Journal of Korea Multimedia Society Vol. 9, No. 7, July 2006(pp. 782-800)

분산 웹 환경에서 이미지 검색을 위한 메타데이타 매핑

권은영[†], 안철범[††], 나연묵[†††]

요 약

이미지 메타 데이타는 이미지 데이터의 내용, 구조, 특징 등을 기술하는 데이타를 말한다. 이러한 메타 데이타는 다양한 시스템 환경에서 서로 공유되며 다양한 표준이 존재한다. 이러한 환경에서 사용자가 요구하는 정확한 데이타를 검색하는 것은 매우 중요하다. 본 논문에서는 분산 환경에서 다양한 표준을 가지는 이질적인 이미지 메타데이타의 상호 운용 시스템 구조를 제안하고, 보다 정확하고 효율적인 검색을 위하여 메타데이타 매핑 방법을 이용한 질의 확장 방안을 제안한다. 제안한 매핑 방법에서 비표준 메타데이타를 사용하는 제공자는 MDR내의 기준 표준인 VRA 표준과 매핑된다. 또한, 불필요한 매핑 연산을 줄이고 MDR에 새로운 표준을 쉽게 추가하기 위해서 VRA 표준을 기준으로 표준간의 매핑을 수행한다. 제안된 확장 질의 방법을 사용하면 다양한 메타데이타 스키마와 표준 간의 관계를 활용하여 검색의 정확도를 높일 수 있다. 본 논문에서 제안된 방법의 유용성을 보이기 위해 시스템 프로토타입을 구현하고 확장 질의 처리에 대한 검색 성능 평가를 수행하였다.

Metadata Mapping for Image Retrieval in Distributed Web Environments

Eun-Young Kwon[†], Chul-Bum Ahn[††], Yun-Mook Nah[†††]

ABSTRACT

Image metadata is used to describe the contents, structures, and features of a given image data. There exist lots of image metadata-related standards. Such image metadata need to be shared to efficiently search the correct images in Web Service-enabled distributed environments. In this thesis, we propose an interoperable system architecture for utilizing image metadata and query extension methods using metadata mapping for correct and efficient image searching. In the proposed methods, the provider schema using non-standard image metadata are mapped into the schema based on the VRA standard, which is used as a basic and intermediary standard in MDR(MetaData Registry).-We support mapping operations between the VRA standard and other standards, to reduce the number of unnecessary mapping operations and to enable easier addition of schema based on new standards. Using the proposed query processing methods, which utilize image metadata and overcome their heterogeneity by metadata mapping, it is possible to severely improve the correctness of image retrieval. To show the usefulness of our schemes, we implemented a prototype system and performed retrieval performance evaluation on extended query processing.

Key words: Image Metadata(이미지 메타데이타), MDR, Web Service(웹 서비스), Metadata Mapping(메타데이타 매핑), Image Retrieval(이미지 검색)

※ 교신저자(Corresponding Author) : 안철범, 주소 : 서울시 용산구 한남동 단국대학교 제1공학관 518호(140-714), 전화 : 016-726-6616, FAX : 02-799-1080, E-mail : ahn555@dankook.ac.kr
접수일 : 2005년 11월 16일, 완료일 : 2006년 3월 13일
[†] 준회원, 단국대학교 전자컴퓨터공학부 (E-mail : love7463@korea.com)
[††] 준회원, 단국대학교 전자컴퓨터공학부
[†††] 정회원, 단국대학교 전자컴퓨터공학부 (E-mail : ymnah@dku.edu)
※ 이 연구는 2005학년도 단국대학교 대학연구비의 지원으로 연구되었음

1. 서 론

정보 통신 기술과 웹 서비스의 급속한 발달로 다양한 곳에서 발생되는 데이타들을 네트워크 환경의 제약을 받지 않고 서로 다른 환경 간에 교환하는것이 가능하게 되었다. 이러한 환경에서 생성되어 검색 대상이 되는 데이타들은 여러 곳에 분산되어 있고 이질적인 형태를 가지며 그양도 방대하다. 그동안 분산 환경에서 각기 다른 형태와 의미를 가진 대량의 데이타들을 정확하게 검색하기 위해 스파이더, 로봇 프로그램 등을 이용해 검색하는 방법이 사용되어 왔지만 현실적으로 원하는 데이타를 정확하게 가져오기는 어려운 실정이다. 또한, 이러한 정보를 효율적으로 관리하기 위해 메타데이타(metadata)에 대한 연구가 활발히 진행되고 있으며, 메타데이타를 이용한 검색 방식 중에서 메타데이타에 대해 전문 검색(full text retrieval)을 수행하는 방식은 가장 단순하고 보편적인 방법이기는 하지만 검색의 양이 방대하고 정확성이 떨어지는 한계점이 있다. 이러한 한계점을 극복하고 필요한 정보 자원에 대해 정확히 접근하기 위해 다양한 분야에서 메타데이타 표준에 대한 연구를 진행하여 왔으며, 대표적인 표준안으로 DC (Doublin Core), VRA(Visual Resources Association), CDWA(Categories for the Description of Works of Art), MPEG-7 등이 제안되었다[1, 2, 3, 4]. 그러나 이들 표준안 역시 전체적인 골격은 유사하나 표현형식이 각기 다르므로 웹과 같은 분산 환경에서 메타데이타 상호간의 호환성 문제를 해결할 필요가 있다. 최근 국가적으로 보존 및 이용가치가 있는 교육, 문화, 역사 등의 자료에 대한 디지탈화가 활발히 추진되고 있으며 이러한 디지탈화된 정보를 효율적이고 정확하게 검색하기 위한 관련 표준의 제정, 메타데이타간 상호 운용성 보장에 관한 연구가 중요한 과제로 부각되고 있다.

본 논문에서는 이미지를 기술하는 다양한 메타데이타 표준 상호간의 이질성 문제를 해결하고 이미지 검색의 정확도를 높이는 방안에 대하여 기술한다. 본 논문의 구성은 다음과 같다. 2장에서는 이미지 메타데이타 관련 표준에 대하여 소개하며 3장은 분산 환경에서 이질의 데이타를 효율적으로 교환할 수 있는 시스템 구조와 MDR(Metadata Registry)을 이용한 메타데이타 매핑 기법에 대하여 설명한다[5,6]. 4장에서는 본 논문에서 제안한 이미지 메타 데이타 매핑을 이용한 질의 처리 시스템 프로토타입의 구현과 그 결과를 보이고 끝으로 5장에서 결론을 맺는다.

2. 이미지 메타데이타 관련 연구

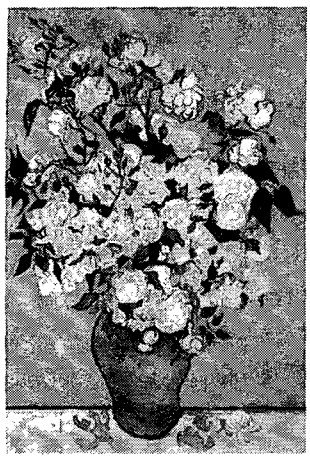
2.1 이미지 메타데이타와 관련 표준안

메타데이타는 데이타에 대한 내용, 구조, 특징 등을 기술한 데이타로서 웹 상의 모든 자원을 정확하고 빠르게 검색하기 위한 데이타이다. 이미지 메타데이타는 이미지에 대한 제반 정보를 기술하기 위한 데이터이다.

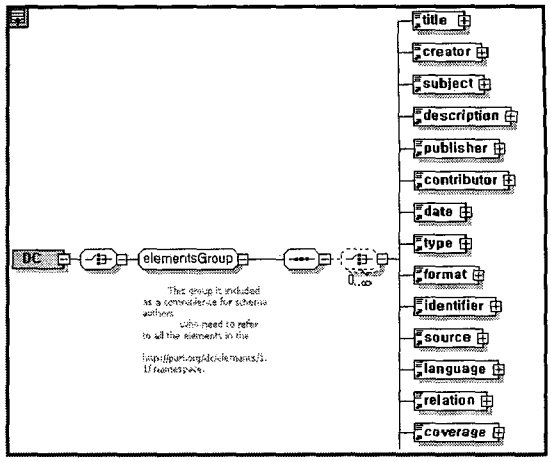
그림 1의 (b)는 이미지 데이타에 대한 정보를 DC 표준을 사용하여 XML 문서로 표현한 예이다. 이미지 메타데이타를 그림 1(c)와 같이 DC 표준을 따르는 XML 문서로 표현하려면 DC에 대한 XML schema를 그림 1(b)처럼 정의하여야 한다. 그림 1(c)에서는 그림 1 (a)의 제목이 "Flowers", 저작자가 "Van Gogh's", 저작권이 "Neue Pinakothek"에 있으며 "1998년"에 제작되었음을 나타낸다.

표준화된 메타데이타를 제공하기 위해 많은 표준들이 제시되어 왔으며 그 중에서 DC(Doublin Core)는 웹 상의 자원(resource)을 기술하기 위한 가장 포괄적인 메타데이타 표준으로 대표적인 15개의 요소 집합(element set)을 제공한다. VRA(Visual Resources Association)는 예술, 조각, 건축, 문학작품과 같은 이미지 들을 기술하기 위한 메타데이타 표준으로 현재 VRA Core 3.0에서는 이미지를 설명하는 대표적인 17개의 요소들과 각 요소별 속성들을 제공하고 있다. CDWA(Categories for the Description of Works of Art)는 미술 작품 도메인의 기술을 위한 메타데이타 표준안이며 대표적인 27개의 요소와 각 요소에 대한 다양한 속성들로 이루어져 있다. DC, VRA, CDWA 표준이 제공하는 각 요소들은 공통된 부분이 많으며 서로 연관성을 가지고 있어 거의 모든 요소들에 대한 상호 매핑이 가능하다.

데이타 공유 목적으로 데이타의 의미(semantic), 구문(syntax), 표현(representation)을 일치 시키기 위하여 국제 표준화 기구인 ISO/IEC JTC1/SC32/WG2 (Metadata Working Group)에서는 MDR (Metadata registry)에 대한 표준화 작업을 진행하였



(a)



(b)

```
<?xml version="1.0" encoding="UTF-8"?>
<dc:COMMON xmlns:dc = "http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:xsi= "http://www.w3.org/2001/XMLSchema-instance" xsi:SchemaLocation="./dc.xsd">
    <dc:TITLE> Flowers </dc:TITLE>
    <dc:CREATOR> Van Gogh's </dc:CREATOR>
    <dc:RIGHTS> Neue Pinakothek </dc:RIGHTS>
    <dc:DATE> Year 1998 </dc:DATE>
</dc:COMMON>
```

(c)

그림 1. 이미지 데이타의 메타데이타 : (a) 이미지 데이타, (b) DC 표준의 XML 스키마, (c) DC 표준을 사용한 이미지 메타데이타

고 그 표준으로 ISO/IEC 11479(1-6)를 제정하였다. MDR은 메타데이터의 등록과 인증을 통하여 표준화된 메타데이터를 유지 관리하며, 메타데이터의 명세와 의미 공유를 목적으로 한다. 따라서 MDR은 데이터 공유를 위한 기본 틀로서 공유를 위한 기본요소로 데이터 요소를 사용하고, 이러한 데이터 요소가 그 내포적인 의미를 충분히 나타내어줄 수 있는 구조 내에 등록하여 각 응용에 활용할 수 있다. 즉, MDR은 메타데이타 표준안을 저장관리하기 위한 프레임워크다. MDR을 이용하면 다음과 같은 기능을 제공할 수 있다.

