 시스템의 구현 결과 및 평가를 보이며 마지막으로 결론을 맺는다.

II. 관련 연구

2.1 온톨로지 기반 검색

현재 온톨로지 검색 분야의 연구는 다양한 분야에서 진행되고 있다. 이는 온톨로지 언어의 기초인 마크업(markup)언어가 폭넓은 확장성을 가지고 있고 문서정보를 논리적으로 표현할 수 있는 장점 때문이다. 온톨로지 기반의 검색엔진의 경우 기본적으로 특정 분야의 온톨로지를 구축하고 추론엔진과 기타 지식베이스를 구성해야 한다. 온톨로지 기반의 추론을 이용한 시맨틱 검색에 관한 기존 연구에는 Semantic Search System using Ontology-based Inference[6]가 있다. 이 연구에서는 사용자의 질의문과 메타데이터의 속성이 정확하게 매칭이 되는 부분을 검색하고 온톨로지의 계층구조와 공리로 표현된 부분을 규칙기반의 추론을 통해 검색을 한다. 이는 검색 질의자의 입력 질의문에서 정확하게 일치되는 부분이 없을 경우에 규칙기반의 추론을 통해서 사용자의 질의문을 확장시켜 검색하는 방법으로 사용자의 질의어 정보가 온톨로지가 내포하고 있는 클래스 중 어느 곳에 위치하고 있는지를 확장시켜 검색하는 장점이 있다. 또한 구축대상이 유사한 영역의 연구에는 An Art Image Retrieval System[7]이 있으며 이 연구는 RDFs/RDF를 사용하여 지식베이스를 구축하였으며 물을 표현 하기위해 RuleML을 사용하였고 온톨로지 검색 질의 언어로는 RDQL(RDF Data Query Language)를 사용하였다. 시스템의 구조는 Knowledge base에 fact와 rule을 저장하고 사용자의 질의가 발생하면 질의에 대한 답은 추론엔진을 이용하여 결과를 제시하였으나, 가장 기본적인 온톨로지 언어인 RDFs/RDF를 사용하여 Knowledge를 구축하였기 때문에 온톨로지 언어의 기본 구조적 형태인 계층적 관계만을 이용하였고 질의어가 모호거나 지식베이스에 있지 않은 동의어와 같은 질의를 할 경우에는 SameAs와 같은 공리를 해석해 낼 수 없었다.

이 논문에서는 다양한 공리관계를 지원해 주는 온톨로지 언어인 OWL을 사용하였으며 질의정보에 대한 상대적인 위치정보를 파악하여 질의자에게 정보들을 제공하는 방식으로 검색엔진을 구현하였다.

2.2 스케치 기반 이미지 검색

스케치 기반 검색은 내용기반 이미지 검색의 한 종류이며, 사용자 인터페이스만 다를 뿐, 그 구현 방법에 있어서는 거의 차이가 없다. 일반적인 내용기반 이미지 검색은 영상의 형상, 색, 질감 등 영상 자체가 지니고 있는 고유한 시각 특징을 추출하여 이를 색인하고, 그 색인정보를 사용하여 원하는 이미지를 검색하는 방법이다[8]. 스케치 기반 검색의 기존 연구로는 Smith가 제안한 VisualSEEK[4]과 Wang[9], 이상봉[10] 등이 주로 색상과 질감 정보를 다루었으나, 대부분 무채색 계열인 유물을 대상으로 할 경우에는 색상이나 질감 정보보다 형상 정보가 보다 적절하다고 판단된다. 형상 특징 정보를 다루는 방법은 크게 영역기반(region-based)과 경계기반(boundary-based) 방법들로 나뉜다[11]. 영역기반 방법들 중 대표적인 방법은 불변 모멘트(moment invariant)[12][13]를 사용하여 형상 특징 정보를 추출한다. 이는 영상의 특징벡터 값들이 회전과 크기 변화에 불변인 성질을 가져 검색 시스템에서 많이 사용되나 전체적으로 영상을 나타내는데 있어서는 외곽에 대한 정보가 불충분하여 검색 효율이 떨어지는 것으로 알려져 있다[14]. 경계기반 방법들에는 rectilinear shapes[15], polygonal approximation [16], 푸리에 서술자(fourier descriptors)[17][18][19] 그리고 웨이블릿 서술자(wavelet descriptors)[20] 등을 사용한다. 특히, 푸리에 서술자는 영상 내의 객체 경계를 푸리에 변환을 통해 그 객체의 형상을 표현하는 것으로, 불변 모멘트의 부족한 인식률을 극복할 뿐만 아니라 잡음(noise)에 민감하지 않다는 등의 장점을 가지고 있다[21]. 최근에는 [20], [22]이 제안한 웨이블릿 서술자를 이용한 형상 기반 방법이 소개되었다. 일반적으로 웨이블릿 변환(wavelet transform)은 푸리에 변환(fourier transform)에 대해 다 해상도 분석(multi resolution analysis)이 가능하다는 장점을 가지나 공간 분해능(spatial resolution)이 증가할수록 주파수 분해능(frequency resolution)이 떨어지는 경향이 있어 저주파수 대역(low frequencies)의 계수(coefficients)는 거의 쓰이지 못했으며, 유사도 측정 시 계산량이 많아 on-line 검색시스템에는 부적합하다는 주장이 거론되었다[21]. 따라서 이 논문에서는 형상 정보 추출을 위해 푸리에 서술자를 사용하며, 이에 대한 설명은 4장에서 자세히 다루기로 한다.