- 하위의 데이터베이스를 통합하기 위한 표준 정보제공
- 데이터요소와 하위 데이터베이스 스키마 사이의 사상 정보 유지
- 사용자 질의에 사상되어 있는 개별 데이터베이스 스키마 정보관리

2.2 이미지 메타데이타 통합 방법

웹 서비스란 모든 종류의 정보 시스템을 통합, 연계 및 정보 시스템내에 존재하는 문서, 데이타베이스, 응용소프트웨어등을 공동으로 사용할 수 있도록 지원하는 기술이다. 그림 2는 기존의 웹 서비스 방식을 사용하는 경우를 보이고 있다. 그림 2(a)의 일반적인 웹서비스 방식의 경우 클라이언트에서는 특정 지정된 서버로만 질의를 전송할 수 있고 이 경우 구조적으로나 의미적으로 이질성을 갖는 데이타에 대해서는 질의를 수행할 수 없는 문제점이 있다. 그림 2(b)의 전문 검색을 이용한 웹서비스 방식의 경우 구조적, 의미적 이질성을 갖는 데이타에 대해서 좀 더 정확한 질의를 수행하기 위해 문서 전체에 대해서 전문검색(full text search)을 수행하게 되며 구조적으로 다른 형태를 갖는 문서에 대해서도 검색이 가능하지만 사용자 의도에 맞는 정확한 검색은 어렵다. 즉, 검색 결과의 값이 검색하려는 값과 동일하지만 사용자가 의도하지 않은 의미가 다른 경우도 검색되어 질 수 있다. 예를 들어, 사용자가 '작가(creator)가 고흐인 그림을 검색하라' 라고 질의했을 경우 전문검색을 수행하면 '고흐'라는 값을 가지고 전문검색을 수행하므로 그림 2(b)의 서버 A와 같이 고흐의 작품이 아닌 결과가 포함되어 정확성이 감소될 수 있다.

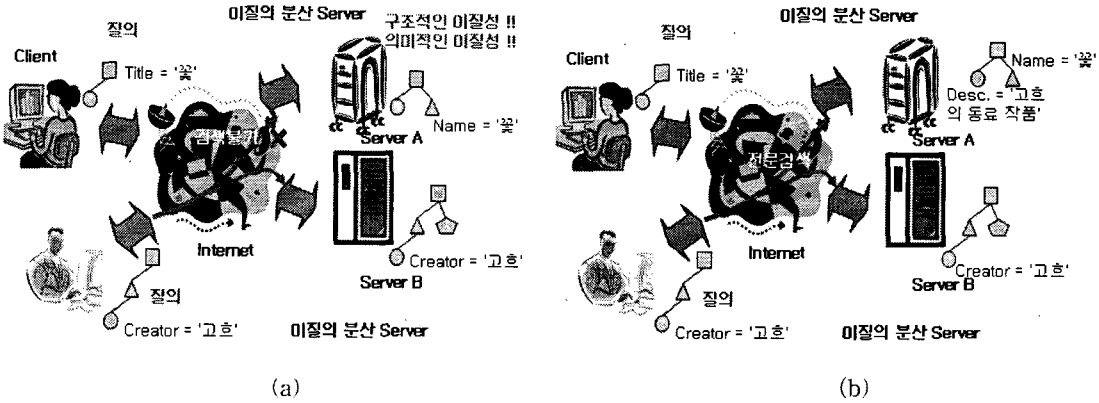


그림 2. 기존의 웹서비스 방식 : (a) 일반적인 웹서비스 방식, (b) 전문검색을 이용한 웹 서비스 방식

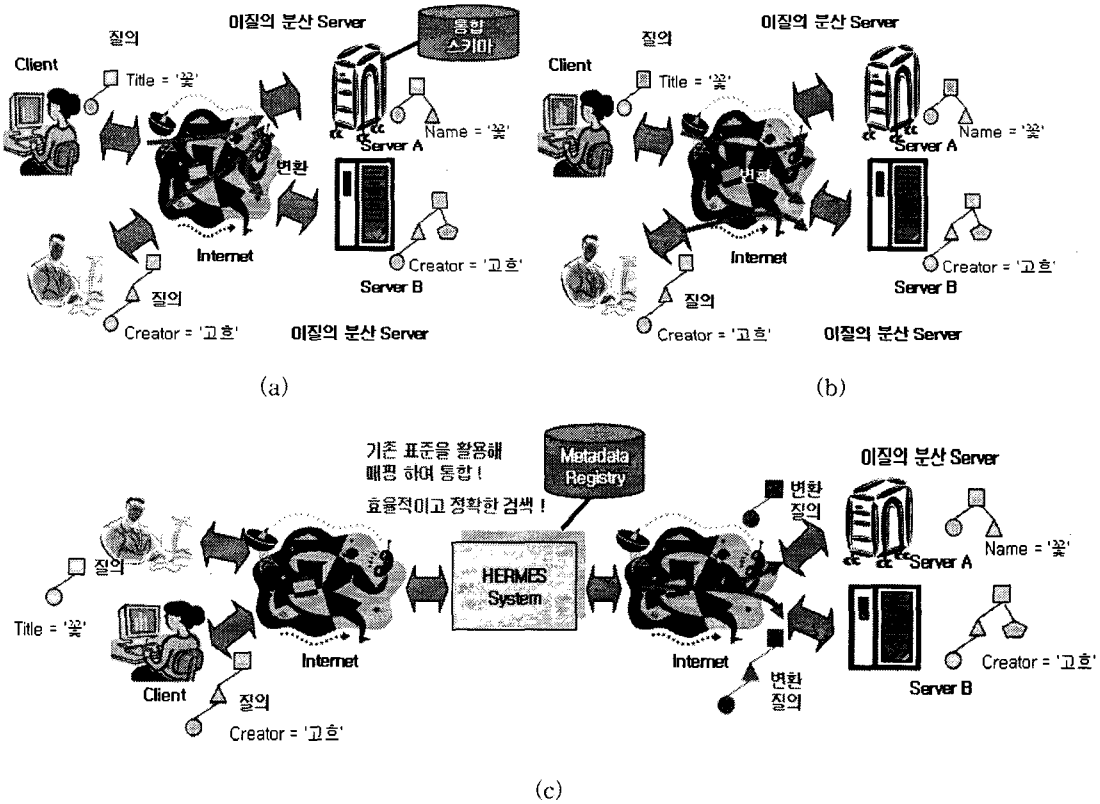


그림 3. 메타데이타 통합을 이용한 웹서비스 방식 : (a) 스키마 통합을 이용한 방식, (b) 일대일 매핑을 이용한 방식, (c) MDR을 이용한 방식

이러한 메타데이타간의 이질성 문제를 해결하여 상호 운용성을 높이기 위한 방법으로는 그림 3 (a)에서와 같이 질의 메타데이타에 대한 통합 스키마를 통해 공유 하는 방식과 그림 3 (b)에서처럼 1:1로 메타데이타를 직접 매핑하는 방식이 있다. 통합 스키마 매핑 방식을 이용해 이질성을 해결하는 경우 관련

된 스키마 구조가 변화될 때 일관성 유지가 어려운 문제가 있으며 일대일 매핑의 경우 높은 정확성을 가지지만 임시방편적인 방법으로 확장성에 문제가 있다. 그림 3 (c)는 메타데이타 공유를 위한 표준이 있는 경우 기준이 되는 표준들을 저장관리하기 위한 ISO/IEC의 MDR을 이용한 방법이다. MDR을 이용한 방법의 경우 MDR내에 관련된 스키마 정보를 등록하여 사용하는 방법으로 통합스키마 방법에서의 관련된 스키마 변화에 따른 일관성 유지 문제나 일대일 매핑에서의 확장성 문제를 해결할 수 있는 방법이다. 그림 3 (c)는 MDR을 사용하여 메타데이타 상호간의 이질성 문제를 해결하고 주어진 질의를 확장 처리하기 위해 제안된 시스템의 구조이다.

3. 메타데이타 매핑을 이용한 확장 질의처리

3.1 HERMES 질의 처리 시스템

본 장에서는 분산된 이미지 데이타를 효율적으로 저장 검색하기 위한 시스템인 HERMES(tHe Retrival framework for visual MEdia Service) 시스템[7]의 Query Handler, Service Finder의 Metadata Mapping Mananger, Service Registration Handler 모듈에 대해 설명하고, 분산된 환경에서 이질성을 갖는 데이타에 대해 정확한 질의를 수행하기 위한 매핑 및 질의 확장을 이용한 질의 처리 방안에 대하여 기술한다.

3.1.1 전체 시스템 구조

본 논문에서 적용한 분산 이미지 검색 시스템은 브로커와 제공자로 구성되며 HERMES라 한다. 이 프레임워크에서 브로커는 분산된 이미지의 지능적인 검색을 위해 서비스 제공자 정보를 관리하고, MDR를 이용해 분산된 메타데이타간의 이질성 문제를 해결한다. 그림 4는 HERMES 브로커의 구조를 보이고 있다.

이 구조에서 Query Handler는 사용자의 질의를 처리하여 결과를 전달하는 역할을 수행하며, Service Finder는 질의의 도메인에 따라 최적의 제공자 리스트를 결정한다. Service Finder는 Service Inference Engine과 Metadata Mapping Manager로 구성되며 Service Inference Engine에서는 Service Ontology와 Service Registry를 이용해서 제공자 리스트를 결정한다. Metadata Mapping Manager에서는 질의 처리 시 질의를 확장하는 기능을 담당한다. Service Registration Handler는 제공자의 각 service를 등록 관리하며 각 제공자 정보를 등록 시 제공자에서 전송된 지역 스키마 정보와 매핑 문서를 저장한다. Feature Handler에서는 특징 추출과 관련된 Feature Extraction Service를 선택하고 실행을 요청한다. 본

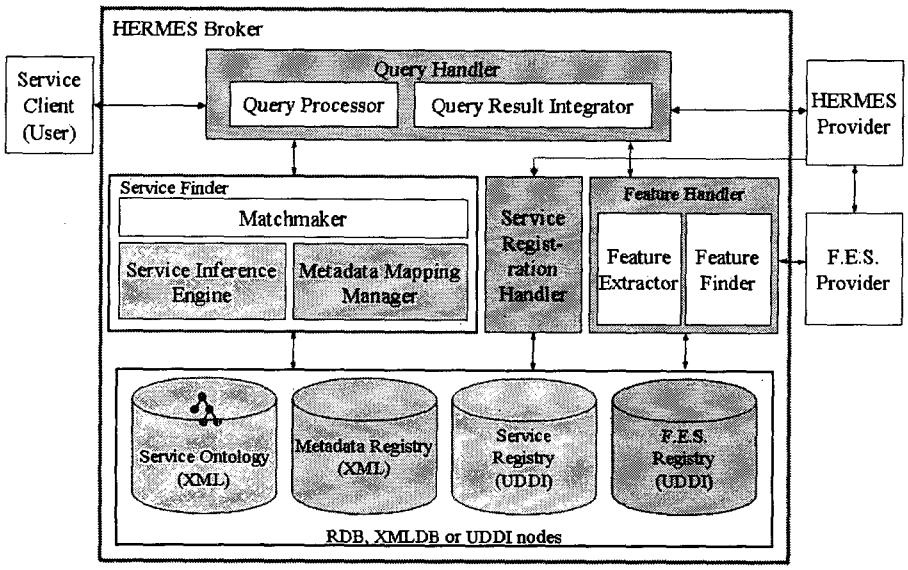


그림 4. HERMES 브로커 시스템 구조

논문에서는 제공자 스키마 정보와 매핑 정보를 저장 관리하는 Service Registration Handler와 질의 시 매핑을 이용한 확장을 담당하는 Metadata Mapping Manager에 대해서 다룬다.

3.1.2 메타데이타 매핑 관리자

메타데이타 매핑은 동일 이미지에 대하여 서로 다른 구조와 의미로 표현된 메타데이타 정보를 상호 교환할 수 있도록 하며, 사용자의 질의도 쉽고 정확하게 할 수 있도록 유도한다. 표 1은 메타데이타 매핑을 이용한 질의처리 시스템으로 2단계 매핑 스키마 레지스트리와 본 논문에서 제안한 HERMES 구조에서의 질의 처리 부분을 비교한 것이다. 2단계 매핑 스키마 구조는 지역 데이타베이스 간의 시스템 통합 방법이다[10]. 첫 번째 단계에서는 요소 매핑을 수행하여 매핑 테이블에 저장하고 두 번째 단계에서는 추가 속성 매핑을 수행하여 추가 매핑 테이블에 저장하는 방식으로 데이타의 변화에 대해서 좀 더 유연하게 대응하여 정확한 질의가 가능한 방법이다.

본 논문에서 제안한 HERMES 시스템에서는 질의 대상 도메인에 따라 필터링하는 기능을 지원하며 미리 정의된 매핑 정보를 이용해 질의를 확장하는 기능이 있다. 질의에 대한 도메인 필터링은 미리 정의된 도메인 정보를 이용하여 질의 시 입력된 도메인과 비교하는 방식으로 수행된다. 또한, 제공자 측의 데이타의 형태를 모두 XML로 가정하고 Xpath를 이용한 XML 질의를 수행할 수 있다.

그림 5는 표 1에서 기술한 두 시스템에서 사용하는 매핑 방법을 비교하는 그림이다. 2단계 매핑 스키마 레지스트리 에서는 그림 5의 (a)와 같이 지역 스키마 별로 각각 관련된 하나의 표준을 생성 및 선택하여 매핑하는 방식이며, MDR내의 여러 표준안 간의 관계는 표현할 수 없다. 그러므로 비표준 스키마뿐만 아니라 표준안을 사용하는 지역스키마의 경우에도 매번 MDR 내의 표준안 정보와 매핑해야 하는 문제점이 있다.

표 1. 메타데이타 매핑을 이용한 질의처리 시스템 비교

	2단계 매핑 스키마 레지스트리	HERMES
MDR 사용여부	사용함	사용함
XML 사용여부	사용안함	사용함
여러 표준간의 매핑	지원안함	지원함
초기 매핑 비용	높음	낮음
도메인 필터링여부	지원안함	지원함
질의 확장 기능	지원안함	지원함
XML 질의 지원 여부	지원안함	지원함

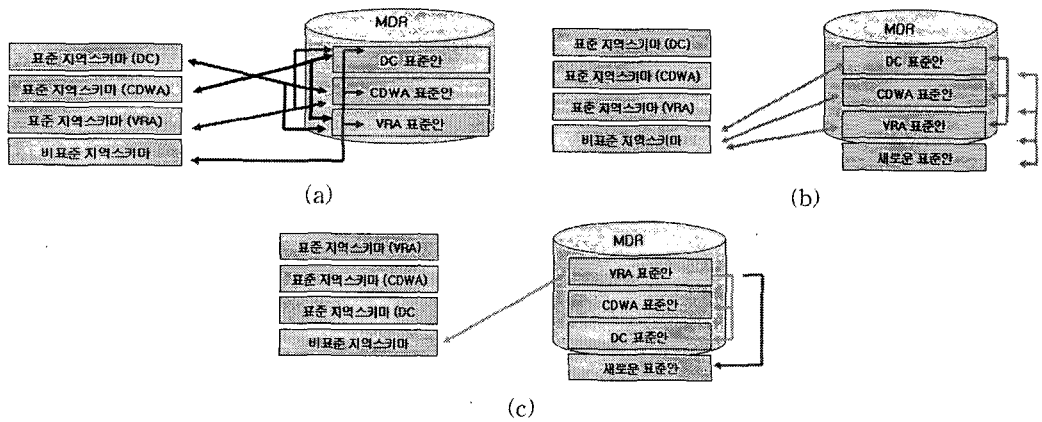


그림 5. 질의 시스템에서 매핑 방법 비교 : (a) 일대일 매핑 방법, (b) 표준 간 매핑을 이용한 매핑 방법, (c) HERMES에서의 매핑

그림 5 (b)에서는 이러한 문제점을 해결하기 위해서 MDR 내의 표준안 간의 연결 관계를 표현할 수 있도록 한 표준이 다른 모든 표준과 연결되는 방법이다. 이 방법에서는 표준 간의 관계가 명확히 표현될 수는 있지만 새로운 표준이 추가 될 경우 다른 표준들에서도 연결 관계를 모두 수정해야 하는 문제점이 있다. 표준안을 사용하는 지역 스키마의 경우 별도의 매핑 과정이 필요 없으며, 비표준 스키마의 경우 에는 MDR 내의 모든 표준들과 매핑을 수행해야 하므로 매핑의 효과는 높지만 매핑 연산이 너무 많아져 현실성이 없다고 볼 수 있다. 본 논문에서 제안한 방법은 그림 5 (c)에서와 같이 MDR 내의 표준 중에서 기준 표준을 정하고 이 기준 표준과 다른 표준과의 매핑을 수행한다. 또한 MDR 내의 모든 표준들은 기준 표준과 미리 매핑이 되어 있다. 이러한 방법으로 표준안을 매핑하면 표준을 사용하는 모든 지역 스키마와 MDR내의 표준안 정보간의 관계를 매번 매핑하는 일대일 매핑에서의 문제점을 현실성 있게 해결할 수 있으며 기준 표준과의 관계만을 표현하면 되므로 새로운 표준을 쉽게 추가할 수 있는 장점이 있다.

3.2 이미지 메타데이타 매핑을 위한 데이타 구조

이미지 메타데이타 간의 매핑 정보를 XML 문서 형태로 표현하기 위해서는 XML 문서를 위한 XML 스키마가 필요하다. 그림 6은 표준을 사용하지 않는 비표준 제공자의 스키마와 MDR내의 표준과 매핑한 결과를 XML 문서 형태로 기술하기 위한 XML 스키마이다. 그림 6에서 'Mapping'은 최상위 요소이며 'Mapping' 요소 하부에는 각각의 매핑 정보를 'Map' 요소 하위에 기술한다. 'Map' 요소 하위에는 제공자 스키마 요소를 기술하는 'Source' 부분과 MDR에 등록된 표준부분을 기술하는 'Target'으로 되어있다. 'SourceElement'에는 'Source'의 요소를 표현하는 'Element'와 속성을 기술하는 'Property' 요소가 있다. 'TargetStandard' 요소는 'Target'의 표준 정보를 기술하고 'TargetElement'내에는 'Target'의 요소를 표현하는 'Element' 요소와 속성을 기술하는 'Property' 요소가 있다. 즉, 해당 'Source' 요소 혹은 속성이 'Target' 요소 혹은 속성에 매핑 된다는 것이다.

그림 7은 MDR 표준을 이용해 표준을 XML 형태로 기술하기 위한 XML 스키마이다. MDR의 기본적인 항목들 외에 VRA 표준을 기준으로 다른 표준 들 간의 관계를 표시하기 위해 데이타 요소 관계(DataElementRelationship) 항목을 추가하였으며 데이타 요소 관계의 하위 요소에서는 VRA 외에 다른 표준들 DC 와 CDWA 요소와 매핑 되는 요소들을 기술해준다.