III. 온톨로지 검색

3.1 온톨로지 구축

온톨로지 검색시스템을 구현하기 위해서 메타데이터의 의미를 정의해 놓은 온톨로지를 설계하였다. 유물검색 시스템에 초점을 맞추어 온톨로지를 설계하였으며 유물 객체의 수는 총 200개로 선정하였다. 유물 객체들에 대한 온톨로지의 최상위 클래스의 구분은 클래스를 2가지로 분류하여 설계하였으며 온톨로지 작성 언어로는 OWL을 사용하였다. 온톨로지 구축도구로는 Protege 3.0[23]을 사용하여 OWL 기반의 온톨로지 모델을 (그림 1)과 같이 생성하였다. 온톨로지의 구분은 유형 클래스, 문화재 지정종목명칭(국보, 보물)에 관한 클래스로 구분하였다. 각 클래스들은 property, domain, range를 가지고 온톨로지를 표현하였으며 클래스(class)와 서브클래스(subclass)의 형태로 카테고리를 구분한다. 온톨로지 메타데이터의 생성은 수동적으로 입력하였으며 생성된 메타데이터는 각 클래스의 정보와 클래스간의 관계를 표현해주고 있으며 이를 통해 메타데이터 상호간의 의미적인 관계를 나타내게 된다.



그림 1. 유형과 지정종목 온톨로지의 예
Fig. 1 An Example of Type and Item Ontology

3.2 검색엔진 설계

이 연구에서 제안하는 온톨로지 기반의 검색방법은 질의자의 의도를 파악하기 위하여 자동으로 생성되는 질의문을 확장시켜 질의하게 된다. 생성된 온톨로지를 검색하기 위해서는 온톨로지 모델을 트리플 형식으로 변환해야 하며 변환된 트리플 형태의 온톨로지를 통해 검색엔진의 질의를 할 수 있다. 검색엔진의 검색방법은 질의자가 알고자 하는 정보의 키워드를 입력하게 되면 검색엔진은 온톨로지 모델을 생성하여 트리플 형태의 온톨로지로 정제시켜 RDQL[3] 질의를 통해 해당 클래스를 찾는다. 검색 엔진은 검색된 클래스를 근거로 현재 클래스의 카테고리를 알아내기 위해 현재 클래스와 연결된 상위, 동급, 하위 노드를 검색하게 된다. 검색엔진은 super-class, sub-class 또는 동급수준의 클래스를 검색하고, 만일 super-class가 또 다른 sub-class에 속한 sub-class라면 재귀적으로 super-class를 검색하여 최상위 클래스까지 검색한다. sub-class도 super-class 검색방법과 마찬가지로 형태로 검색한다. 검색 엔진은 최종 검색된 계층구조의 class 정보와 메타데이터 정보를 사용자에게 출력하여 사용자가 원하는 메타정보를 선택하게 된다[24].

IV. 스케치 기반 이미지 검색

4.1 전처리 과정

다른 이미지 검색과 달리 스케치 기반 이미지 검색은 사용자의 스케치 능력에 따라 그 성능이 좌우되며 이미지가 복잡할수록 특징의 적절한 표현이 어렵다는 단점을 가지고 있다. 따라서 사용자는 대상물의 윤곽선 정도만 스케치 할 수 있다고 판단하여 검색대상이 될 이미지를 단순화하는 전처리 과정이 필요하다. (그림 2)는 크게 4단계를 가지는 전처리 과정을 보여준다.



그림 2. 이미지 전처리 과정
Fig. 2 Image preprocessing

먼저 각 크기가 다른 대상 이미지를 128×128 화소의 크기로 변환(Resize)한 다음, 이를 그레이(Gray) 변환하게 된다. 그레이 변환을 하는 이유는 질의가 될 스케치 이미지가 이진 영상만을 사용하기에 칼라 색상은 의미가 없기 때문이다. 특히, 이 논문에서 대상으로 하는 문화재 유물의 경우 대부분이 단조로운 무채색이기에 색상분포를 고려한 특징 추출은 고려하지 않았다. 그레이 변환된 이미지는 윤곽선 검출기(Canny Edge Detector)로 윤곽선을 얻으며 이는 스케치 이미지를 감안하여 클로우징(Closing) 연산을 통해 좀 더 영상을 단순화하였다.

4.2 형상 특징 추출 과정

전처리 과정을 마친 이미지와 질의로 제시된 스케치 이미지는 푸리에 서술자(fourier descriptors)로 표현하기 위해 수식 (4.1)과 같이 복소수 형태로 표현하였다[18].

$$z(k) = (x_k - x_c) + j(y_k - y_c) \dots\dots\dots (4.1)$$

여기서 $(x_k, y_k), k=0, 1, 2, \dots, N-1$ 는 이미지 객체의 경계면 좌표(boundary coordinates), (x_c, y_c) 는 이미지 객체의 중심(centroid), N 은 이미지의 화소크기가 된다. 이에 대한 이산 푸리에 변환(DFT : Discrete Fourier Transform)은 식 (4.2)와 같다.

$$F_n = \frac{1}{N} \sum_{k=0}^{N-1} z(k)e^{-j2\pi nk/N} \dots\dots\dots (4.2)$$

여기서 $F_n, n=0, 1, 2, \dots, N-1$ 은 $z(k)$ 에 대한 푸리에 계수(fourier coefficients)이다. 이 푸리에 서술자에서 푸리에 계수의 절댓값 즉, $|F_n|$ 을 사용하여 크기가 L 이고 정규화(normalization)된 회전 불변(rotation invariant) 특징 벡터를 윤곽선 푸리에 방법(Contour Fourier method)[18]을 통해 구한다. 이는 식 (4.3)에서 보여준다.

$$x = \left[\frac{|F_{-(L/2-1)}|}{|F_1|}, \dots, \frac{|F_{-1}|}{|F_1|}, \frac{|F_2|}{|F_1|}, \dots, \frac{|F_{L/2}|}{|F_1|} \right]^T \dots (4.3)$$

4.3 유사도 측정

유사도는 질의 스케치 이미지와 검색 대상 이미지가 얼마나 유사한가를 측정하는 것으로 식 (4.4)와 같이 정의되며 유클리디언 거리(Euclidean distance)를 사용하였다.

$$D(I, J) = \left(\sum_{i=0}^{N-1} |x_i(I) - x_i(J)|^2 \right)^{1/2} \dots\dots\dots (4.4)$$

여기서 $x_i(I)$ 와 $x_i(J)$ 는 각각 질의 스케치 이미지와 검색 대상 이미지 특징 벡터의 i 번째 원소를 의미한다.

V. 전체 시스템 구조

제안한 시스템의 구조는 사용자의 다양한 질의 형태를 고려하여 온톨로지 기반 텍스트 검색과 스케치 기반 이미지 검색으로 나누어 설계하였는데 전체 시스템의 구조는 (그림 3)과 같다. 사용자의 질의가 텍스트 형태일 경우에는 온톨로지 기반 검색을, 형상일 경우에는 스케치 기반 이미지 검색을 하게 된다.