그림 7 (a)에서 표준안 요소를 기술 하기위해서는 요소의 아이디(DataElementID)와 요소명(DataElementName), 요소에 대한 설명(DataElementDescription)을 기술한다. 값의 도메인(ValueDomain)은 요소의 데이타 타입과 코드 셋을 나타낸다. 데이타 요소 개념(DataElementConcept)은 요소가 속한 표준안에 대한 표준안 아이디(ObjectClassID)정보와 각 요소에 포함된 속성(Property)들을 표현하며 또한 질의 확장을 위해 VRA 표준 요소와 다른 표준과의 매핑을 기술하기 위해 데이타 요소 관계(DataElementRelationship)가 있다. 그림 7 (b)

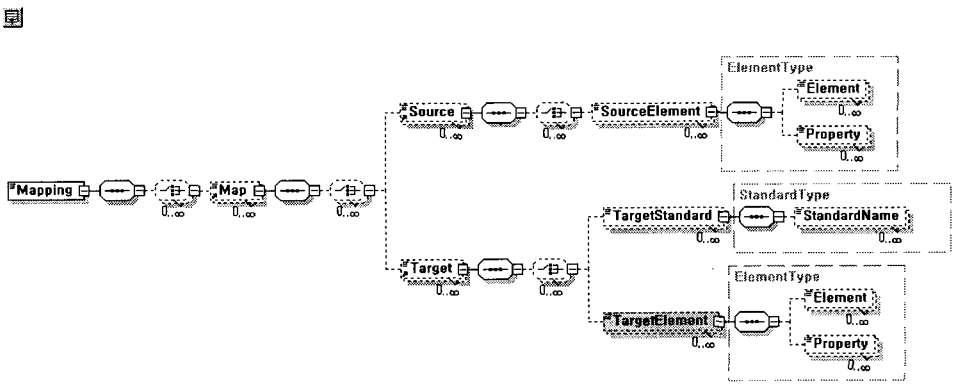
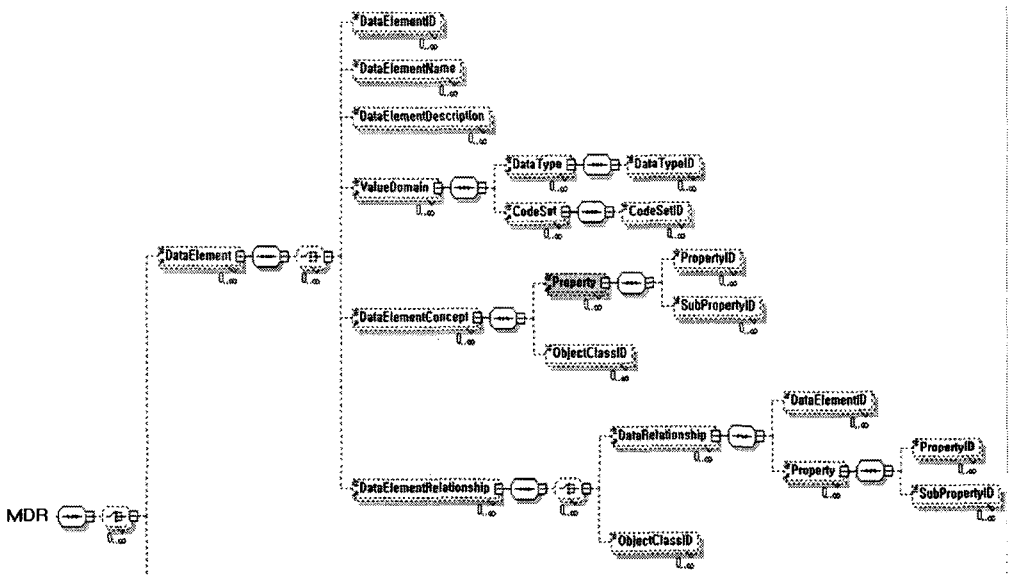
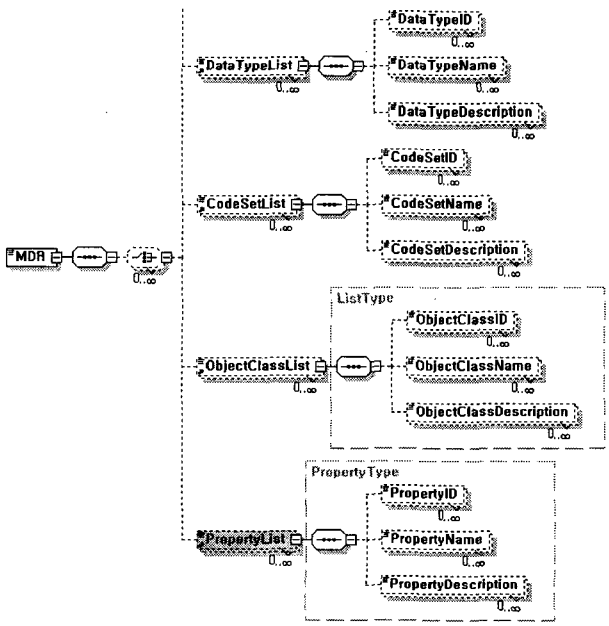


그림 6. 매핑을 위한 XML 스키마

는 각각의 Element를 기술하는데 필요한 값의 도메인과 데이타 요소개념의 내의 구성 요소인 데이타 타입, 속성, 표준안 아이디에 대한 정보를 기술하기 위한 부분이다. XML 문서 형태나 텍스트 형태로 표현되는 이미지 메타데이타 정보는 관계형 테이블에 저장된다. 그림 8은 본 논문에서 제안한 질의 처리 방안을 위한 데이타베이스 테이블이다. 제공자 테이블은 제공자의 스키마를 저장하기 위한 테이블이며, 아이디와 스키마를 소유한 제공자의 URL 주소, 스키마 문서, 적용 표준안, 스키마 버전, 도메인 아이디를



(a)



(b)

그림 7. MDR의 표준을 위한 XML 스키마 : (a) 데이타 요소를 기술하는 부분, (b) 데이타 요소를 기술하기 위한 다른 부분들

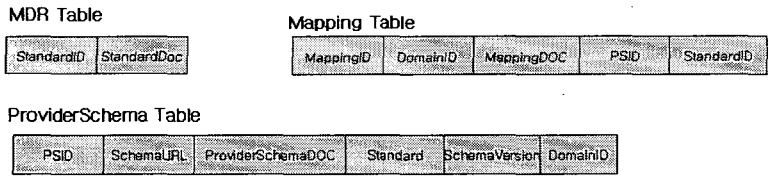


그림 8. 관련 DB 테이블

저장한다. MDR 테이블에는 표준 문서들을 저장하며 매핑 테이블에는 비표준을 쓰는 제공자 스키마와 MDR내의 VRA 표준과의 매핑 정보를 저장한다.

3.3 이미지 메타데이타 매핑 방안

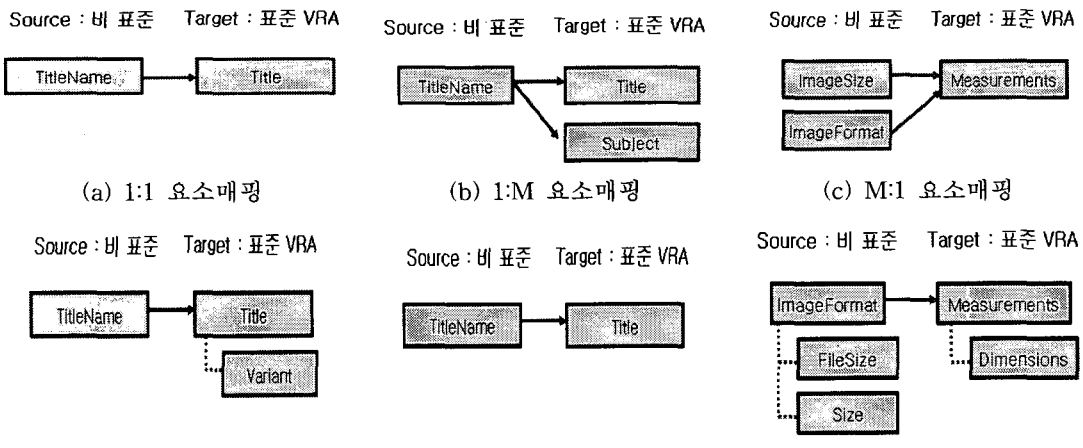
본 논문에서는 이미지 메타데이타 매핑을 위한 기준 표준 메타데이타로 VRA를 선택하였다. DC는 웹상의 모든 자원을 기술하는 범용 표준으로 너무 포괄적인 반면 CDWA 표준은 특정(Art) 도메인의 이미지를 대상으로 하므로 지나치게 지역적인 특성을 가진다. 따라서 웹상의 다양한 이미지를 기술하는 메타데이타 표준으로 DC보다는 구체적이고 CDWA 보다는 일반적인 VRA를 기준 표준으로 선정하였다. 또한 VRA Core 3.0의 경우 DC와 CDWA 표준 요소들에 대한 매핑 정보를 제공해주는 이점도 있다. 본 논문에서 제안하는 매핑 방식은 제공자와 MDR내의 VRA간에 이루어지며, MDR 내에 있는 표준을 따르는 제공자인 경우 매핑을 수행할 필요가 없으며 표준을 따르지 않는 경우에 대해서만 MDR의 기준 표준인 VRA와 요소 매핑 혹은 요소 내 속성 매핑을 수행한다. MDR 내에서는 여러 표준 들을 VRA를 기준으로 이들 간의 관계를 표현하기 위한 매핑을 수행하며 질의 처리 시에 질의를 확장하는 역할을 한다.

3.3.1 비표준과 VRA 표준간의 매핑

이미지 메타데이타 요소들 간의 매핑을 하기위한 매핑 방안으로 표준을 사용하지 않는 제공자와 VRA 표준간의 매핑과 MDR 내의 표준안들 간의 매핑 방법으로는 요소매핑, 속성매핑 방법이 있으며 요소매핑은 일대일, 일대다, 다대일 방식으로 매핑되며 속성매핑 또한 일대일, 일대다, 다대일 방식으로 매핑한다.

[정의 1] 1:1 요소 매핑 : 일대일 요소 매핑은 제공자 비표준 스키마에 있는 하나의 요소가 MDR내의 VRA 표준 요소 하나와 매핑되는 경우이다. 예를 들어, 그림 9 (a)에서와 같이 비표준 스키마에 있는 하나의 요소'TitleName'이 VRA 표준 요소들 중 하나의 요소 'Title'과 일치하여 일대일로 매핑된다.

[정의 2] 1:M 요소 매핑 : 일대다 요소 매핑은 제공자의 비표준 스키마의 하나의 요소와 MDR에 저장



(a) 1:1 요소매핑 (b) 1:M 요소매핑 (c) M:1 요소매핑

(d) 1:1 속성매핑 (e) 1:M 속성매핑 (f) M:1 속성매핑

그림 9. 요소 및 속성 매핑 예

되어 있는 VRA 표준 요소 여러 개와의 매핑되는 경우이다. 그림 9의 (b)는 비표준 스키마의 요소 'TitleName'은 두 개의 속성 'Main Title'과 'Sub Title'로 이루어져 있으며, 'Main Title'은 VRA 표준 요소의 'Title' 요소와 같은 의미이며 'Sub Title'은 VRA 표준 요소의 'Subject' 요소와 같은 의미임을 보여준다. 따라서 비표준 스키마의 하나의 요소 'TitleName'가 표준 요소들 중 여러 개의 요소 'Title'과 'Subject'와 일치하여 일대다로 매핑된다.