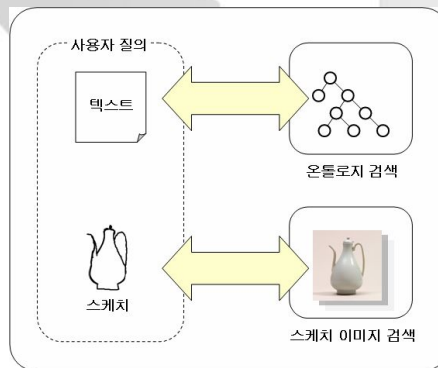


그림 3. 전체 시스템 구조
Fig 3. System Architecture

먼저, 온톨로지 기반 검색 시스템의 상세 구조는 (그림 4)와 같으며 실행 절차는 다음과 같다.

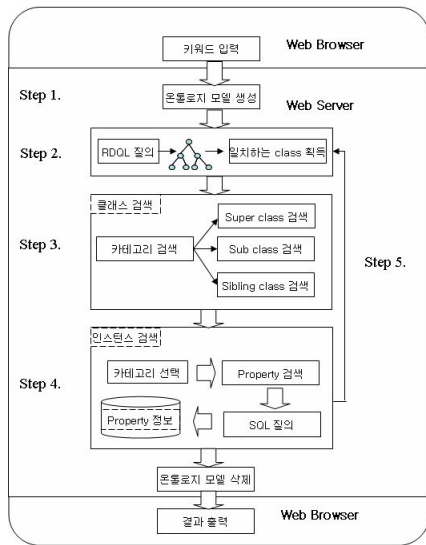


그림 4. 온톨로지 기반 검색
Fig. 4 Ontology-based Information Retrieval

- [Step 1] 클라이언트가 키워드를 입력하게 되면 서버는 검색을 위해 온톨로지 모델을 검색엔진에 로드시켜 주게 된다.
- [Step 2] 로딩된 온톨로지는 질의를 하기 위해 자동으로 RDQL 질의문을 생성하여 해당 클래스를 획득하게 된다.
- [Step 3] 검색엔진은 질의문을 통해 해당클래스와 연관성이 있는 상위, 동위, 하위 클래스를 검색한다.
- [Step 4] 질의자가 원하는 카테고리를 선택하게 되면 해당 속성정보를 데이터베이스에서 획득하는 단계이다. 검색엔진은 인스턴스를 생성하여 해당 속성정보를 획득하게 된다.
- [Step 5] 만일 질의자가 현재 카테고리와 연결되어 있는 다른 정보를 검색할 경우 다시 Step 1 으로 넘어가게 되고 그렇지 않으면 사용자에게 최종 검색 결과를 출력한다.

또한 스케치 기반 이미지 검색의 상세 구조는 (그림 5)와 같으며, 실행 절차는 다음과 같다.

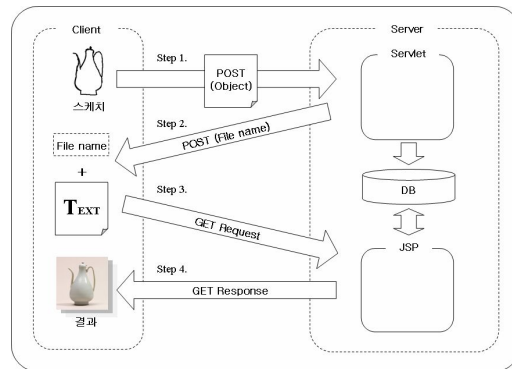


그림 5. 스케치 기반 이미지 검색
Fig. 5 Sketch-based Image Retrieval

- [Step 1] 사용자는 자바 애플릿(Java Applet)으로 작성된 클라이언트(Client)에 직접 질의 영상을 스케치한다. 이 영상에서 형상 특징 정보를 추출하여 이를 객체 직렬화(Object Serialization)를 통해 서버로 보내진다.
- [Step 2] 서블릿(Servlet)에서는 받은 형상 특징 정보를 파일로 생성하고 이를 DB에 저장하며, 임의로 생성된 파일명(unique temporary file name)을 다시 클라이언트로 보낸다.
- [Step 3] 내부적으로 파일명을 통해 다시 서버에 질의를 하게 되는데, 경우에 따라 텍스트 질의 정보와 함께 서버로 재전송된다.
- [Step 4] JSP에서는 질의와 가장 유사한 이미지를 검색한 후에 그 결과를 클라이언트에게 전달한다.

VI. 구현 및 평가

6.1 시스템 구현

전체 시스템의 구현 화면은 (그림 6)과 같다. 검색 시스템의 사용자 인터페이스는 사용자가 직접 형상을 그릴 수 있는 스케치 영역과 관련 도구 그리고 검색어 입력창으로 구성되어 있다.

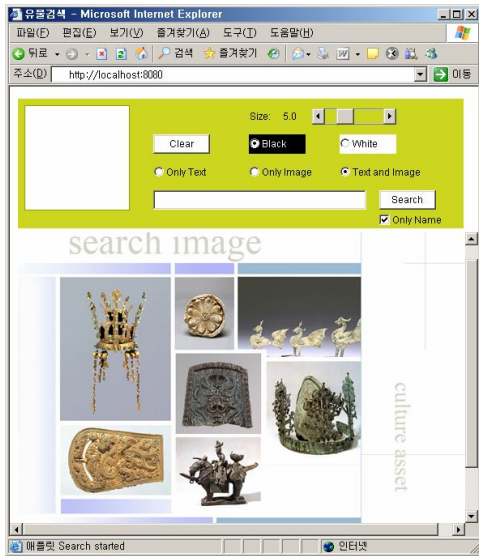


그림 6 검색 시스템 초기화면
Fig. 6 The initial display of our retrieval system

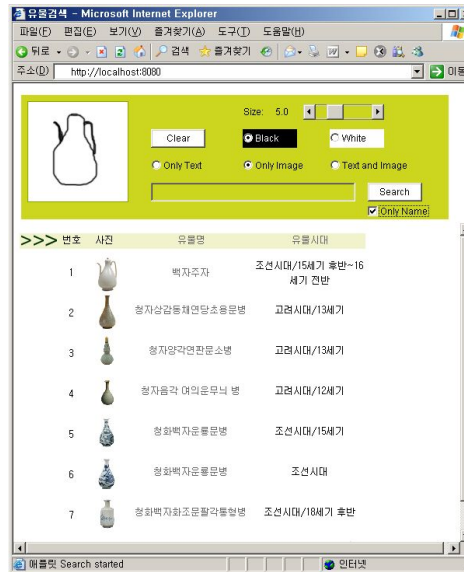


그림 8 스케치 기반 검색결과
Fig. 8 The result of sketch-based retrieval

사용자는 온톨로지 기반 검색과 스케치 기반 검색을 각각 따로 하거나 동시에 검색할 수 있도록 설계되었다.

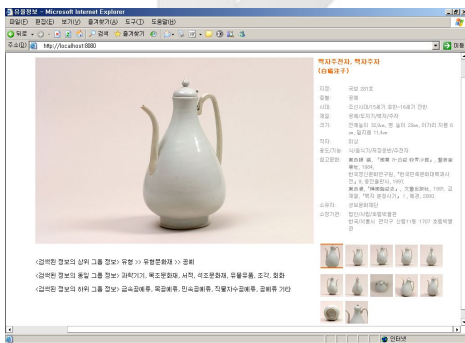


그림 7. 온톨로지 기반 검색 결과
Fig. 7 The result of ontology-based retrieval

온톨로지 기반 검색 결과는 (그림 7)에서 보여주고 있다. 검색창을 이용하여 ‘공예’를 검색을 하게 되면 온톨로지 상의 공예와 관련된 클래스 정보를 사용자에게 제시해 준다. 온톨로지 내에 공예와 연관된 상위, 동급, 하위 클래스를 제시하여 사용자의 질의와 연관된 정보를 제공해 준다. 공예 정보와 관련된 정보는 상위 클래스의 유형과 동급 클래스인 과학기, 목조문화재, 서적, 석조 문화재, 유물·유품 등이 있으며 하위 클래스정보에는 금속공예류, 목공예류, 민속공예류, 직물자수공예류 등이 제시가 되며 사용자는 이 정보들 중에서 원하는 클래스 정보를 선택하게 되면 검색 엔진은 해당 클래스의 인스턴스 정보를 통해 자세한 속성정보를 가져오게 된다. 속성정보의 검색은 데이터베이스 질의를 통해 제공하며, 최종 검색된 정보와 관련이 있는 클래스 정보들도 질의자에게 제공한다.