[정의 3] M:1 요소 매핑 : 다대일 요소 매핑은 제공자 비표준 스키마에 있는 여러 개의 요소와 MDR내에 저장되어 있는 VRA 표준 요소 하나가 매핑되는 경우이다. 그림 9의 (c)는 비표준 스키마의 요소 'ImageSize'와 'ImageFormat'이 VRA 표준요소의 'Measurements' 요소와 일치함을 보여준다. 즉, 비표준 스키마의 여러 개의 요소 'ImageSize'와 'Iamge-Format'는 표준 요소들 중 하나의 요소 'Measurements'와 일치하므로 다대일로 매핑된다.

[정의 4] 1:1 속성 매핑 : 일대일 속성 매핑은 제공자 비표준 스키마에 있는 하나의 요소가 MDR내의 VRA 표준 요소에 포함된 속성 하나와 매핑 되는 경우이다. 그림 9 (d)에서와 같이 비표준 스키마의 하나의 요소 'TitleName'이 VRA내의 표준 요소들 중 요소 'Title'안에 있는 하나의 속성 'Variant'와 일치하여 일대일로 매핑된다.

[정의 5] 1:M 속성 매핑 : 일대다 속성 매핑은 제공자 비표준 스키마에 있는 하나의 요소가 MDR내의 VRA 표준 요소에 포함된 속성 여러 개와 매핑 되는 경우이다. 그림 9의 (e)와 같이 비표준 스키마의 하나의 요소 'Title'이 VRA 표준 요소의 'Title' 요소와 일치 하며 'Title' 요소의 속성들과도 모두 매핑된다. 따라서, 비표준 스키마의 요소 'Title'이 표준 요소들 중 Title' 요소에 포함되어 있는 모든 속성과 일대다로 매핑된다.

[정의 6] M:1 속성 매핑 : 다대일 속성 매핑은 제공자 비표준 스키마에 있는 요소의 여러 속성이 MDR내의 VRA 표준 요소에 포함된 하나의 속성과 매핑되는 경우이다. 예를 들어, 그림 9의 (f)에서와 같이 비표준 스키마의 요소 'ImageFormat'안에 있는 속성 'Image-Format'과 'Size'가 VRA 표준 요소인 <Measurements>의 안에 있는특정 하나의 속성 'Dimension'과 매핑된다. 따라서, 비표준 스키마의 요소 'ImageFormat'안에 있는 'FileSize'와 'Size'는 표준 요소들 중 하나의 요소 'Measurements' 안에 있는 하나의 속성인 'Dimensions'과 다대일로 매핑된다.

그림 10은 비표준을 사용하는 제공자의 스키마와

```
<Mapping xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Map>
    <Source>
      <SourceElement>
        <Element>TitleName</Element>
        <Property>Main Title</Property>
        <Property>Sub Title</Property>
      </SourceElement>
    </Source>
    <Target>
      <TargetStandard>
        <StandardName>VRA </StandardName>
      </TargetStandard>
      <TargetElement>
        <Element>TITLE</Element>
      </TargetElement>
      <TargetElement>
        <Element>SUBJECT</Element>
      </TargetElement>
    </Target>
  </Map>
  <Map>
    <Source>
      <SourceElement>
        <Element>CreatorName</Element>
        <Property>First Name</Property>
        <Property>Last Name</Property>
      </SourceElement>
    </Source>
    <Target>
      <TargetStandard>
        <StandardName>VRA </StandardName>
      </TargetStandard>
      <TargetElement>
        <Element>CREATOR</Element>
        <Element>Personal name</Element>
      </TargetElement>
    </Target>
  </Map>

  <Map>
    <Source>
      <SourceElement>
        <Element>Style</Element>
      </SourceElement>
    </Source>
    <Target>
      <TargetStandard>
        <StandardName>VRA </StandardName>
      </TargetStandard>
      <TargetElement>
        <Element>STYLE/PERIOD</Element>
        <Property>Style</Property>
      </TargetElement>
    </Target>
  </Map>
  <Map>
    <Source>
      <SourceElement>
        <Element>Description</Element>
      </SourceElement>
    </Source>
    <Target>
      <TargetStandard>
        <StandardName>VRA </StandardName>
        <StandardID>VRA001</StandardID>
      </TargetStandard>
      <TargetElement>
        <Element>DESCRIPTION</Element>
      </TargetElement>
    </Target>
  </Map>
</Mapping>
```

그림 10. 비표준과 표준과의 매핑 예

VRA 표준과 매핑을 수행한 결과 XML 문서이다.

3.3.2 MDR 내의 표준안간의 매핑

MDR내에서 표준안 간의 매핑은 VRA 표준을 기준으로 요소 매핑과 속성 매핑이 있으며 각각 일대일과 일대다 그리고 다대일 관계가 있으며 3.3.1에서 기술한 비표준 스키마와 VRA 표준간의 매핑 규칙이 그대로 적용되며 양쪽 방향에서 적용된다는 차이점이 있다. 그림 11에서와 같이 VRA의 'Title' 요소는 DC의' Title' 요소와 매핑 되며, CDWA 'Titles' 요소의 'Text' 속성과 매핑 된다. VRA의 'Creator' 요소는 DC의 'Creator', 'Contributor'요소와 매핑 되며, CDWA의 'Creation' 요소의 'Names' 속성과, 'Qualifier' 속성 그리고 'Roles' 속성과 매핑된다.

그림 12는 MDR 내에서 표준안들과의 매핑을 XML 문서로 작성한 것이며 각 요소들에 대한 매핑은 <DataElementRelationship>.. </DataElementRelationship> 요소에서 매핑되며, VRA 표준의 'Title'은 DC의 'Title'과 CDWA의 'Title or Names' , 'Title or Names.Text', 'Related Visual Documentation.

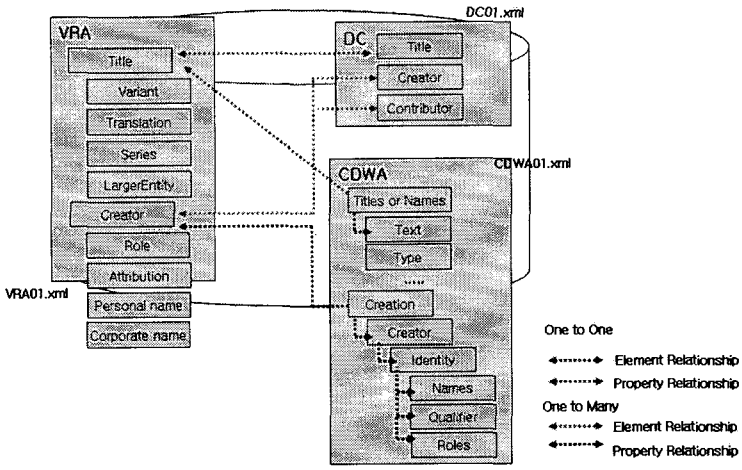


그림 11. MDR 내에서의 표준안 매핑

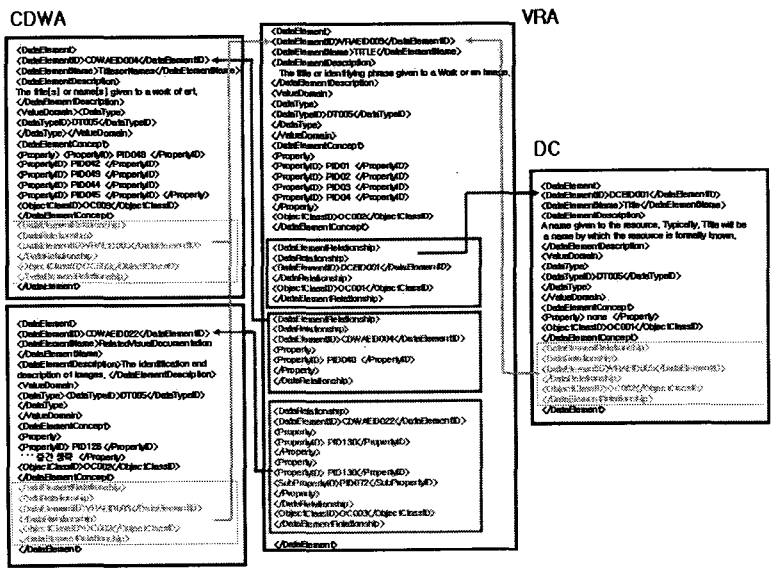
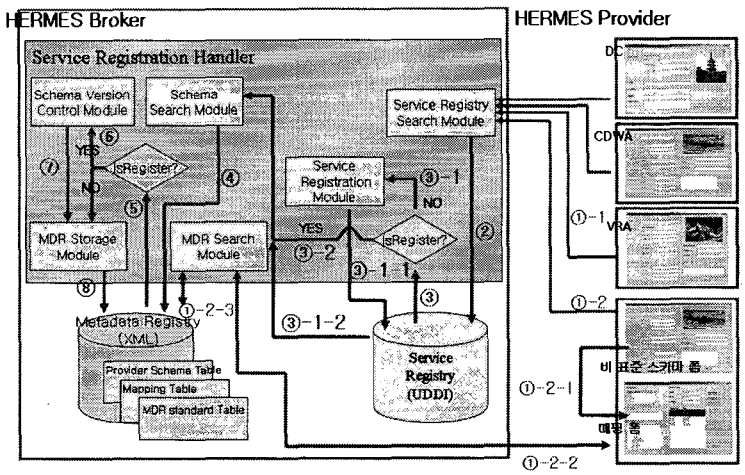


그림 12. MDR 내에서의 표준안 매핑 예

View', 'Related Visual Documentation. View. Indexing Terms'과 매핑 되어졌음을 알 수 있다.

3.4 제공자 스키마 등록 및 질의 처리 흐름

정확한 질의 처리를 위해 제공자 측의 스키마를 등록하고 브로커에서 제공하는 매핑 사용자 인터페이스를 이용하여 매핑 문서를 만들기 위한 절차가 필요하며, 사용자로 부터 입력된 질의를 처리하기위해서 매핑된 정보를 이용하여 질의를 확장하고 각 제공자에 알맞은 Xpath 형태로 변환하여 전송하기 위한 과정이 필요하다. 그림 13은 제공자 측의 스키마를 HERMES 브로커에 등록하기 위한 절차를 나타낸다. 그림 13 (a)에서와 같이 제공자에서는 해당 제공자의 스키마 요소 및 속성 리스트를 VRA 표준의 표준 혹은 속성에 사용자가 매핑 하여 브로커로 전송한다. 전송된 스키마와 매핑 정보는 등록 여부를 검사 후 스키마는 제공자 스키마 테이블에 정보를 저장하고 매핑정보는 매핑 테이블에 저장한다. 그림 13 (b)는 등록 절차에서 각 모듈이 수행하는 작업에 대한 설명이다. 그림 14는 클라이언트로부터 입력된 질의를 매핑을 이용해 질의를 확장 전송하여 검색하는 절차를 나타낸다. 그림 14 (a)에서와 같이 입력된 질의는 분석되어서 도메인 필터링을 거친 후 질의와 관련된 제공자에게로 알맞은 Xpath 형태로 변환되어 전송된다. 각 제공자에서는 Xpath를 받아서 질의를 처리한 후 다시 브로커로 전송하며 브로커에서는 수합하여 사용자에게 결과를 보여준다. 그림 14 (b)는 질의 절차에서 각 모듈이 수행하는 작업에 대한 설명이다.