스케치 기반의 검색은 사용자가 형상정보를 기억하고만 있을 경우에 검색을 가능하게 하는 방법이므로 사용자가 기억하고 있는 유물의 형상정보를 자바기반의 애플릿 영역에 스케치함으로써 검색을 하게 된다. (그림 8)은 주전자 모양의 도자기를 스케치하여 검색한 결과로 가장 유사한 형상을 가진 유물 순으로 정렬된 결과를 보여주고 있다.

6.2 평가

아직까지 온톨로지 검색 시스템에 관한 정량적인 성능평가에 관한 연구는 거의 보고되지 않았고 추가적으로 검색된 부분에 대한 평가가 이루어지기 어렵다[6]. 따라서 제안한 시스템의 효율성을 평가하기 위해 검색의 정확도가 아닌 사용자에게 몇 가지 문제(예: “조선시대에 쓰인 천문관측 도구 중 창경궁에 소장되어있고 그 재질이 석재인 것은?”)를 검색하게 하여 원하는 결과를 얻는 데까지 걸리는 시간을 측정하는 방법으로 시간적 효율성을 측정해 보았다.

실험은 재학 중인 학부생 120명을 대상으로 이들을 임의적으로 60명씩 A, B 2개의 피실험군으로 나누고 유물관련 문제 6개를 각각 A군은 기존 텍스트 검색 방식으로 구현된 시스템을 B군은 제안한 시스템을 사용하게 하여 측정하였다. 실험 결과는 <표 1>과 같으며 온톨로지 검색이 상대적으로 나은 결과를 보여주었는데(t-검정, 유의확률: 0.0237), 이는 추가 검색과 같은 부가 작업을 다소 적게 했다고 해석될 수 있다.

표 4. 문항별 평균 소요 시간(단위: 초)
Table 1. Average time per quiz(sec.)

문항 번호		1	2	3	4	5	6	합계
실험군	A	58.3	48.6	34.7	68.1	67.8	52.1	329.6
	B	41.4	31.8	35.7	43.9	53.6	41.6	248.0

VII. 결론

제안한 혼합형 질의 방법에 의한 유물검색 시스템은 기존 텍스트 검색이 정확한 키워드를 요구한다는 점과 복잡한 질의의 경우 추가검색이 비교적 많이 요구된다는 단점을 보완하기 위하여 RDQL 질의어를 통해 카테고리의 상대적인 위치 정보를 파악하고 데이터베이스에 저장된 온톨로지의 속성 정보를 사용하여 보다 나은 결과를 보여주었다. 또한 사용자가 기억하고 있는 유물에 관한 정보의 형태(텍스트, 형상 등)에 상관없이 다양한 질의방법을 선택할 수 있도록 스케치 기반 이미지 검색기능을 추가하여 보다 향상된 사용자 인터페이스를 제공하였다.

추후 과제로는 온톨로지 기반 검색의 효율성을 보다 객관적으로 평가할 수 있는 방법에 관한 연구가 요구되어진다.

참고문헌

- [1] OWL Web Ontology Language Overview, <http://www.w3.org/TR/owl-features/>.
- [2] OWL Web Ontology Language Use Cases and Requirements, <http://www.w3.org/TR/webont-req/>.
- [3] Jena - A Semantic Web Framework for Java, <http://jena.sourceforge.net/>.
- [4] J. R. Smith, S. F. Chang, "VisualSEEK: a fully automated contentbased image query system", ACM International Multimedia Conference, pp. 87-98, 1996.
- [5] 임미영, 김형준, 김경수, 김희정, 하명환, 정병희, 김희을, "내용 기반 영상 검색을 위한 개선된 질의 방법", 대한전자공학회 하계종합학술대회, 제26권, 제1호, pp.1665-1668, 2003.
- [6] 이상범, 박영택, "온톨로지기반 추론을 이용한 시맨틱 검색 시스템", 정보과학회논문지, Vol. 23, No. 3, pp.202-214, 2005.
- [7] 한상진, 조우상, 이복주, "온톨로지 추론 엔진을 이용한 미술 작품 검색 시스템", 한국정보과학회 가을 학술발표논문집, Vol. 31, No. 2, pp.139-141, 2004.
- [8] 김봉기, 김홍준, 김창근, "특징기반 계층적 영상 검색 시스템의 구현", 한국 OA학회 논문지, 제5권, 제2호, pp.60-70, 2000.
- [9] J. Z. Wang, G. Wiederhold, O. Firschein, and S. X. Wei, "Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability", 4th International Forum on Research and Technology Advances in Digital Libraries, pp.13-24, 1997.
- [10] 이상범, 고병철, 변혜란, "스케치 질의를 통한 웹기반 영상 검색과 분류 시스템", 한국정보과학회논문지: 소프트웨어응용, 제30권, 7-8호, pp.703-712, 2003.
- [11] D. S. Zhang, G. Lu, "A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures", In Proc. of International Conference on Intelligent Multimedia and Distance Education, pp.1-9, 2001.

- [12] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", "IEEE Trans. Information Theory", Vol. 8, pp.179-187, 1962.
- [13] L. Yang, F. Algregrtsen, "Fast computation of invariant geometric moments: A new method giving correct results", Proc. IEEE Int. Conf. on Image Processing, pp. 201-204, 1994.
- [14] A. Jain, A. Vailaya, "Shape based retrieval: a case study with trademark image databases", Pattern Recognition, Vol. 31, No. 9, pp.1369-1390, 1998.
- [15] H. V. Jagadish, "A Retrieval Technique for Similar Shapes", SIGMOD Conference, pp.208-217, 1991.
- [16] E. M. Atkin, L. P. Chew, D. P. Huttenlocher, K. Kedem, and J. S. B. Mitchell, "An Efficiently Computable Metric for Comparing Polygonal Shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 3, pp.209-216, 1991.
- [17] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of Affine-Invariant Fourier Descriptors to Recognition of 3-D Objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 7, pp.640-647, 1990.
- [18] H. Kauppinen, T. Seppänen, and M. Pietikäinen, "An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 2, pp.201-207, 1995.
- [19] D. Rafiei, A. O. Mendelzon, "Efficient retrieval of similar shapes", VLDB Journal, Vol. 11, No. 1, pp.17-27, 2002.
- [20] H. S. Yang, S. U. Lee, "Recognition of 2D Object Contours Using Starting-Point-Independent Wavelet Coefficient Matching", Journal of Visual Communication and Image Representation, Vol. 9, No. 2, pp.171-181, 1998.
- [21] D. S. Zhang, G. Lu, "A Comparison Of Shape Retrieval Using Fourier Descriptors And Short-Time Fourier Descriptors", In Proc. of The Second IEEE Pacific-Rim Conference on Multimedia, pp.855-860, 2001.
- [22] Q. M. Tieng, W. W. Boles, "Recognition of 2D Object Contours Using the Wavelet Transform Zero-Crossing Representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 8, pp.910-916, 1997.
- [23] Protege, <http://protege.stanford.edu/>.
- [24] 백승재, 천현재, 이흥철, "문화재 정보의 온톨로지 기반 검색시스템", 한국컴퓨터정보학회, Vol. 10, No. 3, pp.228-236, 2005.

저자 소개



천 현재

1997년 2월 인천대학교 산업공학
학사

1999년 2월 고려대학교 산업공학
석사

1999년~현재 고려대학교 산업시스
템정보공학과 박사과정
<관심분야> Semantic Web, CRM



백 승 재

2004년 2월 단국대학교 산업공학
학사

2004년~현재 고려대학교 산업시스
템정보공학과 석사과정
<관심분야> Semantic Web, SCM



이 흥 철

1983년 2월 고려대학교 산업공학
학사

1988년 2월 Univ. of Texas
산업공학 석사

1993년 2월 Texas A&M
Univ. 산업공학 박사

현재 고려대학교 산업시스템정보
공학과 교수

<관심분야> 웹 기반 생산 및 물류
정보시스템, Simulation,
SCM