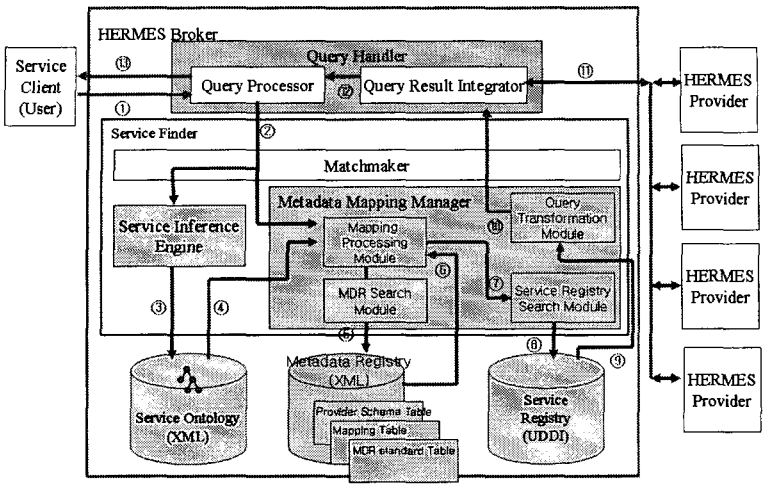


(a)

① 표준을 따르는 경우와 표준을 따르지 않는 경우에 따라서,
①-1 표준을 따르는 경우 제공자 정보만 브로커로 전송
①-2 표준을 따르지 않는 경우 다음과 같은 정보를 가지고 사용자가 매핑처리 후 제공자 정보, 스키마와 매핑 정보를 브로커로 전송
①-2-1 현재 스키마의 요소 및 속성 리스트를 매핑 폼에서 출력
①-2-2 MDR 내의 VRA 표준에 대한 요소 및 속성 리스트 조회
①-2-3 MDR의 VRA 정보를 검색해 제공자의 매핑 폼에 출력
② 현재 등록을 요청하는 제공자 정보를 Service Registry에서 검색
③ 제공자 정보가 등록되어있는지 검사
③-1 등록되어 있지 않다면, ③-1-1 제공자 정보를 등록
③-1-2 등록 정보, 전송된 스키마와 매핑 정보를 Schema Search Module로 보냄
③-2 등록되어 있다면 등록정보, 전송된 스키마와 매핑정보를 Schema Search Module로 보냄
④ 제공자 스키마 테이블에서 해당 스키마를 검색
⑤ 현재 스키마가 기존에 등록되어 있는 지 검사
⑥ 검색된 스키마 정보가 등록 되어 있지 않다면, MDR Storage Module로 스키마와 매핑정보를 보냄.
등록되어 있다면, Schema Version Control Module로 등록정보 와 스키마, 매핑 정보를 보냄
⑦ 스키마 정보에서 버전을 증가시켜 MDR Storage Module로 등록정보와 스키마, 매핑정보를 보냄
⑧ 제공자의 스키마와 관련정보를 제공자 스키마 테이블에 저장하고, 매핑 정보를 매핑 테이블에 저장

(b)

그림 13. 제공자 스키마 등록 절차 : (a) 등록 절차에 대한 흐름도, (b) 각 등록 절차에 대한 설명



(a)

① 질의가 입력 되면 도메인 정보(DomainID)와 질의(요소, 키워드)를 Query Processor로 보냄
② 입력된 도메인 정보를 Service Inference Engine으로 보내고, 질의는 Mapping Processing Module로 보냄
③ 도메인 정보를 Service Ontology에서 검색
④ 입력된 질의 와 확장된 도메인 정보를 Mapping Processing Module로 보냄
⑤ 검색 대상이 되는 매핑 문서와 표준 문서를 MDR에서 검색하기위한 정보를 MDR로 보냄
⑥ 검색된 MDR Table의 표준 정보와 제공자 스키마 테이블 정보를 Mapping Processing Module로 보냄
⑦ 매핑 처리 후 관련된 질의, 도메인 정보, 요소 정보를 Service Registry Module로 보냄
⑧ 해당 요소를 포함한 제공자의 정보를 Service Registry에서 검색
⑨ 질의, 도메인정보, 요소 정보, 검색된 제공자 정보를 Query Transformation Module로 보냄
⑩, ⑪ 질의를 요소 정보를 이용해 해당 제공자의 형식에 맞게 Xpath로 표현해 Query Result Integrator로 보냄
⑫ 질의를 제공자로 전송 후 결과를 받아 수합 후 Query Processor로 보냄
⑬ 질의 결과를 리스트 형태로 사용자에게 반환

(b)

그림 14. 매핑을 이용한 질의 처리 절차 : (a) 매핑을 이용한 질의 처리 흐름도, (b) 각 등록 절차에 대한 설명

4. 이미지 메타데이타 매핑을 이용한 확장 질의 처리 프로토타입

4.1 프로토타입 구현 환경 및 사용자 인터페이스

본 논문에서 제안한 메타데이타 매핑을 이용한 질의 처리 방안의 실험 평가를 위해 프로토타입을 구현하였다. 본 논문에서 구현한 프로토타입은 하나의 브로커와 8개의 제공자를 가정하고 구현하였으며 브로커 측의 OS는 MS windows 2003을 사용하였고 MS visual basic .NET 으로 프로그래밍하였다. 관련된 XML 문서 데이타를 저장하기 위한 DBMS로는 Oracle 9i를 사용하였다. 제공자 측의 프로그램은 MS windows XP에서 MS visual basic .NET으로 구현되었으며 DBMS 역시 Oracle 9i를 이용하였다. 브로커 측의 프로그램은 MDR에 표준 정보를 등록하기 위한 프로그램과 등록 프로그램, 제공자 측에서 메타데이타 매핑 처리를 수행하기 위한 매핑 프로그램, 그리고 질의를 입력받아 확장해서 처리하는 질의 처리 프로그램으로 구성된다. 표 2는 본 실험을 위하여 구축한 8개의 서로 다른 서비스 제공자 정보를 보여준다. Provider 3, 6, 7, 8은 동일한 Art 도메인에 대한 이미지를 제공하지만 이미지를 기술하는 메타데이타는 서로 상이하다.

본 논문에서 제안한 메타데이타 매핑은 사용자에 의해서 수행되며 매핑을 수행하는 사용자를 위한 사용자 인터페이스가 필요하다. 그림 15는 본 논문에서 제안한 메타데이타 매핑을 위한 사용자 인터페이스 화면이다. 그림 15에서와 같이 왼쪽 부분은 제공자 측의 사용자가 MDR내에 있는 VRA표준의 요소 와 속성들 각각의 설명정보를 제공하여 선택할 수 있게

표 2. 서비스 제공자 정보

제공자 명	도메인	메타데이타 표준	데이타 수
Provider 1	Art/Sculpture	CDWA	150
Provider 2	Photo/Vehicle	DC	200
Provider 3	Art/Painting	VRA	200
Provider 4	Photo/Plant	VRA	150
Provider 5	Art/Sculpture	비표준	100
Provider 6	Art/Painting	DC	300
Provider 7	Art/Painting	비표준	150
Provider 8	Art/Painting	CDWA	150

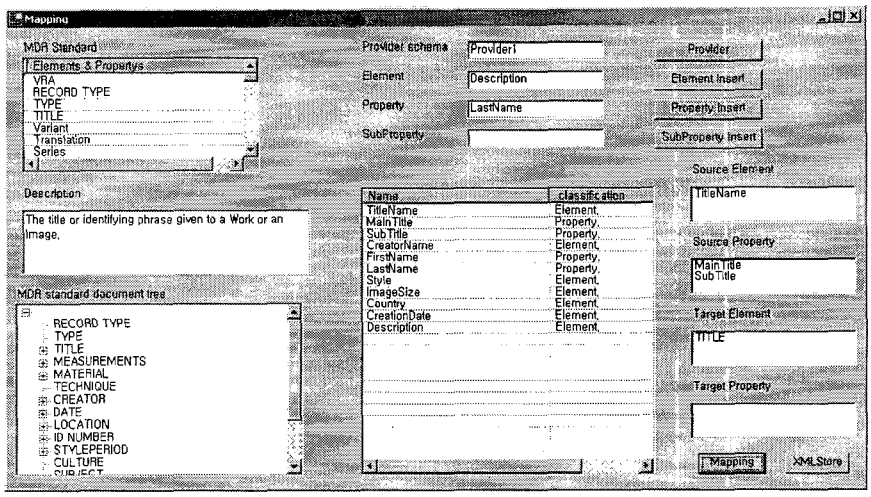
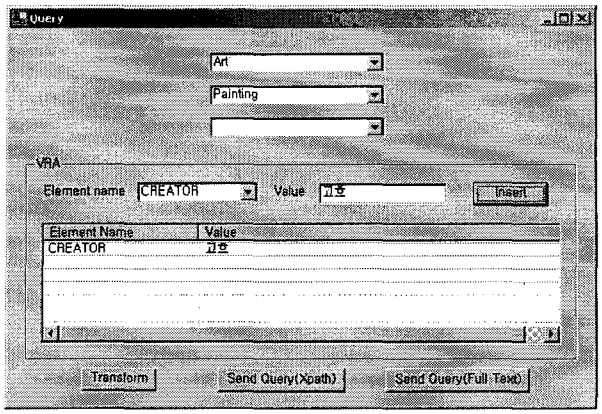


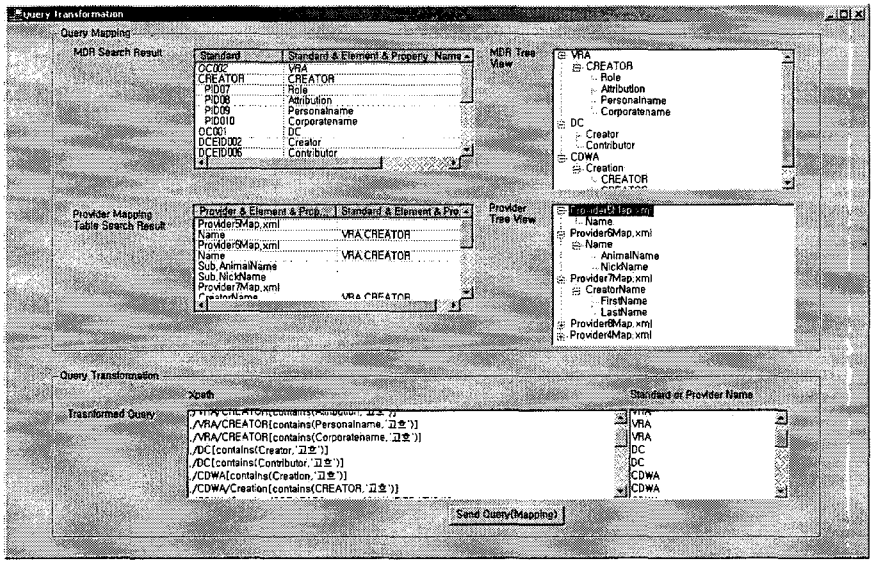
그림 15. 메타데이타 매핑 사용자 인터페이스

되어있고 VRA표준 요소들과 속성들에 대한 구조 정보는 왼쪽 하단의 MDR 표준 문서 트리에서 볼 수 있다. 중앙 부분은 제공자 스키마에 대한 부분이며 이미지 입력 폼에 대한 스키마 정보를 입력하고 입력된 정보를 리스트로 보여주고 있으며 매핑 시에는 제공자 스키마 리스트의 하나의 요소와 그 요소에 해당하는 속성들을 선택하고 매핑되는 MDR의 VRA표준 요소와 속성을 선택하면 오른쪽 Source 부분과 Target 부분으로 나뉘어 선택된 요소들을 보여주고 매핑 버튼을 누르면 XML방식으로 매핑문서가 만들어지며 이러한 방식으로 모든 요소들을 추가하여 매핑문서를 생성한다.

질의 처리를 수행하기 위해서는 질의를 입력 받기 위한 사용자 인터페이스와 입력된 질의에 대한 매핑을 수행하여 확장된 질의문을 생성하여 서비스 제공자에게 전송하는 사용자 인터페이스가 필요하다. 그림 16은 본 논문에서 제안한 질의 확장을 구현한 화면이며 질의를 입력하기 위한 화면과 입력된 질의를 제안된 방법으로 확장한 화면이다. 확장된 요소들은 요소들의 형태에 따라서 XPath로 변환된다. 그림 16 (a)에서는 검색 하려는 대상 도메인을 선택하고 요소명과 값을 선택하면 그림 16 (b)에서와 같이 매핑한 결과를 출력하고 각 요소의 형태에 따라서 XPath로 변환한다. 그림 16 (a)에서 입력된 질의는 'CREATOR' 요소의 값이 '고흐'인 질의이며 입력된 질의는 그림 16 (b)에서와 같이 MDR에서 각 표준의 관련 요소로 확장되며 도메인이 일치하는 제공자의 매핑 문서에서도 관련 요소로 확장된다. 확장된 질의들은 해당 요소의 형태에 따라서 XPath로 변환되어 각 제공자로 전송된다. 예를 들어, MDR 내의 VRA 표준에서



(a)



(b)

그림 16. 질의 입력 및 확장 변환 : (a)도메인 선택 및 질의 입력, (b) 질의 확장 및 변환

는 'CREATOR'요소의 여러 속성으로 매핑 되는데 'PersonalName'속성의 경우 XPath로 표현하면 './VRA/CREATOR[PersonalName=고흐]'가 된다.

4.2 검색 성능 평가

본 논문에서 제안한 메타데이타 매핑을 이용한 확장 질의 처리 방안의 유용성을 보이기 위하여 매핑 기법을 적용하지 않은 질의 처리 방법과 일반적인 텍스트 검색에 이용되는 전문 검색 질의 처리 방식과의 비교를 수행하였다. 검색 결과에 대한 성능 평가는 일반적으로 많이 사용되는 Recall과 Precision을 이용하여 측정하였으며 각각에 대한 비율 계산은 식 1과 같다.

$$Recall = \frac{R_r}{T} \qquad \Pr ecision = \frac{R_r}{T_r} \qquad \text{(식 1)}$$

여기서 T는 검색 대상 데이터베이스에서 주어진 질의와 관련된 전체 항목의 개수를 의미하며, T_r은 검색된 전체 항목의 수, R_r은 검색된 항목 중에서 질의와 관련된 항목의 수를 말한다. 표 3은 본 실험에서 사용된 질의 유형에 따른 3가지 질의 처리 방식의 검색 결과를 보여준다. 질의 유형의 Q1은 질의 요소의 수가 1개, Q2는 2개, Q3는 3개인 질의 패턴을 말하

표 3. 질의처리방식에 대한 검색 결과 비교

질의유형	질의 예	질의처리방식	R_r	T_r	T	Recall	Precision
Q1	creator='gogh'	NQ	20	20	60	33	100
		FQ	60	98	60	100	61
		EQ	60	60	60	100	100
Q2	creator='gogh' date=1888	NQ	8	8	23	35	100
		FQ	23	30	23	100	77
		EQ	23	23	23	100	100
Q3	creator='gogh' date=1888 title='Sun Flower'	NQ	1	1	5	33	100
		FQ	5	6	5	100	83
		EQ	5	5	5	100	100

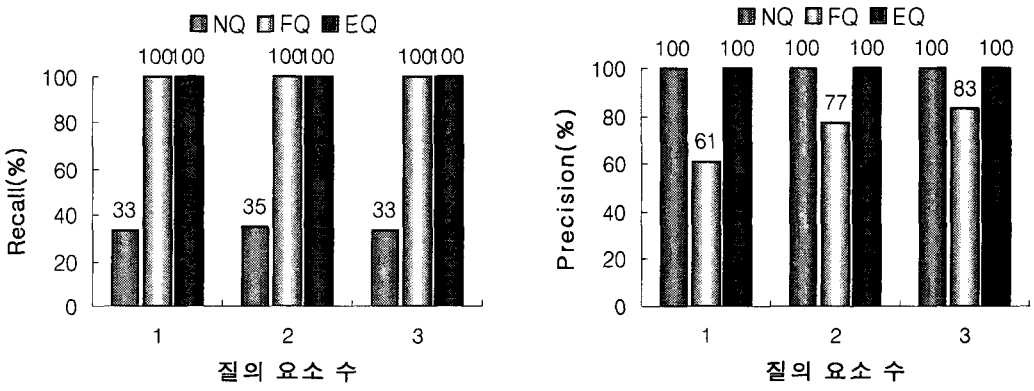


그림 17. 질의처리방식에 대한 검색 결과그래프 : (a)질의처리방식에 대한 Recall 비교, (b)질의처리방식에 대한 Precision 비교

며 질의처리방식의 NQ, FQ, EQ는 각각 확장되지 않은 질의(NQ: none extended query), 전문검색질의(FQ: full-text search query), 매핑을 이용한 확장된 질의(EQ: extended query)를 의미한다.

그림 17은 표 3의 3가지 질의처리 방식의 검색 결과에 대한 Recall과 Precision 비율을 그래프로 도식화하여 표현한 것이다. 질의 확장을 하지 않은 경우(NQ)의 검색 결과는 확장을 한 경우(EQ)에 비해 Precision은 차이가 없으나 Recall 비율이 현저히 떨어짐을 알 수 있다. Recall 비율이 낮다는 말은 주어진 질의와 관련된 이미지를 제대로 검색하지 못하였음을 의미한다. 이에 반해 전문 검색 질의(FQ)를 수행한 경우는 Recall 비율은 높지만 상대적으로 Precision 비율이 낮음을 알 수 있다. 즉, 검색 결과 주어진 질의와 관련된 대부분의 이미지를 검색하지만 질의와 무관한 이미지도 검색 결과 상당수 포함되어 있음을 뜻한다. 결국 메타데이타 매핑을 이용한 확장 질의 수행 결과에 비해 Precision 비율이 현저히 떨어짐을 확인할 수 있다. 또한 전문 검색 질의의 경우에는 질의 요소 수의 증가에 비례하여 Precision 비율이 증가하고 있음을 알 수 있다. 그 이유는 질의 요소 수가 늘어날수록 질의를 만족시키는 키워드의 수도 증가하며 이들 키워드가 질의와 무관한 데이터를 걸러내는 제약 조건으로 작용하기 때문이다. 따라서 본 논문에서 제안한 질의 확장방법을 이용하면 Recall과 Precision 비율을 높여 전체 검색 시스템의 성능을 향상시킬 수 있다.

4.3 질의 수행 결과

그림 18은 본 논문에서 제안한 방법을 이용하여 확장 질의를 수행한 경우와 확장하지 않고 질의를 수행한 경우의 결과를 보여준다. VRA 표준을 따르는 Provider 3과 비표준 메타데이타 스키마를 가지는 Provider 7에 각각 그림 제목이 모나리자인 이미지가 하나씩 저장되어 있을 경우, Provider 3에서는 그림

제목이 '<TITLE><Variant> 제목</Variant></TITLE>' 형식의 태그로 저장되며, Provider 7에서는 '<TitleName><MainTitle> 제목 </MainTitle></TitleName>' 형태의 다른 구조로 저장된다. 따라서 메타데이타 매핑이 수행되지 않은 경우에는 사용자 질의 폼이 VRA 표준 요소를 기준으로 만들어 졌기 때문에 두 제공자 모두에게 './VRA/TITLE [contains(Variant,'모나리자')]'와 같은 VRA 표준 형식의 질의가 전송된다. 그 결과 그림 18(a)에서와 같이 VRA 표준에 맞는 Provider 3의 모나리자 이미지만 검색된다. 그러나 메타데이타 매핑 수행 결과, 질의가 각 제공자의 스키마 형식에 적합하게 확장된 경우에는 Provider 3에는 VRA 표준 형식의 질의가 Provider 4에는 './Provider7/TitleName [contains (MainTitle,'모나리자')]'와 같은 비표준 메타데이타 스키마 형식으로 질의가 확장되어 각각 전송된다. 그림 18의 (b)는 확장 질의 처리 결과 Provider 3과 Provider 7로부터 제목이 모나리자인 이미지가 검색되었음을 보여준다. 즉, 본 논문에서 제안한 확장된 XPath 질의를 사용하면 스키마 구조의 이질성으로 인한 문제를 해결함으로써 Recall 비율을 높여 전체 시스템의 검색 성능을 향상시킴을 알 수 있다.

기존에 텍스트 검색에 많이 활용되는 전문검색(full-text search) 방법과 제안된 방법의 경우에도 정확성 면에서 많은 차이가 있을 수 있다. 그림 19는 전문 검색을 사용하여 질의를 수행한 경우와 본 논문에서 제안한 확장된 XPath 질의를 수행한 결과에 대한 화면이다.

그림의 제목이 자화상인 이미지를 찾는 질의가 주어질 때, 그림 19의 (a)와 같이 전문 검색의 경우에도 이미지 메타데이타에 대한 XML 문서로부터 자화상이라는 질의 값을 포함하는 문서를 찾을 수 있다. 그러나 그림 19의 (b)와 같이 다른 항목에 자화상이란 단어가 포함되어 있을 경우 제목이 '타히티의 여인' 이라는 질의와는 무관한 잘못된 그림을 검색할 수도 있다. 이처럼 전체 문서에 대한 검색은 질의와 무관한 결과를 다수 포함함으로써 전체 시스템의 검색 효율(정확도)이 떨어지는 단점이 있다. 반면 본 논문에서 제안한 확장 질의 방법을 사용하면 이러한 문제점을 해결할 수 있다. 그림 19의 (c)는 질의 확장을 이용한 Xpath 질의에 대한 검색 결과를 보여준다.

5. 결 론

메타데이타의 상호 운용성과 정확한 검색은 이종의 플랫폼으로 구성되어 있는 분산 환경에서 중요한 사항이다. 웹 문서의 표준인 XML은 서로 다른 시스템 환경에서 만들어진 정보를 효율적으로 교환하는 표현 방법이며, MDR은 이질의 이미지 메타데이타를 상호공유하기 위해서 관련된 표준안을 저장 관리하기 위한 방법이다.

본 논문에서는 이질의 분산 환경에서 이미지 메타데이타에 대한 정확한 검색을 수행하기 위해서 비표준 지역 스키마와 MDR내의 기준 표준인 VRA 표준

(a) (b)

그림 18. 질의 확장 여부에 따른 검색 결과 : (a) 확장하지 않은 XPath 질의 결과, (b) 확장한 XPath 질의 결과

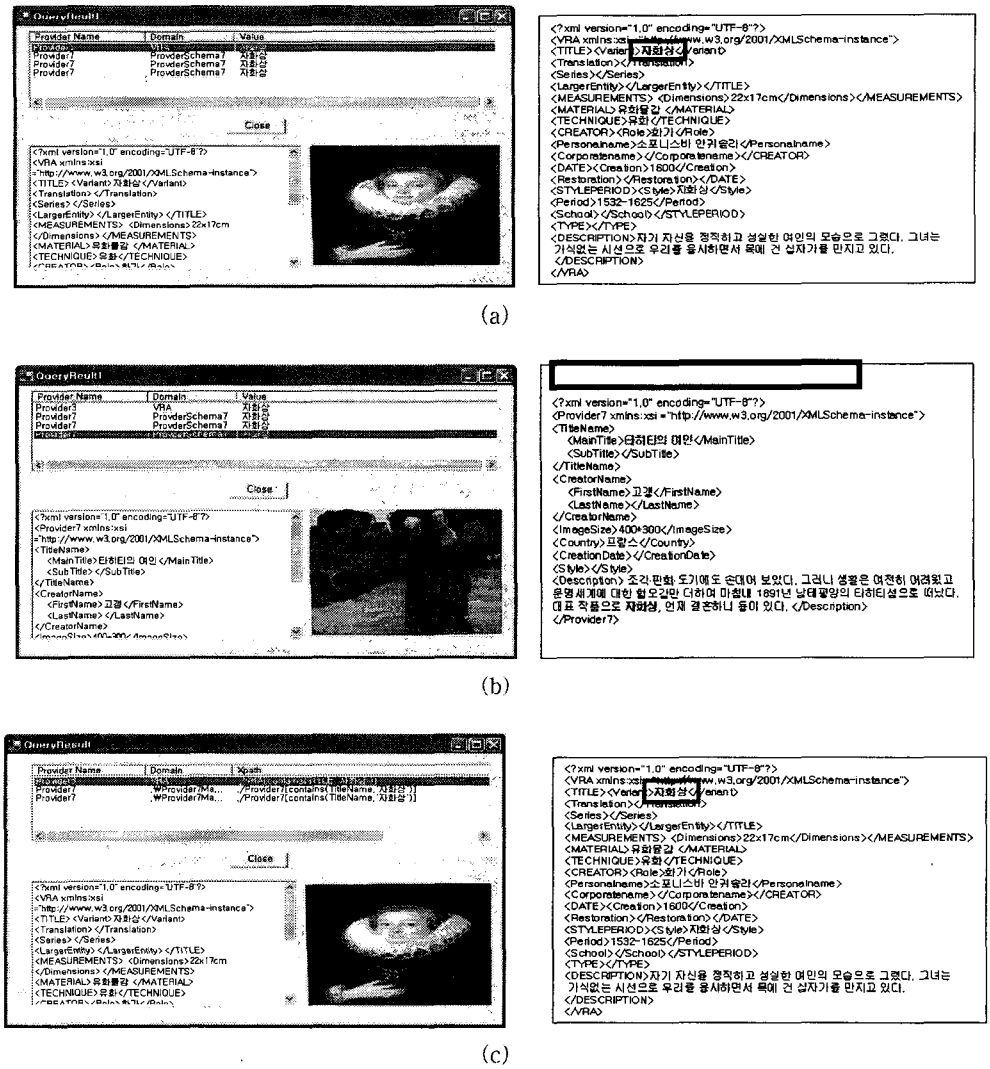


그림 19. 전문 검색과 제안된 방법의 질의 결과 : (a) 전문 검색에서 TITLE 요소에 질의 값을 포함한 결과, (b) 전문 검색에서 다른 항목에 질의 값을 포함한 결과, (c) 질의 확장을 이용한 XPath 질의 결과

과 매핑을 수행하는 방안을 기술하였다. 또한 불필요한 매핑 연산을 줄이고 새로운 표준을 쉽게 추가하기 위해 MDR 내에서 VRA 기준 표준과 다른 표준안 간의 매핑 방법에 대해서 설명하였다. 제안된 매핑 방법을 이용한 결과는 XML문서 형태로 표현되어 저장되며 질의 시에 질의를 확장하는데 활용된다. 확장된 질의는 해당 정보를 제공하는 각 제공자로 전송되기 위해 각 제공자의 메타데이타 구조에 적합한 XPath 형태로 변환되어 전송된다. 또한, 본 논문에서 제안한 이미지 메타데이타 매핑을 이용한 확장 질의 처리의 유용성을 보이기 위해 매핑 및 질의 확장 프로토타입을 구현하였다. 또한 검색 성능 평가를 위하여 질의처리 결과에 대한 Recall과 Precision 비율을 측정하였다. 실험 결과 이미지 메타데이타 매핑을 이용한 질의 확장을 통해 서로 다른 메타데이타 스키마를 가지는 제공자로부터 주어진 질의에 대한 정확한 검색 결과를 얻을 수 있으며, Recall과 Precision 비율을 높임으로써 전체 시스템의 검색 성능을 향상시킨다.

본 논문에서 제안한 매핑 방안과 질의 확장 방법은 분산 웹 서비스 환경에서 정확한 검색을 위한 요소 확장에 이용될 수 있다. 향후 연구 과제로는 비표

준 또는 다른 표준에서는 정의되지만 VRA에서 정의되지 않은 요소들에 대한 매핑 및 처리 방안에 대한 추가적인 연구가 필요하며 다양한 형태의 질의를 지원하는 방안과 웹 환경에서의 질의 시스템을 구현하는 방안 등을 들 수 있다.

참 고 문 헌

[1] Doublin Core Organization, Dublin Core Metadata Initiative, http://dublincore.org/index.shtml/.

[2] Visual Resources Association Data Standards Committe, VRA Core Categories Ver. 3.0, http://www.vraweb.org/vracore3.htm.

[3] ISO IEC JTC1/SC29/WG11, "Overview of the MPEG-7Standard," MPEG2001/N4031, Singapore, March 2001.

[4] Art Information Task Force(AITF), Categories for the Description of Works of Art, http://www.getty.edu/research/institute/standards /cdwa/index.html.

[5] Rachel H. and Harry W., "A Metadata Registry for the Semantic Web," *D-Lib Magazine*, Vol.8(5), 2002.

[6] ISO/IEC 11179, Information Technology Metadata Registries, Part 1-6, 2003.

[7] Yunmook Nah, Bogju Lee, and Jungsun Kim, "Visual Media Retrieval Framework using Web Service," *LNCS 3597*, pp.361-365, 2005.

[8] Jane Hunter, "MetaNet - A Metadata Term Thesaurus to Enable Semantic Interoperability Between Metadata Domains," *Journal of Digital Information*, 1(8), 2001.

[9] Jon D. et. al., *PROFESSIONAL XML schemas*, Information Publishing Group, 2002.

[10] 김미경, "메타데이터 레지스트리를 이용한 지역 데이터베이스 간 질의처리 방법," 고려대 석사학위논문, 2001.

[11] Ayesh M., Design XML schemas using UML, http://www-128.ibm.com/developerworks/xml/library/x-umlschem/index.html.

[12] 나연묵, "의미 기반 검색을 위한 이미지 내용 모델링," *정보과학회논문지: 데이타베이스*, 30권, 제2호, pp. 145-156, 2003년 4월.

[13] 나홍석, "XML 문서 공유를 위한 MDR 기반의 메타데이터 상호운용성 프레임워크," 고려대 박사학위논문, 2004.

권 은 영

2002년 단국대학교 전자계산학과 (공학사)
2005년 단국대학교 대학원 전자컴퓨터공학과 (공학석사)
2005년~현재 중외정보기술 연구원
관심 분야 : 데이타베이스, 멀티미디어 정보검색, 의료정보시스템

안 철 범

1993년 *성균관대학교* 생물학과 (이학사)
2001년 단국대학교 대학원 컴퓨터공학과 (공학석사)
2001년~현재 단국대학교 대학원 전자컴퓨터공학과 (박사수료)
관심 분야 : 데이타베이스, 멀티미디어 정보검색, 시맨틱 웹, U-Healthcare

나 연 묵

1986년 서울대학교 컴퓨터공학과 (공학사)
1988년 서울대학교 대학원 컴퓨터공학과 (공학석사)
1993년 서울대학교 대학원 컴퓨터공학과 (공학박사)
1993년~현재 단국대학교 전자컴퓨터공학부 컴퓨터공학전공 교수
1991년 IBM T. J. Watson 연구소 객원연구원
2001년~2002년 *University of California, Irvine* 객원교수
관심 분야 : 데이타베이스, 데이타 모델링, 데이타베이스 설계, 객체지향 데이타베이스, 멀티미디어 데이타베이스, 멀티미디어 정보 검색, 시맨틱 웹, 이동객체 데이타베이스, 대용량 분산 데이타베이스